ELECTROMAGNETIC TARGET RECOGNITION FOR LOSSY AND
DISPERSIVE DIELECTRIC OBJECTS: APPLICATIONS TO BREAST TISSUE
CLASSIFICATION AND TUMOR DETECTION PROBLEM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BAŞAK IŞIK BARUT

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JULY 2014

Approval of the thesis:

ELECTROMAGNETIC TARGET RECOGNITION FOR LOSSY AND DISPERSIVE DIELECTRIC OBJECTS: APPLICATIONS TO BREAST TISSUE CLASSIFICATION AND TUMOR DETECTION PROBLEM

submitted by **BAŞAK IŞIK BARUT** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen   _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Gönül Turhan Sayan   _____
Head of Department, **Electrical and Electronics Engineering**

Prof. Dr. Gönül Turhan Sayan   _____
Supervisor, **Electrical and Electronics Eng. Dept., METU**

Examining Committee Members:

Prof. Dr. Mustafa Kuzuoğlu   _____
Electrical and Electronics Engineering Dept., METU

Prof. Dr. Gönül Turhan Sayan   _____
Electrical and Electronics Engineering Dept., METU

Prof. Dr. Kemal Leblebicioğlu   _____
Electrical and Electronics Engineering Dept., METU

Prof. Dr. Gülbin Dural   _____
Electrical and Electronics Engineering Dept., METU

Emre Ergin, M.Sc. EE   _____
Turkish Aerospace Industries, TAI

Date: 02.07.2014

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited referenced all material and results that are not original to this work.

Name, Last name: BAŞAK IŞIK, BARUT

Signature:

iv

ABSTRACT

ELECTROMAGNETIC TARGET RECOGNITION FOR LOSSY AND DISPERSIVE DIELECTRIC OBJECTS: APPLICATIONS TO BREAST TISSUE CLASSIFICATION AND TUMOR DETECTION PROBLEM

BARUT, BAŞAK IŞIK

M.S., Department of Electrical and Electronics Engineering

Supervisor: Prof. Dr. Gönül TURHAN SAYAN

July 2014, 110 pages

The aim of this thesis is to understand the fundamental concepts behind two different electromagnetic target recognition (EMTR) techniques and to extend their applications to the problem of classifying lossy and dispersive objects embedded either in air or in another lossy and dispersive medium. The EMTR techniques, which use either Wigner Distribution-Principal Component Analysis (WD-PCA) approach or Multiple Signal Classification (MUSIC) Algorithm approach, are both based on the Singularity Expansion Method (SEM). They use wide-band scattered data in resonance region to extract object features which are related to complex natural resonance (CNR) frequencies. Signal processing tools such as WD-PCA and MUSIC algorithm are used to extract object features. These two EMTR techniques are already applied in literature to the classification of targets made of perfect conductors and/or perfect dielectrics with success.

In this thesis, application of aforementioned techniques will be studied for two different but closely related problems. First, classification of different lossy and dispersive breast tissue samples will be demonstrated when they are embedded in air. Secondly, the problem of breast tumor detection will be illustrated where both the object (tumor) and the surrounding medium (breast tissue) are lossy and dispersive. The Cole-Cole parameters will be used to model realistic tumor tissues as well as realistic breast tissues having different fat (adipose) content. Difficulty of recognition of tumors embedded in different types of breast tissues will be investigated for different tumor sizes by designing proper target classifiers using both EMTR techniques.

Keywords: Electromagnetic target recognition, lossy and dispersive media, WD-PCA method , MUSIC algorithm, breast tumor detection.

ÖZ

KAYIPLI VE DAĞITICI DİELEKTRİK CİSİMLER İÇİN ELEKTROMANYETİK HEDEF TANIMA: MEME DOKUSU SINIFLANDIRMA VE TÜMÖR TESPİT PROBLEMİNE UYGULAMALAR

BARUT, BAŞAK IŞIK

Yüksek Lisans, Elektrik Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Gönül TURHAN SAYAN

Temmuz 2014, 110 sayfa

Bu tezin amacı, iki ayrı elektromanyetik hedef tanıma (EMHT) tekniğinin temel kavramlarını anlamak ve bu tekniklerin uygulamalarını, hava içerisinde ya da kayıplı ve dağıtıcı bir başka ortam içerisinde gömülü olan ve kendisi de kayıplı ve dağıtıcı olan dielektrik cisimlerin tespiti problemine genişletmektir. WD-PCA ve MUSIC Algoritma yöntemlerini kullanan bu EMHT tekniklerinin ikisi de Tekillik Açılım Metoduna (SEM) dayalıdır. Bu teknikler, karmaşık değerlikli doğal rezonans (CNR) frekans bazlı hedef öznitelikleri elde etmek için rezonans bölgesinde geniş bantlı saçınım verilerini WD-PCA ve MUSIC algoritması gibi araçlar kullanarak işlerler. Her iki EMHT tekniği de, mevcut literatürde, mükemmel iletkenler ve/veya mükemmel dielektriklerden yapılmış hedeflerin sınıflandırılması problemlerine başarı ile uygulanmıştır.

Bu tezde, yukarıda bahsedilen tekniklerin iki farklı ancak yakından ilişkili probleme uygulanması üzerinde çalışılmıştır. Önce hava içine gömülü olan, kayıplı ve dağıtıcı

özellikteki farklı meme doku örneklerinin sınıflandırılması gösterilecektir. İkinci olarak, hem objenin (tümör) hem de çevresindeki ortamın (meme dokusu) kayıplı ve dağıtıcı olduğu, meme tümörü tespit problemi incelenecektir. Gerçekçi tümör dokularının ve farklı yağ doku oranlarına sahip gerçekçi meme dokularının modellenmesi için Cole-Cole parametreleri kullanılmıştır. Farklı tipteki meme dokuları içerisine yerleşmiş farklı büyüklükteki tümörlerin tanınmasındaki zorluk derecesi, her iki EMHT tekniği de kullanılarak tasarlanan uygun hedef sınıflandırıcılar yardımı ile incelenecektir.

Anahtar Sözcükler: Elektromanyetik hedef sınıflandırma, kayıplı ve dağıtıcı ortam, WD-PCA metod, MUSIC algoritması, meme tümörü tespiti.

To My Family

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Electromagnetic target recognition (EMTR) is an interdisciplinary area of research that makes use of electromagnetic theory, system theory and signal processing techniques. EMTR techniques to be used in a given application vary within a broad range depending upon the physical and geometrical properties of the object to be recognized, measurement setup and frequency band available for data collection. Additionally, noise conditions as well as special constraints of the problem such as the need for real-time data processing, decision speed or data storage capacity are important considerations affecting the choice of a proper EMTR technique.

There are three basic frequency regions calling for completely different EMTR techniques, which are the Rayleigh region in the low frequency end, the geometrical optics region in the high frequency end, and the resonance region in between for which the largest line-of-sight dimension D of the object must be in the range from $0.1\,\lambda$ to $10\,\lambda$, as a rule of thumb, where λ is the wavelength. Atmospheric scattering from tiny particles is a typical example for Rayleigh scattering [1] where wavelength is much larger than the sizes of scattering particles. On the other extreme, the wavelength is much smaller than the overall size of the scattering objects in the geometrical optics region applications as in the case of X-band radar signals scattered by an airplane. In the resonance region, on the other hand, target dimensions are comparable with the wavelength, and the scattered data provide detailed information about the material composition and geometry of the objects.

Polarization and aspect dependency of electromagnetic scattered data is a well-known source of difficulty which may be neglected in the Rayleigh region, but becomes a very serious problem in the geometrical optics regime. In the resonance region, on the other hand, scattered signals are moderately dependent on the choice of aspect and polarization [2, 3]. Therefore, extraction of fused target features under different excitation conditions is a very important intermediate step in classifier design in the resonance region.

A physics-based model of the electromagnetic system function of an object can be obtained in resonance region by using the Singularity Expansion (SEM) method. The SEM models the late-time response (natural response) of the target as the superposition of damped sinusoidal signals associated with complex natural resonance (CNR) frequencies of the target [4-10]. The infinite set of CNR frequencies of a target (also called target's system poles, or simply target poles) defines a target uniquely as they are closely related to the geometrical and physical properties of the object in an aspect and polarization independent manner. In a practical case, scattered electromagnetic signals become available over a finite measurement bandwidth. Therefore, only a finite subset of target poles becomes available for target recognition in real world applications. Calculation of system poles directly using scattered data is known to be highly sensitive to noise [11, 12]. Therefore, instead of directly computing system poles of an object, computation of pole-related object features is more useful in practical cases.

The K-Pulse (Kill-Pulse) method was the first example in literature making use of a pole-related object feature extraction in late 1970's. In his 1981 paper [13], E.M. Kennaugh defined the K-pulse as a time-limited signal of a target whose Laplace spectrum zeros are the same as target's system poles.

Hence, due to pole-zero cancellations, when a target is excited by its own K-pulse, the resulting response of the target becomes time-limited. There are many resonance region methods reported in literature inspired by the K-pulse technique and by the SEM theory. Part of those publications can be classified as "K-pulse based methods" [14-18,47,48] and some others can be grouped under "E-pulse based methods" [19-21] where both methods actually have the same theoretical foundation.

Use of time-frequency representation (TFR) techniques is a relatively new approach in EMTR research with results reported in 2000's [22-25]. Short-time Fourier transform, Gabor transform, Wigner-Ville distribution, Page Distribution, Wavelets are well known TFR techniques [26, 27]. The EMTR approach, which combines the Wigner distribution (WD) technique with the principal component analysis (PCA) [28] for feature extraction, is called the WD-PCA technique as originally introduced by Gönül Turhan-Sayan in 2005 [29]. Shortly after, in 2009, another resonance region EMTR technique named after the multiple signal classification (MUSIC) algorithm [45, 46] is established by Mustafa Seçmen and Gönül Turhan-Sayan [30]. Both of these EMTR techniques have been successfully demonstrated to classify conducting, dielectric or dielectric coated conducting targets with high accuracy [25, 29-38, 44] under both noise-free and noisy conditions. Also, there have been other EMTR applications combining the CNR based target recognition approach with pattern recognition methods [39], neural networks [40-43], statistical techniques [44] or global optimization techniques [47, 48].

In all these applications, targets are assumed to be made of linear, isotropic, homogeneous, loss-free and dispersion-free materials. In the real-world, however, there is no way of avoiding lossy and/or dispersive behaviors of materials. Not only the objects/targets but also the media enclosing them may be lossy and dispersive.

Detection of explosive mines buried in lossy and dispersive soil and the detection of tumors within lossy and dispersive human tissue are well known, practically important and highly challenging applications that can be solved by resonance region EMTR techniques [49-53, 57].

The aim of this thesis is to understand the fundamental concepts behind the WD-PCA and MUSIC algorithm based electromagnetic target recognition techniques mentioned earlier and to extend their applications to the problem of detecting lossy and dispersive objects embedded in a different dispersive and lossy medium, in general. The WD-PCA based EMTR technique and the MUSIC Algorithm based EMTR technique are both based on the Singularity Expansion Method (SEM) and use wide-band scattered data in resonance region to extract complex natural resonance (CNR) related target features by using the signal processing tools such as Wigner distribution (WD), Principal Component Analysis (PCA) and Multiple Signal Classification (MUSIC) algorithm.

In the rest of this thesis, application of aforementioned techniques will be presented for two different but closely related problems. First, classification of different lossy and dispersive breast tissue samples will be demonstrated when they are embedded in air. Secondly, the problem of breast tumor detection will be illustrated where both the object (tumor) and the surrounding medium (breast tissue) are lossy and dispersive. The Cole-Cole parameters will be used to model realistic tumor tissues as well as realistic breast tissues having different fat (adipose) content. Difficulty of detecting tumors embedded in different types of breast tissue will be investigated for different tumor sizes by designing proper target classifiers using both EMTR techniques.

Frequency dependent complex valued permittivity functions of realistic biological tissues (used as targets in the classification design problems of this thesis) are described by the experimentally extracted Cole-Cole parameters [54, 55]. Part of the

permittivity data belongs to healthy breast tissues with different adipose (fat) content (adipose, glandular and fiberconnective type breast tissues) which were obtained during breast reduction and cancer surgeries. Also, data characterizing malignant (invasive or non-invasive ductal and lobular carcinomas) tissues were obtained during cancer surgeries. It is reported by a completed large-scale experimental study that the so called "Cole-Cole model" is useful to define electrical properties of these biological tissues and Debye model offers an excellent fit to Cole-Cole data [56]. As reported in literature [50, 54-56, 58], Cole-Cole model is widely used to describe frequency dependent dielectric properties of the materials including the breast tissues with or without tumors.

In the application parts of the thesis, different object classifiers are designed by using both WD-PCA method and MUSIC algorithm method to classify healthy breast tissues and to detect tumors of different sizes in varying types of breast tissues. Performances of these two target classifier design techniques are investigated on the comparative basis.

The organization of the thesis is as follows:

Chapter 2 presents the theoretical background about permittivity modeling for lossy and dispersive dielectric materials. Debye and Cole-Cole models that are used to define the permittivity characteristics of such dielectric materials are outlined. Then, complex natural frequency (CNR) mechanism and SEM theory are shortly discussed for target recognition problems in resonance region. Next, computation of electromagnetic field scattered by lossy/dispersive dielectric objects is presented using analytical formulations for the spherical target geometry. Then, similar computations are described for lossy/dispersive dielectric spheres coated by lossy/dispersive dielectric spherical shells. Finally, theory and design steps of the MUSIC algorithm based and the WD-PCA based electromagnetic target recognition techniques are outlined to be used in classifier design.

Chapter 3 presents the design steps and performance testing results of the classifiers designed by using both MUSIC algorithm based and the WD-PCA based techniques. Classifier accuracies obtained by both methods are compared.

The concluding remarks and suggestions for future studies are given in Chapter 4.

Finally, Appendix contains a sample MATLAB code written for real-time decision testing as a part of the WD-PCA algorithm based classifier design for lossy dielectric spheres.

CHAPTER 2

THEORETICAL BACKGROUND

2.1. Review of Permittivity Concept for Lossy Dielectrics

The phenomena of electromagnetic (EM) wave propagation is described by Maxwell's equations under proper boundary conditions and for pertinent medium parameters. For linear, non-magnetic media, in particular, permittivity (ε) is the most important medium parameter which is defined in phasor domain by the well known relation $D = \varepsilon\overline{E}$ where $\overline{E}$ is the electric field intensity vector and $\overline{D}$ is the electric field density (electric displacement) vector. Permittivity of a lossy medium is expressed as complex valued quantity (ε_c) given as

$$\varepsilon_c = \varepsilon_0 \varepsilon_r = \varepsilon' - j\varepsilon'' \tag{1}$$

where ε_0 is the permittivity of free space, ε_r is the relative permittivity, the real part ε' is dielectric constant and the imaginary part ε'' is called as dielectric loss factor, While ε' is related to stored electric energy, ε'' is associated with material losses.

Measure of material losses can be expressed in term of loss tangent ($\tan\delta$) that is defined starting with the phasor domain Maxwell's equation

$$\overline{\nabla} \times \overline{H} = \overline{J} + j\omega\varepsilon\overline{E} = \sigma\overline{E} + j\omega\varepsilon\overline{E} = j\omega(\varepsilon - j\frac{\sigma}{\omega})\overline{E} \tag{2.a}$$

in a linear, source-free, lossy medium of conductivity σ and dielectric constant ε.

The term in the parenthesis on the right hand side is defined as complex permittivity (ε_c) as

$$\varepsilon_c = \varepsilon - j\frac{\sigma}{\omega} = \varepsilon' - j\varepsilon'' \qquad (2.b)$$

Hence,

$$\varepsilon' = \varepsilon \qquad (2.c)$$

$$\varepsilon'' = \frac{\sigma}{\omega} \qquad (2.d)$$

and

$$\tan\delta = \frac{|J_c|}{|J_d|} = \frac{|\sigma\overline{E}|}{|j\omega\varepsilon\overline{E}|} = \frac{\sigma}{\omega\varepsilon} = \frac{\sigma/\omega}{\varepsilon} = \frac{\varepsilon''}{\varepsilon'} \qquad (2.e)$$

are defined where J_c is the conduction current density phasor and J_d is the displacement current density phasor. The loss tangent becomes larger as the dielectric material becomes more lossy.

In a lossy medium, propagation constant of an EM wave is given as $\gamma = \alpha + j\beta = j\omega\sqrt{\mu\varepsilon_c}$, where α is the attenuation constant, β is the phase constant and μ is permeability of the propagation medium which is equal to $\mu_0 = 4\pi10^{-7}\,Hm^{-1}$ for nonmagnetic materials.

Using (1) and (2.e) ε_c and γ can be expressed as

$$\varepsilon_c = \varepsilon'(1 - j\tan\delta) \qquad (3)$$

and

$$\gamma = j\omega\sqrt{\mu\varepsilon'(1 - j\tan\delta)} \qquad (4)$$

For a low-loss dielectric material, for example, $\tan\delta$ is much smaller than unity, therefore equation (4) can be approximated as

$$\gamma \cong j\omega\sqrt{\mu\varepsilon'} + \omega(\tan\delta)\sqrt{\mu\varepsilon'}/2 \tag{5}$$

after writing the Binomial expansion of the term $\sqrt{1-j\tan\delta}$ and keeping the first two terms of the expansion as $\tan\delta << 1$ [60]. Real part of propagation constant γ in equation (5) is the attenuation constant (α) which is responsible of the exponential decay of the field strength in the direction of propagation.

Another important parameter of EM wave propagation is the skin depth (penetration depth), that is defined as the distance along the propagation direction at which the field strength is reduced to (1/e) of its initial value. In general, the skin depth equals to the inverse of attenuation constant α.

2.2. Theoretical Model of Dielectric Relaxation

Ability of polarized molecules to align with the applied E-field is described by the relaxation phenomena, which has a significant effect in determining the permittivity of a given material.

After some mathematical derivations based on relaxation phenomena, Debye relaxation equation is obtained as [60]

$$\varepsilon_r = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + j\omega\tau} \tag{6}$$

which exhibits a relaxation frequency centered at $f_c = 1/2\pi\tau$. The parameters ε_s and ε_∞ refer to the permittivity values well below and well above f_c, respectively. The τ is called as relaxation time constant.

This model can be extended to include a static conductivity term $\sigma_s = j\omega\varepsilon_0\varepsilon_s$.

Then, the parameter (ε_s) for static relative permittivity is defined as

$$\varepsilon_s = \sigma_s / j\omega\varepsilon_0 = -j\sigma_s / \omega\varepsilon_0 \tag{7}$$

Inserting (7) in (6) results in

$$\varepsilon_r = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + j\omega\tau} - j\frac{\sigma_s}{\omega\varepsilon_0} \tag{8}$$

Or, ε_r can be expressed as

$$\varepsilon_r = \varepsilon_\infty + \frac{(\varepsilon_s - \varepsilon_\infty)(1 - j\omega\tau)}{1 + \omega^2\tau^2} - j\frac{\sigma_s}{\omega\varepsilon_0} \tag{9}$$

The empirical function of permittivity given below is known as Cole-Cole model

$$\varepsilon(\omega) = \varepsilon^{'}(w) - j\varepsilon^{''}(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + (j\omega\tau)^{1-\alpha}} + \frac{\sigma_s}{j\omega\varepsilon_0} \tag{10}$$

where ε_∞ is the relative permittivity for infinite value of frequency, $\Delta\varepsilon$ is difference between the infinite and static relative permittivity, $\tau(s)$ is the relaxation time constant, $\sigma_s(S/m)$ is the static conductivity term. The value of the parameter α is empirically determined as a correction factor.

When $\alpha = 0$, relative permittivity can be expressed as

$$\varepsilon_r = \varepsilon_\infty + \frac{(\varepsilon_s - \varepsilon_\infty)}{(1 + jf / f_c)} - j\frac{\sigma_s}{\omega\varepsilon_0} \tag{11}$$

$f = \omega / 2\pi$ and $f_c = 1/2\pi\tau$ then $f/f_c = \omega\tau$, i.e. the equation (10) becomes the same as equation (6). This means that, the Cole-Cole equation becomes Debye relaxation equation for $\alpha = 0$.

It is reported as a result of a large scale experimental study that the Cole-Cole model is found useful to define electrical properties of healthy and malignant breast tissues at microwave frequencies and in particular, Debye model offers an excellent fit to the Cole-Cole model [56].

2.3. Use of Complex Natural Resonant Frequencies (CNRs) and Singularity Expansion Method (SEM) Theory in Electromagnetic Target Recognition

Experimental and theoretical studies shows that if a target is excited by a short EM pulse, the late-time scattered signals include identifiable contributions from the complex natural resonances (CNRs) of the target. These CNRs provide valuable information about target properties and can be used as the features of the target in aspect and polarization independent manner [49]. CNR mechanism can be used for feature extraction not only for targets in air but also for buried targets such as tumors in breast tissues or explosive mines buried in soil [49-53, 57].

Singularity Expansion Method (SEM) is formalized by Baum in mid 1970s to establish a mathematical representation for the scattered field of a target in resonance region. This representation gives natural response of a target in terms of target poles and the associated residue functions [4, 54]. The poles are aspect and polarization invariant but the residues vary with the excitation and observation conditions [6]. For this reason, extracted target features must be obtained at different aspects for effective target characterization. However, it is important to form a classifier

database at only a few aspect angles, if possible, for a practical and computationally simple classifier design.

In SEM, impulse response of a target can be expressed as [7, 8]

$$h(t,\theta_d,\varphi_d) = e(t,\theta_d,\varphi_d) + \sum_{n=1}^{\infty} b_n(\theta_d,\varphi_d)e^{\alpha_n t}\cos(\omega_n t + \Phi_n) \tag{13}$$

where $s_n = \alpha_n + j\omega_n$ are complex conjugate poles of the target, b_n is the associated residue for n^{th} pole pair, Φ_n is the phase of nth sinusoidal natural response component. While the signal component $e(t,\theta_d,\varphi_d)$ represents the early time response, the summation of cosine functions stand for the late-time scattered signal.

Time duration needed for an impulse excitation wave to fully pass through the target is called as early time period. In this region, the forced response exists which corresponds to the function $e(t,\theta_d,\varphi_d)$ shown in (13). As forced response is highly dependent on polarization and aspect angle, the early-time portion of the scattered signals need to be avoided in classifier design.

After a late time instant called t_{late} , forced responses which are visible in early time completely disappears and only the resonant natural responses of the target remains [10]. For t > t_{late} , the time domain expression (13) reduces to

$$h(t,\theta_d,\varphi_d) = \sum_{n=1}^{\infty} b_n(\theta_d,\varphi_d)e^{\alpha_n t}\cos(\omega_n t + \Phi_n) \tag{14}$$

where the CNR frequencies $s_n = \alpha_n + j\omega_n$ with $\alpha_n < 0$, are functions of only geometry and electrical properties of the target.

Accordingly, the late time response of the scattered signal can be modeled as a combination of damped sinusoidal waveforms oscillating in time at the target's CNR frequencies by SEM theory.

The late time response is weaker than the early time response. Moreover, the late time response shows a decaying behavior in time. Target poles that have smaller damping coefficients, survive longer through the late time region. Such poles are called dominant poles. Target feature extraction algorithms, which are focused on the CNR mechanism, usually makes use of these dominant poles at selected late time windows [9]. Determination of the late-time instant (t_{late}) may be a difficult problem calling for the solution of complicated optimization problems [47]. Similarly, the late-time design interval is usually determined via an optimization routine [18, 29, 30].

2.4.   Computation of Scattered Fields for Lossy Dielectric Spheres

Spherical objects are canonical targets, which are frequently used in classifier design simulations as their electromagnetic fields can be computed using analytical solutions.

2.4.1.   Lossy Dielectric Sphere Embedded in a Lossless Medium

The problem geometry to be used in this section is given in Figure 2.1. The sphere has radius a, and a frequency dependent complex relative permittivity ε_r. The incident plane wave is assumed to be polarized in x-direction and propagates in z direction. The observation point lies on the yz-plane, at $\phi = \pi/2$, at an angle θ measured from the z-axis.
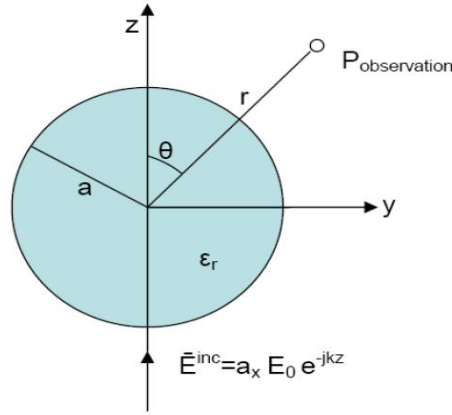
Figure 2. 1 General problem geometry for scattering signals from spherical targets

Homogeneous scalar Helmholtz equation for time dependence $e^{j\omega t}$ is defined as $\nabla^2 u' + k^2 u' = 0$ where $u = u' e^{j\omega t}$. Debye potentials (π) are the solutions of this equation.

In spherical coordinates, the Helmholtz equation is expressed as

$$\frac{1}{r}\left(\frac{\partial^2 r\pi}{\partial r^2}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\pi}{\partial\theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2\pi}{\partial\phi^2} + k^2\pi = 0 \tag{15}$$

where $\pi = R(r)\Theta(\theta)\Phi(\phi)$.

Each of these product functions satisfies well-known ordinary differential equations as

$$\frac{d^2 rR(r)}{dr^2} + \left[k^2 - \frac{n(n+1)}{r^2}\right]rR(r) = 0 \tag{16}$$

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta(\theta)}{d\theta}\right) + \left[n(n+1) - \frac{m^2}{\sin^2\theta}\right]\Theta(\theta) = 0 \tag{17}$$

$$\frac{d^2\Phi(\phi)}{d\phi^2} + m^2\Phi(\phi) = 0 \tag{18}$$

where m can assume only the integer values m=0, ±1, ±2 ... due to azimuthal symmetry of the problem.

The solutions of (16) are Ricatti-Bessel functions

$$\psi_n(kr) = \left(\frac{\pi kr}{2}\right)^{\frac{1}{2}} J_{n+\frac{1}{2}}(kr) \tag{19}$$

$$\chi_n(kr) = -\left(\frac{\pi kr}{2}\right)^{\frac{1}{2}} N_{n+\frac{1}{2}}(kr) \tag{20}$$

where $J_{n+\frac{1}{2}}(kr)$ and $N_{n+\frac{1}{2}}(kr)$ are the half integral order Bessel and Neumann functions.

Linear Combination of (19) and (20) gives the general solution as

$$\zeta_n(kr) = \psi_n(kr) + i\chi_n(kr) = \left(\frac{\pi kr}{2}\right)^{\frac{1}{2}} H_{n+1/2}^{(2)}(kr) \tag{21}$$

where $H_{n+1/2}^{(2)}(kr)$ is the half integral order Hankel function of second kind and k is the wave number.

The solutions of (17) are associated Legendre polynomials

$$\Theta = P_n^{(m)}(\cos(\theta)) \tag{22}$$

and the solutions of (18) are $\sin(m\phi)$ and $\cos(m\phi)$ functions.

While the spherical medium in Figure 2.1 is characterized by a complex valued wave number k_1, the outside medium is characterized by a real valued wave number k_2.

Then, the relative refractive index (m) is defined as

$$m = \frac{k_1}{k_2} = m_1 k_0 / m_2 k_0 = m_1 / m_2 \tag{23}$$

Using the solutions mentioned above and the formulas from (3.3.9) to (3.3.14) provided by the reference [59] (which gives the components of the field vectors (E, H) in terms of Debye potentials (π_1, π_2) in spherical coordinates), the general solution for the scattered wave can be expressed as

$$r\pi_1^s = \frac{-1}{k_2^2} \sum_{n=1}^{\infty} i^{n-1} \frac{2n+1}{n(n+1)} a_n \zeta_n(k_2 r) P_n^{(1)}(\cos\theta)\cos\phi \tag{24}$$

$$r\pi_2^s = \frac{-1}{k_2^2 \kappa_2^{(2)}} \sum_{n=1}^{\infty} i^{n-1} \frac{2n+1}{n(n+1)} b_n \zeta_n(k_2 r) P_n^{(1)}(\cos\theta)\sin\phi \tag{25}$$

where $\kappa_2^{(2)} = ik_0$ and $P_n^{(1)}(\cos\theta)$ is the associated Legendre function of the first kind.

Tangential components of E and H must be continuous across the spherical surface at $r = a$, in accordance with the EM boundary conditions. In terms of Debye potentials, the equalities expressing the boundary conditions are given in the formulas from (3.3.38) to (3.3.41) in reference [59] resulting in the equations

$$m\left[\psi_n^{'}(k_2 a) - a_n \zeta_n^{'}(k_2 a)\right] = c_n \psi_n^{'}(k_1 a) \tag{26}$$

$$\left[\psi_n^{'}(k_2 a) - b_n \zeta_n^{'}(k_2 a)\right] = d_n \psi_n^{'}(k_1 a) \tag{27}$$

$$\left[\psi_n(k_2 a) - b_n \zeta_n(k_2 a)\right] = c_n \psi_n(k_1 a) \tag{28}$$

$$m\left[\psi_n(k_2 a) - b_n \zeta_n(k_2 a)\right] = d_n \psi_n(k_1 a) \tag{29}$$

where $k_0 = 2\pi / \lambda_0$ is the wave number in free space.

By solving the equations from (26) to (29) simultaneously, we get

$$a_n = \frac{\psi_n(\alpha)\psi_n^{'}(\beta) - m\psi_n(\beta)\psi_n^{'}(\alpha)}{\zeta_n(\alpha)\psi_n^{'}(\beta) - m\psi_n(\beta)\zeta_n^{'}(\alpha)} \tag{30}$$

$$b_n = \frac{m\psi_n(\alpha)\psi_n^{'}(\beta) - \psi_n(\beta)\psi_n^{'}(\alpha)}{m\zeta_n(\alpha)\psi_n^{'}(\beta) - \psi_n(\beta)\zeta_n^{'}(\alpha)} \tag{31}$$

with $\alpha = k_2 a = \dfrac{2\pi a}{\lambda} = 2\pi n_2 a / \lambda_0$, $\beta = k_1 a = \dfrac{2\pi n_1 a}{\lambda_0} = m\alpha$

Then, Debye potentials of the scattered wave at equations (24) and (25) can be calculated.

At far zone, the equation (24) and (25) are reduced to the equations (3.3.52) and (3.3.53) in reference [59] and the scattered wave which becomes a transverse wave as a result of the rapid decay of the longitudinal component can be defined as

17

$$E_\phi = \frac{H_\theta}{m_2} = -\frac{i\exp(-ik_2 r)}{k_2 r}\sin\phi\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\left\{a_n\frac{P_n^{(1)}(\cos\theta)}{\sin(\theta)}+b_n\frac{dP_n^{(1)}(\cos\theta)}{d\theta}\right\} \quad (32)$$

$$E_\theta = \frac{-H_\phi}{m_2} = -\frac{i\exp(-ik_2 r)}{k_2 r}\cos\phi\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\left\{a_n\frac{dP_n^{(1)}(\cos\theta)}{d\theta}+b_n\frac{P_n^{(1)}(\cos\theta)}{\sin(\theta)}\right\} \quad (33)$$

The transverse components of the field vectors ($E_\theta, E_\phi, H_\theta, H_\phi$) decay with λ/r but the radial component E_r decay with $(\lambda/r)^2$ so that radial components becomes negligible at far zone [59].

2.4.2. Lossy Dielectric Sphere Coated by Another Lossy Dielectric Shell and Embedded in a Lossless Medium

The problem geometry to be used in this section is given in Figure 2.2 where Region 1 and Region 2 are lossy dielectric media and Region 3 is air. The radius of core (inner spherical region) is a and the outher radius of coated sphere is b. There are three wave numbers k_1, k_2 and k_3 for core, shell and outside medium, respectively.
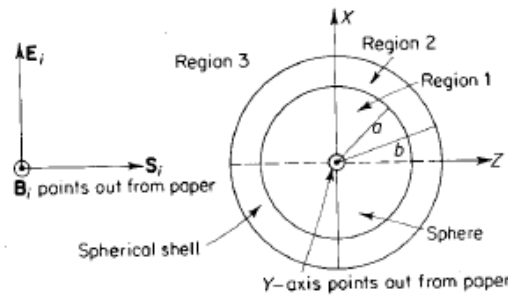
Figure 2. 2 General problem geometry for scattering signals from coated spherical targets

18

In this case, the general solution of the scattered wave is obtained in terms of the Debye potentials corresponding to the TM and TE waves which can be expressed as

$$r\pi_1^s = \frac{-1}{k_3^2} \sum_{n=1}^{\infty} i^{n-1} \frac{2n+1}{n(n+1)} a_n \zeta_n(k_3 r) P_n^{(1)}(\cos\theta)\cos\phi \qquad (34)$$

$$r\pi_2^s = \frac{-1}{k_3^2 \kappa_2^{(3)}} \sum_{n=1}^{\infty} i^{n-1} \frac{2n+1}{n(n+1)} b_n \zeta_n(k_3 r) P_n^{(1)}(\cos\theta)\sin\phi \qquad (35)$$

Using the equations from (5.1.9) to (5.1.32) in reference [59] and by imposing the boundary conditions, the scattering coefficients a_n and b_n can be calculated as

$$a_n = \frac{[\psi\psi]_{n,\alpha}'[\chi\psi]_{n,v} - [\chi\psi]_{n,\alpha}[\psi\psi]_{n,v}'}{[\psi\psi]_{n,\alpha}'[\chi\zeta]_{n,v} - [\chi\psi]_{n,\alpha}[\psi\zeta]_{n,v}'} \qquad (36)$$

$$b_n = \frac{[\psi\psi]_{n,\alpha}''[\chi\psi]_{n,v} - [\chi\psi]_{n,\alpha}[\psi\psi]_{n,v}''}{[\psi\psi]_{n,\alpha}''[\chi\zeta]_{n,v} - [\chi\psi]_{n,\alpha}[\psi\zeta]_{n,v}''} \qquad (37)$$

where $k_3 a = \dfrac{2\pi a}{\lambda} = \alpha$, $k_3 b = \dfrac{2\pi b}{\lambda} = v$ [59].

2.5. Mathematical Review of the WD-PCA and MUSIC Algorithm Based Target Classifier Design Techniques

The WD-PCA and MUSIC Algorithm based target recognition methods have been already applied to classify objects which are made of perfect conductors and perfect dielectrics.

In this thesis, these two techniques are further applied for the recognition of lossy and dispersive objects. Before presenting the applications, the theoretical outline of these two methods will be given in this section.

2.5.1.   Review of WD-PCA Based Target Classification Method

WD-PCA method use Wigner Distribution (WD) outputs instead of directly using time domain scattered signals to extract target poles. The Wigner Distributions (WD) characterize the distribution of the target's natural resonance-related scattered energy over a selected late time segment of joint time-frequency plane. This method was developed and demonstrated by G.Turhan- Sayan [25, 29].

The classifier design procedure begins with constructing moderately aspect dependent feature vectors called Late Time Feature Vectors (LTFVs) by computing Wigner distributions for all candidate targets at chosen design aspects over a common late-time interval by using scattered data over a common frequency band.

Wigner distribution output gives an approximated energy density function over the joint time-frequency plane [26]. The Wigner distribution $W_x(t, f)$, of a real-valued scattered time signal x(t) is expressed as

$$W_x(t,f) = \int_{-\infty}^{\infty} e^{-j2\pi f \tau} x\left(t + \frac{\tau}{2}\right) x *\left(t - \frac{\tau}{2}\right) d\tau \tag{38}$$

with the properties

$$\int_f W_x(t,f) df = |x(t)|^2 \quad \text{and} \quad \int_t W_x(t,f) dt = |X(f)|^2 \tag{39}$$

20

where the superscript (*) denotes complex conjugation, $|x(t)|^2$ is the signal's instantaneous power, $X(f)$ is the Fourier transform of the signal x(t) and $|X(f)|^2$ represents spectral energy density of the signal [26].

Due to the uncertainty principle, it is not possible to have infinite resolution both in time and in frequency simultaneously [27]. So, WD outputs contain strong oscillatory terms with non-physical negative values which may seriously deteriorate identification capacity of the classifier. These negative values, which can not be interpreted as energy density terms, are replaced by zeros to improve the classifier performance [25]. Accordingly, modified WD can be expressed as

$$\tilde{W}_x(t,f) = \frac{W_x(t,f) + abs(W_x(t,f))}{2} \tag{40}$$

The discrete WD output is a matrix of size $(N \times N/2)$ where N is the number of sample points in time.

Modified WD outputs are processed for a given target over a chosen late-time interval at selected aspect angles. Total time span T_0 is divided into Q equal length non-overlapping time bands. The amount of energy contributed to each subinterval at a sample frequency f_m is defined as

$$E_q(f_m) = \int_{(q-1)\Delta}^{q\Delta} \tilde{W}_x(t,f)dt \quad \text{for} \quad q = 1,2,3,...Q \tag{41}$$

where $\Delta = T_0/Q$, $M = 1,2,....,N/2$ and $f_m = (m-1)/T_0$.

Then, in vector form, the spectral distribution of the partial signal energy in the subinterval q given as

$$\overline{E}_q = [E_q(f_1) \quad E_q(f_2) \quad ... \quad E_q(f_{N/2})] \tag{42}$$

So, the total partitioned energy density vector $\bar{E}$ which has length of $(N/2) \times Q$ can be formed as

$$\bar{E} = \begin{bmatrix} \bar{E}_1 & \bar{E}_2 & \bar{E}_3 & \dots & \dots & \bar{E}_Q \end{bmatrix} \tag{43}$$

As it can be seen obviously, using E instead of the WD output matrix provides a dimensional reduction by the factor $2N/Q$.

To enhance classifier performance, two successive time bands $\bar{E}_{q*}$ and $\bar{E}_{q*+1}$, are chosen to create late time feature vectors (LTFVs) of the targets instead of using $\bar{E}$ vectors as a whole [29]. Therefore, length of final feature vector becomes $2 \times \dfrac{N}{2} = N$ which is equal to length of discrete scattered time domain signal.

Then, to select the optimum late time interval q*, maximization of the Correct Classification Factor (CCF) is needed. This factor is defined as

$$CCF(q) = \frac{1}{M_{tar}K_2} \sum_{i,j} r_{i,j}^{matched} - \frac{1}{(M_{tar}^2 - M_{tar})K^2} \sum_{i,j} r_{i,j}^{mismached} \tag{44}$$

where M_{tar} is the number of targets and K is the number of reference aspects.

CCF is computed for a selected Q value. The main idea is selecting q value which makes CCF maximum. The term $r_{i,j}^{matched}$ defines the correlation coefficient between any two LTFVs that belong to the same target at different reference aspects and $r_{i,j}^{mismatched}$ defines the correlation coefficient between any two LTFVs that belongs to different targets.

After that, the LTFVs that are created at a selected optimum late time interval for all library targets and all reference aspects are fused by the well-known statistical method called principal component analysis (PCA) for each target.

The PCA method is introduced by Pearson in 1901 [28]. In the WD-PCA method, the PCA method is used to reduce the aspect sensitivity problem in target recognition. Use of PCA method for multi-aspect feature fusion is suggested by G.Turhan-Sayan [29] for the first time to obtain a single characteristic feature vector representing a given target. The fused feature vector (FFV) is obtained by combining the information in to characterize the given target over a broad range of aspects.

In the implementation of the PCA procedure a real valued matrix that contains LTFVs (of size $1xN$ each) at K different aspects is constructed as

$$F^T = \begin{bmatrix} \bar{e}_1^T & \bar{e}_2^T & \dots & \bar{e}_K^T \end{bmatrix}_{KxN} \tag{45}$$

where T denotes the transpose operator. Then, the covariance matrix of F^T, which is a symmetric positive definite matrix of size KxK, is computed as

$$S_F = \begin{bmatrix} s_1^2 & s_{12} & \dots & s_{1K} \\ s_{21} & s_2^2 & \dots & s_{2K} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ s_{K1} & s_{K2} & \dots & s_K^2 \end{bmatrix} \tag{46}$$

where the diagonal entries (s_i^2) represent variance of the feature vector $\bar{e}_i$ while the off-diagonal entries $s_{i,j} = s_{j,i}$ represent the covariance between the feature vectors $\bar{e}_i$ and $\bar{e}_j$.

Correlation coefficient $r_{i,j}$ between the feature vectors $\bar{e}_i$ and $\bar{e}_j$ can be written as

$$r_{i,j} = \frac{s_{i,j}}{\sqrt{s_i s_j}} \tag{47}$$

Through a similarity transform, the covariance matrix S_F can be transformed into a diagonal matrix Λ which is expressed as

$$\Lambda = U^T S_F U = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_K \end{bmatrix} \tag{48}$$

where the $U = \begin{bmatrix} u_1 & u_2 & \dots & u_K \end{bmatrix}$ is a modal matrix which is the combination of eigenvectors of the S_F.

The eigenvalues λ_i are solved from $\det(S_F - \lambda I) = 0$ where I is the identity matrix of size KxK. After ordering the computed eigenvalues such that $\lambda_1 > \lambda_2 > \dots > \lambda_K$, the corresponding eigenvectors are solved by $\begin{bmatrix} S_F - \lambda_i I \end{bmatrix} t_i = 0$ for $i = 1,2,\dots K$ where t_i denotes the eigenvectors.

The U matrix is used to make transformation from the correlated feature vectors $\bar{e}_1, \bar{e}_2, \dots \bar{e}_K$ to uncorrelated vectors $z_1, z_2, \dots z_K$. To construct the matrix U, the orthogonal eigenvectors are normalized to obtain orthonormal eigenvectors as $u_i = \dfrac{t_i}{\sqrt{t_i^T t_i}} = \dfrac{t_i}{|t_i|}$.

The matrix Z formed by uncorrelated vectors $z_1, z_2, \dots z_K$ has the same size as the F matrix as

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix} = U^T \begin{bmatrix} \bar{e}_1 - mean(\bar{e}_1)I_N \\ \bar{e}_2 - mean(\bar{e}_2)I_N \\ \vdots \\ \bar{e}_k - mean(\bar{e}_k)I_N \end{bmatrix}_{KxN} \tag{49}$$

where I_N is an all-ones raw vector of length N.

These transformed zero mean vectors $z_i's$ with variance λ are the principle components of the feature matrix F. It is shown in [28] that the first principle component z_1 has the highest correlations with the late time feature vectors $e_i's$. So, the leading component z_1 can usually be used to represent the aspect independent FFV of the target by neglecting the other principle components with sufficiently small variances [29].

But if it is needed, the other principle components can also be included into the process as

$$z_{FFV} = \sum_{i=1}^{K} \overline{\lambda}_i z_i \tag{50}$$

where $\overline{\lambda}_i$ are the proper weighted factors computed as $\overline{\lambda}_i = \dfrac{\lambda_i}{\sum_{k=1}^{K} \lambda_k}$.

The classifier's database is constructed with FFVs of the targets. During the real-time test phase, test targets's LTFV is extracted from the received test signal over the selected optimal late time interval. As the last step of the classifier design, FFVs of the classifier database are compared with the extracted LTFV that belongs to the test signal. The target is recognized by the computed highest correlation coefficient between the test target LTFV and a particular FFV in the classifier database.

2.5.2. Review of MUSIC Algorithm Based Target Classification Method

As mentioned earlier, some target recognition methods based on estimating target poles directly from the scattered signal produce inaccurate results as the SNR of the

scattered data gets lower. However, MUSIC algorithm method uses the effects of poles indirectly and this approach is known to be more robust against low SNR values. MUSIC method is also theoretically inspired by the SEM theory. It creates target's natural-resonance related spectral power distributions, which are called as normalized MUSIC spectrum matrices (MSM's), over complex time-frequency plane.

To be used in the MUSIC algorithm method, late time portion of the scattered signal of a target at a certain aspect angle/polarization combination in the presence of noise can be expressed as

$$z(n) = x(n) + w(n) = \sum_{i=1}^{L/2} b_i(\theta, \varphi) e^{\alpha_i n} \cos(\omega_i n + \Phi_i) + w(n) \tag{51}$$

where $w(n)$ is additive Gaussian noise. The noise free component of the signal, $x(n)$ is expressed in terms of linear combination of L complex exponentials having the target poles $s_i = \alpha_i \pm j\omega_i$ where $i = 1,....,L$. The signal $z(n)$ is sampled at N discrete time points. The b_i's are related complex valued weight coefficients where the signal and noise component are uncorrelated.

First of all, a vector $y(n)$ of length m and its correlation matrix IR are formed to construct MSM's as

$$y(n) = \begin{bmatrix} z(n) & z(n-1) & \dots & z(n-m+1) \end{bmatrix}^T \tag{52}$$

$$IR = E\left\{ y(n)\overline{y(n)}^H \right\} = E\left\{ x(n)\overline{x(n)}^H \right\} + E\left\{ w(n)\overline{w(n)}^H \right\}$$

$$= AE\left\{ cc^H \right\} A^H + \sigma^2 I_{mxm}$$

$$= ACA^H + \sigma^2 I_{mxm} \tag{53}$$

where m satisfies the condition L<m<N. E is the expected value operator, the superscript H denotes the complex conjugate transpose, σ^2 is the variance of the Gaussian noise and I is the unit matrix.

A and C matrixes can be formed as

$$A = \begin{bmatrix} a(s_1) & a(s_2) & \dots & a(s_L) \end{bmatrix} \text{ with } a(s) = \begin{bmatrix} 1 & e^{-1} & \dots & e^{-s(m-1)} \end{bmatrix}^T \tag{54}$$

$$C = E\{cc^H\} = E\left\{ \begin{bmatrix} c_1 e^{s_1 n} & c_2 e^{s_2 n} & \dots & c_L e^{s_L n} \end{bmatrix}^T \begin{bmatrix} c_1^* e^{s_1^* n} & c_2^* e^{s_2^* n} & \dots & c_L^* e^{s_L^* n} \end{bmatrix} \right\}$$

$$= \begin{bmatrix} m_1 & 0 & \dots & 0 \\ 0 & m_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & m_L \end{bmatrix} \qquad \text{since,}$$

$$E\left\{ c_l e^{s_l n} c_k^* e^{s_k^* n} \right\} = E\left\{ \beta_l e^{\varphi_l} e^{s_l n} \beta_k e^{-j\varphi_k} e^{s_k^* n} \right\}$$

$$= \beta_l e^{s_l n} \beta_k e^{s_k n} E\left\{ e^{j(\varphi_l - \varphi_k)} \right\}$$

$$= \begin{cases} m_1 & if \quad k = l \\ 0 & if \quad k \neq l \end{cases} \tag{55}$$

In these expressions c_i's are assumed to be uniformly distributed random phases in interval $[-\pi \ \pi]$.

Then, let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ be eigenvalues of the correlation matrix IR and $\{e_1 e_2, \dots e_L, e_{L+1}, \dots, e_m\}$ be the set of corresponding orthonormal eigenvectors. G=$[e_{L+1} \dots e_1]$ is the eigenvector matrix corresponding to eigenvalues for i=L+1,..,m. Since the rank of the matrix ACA^H is equal to L, m-L of its eigenvalues are zero and the eigenvalues of IR for i=L+1,...,m are all equal to σ^2. Therefore, multiplication of the correlation vector and G vector can be written as

$$IRG = \sigma^2 G \tag{56}$$

By using (53), $IRG = ACA^H G + \sigma^2 G$, so $ACA^H G = 0$ is obtained.

Since the rank of AC is equal to L, it should be non-zero. Then, the true complex frequencies s_i's where i=1,...L are only the solutions of the equation

$$a^H(s)GG^H a(s) = 0 \tag{57}$$

As a result, the non-normalized MUSIC spectrum function $P_{unnorm}(s)$ which has peak values at $s = s_i$ is defined as

$$P_{unnorm}(s) = \frac{1}{a^H(s)GG^H a(s)} \tag{58}$$

where $s = \alpha + j\omega$ is the complex frequency. With a matrix considered discrete values along the $\alpha = \text{Re}\{s\}$ and $\omega = \text{Im}\{s\}$ axes with indices u and v, the normalized version of $P_{unnorm}(s)$ can be approximated as

$$P(u,v) = \frac{P_{unnorm}(u,v)}{\sum_{v=1}^{k_2}\sum_{u=1}^{k_1}(P_{unnorm}(u,v))^2} \tag{59}$$

where,

$$P_{unnorm}(u,v) = \frac{1}{a^H(\alpha_u + jw_v)GG^H a(\alpha_u + jw_v)} \tag{60}$$

and k_1, k_2 are the MSM (P(s)) dimensions.

Target's natural responses may be excited at different strengths at different aspects due to the high aspect dependency of the scattered responses. For the classifier design, MSM's belonging to different targets at different aspects are constructed because of the aspect dependency of these target feature matrices. The target's

FMSM vectors are obtained by fusing its MSMs to solve this aspect dependency problem and to characterize the effects of a large number of CNRs for the target. This process also reduces the redundancy of the classifier's reference feature database.

To estimate an optimum number for m and L some studies are done and choosing m=N/2 and L=m/2 are suggested [44-46]. m can be chosen larger for good resolution, but it is limited by computational complexity and to allow a reliable estimation of the covariance matrix it should not be too close to N. L represents number of estimated damped sinusoidal signals in reference data and according to the previous studies underestimation of L may decrease the classifier performance.

To construct the classifier database, all targets FMSMs are obtained by

$$FMSM_i(u,v) = \frac{P_{sum,i}(u,v)}{\left\|P_{sum,i}(u,v)\right\|} \quad \text{for} \quad \forall i = 1,\ldots,K_T \tag{61}$$

where, $P_{sum,i}(u,v) = \sum_{j=1}^{K_A} P_{i,j}(u,v)$ with $P_{i,j}$ refers to the MSM of target i at reference aspect angle j.

The scattered signal's common time span is divided into overlapping subintervals (Q) and either MSMs or FMSMs are constructed for all these intervals to decide the optimum late time interval.

As an important step for classifier design, the optimum late time interval is chosen where MSMs belong to a target at reference aspects give high correlation with its own FMSM and give low correlation with other target's FMSMs in database. This procedure can be expressed mathematically as

$$CCF(q) = \frac{1}{K_A K_T} \sum_{i=1}^{K_T} r_{i,match} - \frac{1}{K_T(K_T-1)K_A} \sum_{i=1}^{K_T} r_{i,mismatch} \tag{62}$$

with,

$$r_{i,match} = \sum_{j=1}^{K_A} \sum_{v=1}^{k_2} \sum_{u=1}^{k_1} FMSM_i(u,v)P_{i,j}(u,v) \qquad (63)$$

$$r_{i,mismatch} = \sum_{\substack{k=1 \\ k \neq i}}^{K_A} \sum_{j=1}^{K_A} \sum_{v=1}^{k_2} \sum_{u=1}^{k_1} FMSM_i(u,v)P_{k,j}(u,v) \qquad (64)$$

for i=1,....,K_T being the target index, j=1,....,K_A being the reference aspect index.

Finally, classifier database is constructed with FMSMs of all targets which are computed over selected late time interval. To recognize a test target its MSM is computed for selected late time interval and compared with all FMSMs at database. It is expected to give highest correlation result with its own FMSM with the similar logic in WD-PCA algorithm based method [30, 31, 34, 44].

CHAPTER 3

APPLICATIONS AND RESULTS

In recent years, there has been a great deal of interest in breast cancer detection techniques, with research studies showing the contrast between dielectric properties of normal and malignant breast tissues at radio and microwave frequencies [54-57]. In this thesis, dielectric properties of healthy (adipose, glandular and fiberconnective) tissue samples obtained from breast reduction and cancer surgeries (Table 3.1 and Table 3.2, respectively) and malignant (invasive or non-invasive ductal and lobular carcinomas) tissue samples obtained from cancer surgeries (Table 3.3) are used in classifier design and testing. Dielectric parameters of various tissue types are fitted by using the Cole-Cole model based on data obtained by a large scale experimental study [54, 55] where a large number of samples for normal breast tissues and tumors are investigated. Three groups of healthy breast tissues are formed based on their adipose contents: Group 1 contains the tissue samples with 0-30% (low) adipose content, Group 2 contains samples with 31-84% (moderate) adipose content and Group 3 contains samples with 85-100% (high) adipose content to maximize the difference between the groups and minimize the variability within the groups. Also Cole-Cole parameters for a tumor sample are given in Table 3.3 where the median parameters (50%) are reported. As mentioned in Chapter 2, Cole-Cole parameters are used to model the permittivity characteristics of breast tissue and tumor tissue samples which are assumed to be spherical samples of lossy and dispersive dielectrics for computational simplicity.

In this thesis, both WD-PCA and MUSIC Algorithm based target classifier design techniques are used to classify healthy breast tissue samples according to their

adipose (fat) content. Then, classifiers are designed to recognize the presence of tumor in each type of healthy breast tissue. Performances of WD-PCA classifiers and MUSIC Algorithm classifiers are compared for each specific classification problem as to be presented in the rest of this chapter.

Table 3.1 Cole-Cole parameters for healthy breast tissue samples obtained from the breast reduction surgeries

%	Group1 (low)			Group2 (moderate)			Group3 (high)		
	25th	50th	75th	25th	50th	75th	25th	50th	75th
ε_∞	9.941	7.821	6.151	8.718	5.573	5.157	2.908	3.14	4.031
$\Delta\varepsilon$	26.60	41.48	48.26	17.51	34.57	45.81	1.2	1.708	3.654
τ	10.9	10.66	10.26	13.17	9.149	8.731	16.88	14.65	14.12
α	0.003	0.047	0.049	0.077	0.095	0.091	0.069	0.061	0.055
σ_s	0.462	0.713	0.809	0.293	0.524	0.766	0.02	0.036	0.083

Table 3.2 Cole-Cole parameters for healthy breast tissue samples obtained from the cancer surgeries

%	Group1 (low)			Group2 (moderate)			Group3 (high)		
	25th	50th	75th	25th	50th	75th	25th	50th	75th
ε_∞	5.013	7.237	7.816	3.891	6.08	6.381	3.122	3.581	3.882
$\Delta\varepsilon$	40.6	46	50.21	4.113	19.26	32.3	2.133	3.337	5.02
τ	10.16	10.3	10.47	0.082	0.297	0.561	14.27	15.21	12.92
α	0.091	0.049	0.055	13.83	11.47	10.41	0.099	0.052	0.059
σ_s	0.607	0.808	0.899	0.038	0.057	0.081	0.034	0.053	0.103

Table 3.3 Cole-Cole parameters for tumor samples that contain at least 30% malignant tissue

%	25th	50th	75th
ε_∞	7.67	6.749	9.058
$\Delta\varepsilon$	43.92	50.09	51.31
τ	10.7	10.5	10.84
α	0.028	0.051	0.022
σ_s	0.748	0.794	0.899

In Table 3.1, Table 3.2 and Table 3.3, the Cole-Cole parameters are listed within each group (Group 1, Group 2 and Group 3) considering 25th, 50th and 75th percentile subgroups.

3.1. Classifier Designs for Healthy Breast Tissues with Different Adipose Content

In this section, different classifiers are designed with target libraries containing spherical samples of biological tissues belonging to Group 1, Group 2, and Group 3. All these spheres are assumed to have the same radius of 10cm. Frequency-dependent permittivity functions of these biological tissues are generated using their Cole-Cole parameters to be used in the computation of scattered electromagnetic signals. Effectiveness of MUSIC and WD-PCA algorithm based classifier design methods are investigated in classifying healthy breast tissues . The problem geometry used in sample problems have already been explained in part 2.4.1. The same

approach will be used in this section. As the first step, the impulse responses of the targets in response to a plane wave excitation are synthesized analytically [59].

Far field scattered responses of the spherical tissue samples are computed at $\phi = \frac{\pi}{2}$ plane for 12 different values of bistatic aspect angle θ_b, where θ_b =180- θ, in the frequency domain over a bandwidth from almost DC to 19.1 GHz with 512 frequency sample points.

$\theta = $ 15°, 30°, 45°, 60°, 75°, 90°, 105°, 120°, 135°, 150°, 165° and 179° are used as aspect angles. Only four of these aspects (179°,135°,90°,45°) are used as reference aspect angles to construct the classifier's raw database. Scattered data of the other aspect angles are used as test angles for performance testing.

After obtaining the scattered target responses using the calculations indicated in section 2.4.1, time domain responses of spherical targets are computed by using Inverse Fast Fourier Transform (IFFT) of windowed frequency domain data. 1024 sample points are used over a total time span of 26.81 ns. These time domain signals become input to the classifier design algorithms. A summary of classifiers designed in section 3.1 are given in Table 3.4.

Table 3.4 Target Libraries and Number of Targets of the Designed Classifiers

Classifier	Target Library	Number of Targets
C1	0-30% adipose content-healthy breast tissue samples obtained from reduction surgeries **(Table 3.1-Group 1)**	3
C2	85-100% adipose content-healthy breast tissue samples obtained from reduction surgeries **(Table 3.1-Group 3)**	3
C3	0-30% adipose content-healthy breast tissue samples obtained from reduction and cancer surgeries **(Table 3.1-Group 1 and Table 3.2 -Group 1)**	6
C4	0-30% adipose content-healthy breast tissue samples obtained from reduction and cancer surgeries, and a tumor **(Table 3.1-Group 1 , Table 3.2 -Group 1 and tumor)**	7
C5	85-100% adipose content-healthy breast tissue samples obtained from reduction and cancer surgeries **(Table 3.1-Group 3 and Table 3.2 -Group 3)**	6
C6	85-100% adipose content-healthy breast tissue samples obtained from reduction and cancer surgeries, and a tumor **(Table 3.1-Group 3 , Table 3.2 -Group 3 and tumor)**	7
C7	31-84% adipose content-healthy breast tissue samples obtained from reduction and cancer surgeries **(Table 3.1-Group 2 and Table 3.2 -Group 2)**	6
C8	All content-healthy breast tissue samples obtained from reduction surgeries and tumor **(Table 3.1 and tumor)**	10
C9	All content-healthy breast tissue obtained from reduction and cancer surgeries and tumor **(Table 3.1 , Table 3.2 and tumor)**	19

3.1.1. Classifier Design and Test Results with the WD-PCA Method for Lossy and Dispersive Breast Tissue Samples

Nine different classifiers will be designed and tested in this section as follows:

3.1.1.1. Design and Test Results of Classifier C1

All of the classifiers listed in Table 3.4 are designed using the WD-PCA algorithm based method first as explained in section 2.5.1. Figures from 3.1 to 3.13 show the results of important design steps for Classifier C1. Design steps are similar for the other classifiers C2 to C9. The classifier C1 is designed for three targets T1, T2 and T3 which correspond to the healthy breast tissue samples in Group 1 represented within the 25 percentile, 50 percentile and 75 percentile, respectively. The frequency-dependent permittivity curves for T1, T2 and T3 are shown in Figure 3.1 and Figure 3.2. Time domain scattered signals of targets in Classifier C1 (Table 3.5) at the reference aspect angle, $\theta = 179°$, are given in Figure 3.3, Figure 3.4 and Figure 3.5, respectively as examples. Contour plots of modified Wigner distribution outputs for Target-1 (T1), Target-2 (T2) and Target-3 (T3) of C1 at reference aspect angles are given in Figure 3.6, Figure 3.7 and Figure 3.8, respectively. As seen in these figures, targets' Wigner distributions are quite different even though their time signals look very similar. Correct Classification Factor (CCF) plotted against subinterval index q is given in Figure 3.9 where Q=32 subintervals are tested in the design process. This figure is used to determine the optimum late time design interval (q* and q*+1 : combination of selected sequential late time intervals). LTFVs for the target T1 which are obtained for the selected optimum late time design interval at reference angles given in Figure 3.10. FFVs of all three targets in database given in Figure 3.11.

The eigenvalues (λ) of the PCA are obtained as explained in section 2.5.1 for feature fusion purpose. Correlation coefficients computed between the reference late time feature vectors and the principle components are given in Table 3.6 for T3, in Table 3.7 for T2, in Table 3.8 for T1. Correlation coefficients computed between the FFVs of targets and the LTFVs of test targets at all aspects are given in Figure 3.12 as a bar chart. The related contour plots are presented in Figure 3.13.

To design the classifier C1 (Table 3.5), time span of the targets' impulse responses are divided into Q=32 non-overlapping subintervals and q*=12 (q*=12 and q*+1=13 selected) is selected to form the optimal late-time design interval that is 9.21 ns to 10.89 ns. As seen in the Table 3.6, Table 3.7 and Table 3.8, correlation coefficients between reference late time feature vectors and z_1 much larger than others. Therefore, FFVs of the targets are calculated by using only z_1 and neglecting the others.

The success rate of the classifier C1, which is designed by the WD-PCA method, is found to 100 percent. It means that targets of C1 (low-adipose content normal breast tissues obtained from reduction surgeries) can be distinguished from each other perfectly. Classification success of the WD-PCA method for C1 can also be seen in Figure 3.12 and Figure 3.13.

Table 3.5   Cole-Cole parameters of tissue samples in target library of the classifier C1

Classifier-1 (C1)			
%	Target-1 (T1)	Target-2 (T2)	Target-3 (T3)
ε_∞	9.941	7.821	6.151
$\Delta\varepsilon$	26.60	41.48	48.26
$\tau(ps)$	10.90	10.66	10.26
α	0.003	0.047	0.049
$\sigma_s(S\ m^{-1})$	0.462	0.713	0.809
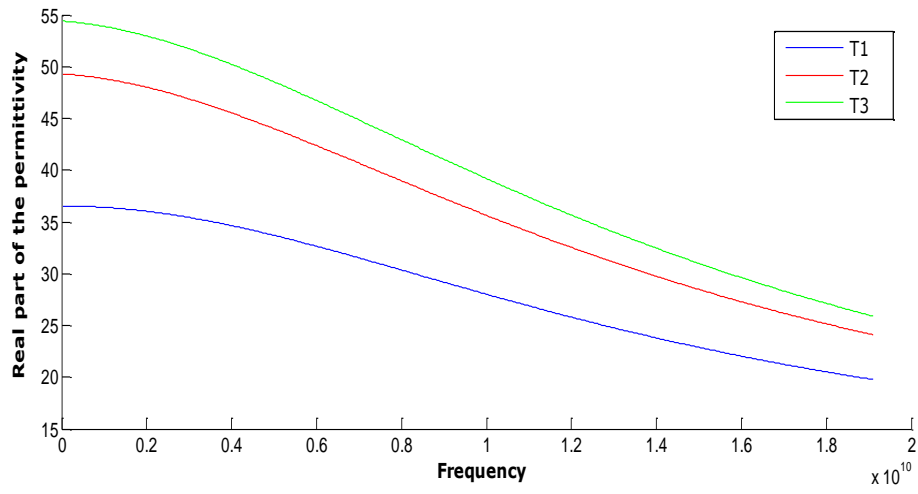
Figure 3.1 Frequency dependent ε' curves for the targets of the classifier C1
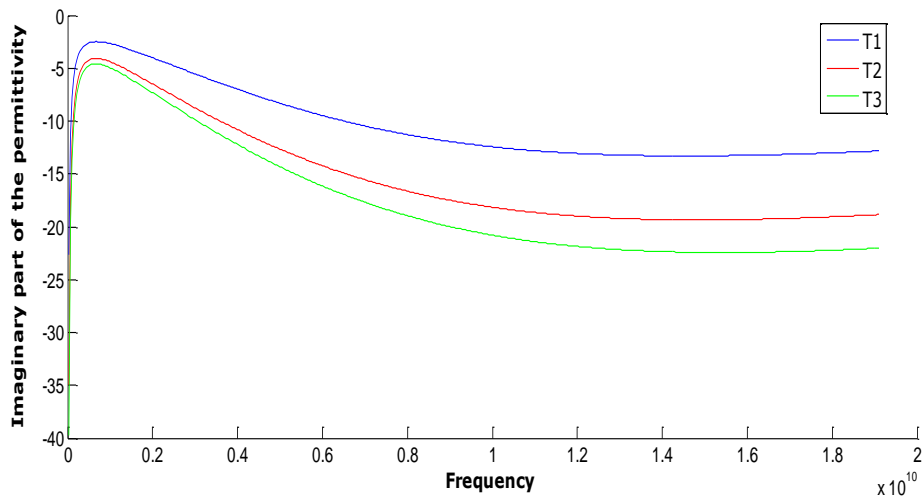
Figure 3. 2 Frequency dependent ε'' curves for the targets of the classifier C1
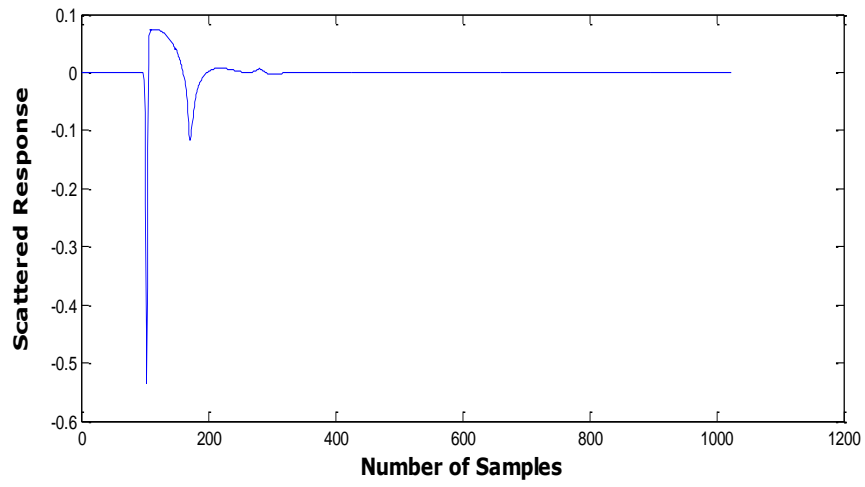
Figure 3.3 Scattered time domain response of the first target T1 at $\theta = 179°$ aspect angle
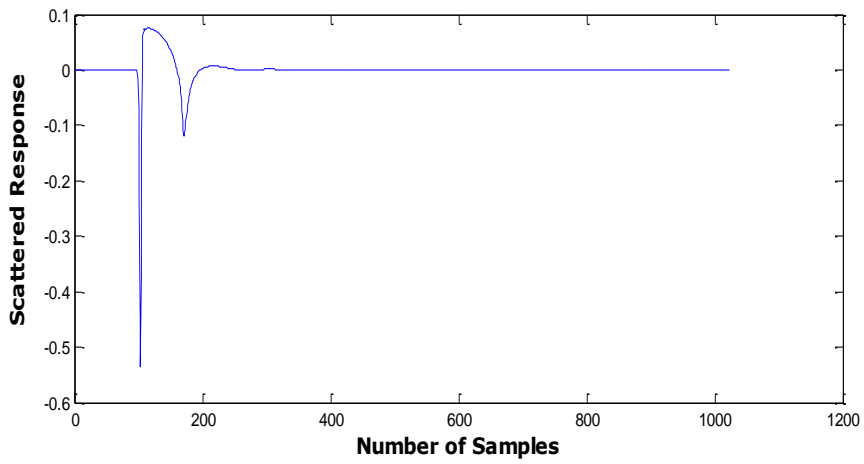
Figure 3.4 Scattered time domain response of second target T2 at $\theta = 179°$ aspect angle
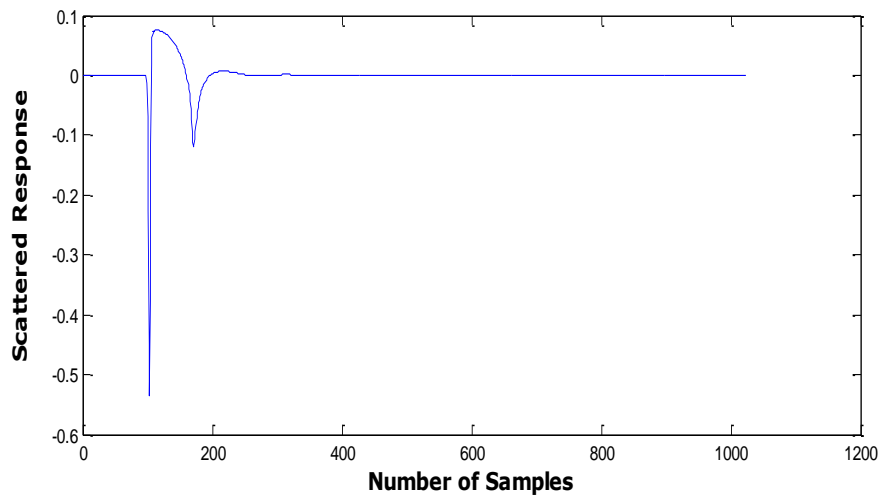
Figure 3.5 Scattered time domain response of third target T3 at $\theta = 179°$ aspect angle
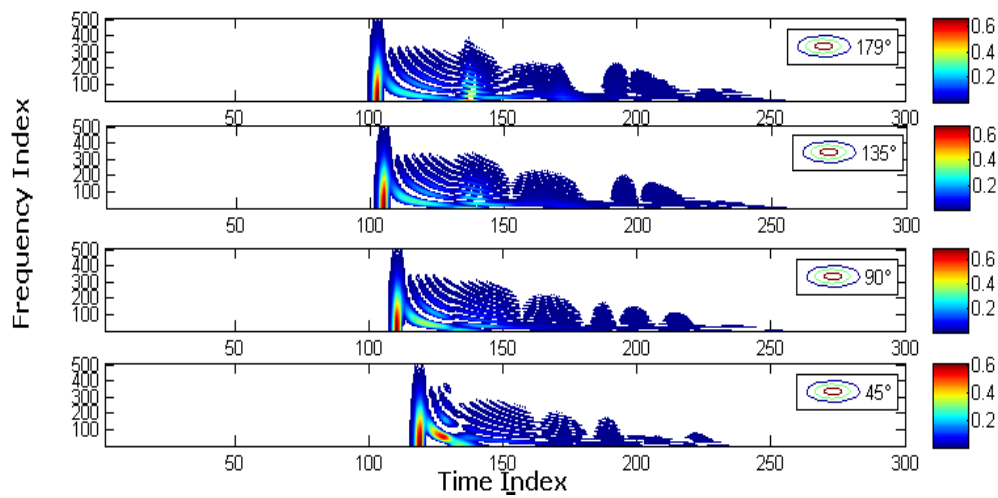
Figure 3.6 Contour plots of the modified Auto-Wigner distribution outputs for the first target T1 at aspect angles 179°, 135°, 90° and 45°
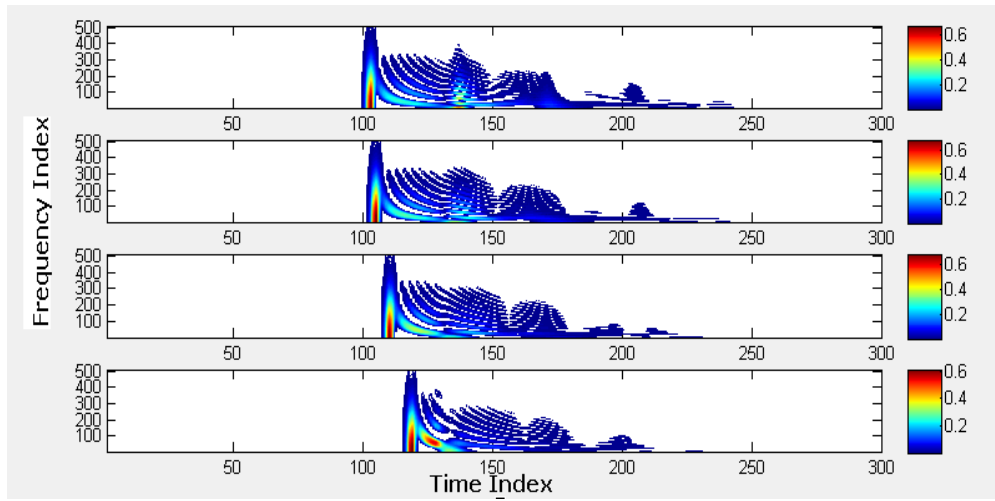
Figure 3.7 Contour plots of the modified Auto-Wigner distribution outputs for the second target T2 at aspect angles 179°, 135°, 90° and 45°
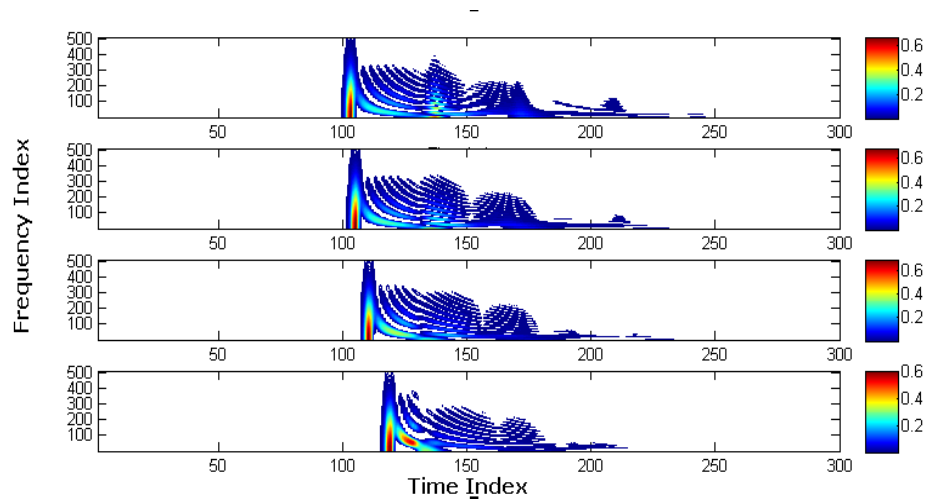
Figure 3.8 Contour plots of the modified Auto-Wigner distribution outputs for the third target T3 at aspect angles 179°, 135°, 90° and 45°

41

Figure 3.9 Correct Classification Factor (CCF) plotted against the late time interval index q for the classifier C1(using the WD-PCA method)
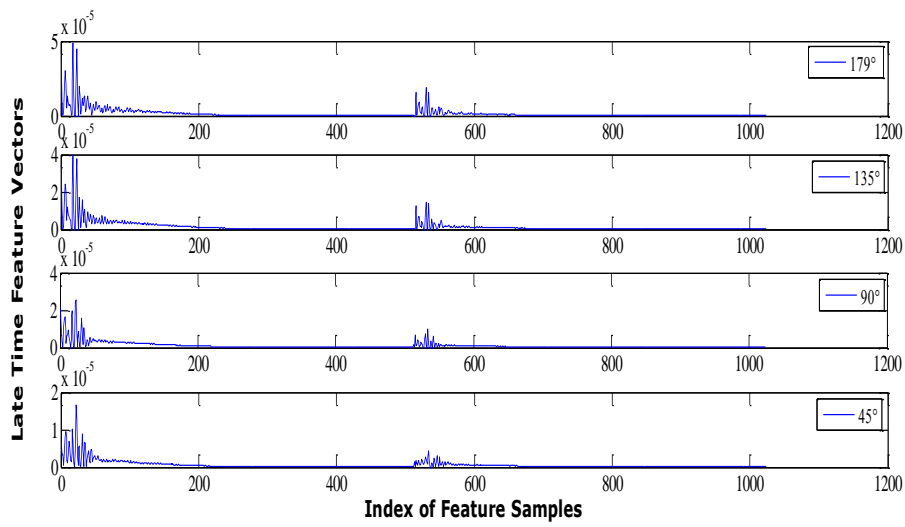
Figure 3.10 LTFVs of the first target T1 at reference aspect angles
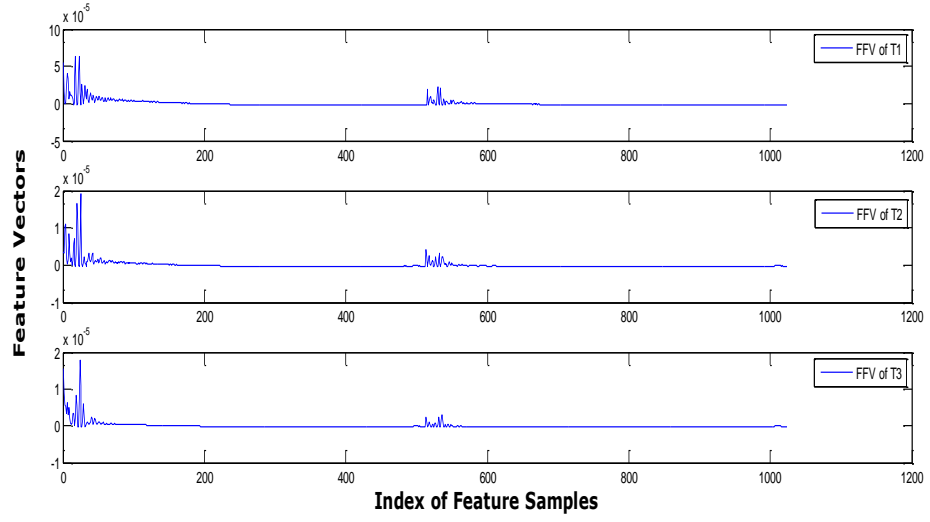
Figure 3.11 FFVs of the all targets in the classifier C1 database

Table 3.6   PCA results for the sample tissue T3 with correlation coefficients computed between reference LTFVs and their principle components together with the associated eigenvalues

Ordered Eigenvalues		$\overline{e1}$ θ_{ref} = 179°	$\overline{e2}$ θ_{ref} = 135°	$\overline{e3}$ θ_{ref} = 90°	$\overline{e4}$ θ_{ref} = 45°
0.1245 e-11	z_1	0.9906	0.9970	0.9796	0.8708
0.0030 e-11	z_2	-0.1141	-0.0343	0.1616	0.4365
0.0012 e-11	z_3	0.0731	-0.0575	-0.1134	0.2245
0.0001 e-11	z_4	-0.0184	0.0391	-0.0377	0.0295

Table 3.7   PCA results for the sample tissue T2 with correlation coefficients computed between reference LTFVs and their principle components together with the associated eigenvalues

Ordered Eigenvalues		$\overline{e1}$ θ_{ref} = 179°	$\overline{e2}$ θ_{ref} = 135°	$\overline{e3}$ θ_{ref} = 90°	$\overline{e4}$ θ_{ref} = 45°
0.1670e-11	z_1	0.9879	0.9965	0.9417	0.8159
0.0061 e-11	z_2	-0.1457	-0.0041	0.3278	0.4486
0.0013 e-11	z_3	0.0477	-0.0678	-0.0506	0.3525
0.0003 e-11	z_4	-0.0249	0.0492	-0.0558	0.0937

Table 3.8   PCA results for the sample tissue T1 with correlation coefficients computed between reference LTFVs and their principle components together with the associated eigenvalues

Ordered Eigenvalues		$\overline{e1}$ θ_{ref} = 179°	$\overline{e2}$ θ_{ref} = 135°	$\overline{e3}$ θ_{ref} = 90°	$\overline{e4}$ θ_{ref} = 45°
0.2725e-10	z_1	0.9898	0.9962	0.9587	0.8539
0.0091e-10	z_2	-0.1314	-0.0158	0.2742	-0.4770
0.0015e-10	z_3	-0.0531	0.0767	0.0427	-0.1890
0.0005e-10	z_4	0.0143	-0.0374	0.0629	-0.0877
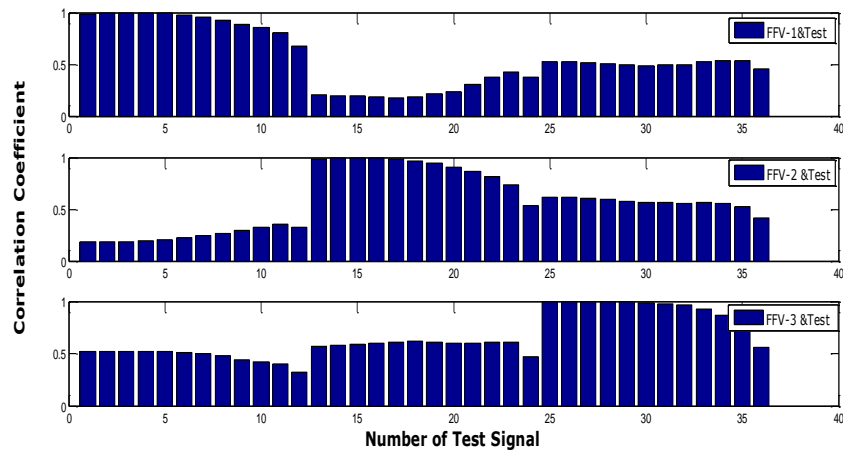
Figure 3.12 Correlation coefficients computed between the FFVs of targets T1, T2 and T3 of the classifier C1 and the LTFVs of all test signals (3 target x 12 angle=36 test signals)
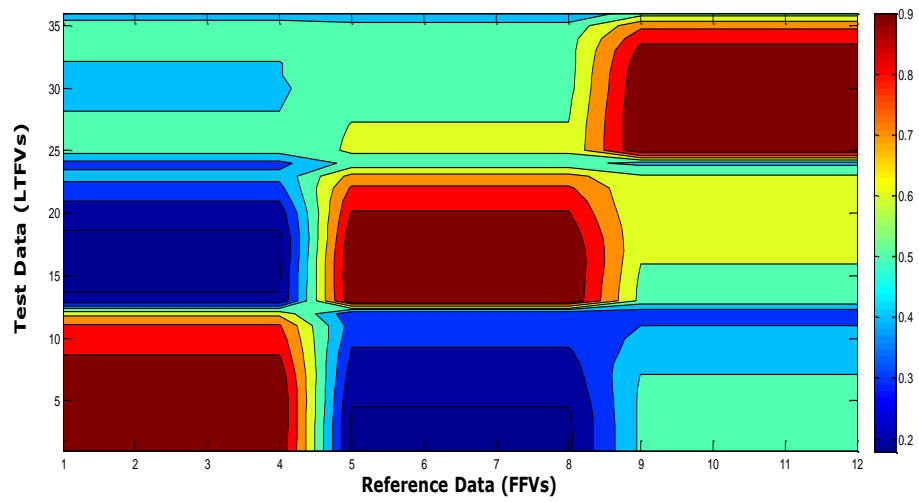
Figure 3.13 Contour plots of the correlation coefficients computed for all possible pairs of test and reference feature vectors for the classifier C1

45

3.1.1.2. Design and Test Results of Classifier C2

When the classifier C2 (see Table 3.9) is designed with Q=16, q*=13 (Figure 3.15), the method gives 91.66 percent accuracy rate. As an example, the scattered time domain signal computed at $\theta = 179°$ for the target T1 with 84-100% adipose content is given in Figure 3.14. The CCF versus late time index bar chart is presented in Figure 3.15. The frequency-dependent permittivity curves for T1, T2 and T3 are shown in Figure 3.16 and Figure 3.17. As seen in Figure 3.18, the target T1 is recognized perfectly at all aspects while T2 and T3 are recognized with smaller safety margins.
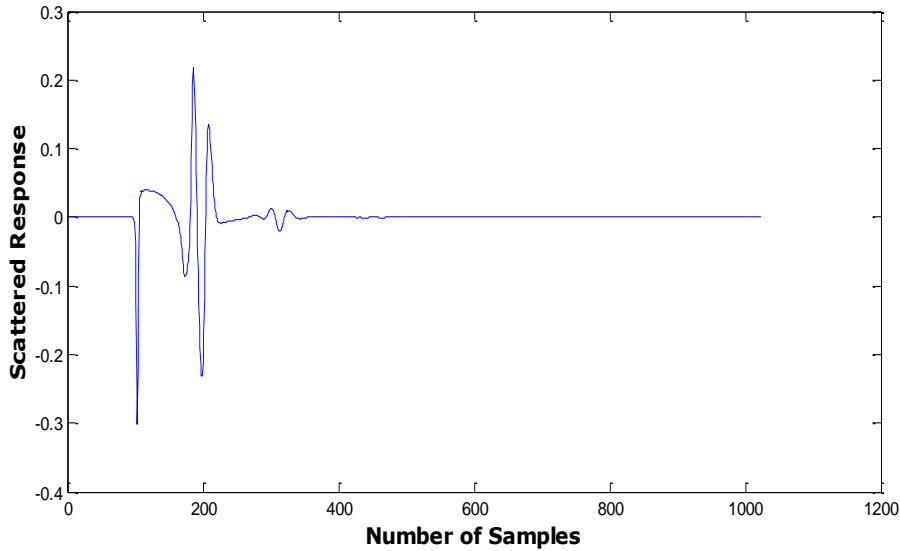
Figure 3.14 Scattered time domain response of the target T1 in the target library of the classifier C2 at 179° aspect angle
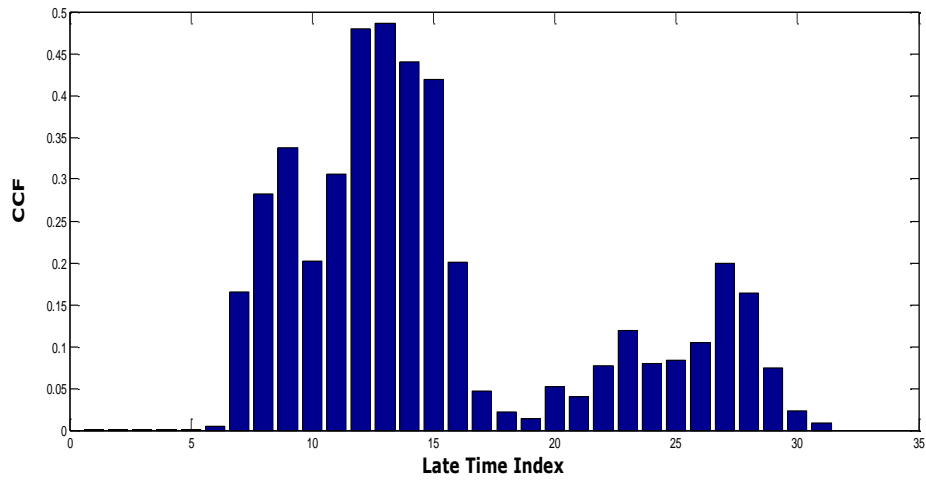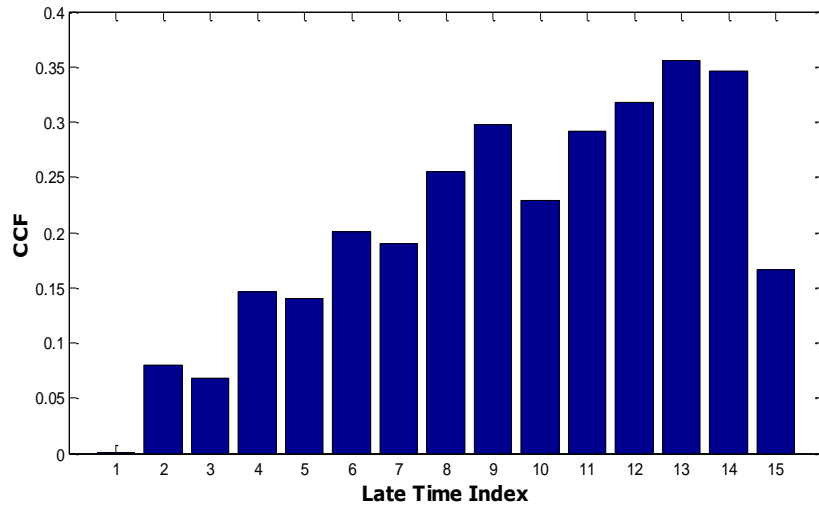
Figure 3.15 CCF plotted against the late time interval index q for the classifier C2
(using the WD-PCA method)
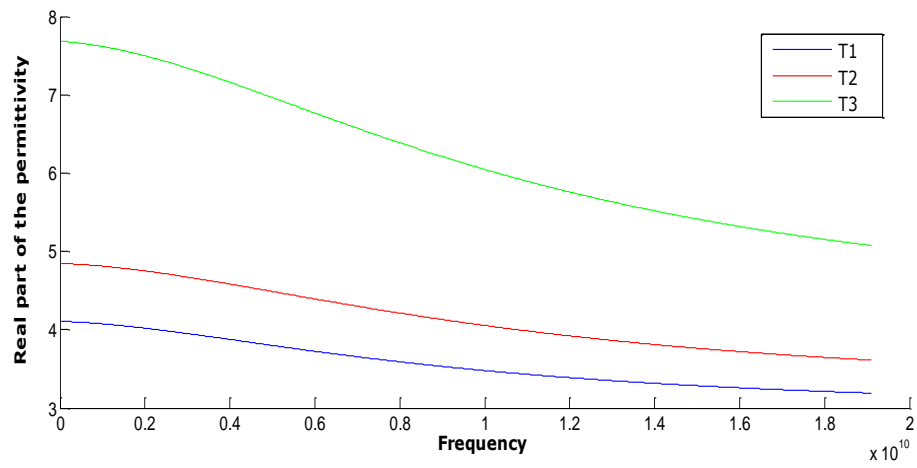
Figure 3.16 Frequency dependent ε' curves for the targets of the classifier C2
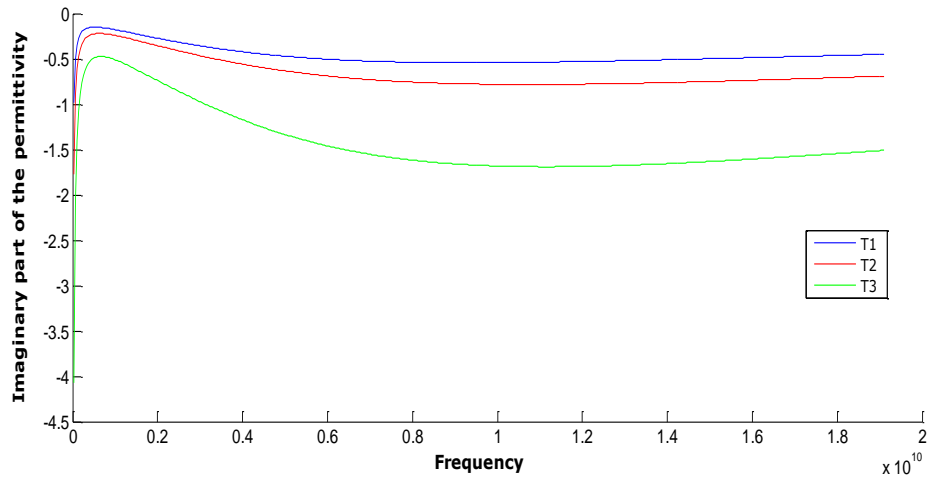
Figure 3.17 Frequency dependent ε'' curves for the targets of the classifier C2

Table 3.9   Cole-Cole parameters of tissue samples in the target library of the classifier C2

Classifier-2 (C2)			
Parameters	**Target-1 (T1)**	**Target-2 (T2)**	**Target-3 (T3)**
ε_∞	2.908	3.14	4.031
$\Delta\varepsilon$	1.2	1.708	3.654
$\tau(ps)$	16.88	14.65	14.12
α	0.069	0.061	0.055
$\sigma_s(S\,m^{-1})$	0.02	0.036	0.083
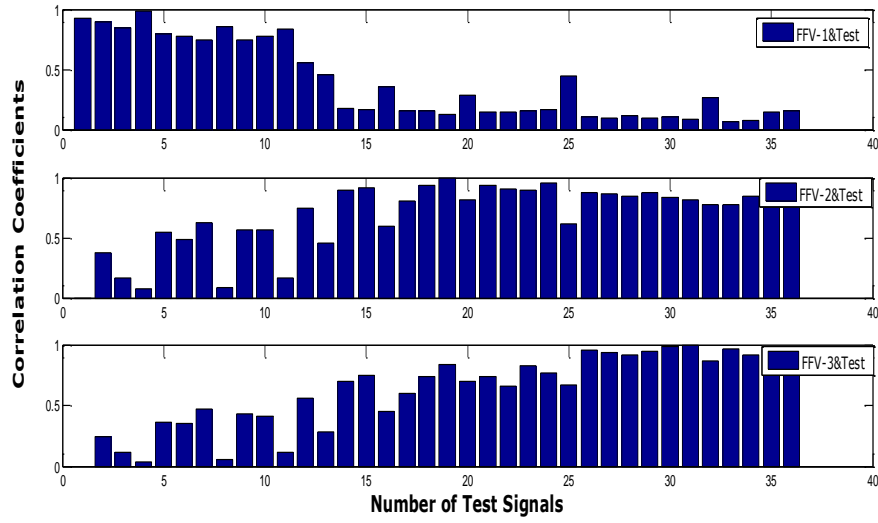
Figure 3.18 Correlation coefficients computed between the FFVs of targets T1, T2 and T3 of the classifier C2 and the LTFVs of all test signals (3 target x 12 angle=36 test signals)

3.1.1.3. Design and Test Results of Classifier C3

According to [55], 0-30% adipose content healthy breast tissue samples obtained from reduction and cancer surgeries have very similar Cole-Cole parameters. However, when the classifier C3 (see Table 3.10) is designed with Q=32, q*=12 (see Figure 3.19), 97.22 percent classification accuracy is obtained to distinguish these six spherical tissue samples with low adipose content.

Contour plot of computed correlation coefficients (Figure 3.20) showed that targets in Table 3.1_Group1 (T1, T2, T3) and the targets in Table 3.2_Group1 (T4, T5, T6) can be distinguished from each other easily within each group but T3 and T5 turn out

to be indistinguishable in the large group. This result is consistent with the results reported in [55]. It is stated in this reference that Cole-Cole parameters of low adipose content healthy breast tissue samples obtained from reduction surgeries and from cancer surgeries are very similar. Based on the test results of classifier C3 we see that if we increase the number of low adipose content samples from three to six, the classifier's accuracy rate is reduced slightly from 100 percent to about 97.22 percent.

Table 3.10   Cole-Cole parameters of tissue samples in target library of the classifier C3

Classifier-3 (C3)						
Parameters	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**
ε_∞	9.941	7.821	6.151	5.013	7.237	7.816
$\Delta\varepsilon$	26.6	41.48	48.26	40.6	46	50.21
$\tau(ps)$	0.003	0.047	0.049	0.091	0.049	0.055
α	0.462	0.713	0.809	0.607	0.808	0.899
$\sigma_s(S\ m^{-1})$	10.9	10.66	10.26	10.16	10.3	10.47
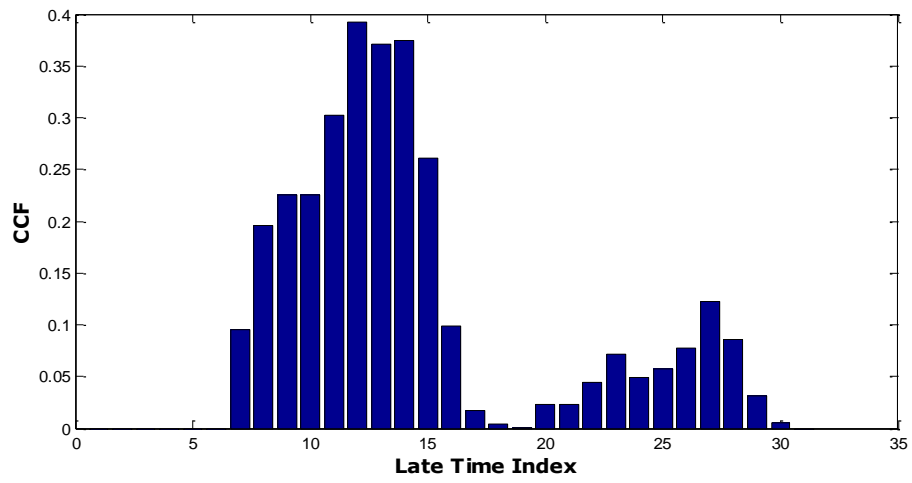
Figure 3.19 CCF plotted against the late time interval index q for the classifier C3 (using the WD-PCA method)
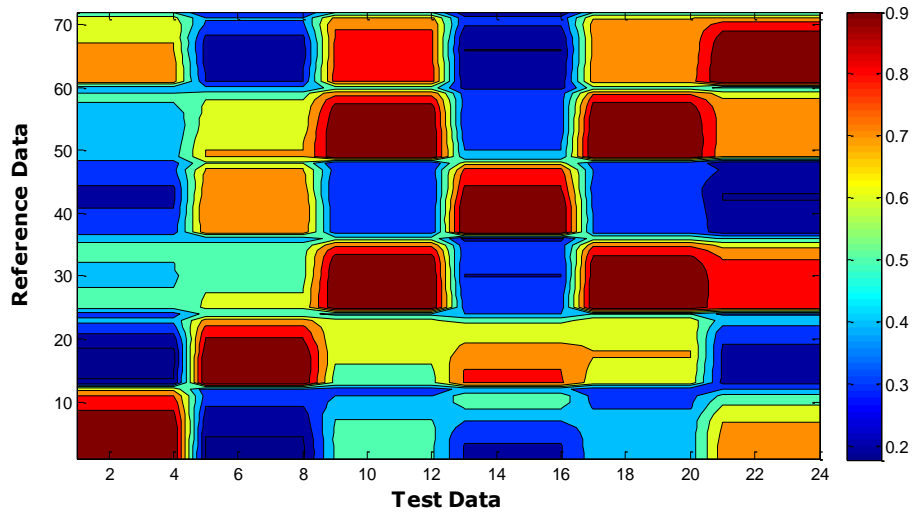

Figure 3.20 Contour plots of the correlation coefficients computed for all possible pairs of test and reference feature vectors for the classifier C3

3.1.1.4. Design and Test Results of Classifier C4

Next, one more tissue sample T7 is added to the target library of classifier C3, and then the classifier C4 is designed with Q=32, q*=12. T7 is characterized by tumor parameters given in Table 3.3. The aim of the classifier C4 is to distinguish the tumor sample from the other samples with 0-30% (low) adipose content. The accuracy rate obtained for C4 is found to be 100 percent. Namely, the tumor tissue can be perfectly distinguished from the other six low-adipose content healthy breast tissue samples. Results presented in Figure 3.22, Figure 3.23 and Figure 3.24 show that tumor sample is more similar to the 0-30% (low adipose content) healthy breast tissue samples obtained from cancer surgeries. Similarity between the scattered field of the tumor sphere and the scattered field of the low adipose content breast tissue spheres can be seen by comparing Figure 3.5 and Figure 3.21. This result also consistent with the results given in [55] which states that the difference between the dielectric properties of the malignant tissues and low adipose content healthy breast tissues is very small, no larger than approximately 10%. As a final information, the accuracy rate of the classifier C4 is found to be 95 percent for these seven spherical tissue samples belonging to low adipose content healthy tissues and the malignant tissue.
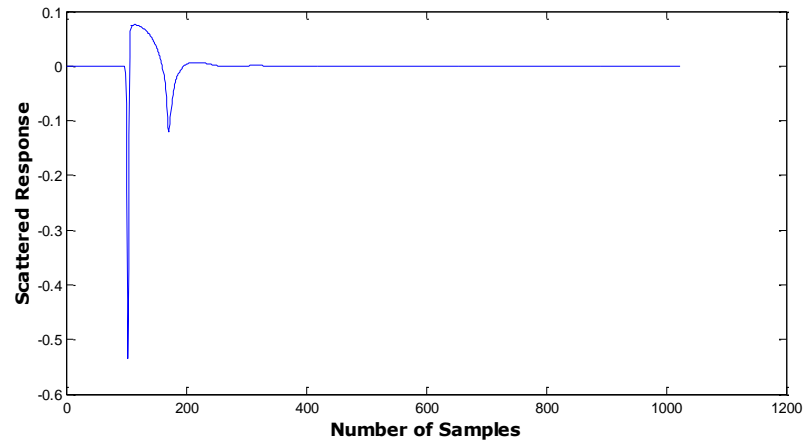
Figure 3.21   Scattered time domain response of the spherical tumor sample at   the aspect angle of $\theta = 179°$
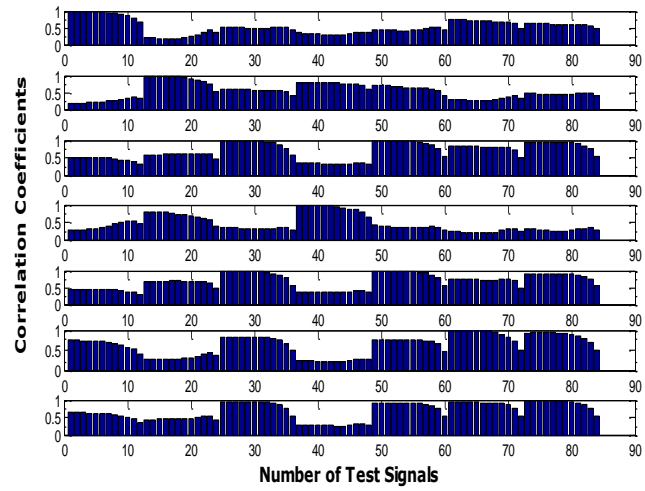
Figure 3.22 Correlation coefficients computed between the FFVs of targets (T1, T2, T3, T4, T5, T6 and T7 in order) and all the LTFVs of test signals (7 target x 12 angle=84 test signals) for the classifier C4
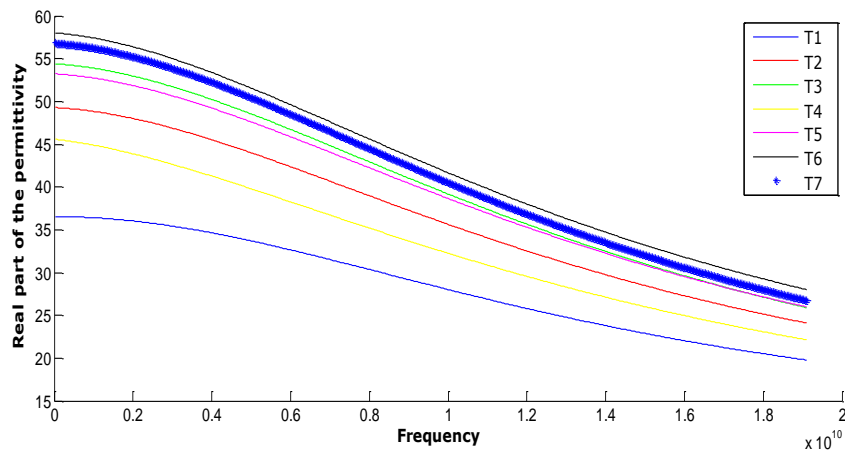
Figure 3.23 Frequency dependent ε' curves for the targets of the classifier C4
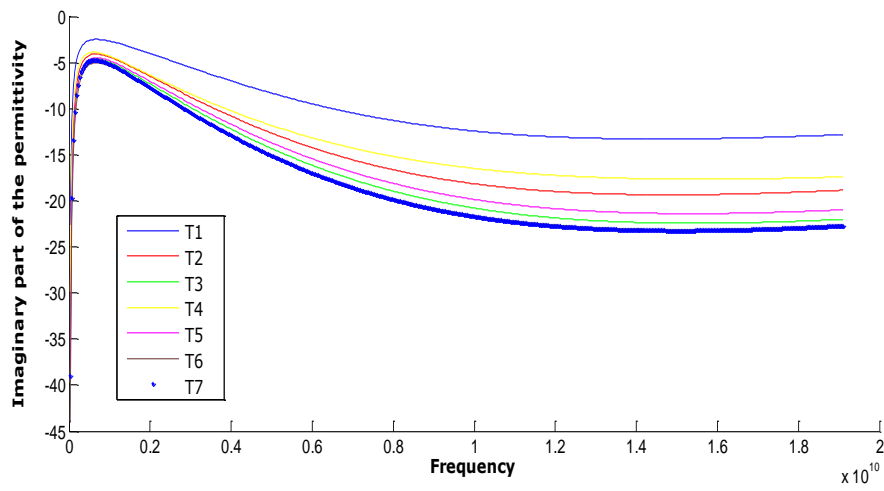
Figure 3.24 Frequency dependent ε'' curves for the targets of the classifier C4

3.1.1.5. Design and Test Results of Classifier C5

Classifier C5 is designed for six spherical healthy breast tissue samples that have 85-100% (high) adipose content obtained from reduction and cancer surgeries. By the way, number of targets for C2 is increased three to six. C5 (see Table 3.11) is designed with Q=16, q*=8. Classifier's accuracy rate is reduced from 91.66% to 79.16% sharply. Targets of C5 (and C2) have relatively lower permittivity values (see Figure 3.25 and Figure 3.26) as compared to the targets of C3 (see Figure 3.23 and Figure 3.24).

Table 3.11 Cole-Cole parameters of tissue samples in target library of the classifier C5

Classifier-5 (C5)						
Parameters	T1	T2	T3	T4	T5	T6
ε_∞	2.908	3.14	4.031	3.122	3.581	3.882
$\Delta\varepsilon$	1.2	1.708	3.654	2.133	3.337	5.02
$\tau(ps)$	16.88	14.65	14.12	0.099	0.052	0.059
α	0.069	0.061	0.055	0.034	0.053	0.103
$\sigma_s(S\,m^{-1})$	0.02	0.036	0.083	14.27	15.21	12.92
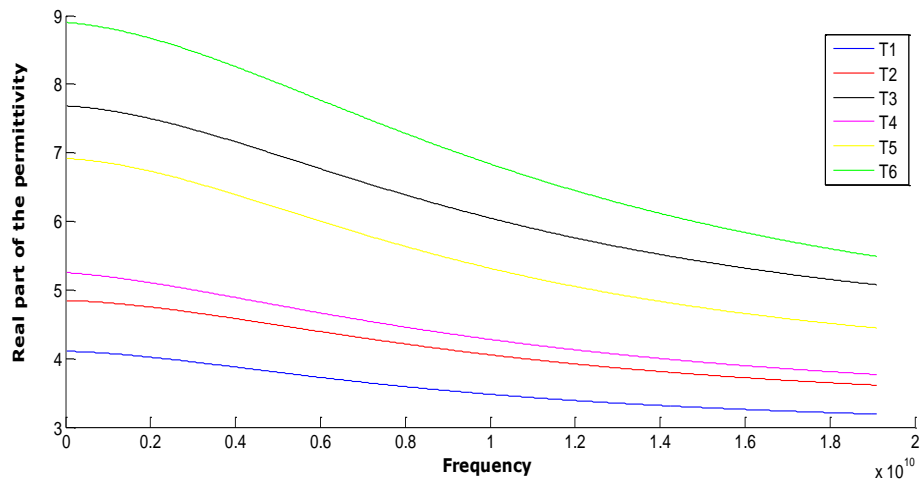
Figure 3.25 Frequency dependent ε' curves for the targets of the classifier C5
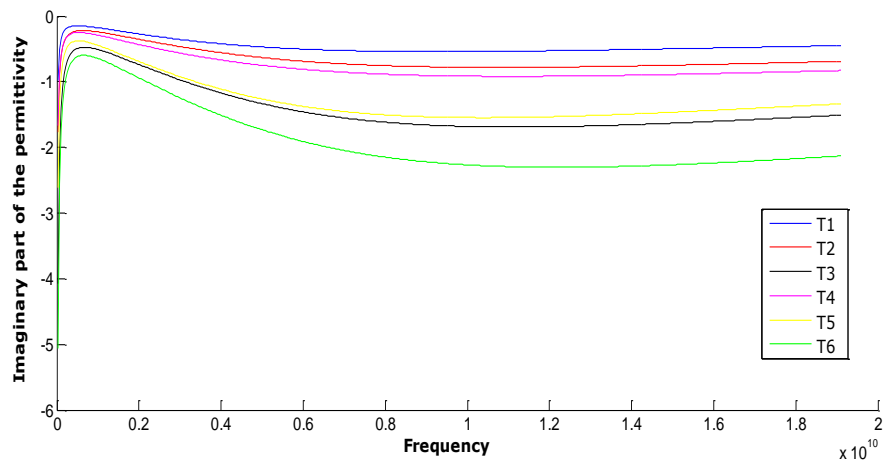
Figure 3.26 Frequency dependent ε'' curves for the targets of the classifier C5

3.1.1.6. Design and Test Results of Classifier C6

Next, we added one more tissue sample T7 to the target library of C5 to design the classifier C6. T7 is the tumor sample whose parameters are given in Table 3.3. The classifier C6, designed with parameters Q=16, q*=8. The purpose is to distinguish the tumor sample from the healthy breast tissue samples with high adipose content. The accuracy rate of the classifier has turned out to be 100 percent.

This means that we can distinguish the tumor sample from the other six samples with high adipose content perfectly. As seen from Figure 3.27 and Figure 3.28, the spherical target T7 which is characterized by tumor parameters is not similar to the high adipose content healthy tissue samples in contrast to the case for C4. This result is also consistent with the findings of [55] which states that, dielectric properties of the malignant tissues are in high contrast to the normal adipose dominated breast tissues, in the order of 10 to 1.
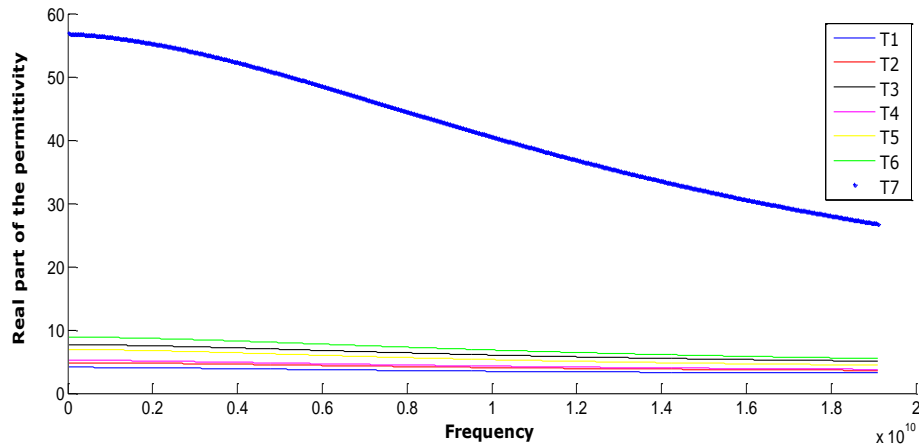
Figure 3.27 Frequency dependent ε' curves for the targets of the classifier C6
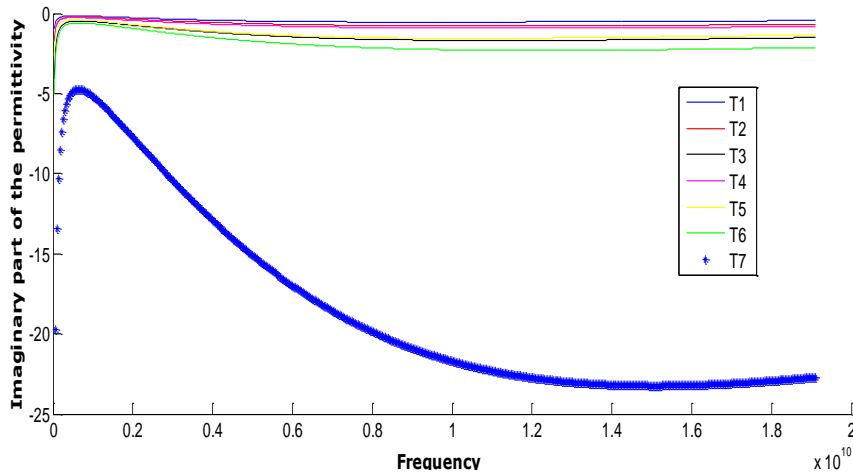
Figure 3.28 Frequency dependent ε'' curves for the targets of the classifier C6

3.1.1.7. Design and Test Results of Classifier C7

C7 is designed to classify the spherical healthy tissue samples with moderate (31-84%) adipose content obtained from reduction and cancer surgeries (see Table 3.4). The parameters Q= 32 and q*=15 (Figure 3.29) are used to design C7. As an example of moderate-adipose content tissue's time domain scattered signal, first target's scattered response at 179° aspect angle is given in Figure 3.30. Accuracy rate of the classifier is found as 84.72% for moderate-adipose content healthy breast tissues.
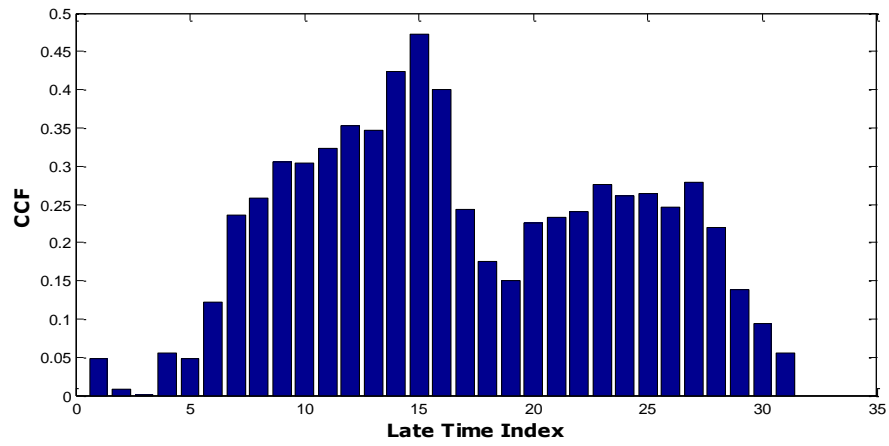
Figure 3.29 CCF plotted against the late time interval index q for the classifier C7 (using the WD-PCA method)
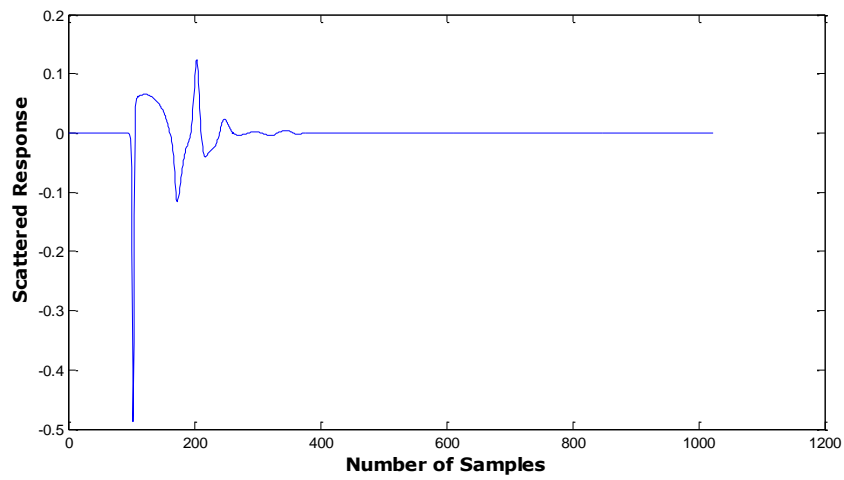
Figure 3.30 Scattered time domain response of the target T1 in the target library of the classifier C7 at 179° aspect angle

3.1.1.8. Design and Test Results of Classifier C8

Next, the classifier C8 is designed to distinguish the tumor sample from different healthy breast tissue samples with adipose contents of 0-30%, 31-84% and 85-100% obtained from breast reduction surgeries. In this design simulation, the parameters Q=32 and q*=14 are selected as optimum design parameters (see Figure 3.31). As mentioned earlier tumor parameters are much more similar to the parameters of 0-30% adipose content normal tissue samples obtained from reduction surgeries than adipose dominated samples. It is also seen in Figure 3.32 and Figure 3.33. In this classifier simulation, the tumor sample is distinguished from the samples of 85-100% adipose content easily. Accuracy rate of the classifier C8, within its target library of 10 spherical samples, is found to be 86.7 percent approximately.
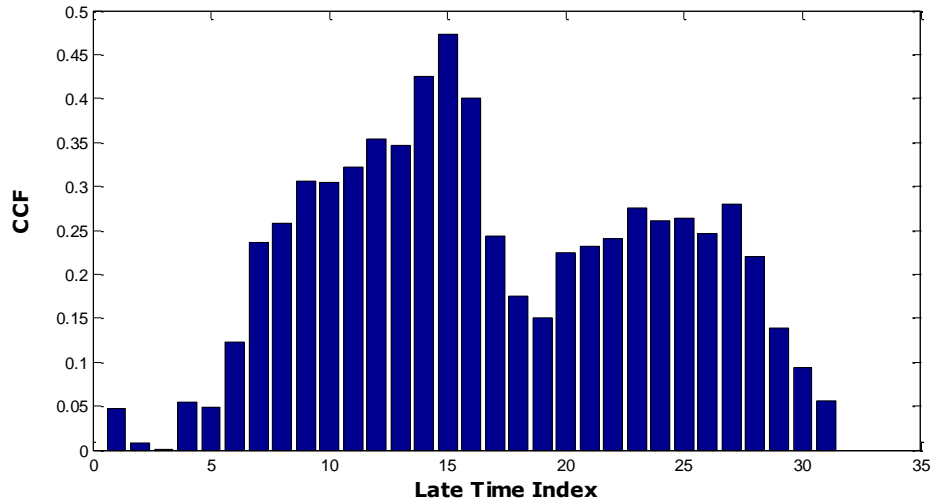
Figure 3.31 CCF plotted against the late time interval index q for the classifier C8 (using the WD-PCA method)
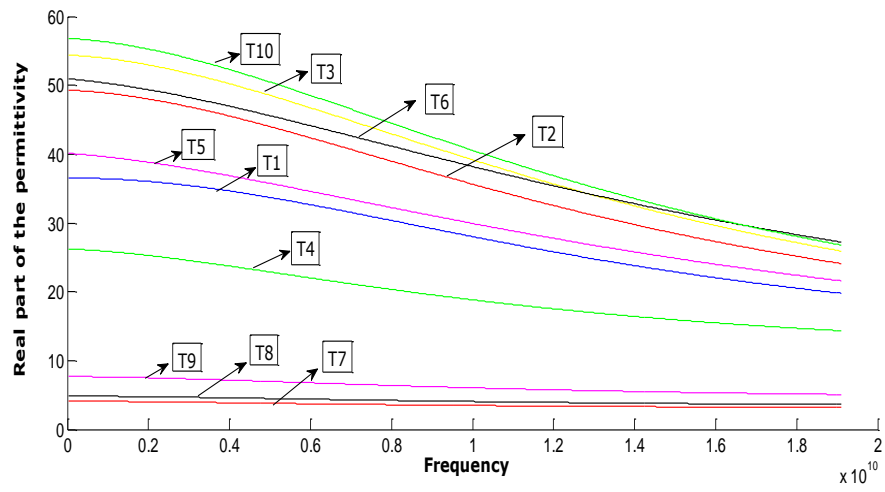
Figure 3.32 Frequency dependent ε' curves for the targets of the classifier C8
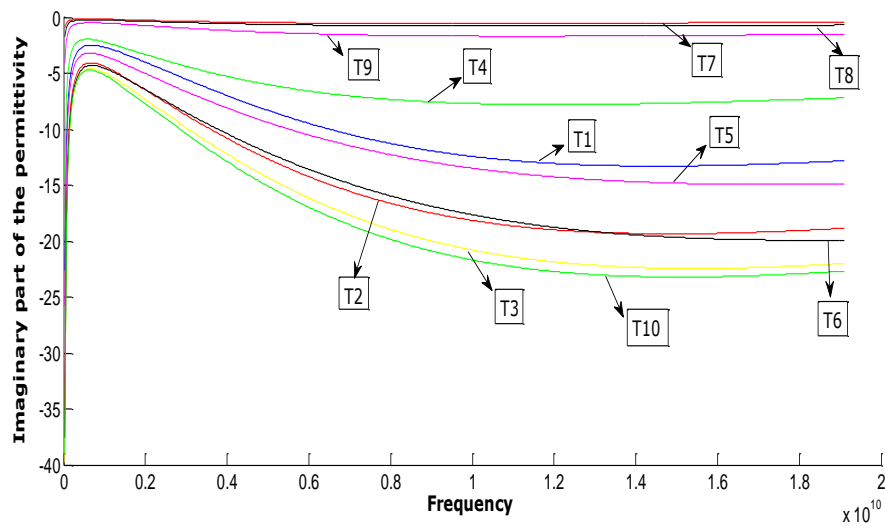
Figure 3.33 Frequency dependent ε'' curves for the targets of the classifier C8

In Figure 3.32 and Figure 3.33, real and imaginary parts of permittivity are plotted against frequency for 10 different samples which are T1, T2, T3 (low-adipose healthy tissues), T4, T5, T6 (moderate-adipose healthy tissues), T7, T8, T9 (high-adipose healthy tissues) and T10 representing the tumor tissue.

3.1.1.9. Design and Test Results of Classifier C9

Finally, all spherical tissue samples are collected to design the classifier C9 with 19 targets using the design parameters Q=32, q*=14. In this design, the aim is to recognize the tumor sample among the whole tissue sample set. The Correct Classification Factor (CCF) bar chart generated for this classifier design is shown in Figure 3.34 and the accuracy rate of the resulting classifier is found to be 97.3 percent.

Therefore, we got a successful high performance result with WD-PCA algorithm based method. The tumor sphere is recognized into others even if they have low or high adipose contents.
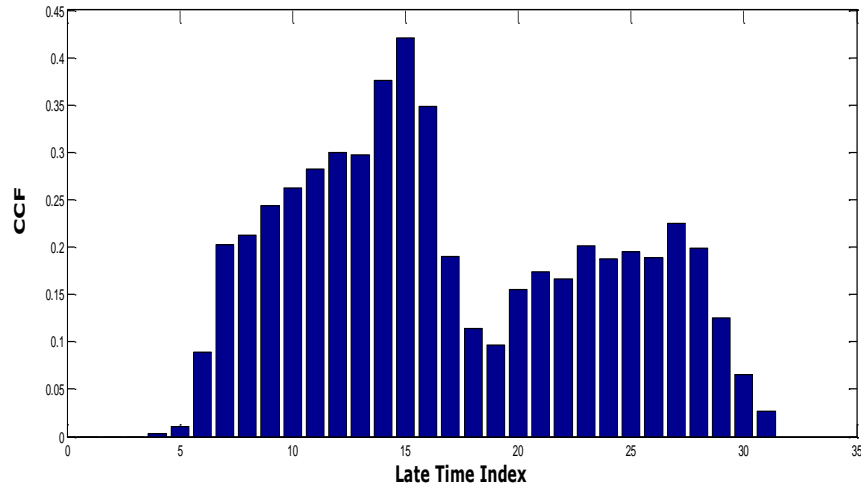
Figure 3.34 CCF plotted against the late time interval index q for the classifier C9 (using the WD-PCA method)

3.1.2. Classifier Design and Test Results with the MUSIC Algorithm Based Method for Lossy and Dispersive Breast Tissue Samples

Classifier designs presented in Section 3.1.1 are going to be repeated in this section using the MUSIC Algorithm based design approach.

3.1.2.1. Design and Test Results for MUSIC Algorithm Based Classifiers to Classify Low, Moderate, High Adipose Content Healthy Breast Tissues and Malignant Tissue Samples

All classifiers indicated in Table 3.4 were designed by WD-PCA method. Now, these classifiers will also be designed using the MUSIC algorithm based method which is explained in section 2.5.2 so we will be able to compare performances of MUSIC and WD-PCA algorithm based methods in the same scenarios. In this subsection, design parameters of the MUSIC algorithm are chosen as N=128, m=64, L=32. As an example of sub-steps of the algorithm, CCF values computed to determine the optimum late time design intervals (q:only one interval used for MUSIC algorithm) are given in Figure 3.35 for Classifier-2. Cole-Cole parameters used for C2 are given in Table 3.9**.** Resulting FMSM features of the sample targets in C2 database are given in Figure 3.36 (a)-(c) and MSM map of the test signal belonging to target T1 of classifier C2 at 165° aspect angle is given in Figure 3.36 (d).

Music Spectrum Matrices (MSMs) are computed at each reference aspects for each subinterval. Then MSMs are superposed as explained in part 2.5.2 for each given target to obtain the FMSMs over each subinterval. To design the classifier C2, time spans are divided into Q=16 overlapping subintervals of length N=128 and q=13 (16.41ns-19.74ns) is selected as optimal late time design interval (Figure 3.35).

Figure 3.35   CCF plotted against the late time interval index q for the classifier C2 (using the MUSIC method)

As seen in the Figure 3.35 the CCF value for q=15 is a little bit higher than the CCF value for q=13 at a later time interval but the classifier performance is known to get worse due to smaller signal to noise ratio levels at very late time intervals. So, q=13 is selected for the optimal late time design interval. For smaller values of q, the CCF values become smaller since these intervals contain forced scattered response components in addition to the superposition of damped sinusoidal CNR components as explained in part 2.3.

The bar chart figures plotting CCF against the late time interval index (q) give an idea to select the optimum late time interval for the classifier design. In general, if we get similar CCF values for some intervals, design procedure is repeated for these intervals to decide the optimum late time interval. As mentioned earlier, the intervals related to very late time regions give unsatisfactory results as the scattered signals (used in classifier design) decay quickly in time. This fact makes the classifier

performance vulnerable to noise effect if very late time intervals are used in the design process.

In the MUSIC algorithm based method, a target is recognized by the similarity of its MSM (obtained from the available test signal) to its own FMSM stored in the classifier database as explained in section 2.5.2. For example, the MSM map shown in Figure 3.36 (d) belongs to target T1 of classifier C2 and it has the highest similarity to the FMSM map given in Figure 3.36 (a) which is in fact the fused MSM feature of T1. The other FMSM features shown in Figures 3.36 (b) and 3.36 (c) belong to targets T2 and T3, respectively. Coordinates of the target in the normalized complex frequency domain and the strength of features at these pole locations can be obtained from these FMSM and MSM features as given in Table 3.12.

(a)
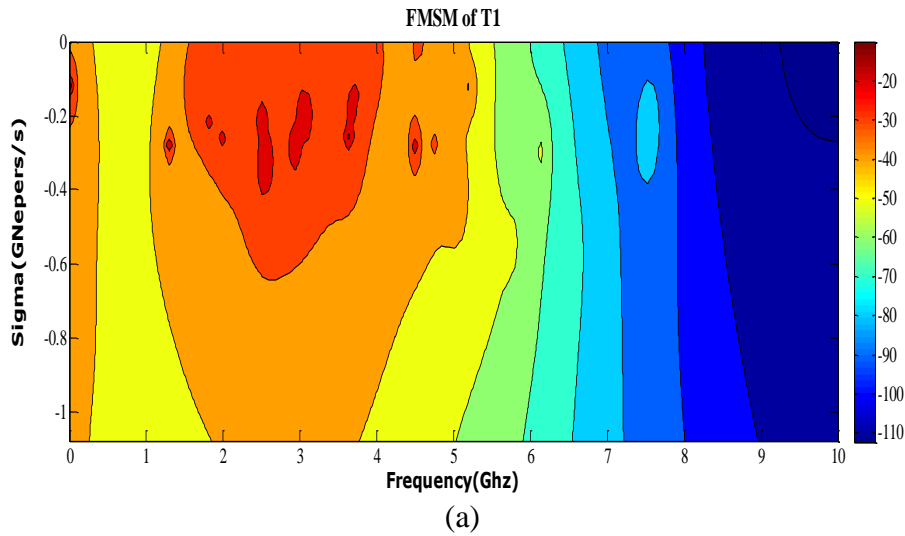
Figure 3.36 (a)-(c) Contour plots for the FMSM features of the targets (tissue samples) T1, T2 and T3 in the feature database of classifier C2 (d) The contour plot for the MSM feature of a test signal which belongs to T1 at the aspect angle $\theta = 165°$

(b)

(c)

Figure 3.36 (continued)

MSM of Test Target
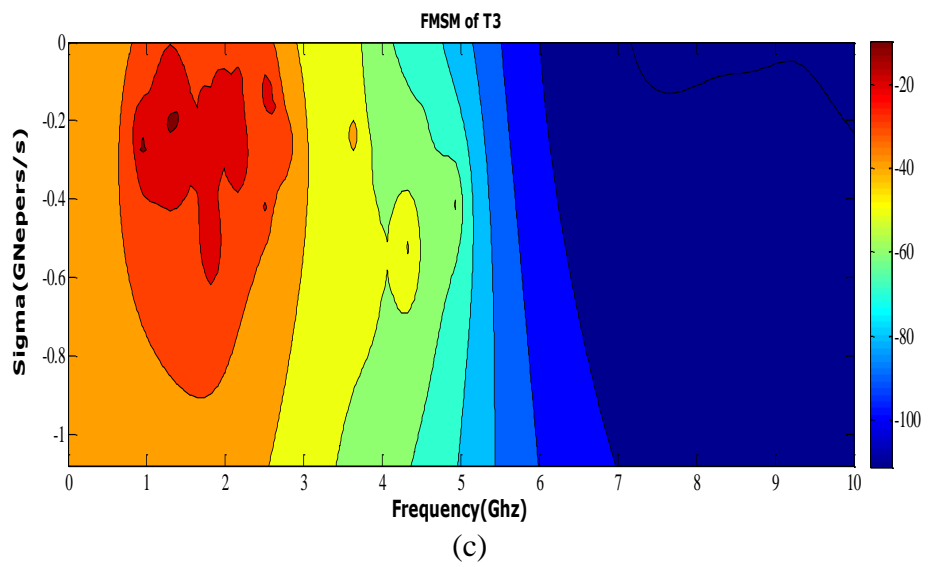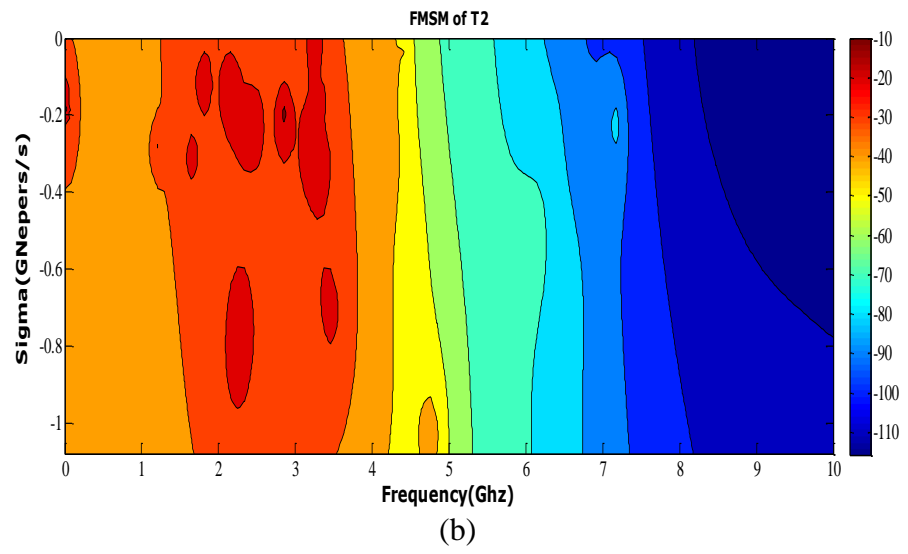
(d)
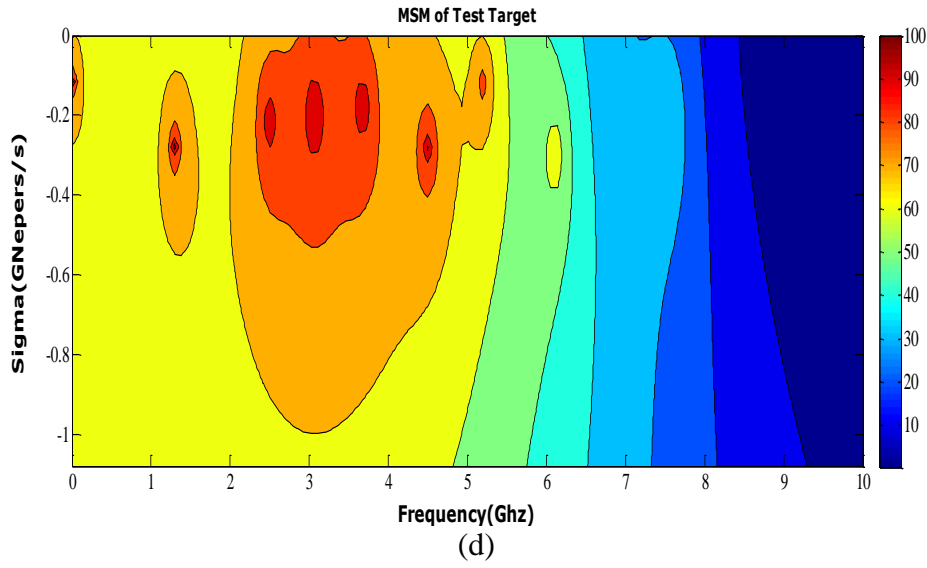
Figure 3.36 (continued)

Table 3.12   Coordinates of the target pole locations in the normalized complex frequency domain $\bar{s} = \dfrac{r}{c}(\sigma + j\omega)$ and the feature strengths (z-values) at these pole locations for the FMSM classifier features and MSM test feature

FMSM_T1	x: 0 y:-0.11 z:-1.17	x: 1.29 y:-0.28 z:-9.8	x: 1.81 y:-0.21 z:-13.8	x: 1.98 y:-0.25 z:-14.5	x: 2.5 y:-0.21 z:-11.7	x: 3 y:-0.20 z:-14
	x: 3.63 y:-0.25 z:-7	x: 3.71 y:-0.18 z:-18	x:4.49 y:-0.28 z:-12.95	x: 4.75 y:-0.27 z:-26.7	x: 4.58 y:-0 z:-29.5	x: 5.18 y:-0.115 z: 29.6
FMSM_T2	x: 0 y:-0.18 z:-2.62	x: 1.64 y:-0.31 z:-11.7	x: 1.81 y:-0.12 z:-15.1	x: 2.33 y:-0.22 z:-12.8	x: 2.33 y:-0.73 z:-17.9	x: 2.85 y:-0.19 z:-7.25
	x: 3.45 y:-0.72 z:-18.33	x: 3.19 y:-0.25 z:-12.2				

68

Table 3.12 (cont'd)

FMSM_T3	x: 0.95 y:-0.27 z:-4.43	x: 1.29 y:-0.20 z:-8.93	x: 2.59 y:-0.16 z:-0.13	x: 2.5 y:-0.41 z:-17.7	x:3.63 y:-0.23 z:-34.6	x: 4.32 y:-0.54 z:-39.3
MSM_Test	x: 0 y:-0.11 z:107.6	x: 1.29 y:-0.28 z:104.2	x: 2.50 y:-0.21 z: 99.6	x: 3 y:-0.20 z: 98.2	x: 3.71 y:-0.18 z: 94.67	x: 4.49 y:-0.28 z:101.4
	x: 5.18 y:-0.115 z: 84.56					

The MUSIC algorithm based method leads to 100 percent correct classification rate for C2. It means that we could distinguish targets T1, T2 and T3 of C2 from each other successfully in an aspect independent manner. In this example, the MUSIC algorithm based method gives higher accuracy rate then WD-PCA based method.

When we design the classifier C1 using the MUSIC algorithm based method with the design parameters Q=16, q=7 over the late time interval [10.05ns-11.72ns], (see Figure 3.37), we get 91.66 percent accuracy rate which is lower than the accuracy rate of 100 percent obtained with the WD-PCA based algorithm for classifier C1.
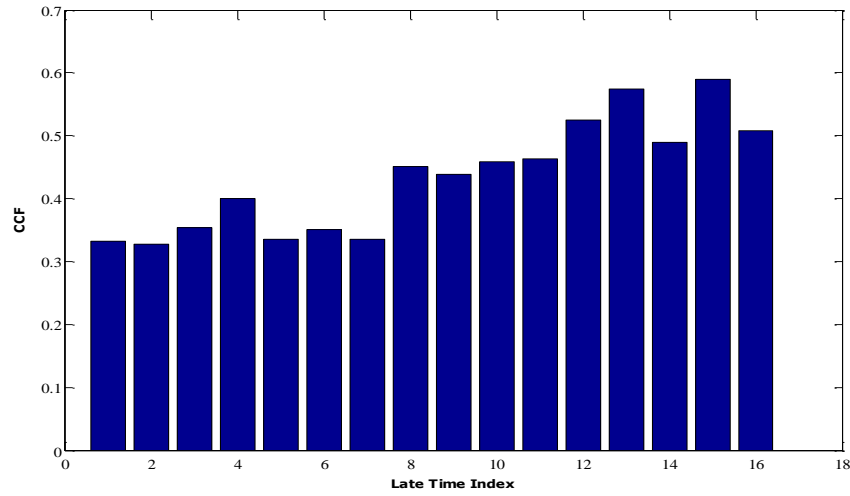
Figure 3.37 CCF plotted against the late time interval index q for the classifier C1 (using the MUSIC method)

Next, the classifier C3 is designed with Q=16, q=9 over the late time interval [13.4 ns-15.08ns], (see Figure 3.38). This classifier contains six spherical samples from the healthy tissue groups of low (0-30%) adipose content obtained during the breast reduction and cancer surgeries. As discussed in section 3.1.1, the accuracy rate of the classifier C3 obtained by WD-PCA method was 97.22 percent. Now, using the MUSIC algorithm based method, the accuracy rate of the classifier C3 is found to be 81.94 percent. Looking at the accuracy rates of classifiers C1 and C3, we may deduce that performance of the MUSIC algorithm based method is worse as compared to the WD-PCA method on low (0-30%) adipose content targets. Furthermore, when we compare the performance results of classifiers C1 and C3, which are both designed with MUSIC method, we may come to the following conclusion: when number of targets increase for this type of tissue samples (low adipose content) classification accuracy decreases in general. For instance, for the classifier C1 the accuracy rate drops from 91.66 percent to 81.94 percent.
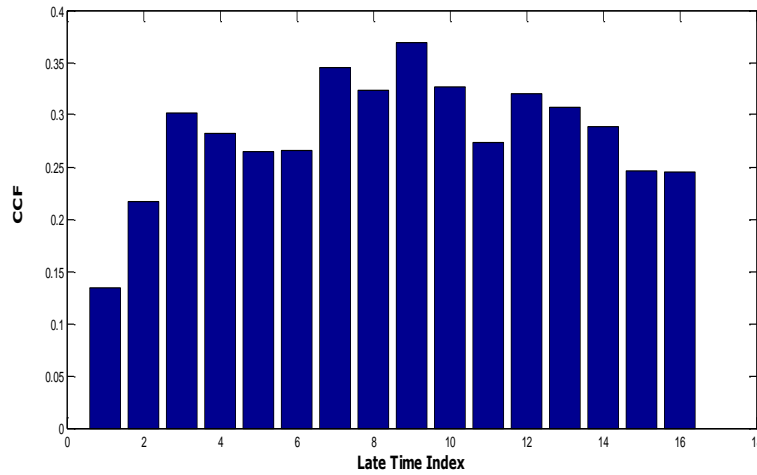
Figure 3.38 CCF plotted against the late time interval index q for the classifier C3 (using the MUSIC method)

Then, the MUSIC algorithm based classifier technique is tested to a spherical tumor tissue sample when it is added into the library of healthy breast tissue samples used in the design of classifier C3. Using the design parameters Q=16 and q=9 is determined (see Figure 3.39) for classifier design.

When the spherical tumor tissue sample is also added to this target library, the correct classification rate is found to be 94.34 percent when the classifier is designed using the MUSIC algorithm method. Earlier, when the classifier was designed by using WD-PCA method, the correct classification rate among all test cases was to be 100 percent as illustrated in section 3.1.1.
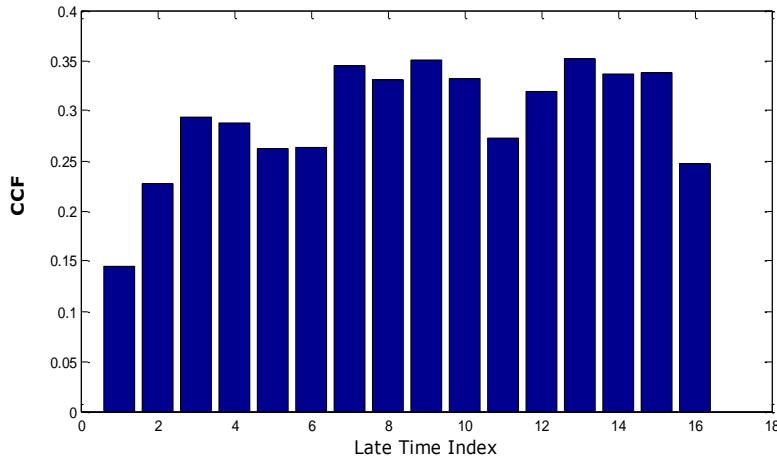
Figure 3.39 CCF plotted against the late time interval index q for the classifier C4 (using the MUSIC method)

Next, the classifier C5 designed with parameters Q=16 and q=11 over the late time interval [16.75ns-18.43ns], (see Figure 3.40). Target library of C5 contains six spherical tissue samples having 85-100 percent adipose content obtained from reduction and cancer surgeries (see Table 3.11). Correct classification rate of the classifier is found to be 88.9 percent. When the classifier was designed by using the WD-PCA method, the accuracy rate was found to be 79.16 percent. From the results of classifiers C2 and C5, it is seen that the MUSIC algorithm based method gives better accuracy rates as compared to the WD-PCA based method on high adipose content targets. Also, by comparing the results obtained for classifiers C2 and C5 (when they are designed by using MUSIC algorithm method) we can also say that, for higher number of targets with high-adipose content the accuracy rate of the classifier decreases from 100 percent to 88.9 percent.
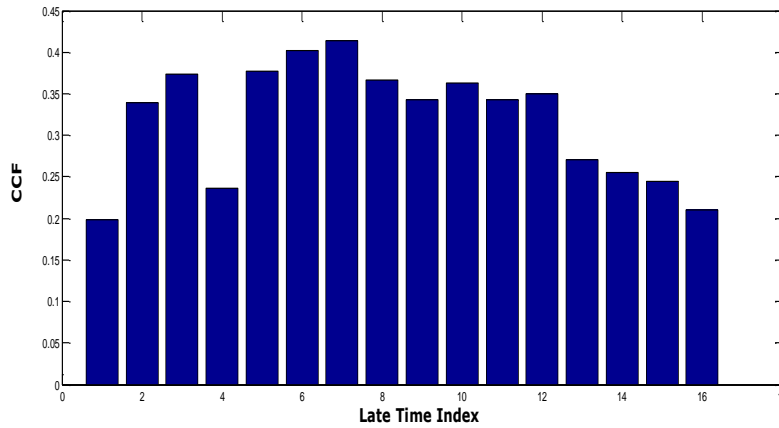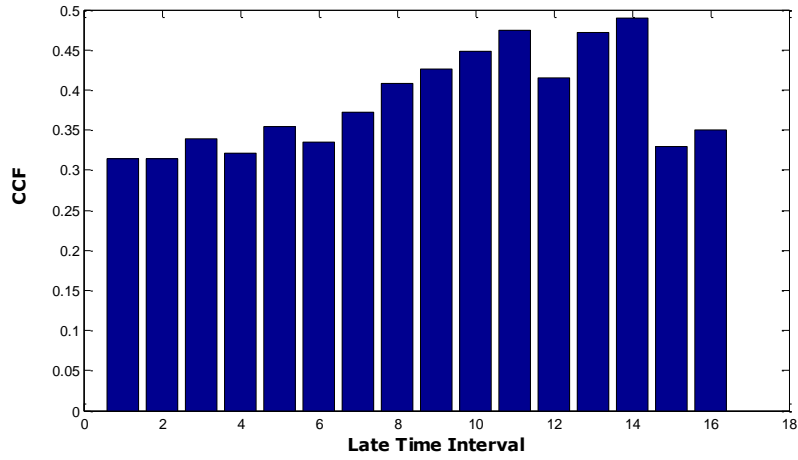
Figure 3.40 CCF plotted against the late time interval index q for the classifier C5 (using the MUSIC method)

Next, the MUSIC algorithm based classifier design technique is tested when a spherical tumor is added to the library of high-adipose content breast tissues. For this purpose, the target library of the classifier C5 is extended to include the tumor tissue as well. The resulting classifier C6 is designed with parameters Q=16, q=13 over [20.1ns-21.78ns] design interval, (see Figure 3.41). It contains six spherical samples having 85-100% adipose content normal breast tissue samples obtained from reduction surgeries and an additional spherical sample of tumor tissue. The accuracy rate of the classifier C6 (with MUSIC method) is found to be 100 percent.

When we examine the results obtained for the classifiers C4 and C6 designed by using either the WD-PCA or the MUSIC method, we can come to the following important conclusion: Recognition of the malignant sample (the tumor sphere) among the high-adipose content breast tissue samples is much easier as compared to identifying the malignant sample among the low-adipose content breast tissue samples. This conclusion looks independent of the classifier design technique used.
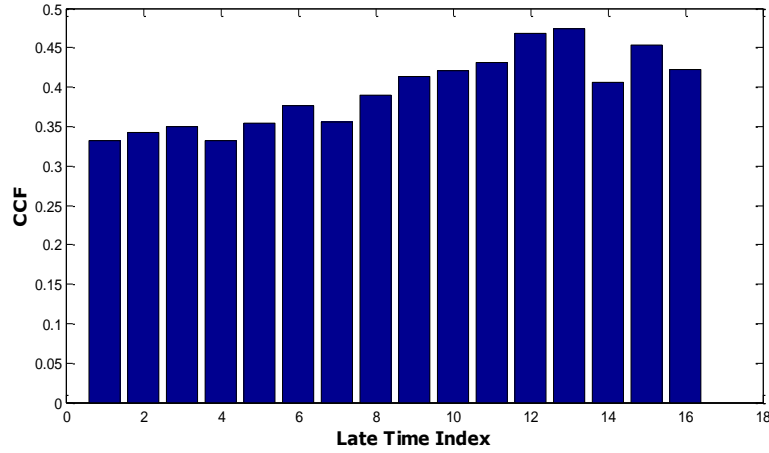
Figure 3.41 CCF plotted against the late time interval index q for the classifier C6 (using the MUSIC method)

For the sake of completeness, the classifier C7 is also designed for six slightly different moderate-adipose content breast tissue samples. The late time design interval [13.4ns-15.08ns] is obtained by using the optimal design parameters Q=16 and q= 9. The same accuracy rate, 84.7 percent, is obtained by the use of both WD-PCA method and the MUSIC algorithm method for this group of tissues having moderate fat content.

3.1.2.2. Design and Test Results for MUSIC Algorithm Based Method Classifiers with Larger Number of Tissue Samples

The classifier C8 is designed with optimal parameters Q=16 and q= 9 leading to the design interval [13.4ns-15.08ns] as obtained by the help of Figure 3.42. The library of C8 contains ten targets; 3 samples with 0-30% adipose content, 3 samples with 31-84% adipose content and 3 samples with 85-100% adipose content (all healthy breast tissue obtained from reduction surgeries) in addition to one tumor sample obtained from cancer surgeries. We got 99.16 percent accuracy rate to distinguish the tumor sphere from others.

In Figure 3.43, FMSM features belonging to samples with different (low, moderate and high level) adipose contents and belonging to the tumor sample are given to show the similarity between the tumor tissue and low (0-30%) adipose content breast tissues. As seen in this figure, the FMSM feature that belongs to the tumor sample (see Figure 3.43(d)) is more similar to the FMSM feature given in Figure 3.43(a) that belongs to the target group of low-adipose content breast tissue, as expected.

Figure 3.42 CCF plotted against the late time interval index q for the classifier C8 (using the MUSIC method)
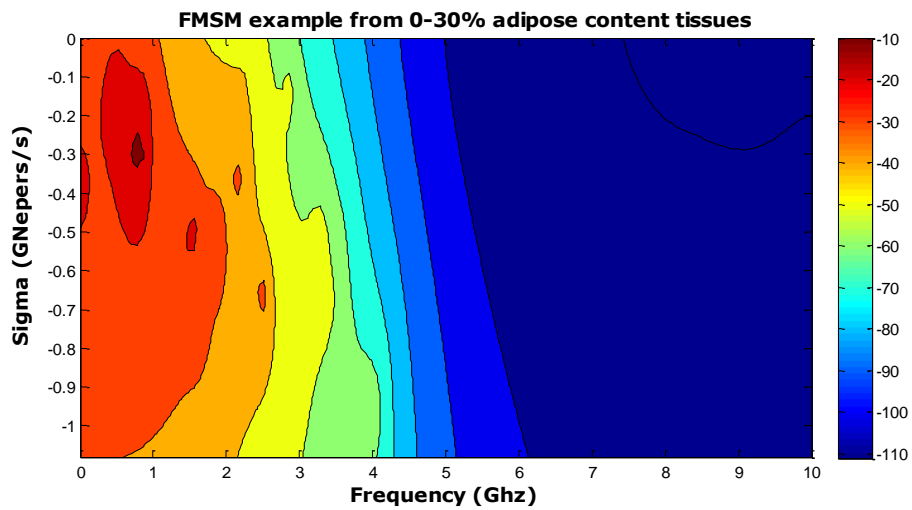
(a)

Figure 3.43 (a)-(c) Contour plots for the FMSM features of healthy breast tissues with low, moderate and high adipose contents (from the database of classifier C8) (d) MSM feature of the tumor tissue

FMSM example from 31-84% adipose content tissues

(b)

FMSM example from 85-100% adipose content tissues
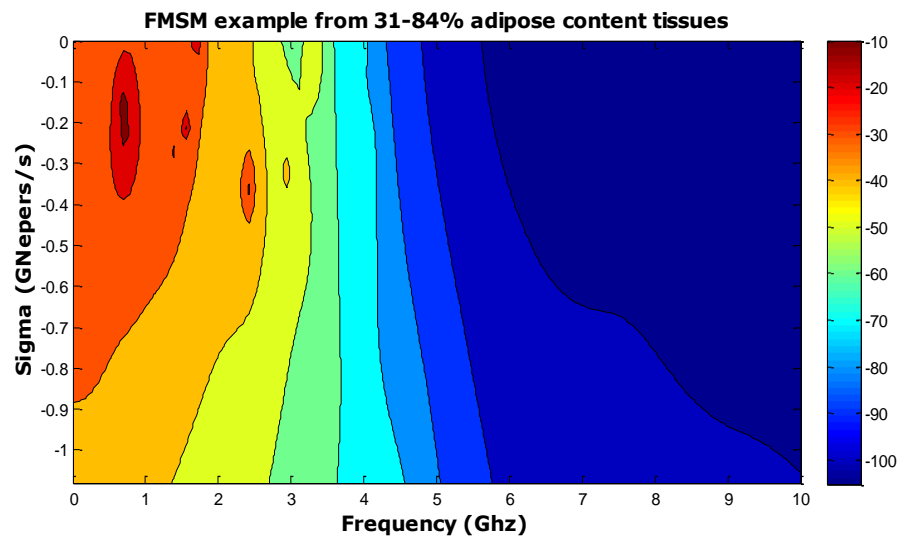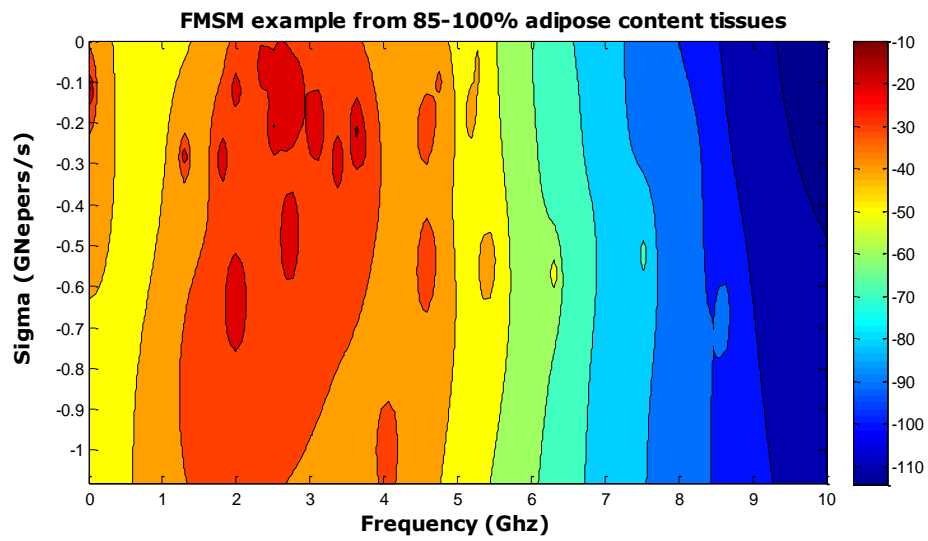
(c)

Figure 3.43 (continued)
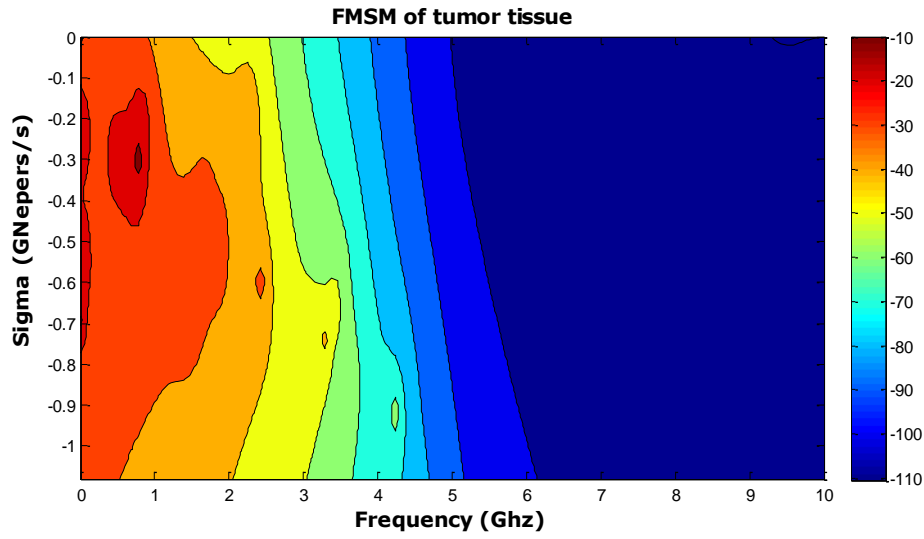
(d)

Figure 3.43 (continued)

Finally, the classifier C9 is designed for 19 different tissue/tumor samples. Parameters of these different tissue groups were given in Table 3.1, Table 3.2 and Table 3.3. Q=16 and q=9 are selected as design parameters. The accuracy rate of this classifier is found to be 97.3 percent.

In summary, the tumor sample is identified successfully among all possible breast tissue samples by using either methods (WD-PCA or MUSIC A algorithm methods) for classifier design.

3.2. Classifier Design and Test Results for Breast Tissue Samples Containing a Tumor at the Center

In this section, we have designed three classifiers (C10, C11 and C12) and repeated the designs for both WD-PCA method and the MUSIC Algorithm based method. The target geometry in all these classifier design simulations, is composed of a spherical core region (the tumor tissue) and the spherical shell coating (the breast tissue). The outer radius of the spherical shell is fixed to be 10cm. Median values of the parameters in Table 3.2 are used to characterize this coating layer. Breast tissue samples obtained from the cancer surgeries are used (instead of those obtained from the breast reduction surgeries) for the worst case design. As shown in section 3.1, tumor tissues are more similar to normal breast tissues obtained during the cancer surgeries as compared to those normal tissues obtained during the reduction surgeries. This makes classification more challenging. In summary, we used 85-100% adipose content normal tissues obtained from cancer surgeries in classifier C10; 31-84% adipose content normal tissues obtained from cancer surgeries in classifier C11 and 0-30% adipose content normal tissues obtained from cancer surgeries in classifier C12 as shell tissues.

The parameters of the tumor tissue, on the other hand, is taken from the 50 percentile column of Table 3.3. Different values for the radius of the tumor sphere are tested to figure out the "tumor detection" performance of the classifiers. The radius of the malignant core sphere is reduced to values as small as 1 mm. The resolution performance of the classifiers is investigated for different types of shell tissues. Success levels of WD-PCA based classifiers and the MUSIC algorithm based classifiers are compared.

Before getting into the details of classifier design and testing simulations, it would be useful to state the nature of the basic electromagnetic problem in this part. All the targets considered in this section are dielectric coated dielectric spheres where the

coating shells also the spherical geometry. Furthermore, both the core region and the coating layer are made of lossy and dispersive materials. It is important to keep in mind that the actual "breast tumor detection" problem is simplified by assuming perfectly spherical geometry for both the tumor and the breast region and only one tumor is assumed to exist at the center of the breast. However, the simulated problem can be easily extended to much more realistic scenarios if real world data become available.

3.2.1. Results for Classifier Design and Testing with WD-PCA Method

In this section, various classifiers will be designed using the WD-PCA Method to detect the presence of a tumor in different types of breast tissues.

3.2.1.1. Detection of a Tumor Embedded in a High-Adipose Content Breast Tissue

The Cole-Cole parameters used in the design of classifier C10 are given in Table 3.13. The tumor radius has been changed from 2.2cm to 2cm, 1.8cm, 1.5cm, 1cm, 8mm, 5mm and 1mm while the shell parameters and outer shell radius remain the same. Q=16 and q*=11 (i.e. q=11 and q=12 are selected in combination) used as design parameters based on Figure 3.44.

In this classifier design problem, for each given value of tumor radius, the classifier library has only two targets: As sphere made of the high adipose content breast tissue

and the same sphere having a tumor core in the middle. The classifier C10 is designed (for each core radius) at the reference aspects as discussed earlier in Chapter 2, and tested at all aspect angles available to compute the accuracy rates listed in Table 3.14.
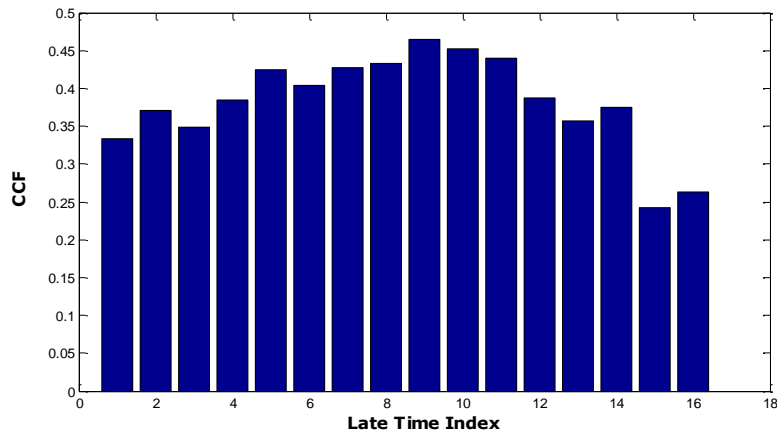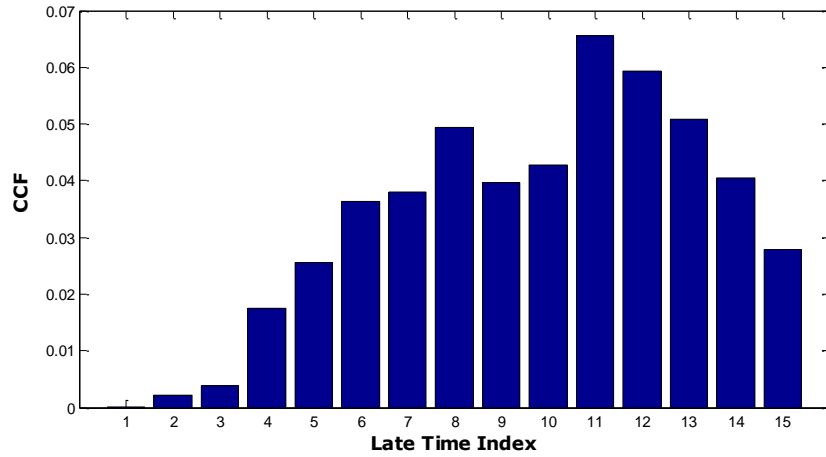

Figure 3.44 CCF plotted against the late time interval index q for the classifier C10 (using the WD-PCA method)

Table 3.13 Cole-Cole parameters used for the core (tumor) and the shell (high-adipose content breast tissue) for the design of classifier C10

Parameters	**Shell (High-Adipose content)**	**Core**
ε_∞	3.581	6.749
$\Delta\varepsilon$	3.337	50.09
τ(ps)	0.052	10.5
α	0.053	0.051
σ_s(S m^{-1})	15.21	0.794

Table 3.14 Accuracy rates for the classifier C10 (based on WD-PCA method) for different tumor sizes

Tumor Radius	Accuracy Rate
2.2cm	100%
2cm	91.66%
1.8cm	91.66%
1.5cm	91.66%
1cm	91.66%
8mm	83.33%
5mm	83.33%
1mm	83.33%

3.2.1.2. Detection of a Tumor Embedded in a Moderate-Adipose Content Breast Tissue

The classifier design and testing task explained in detail for the classifier C10 in section 3.2.1.1 is repeated here for the classifier C11 with the only difference that the spherical shell coating is composed of moderate-adipose content breast tissue here (see Table 3.15). Using the optimal design parameters Q=16 and q*=8 (see Figure 3.45), the accuracy rates for the classifier C11 are obtained and listed in Table 3.16.

Figure 3.45 CCF plotted against the late time interval index q for the classifier C11 (using the WD-PCA method)

Table 3.15 Cole-Cole parameters used for the core (tumor) and the shell (moderate-adipose content breast tissue) for the design of classifier C11

Parameters	Shell (Moderate-Adipose content)	Core
ε_∞	6.08	6.749
$\Delta\varepsilon$	19.26	50.09
$\tau(ps)$	11.47	10.5
α	0.057	0.051
$\sigma_s(S\,m^{-1})$	0.297	0.794

Table 3.16 Accuracy rates for the classifier C11 (based on the WD-PCA method) for different tumor sizes

Tumor Radius	Accuracy Rate
2.2cm	100%
2cm	100%
1.8cm	100%
1.5cm	83.33%
1cm	75%
8mm	75%
5mm	75%
1mm	75%

3.2.1.3. Detection of a Tumor Embedded in a Low-Adipose Content Breast Tissue

Classifier design and testing steps followed for classifiers C10 and C11 are repeated here to design the classifier C12 where the tumor is embedded in a low-fat content breast tissue. The Cole-Cole parameters used for the tumor and breast tissue are tabulated in Table 3.17. The WD-PCA classifier parameters are Q=16 and q*=6 (see Figure 3.46). The classifier accuracy rates obtained for different tumor sizes with multi-aspect testing are listed in Table 3.18.

Figure 3.46 CCF plotted against the late time interval index q for the classifier C12 (using the WD-PCA method)

Table 3.17 Cole-Cole parameters used for the core (tumor) and the shell (low-adipose content breast tissue) for the design of classifier C12

Parameters	Shell (Low-Adipose content)	Core
ε_∞	7.237	6.749
$\Delta\varepsilon$	46	50.09
$\tau(ps)$	0.049	10.5
α	0.808	0.051
$\sigma_s(S\,m^{-1})$	10.3	0.794

Table 3.18 Accuracy rates for the classifier C12 (based on the WD-PCA method) for different tumor sizes

Tumor Radius	Accuracy Rate
2.2cm	66.6%
2cm	58.3%
1.8cm	58.3%
1.5cm	58.3%
1cm	58.3%
8mm	58.3%
5mm	58.3%
1mm	58.3%

As seen from Table 3.14, 3.16 and 3.18, where the accuracy rates of the WD-PCA based classifiers are listed, the tumor can be distinguished within the adipose-dominated tissues successfully even if it is very small. When the adipose content of the shell tissue decrease, distinguishing the tumor becomes more difficult. This results are consistent with the conclusions stated in [55] and also consistent with the results obtained in section 3.1 where we have already shown that the tumor parameters are very similar to the parameters of low adipose content breast tissues.

3.2.2.　　Results for Classifier Design and Testing with MUSIC Algorithm Based Method

The classifiers C10, C11 and C12 designed and tested in part 3.2.1 are going to be designed by MUSIC algorithm based method now to compare the performances of these two different methods regarding the tumor detection problem.

3.2.2.1. Detection of a Tumor Embedded in a High-Adipose Content Breast Tissue

The classifier C10 is designed by using the MUSIC algorithm approach with the data given in Table 3.13 and using the design parameters Q=16 and q=10 (see Figure 3.47). Accuracy rates obtained for the resulting classifier C10 are provided in Table 3.19 for different tumor sizes.
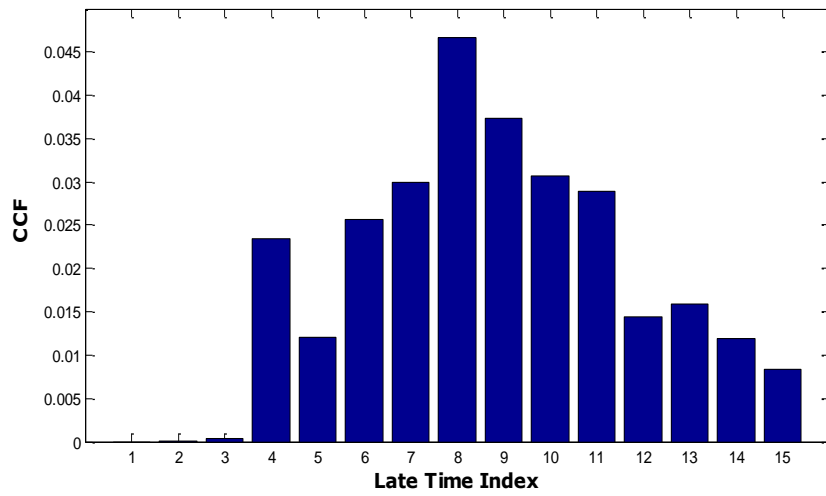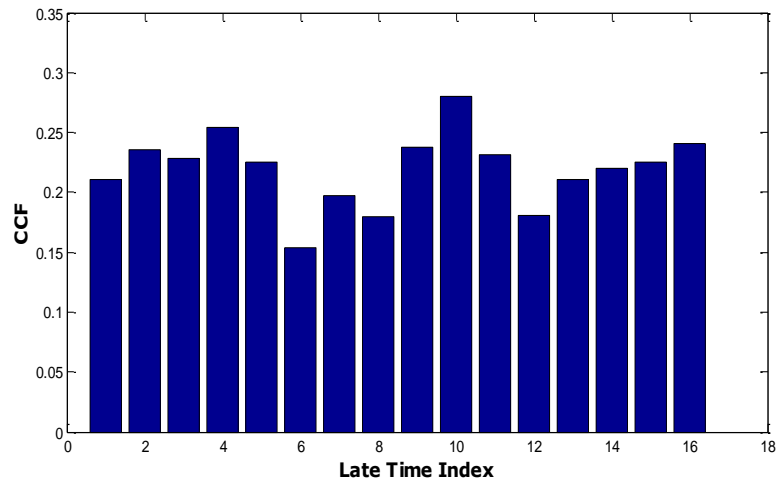

Figure 3.47 CCF plotted against the late time interval index q for the classifier C10 (using the MUSIC method)

Table 3.19 Accuracy rates for the classifier C10 (based on the MUSIC method) for different tumor sizes

Tumor Radius	Accuracy Rate
2.2cm	100%
2cm	100%
1.8cm	91.66%
1.5cm	91.66%
1cm	91.66%
8mm	83.3%
5mm	83.3%
1mm	66.6%

3.2.2.2. Detection of a Tumor Embedded in a Moderate-Adipose Content Breast Tissue

The classifier C11 is designed, by using the MUSIC algorithm based method and the Cole-Cole parameters given in Table 3.15. Q=16, q=5 are taken as optimum design parameters (see Figure 3.48). The accuracy rates obtained for the resulting classifier are listed in Table 3.20 for different tumor sizes.
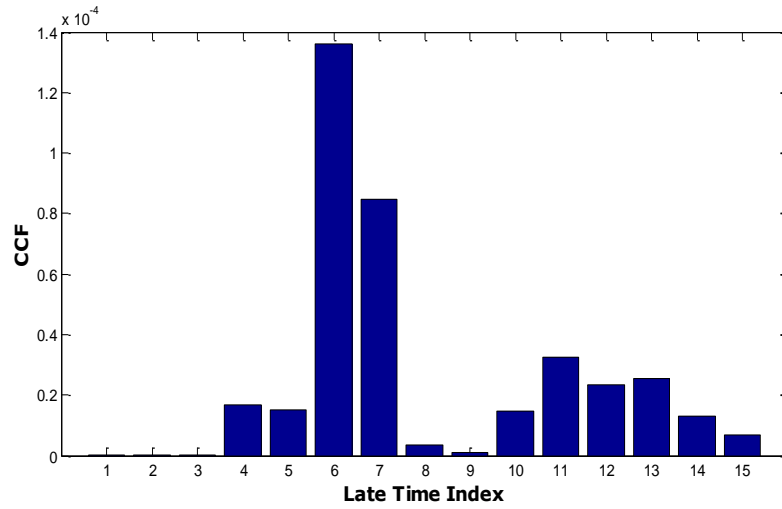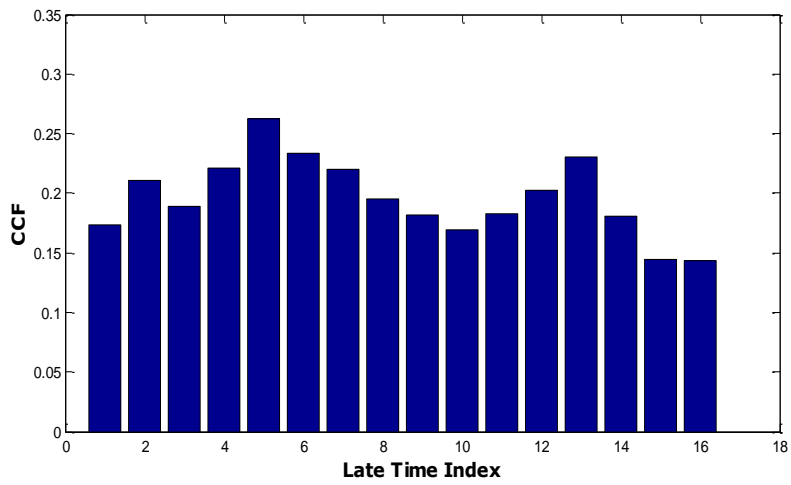
Figure 3.48   CCF plotted against the late time interval index q for the classifier C11 (using the MUSIC method)

Table 3.20 Accuracy rates for the classifier C11 (based on the MUSIC method) for different tumor sizes

Tumor Radius	Accuracy Rate
2.2cm	91.66%
2cm	91.66%
1.8cm	91.66%
1.5cm	91.66%
1cm	91.66%
8mm	91.66%
5mm	75%
1mm	75%

3.2.2.3. Detection of a Tumor Embedded in a Low-Adipose Content Breast Tissue

The classifier C12 is designed by using the MUSIC algorithm based method and the Cole-Cole parameters given in Table 3.17. Q=16, q=13 are used as design parameters (see Figure 3.49). The accuracy rates obtained for the resulting classifier are listed in Table 3.21 for different tumor sizes.
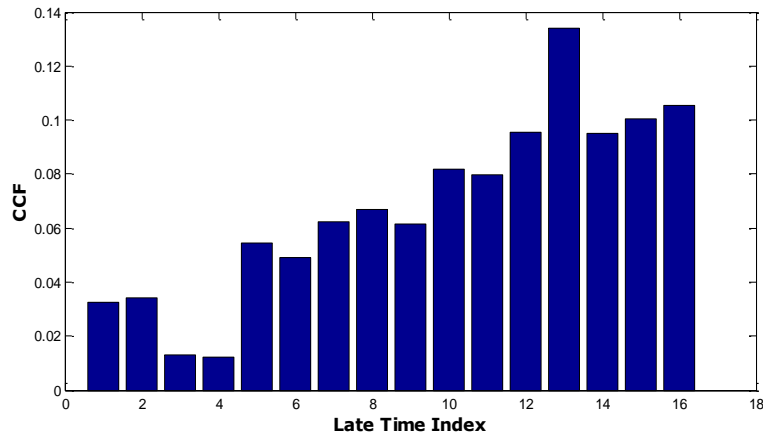
Figure 3.49 CCF plotted against the late time interval index q for the classifier C12 (using the MUSIC method)

Table 3.21 Accuracy rates for the classifier C12 (based on the MUSIC method) for different tumor sizes

Tumor Radius	Accuracy Rate
2.2cm	91.66%
2cm	83.3%
1.8cm	83.3%
1.5cm	66.6%
1cm	58.8%
8mm	58.8%
5mm	58.8%
1mm	58.8%

All the accuracy rates computed so far for different tumor sizes (based on the use of both classifier design methods, WD-PCA and MUSIC algorithm approaches, are plotted against the tumor (core) radius in Figure 3.50, Figure 3.51 and Figure 3.52 for the classifiers C10, C11 and C12, respectively.
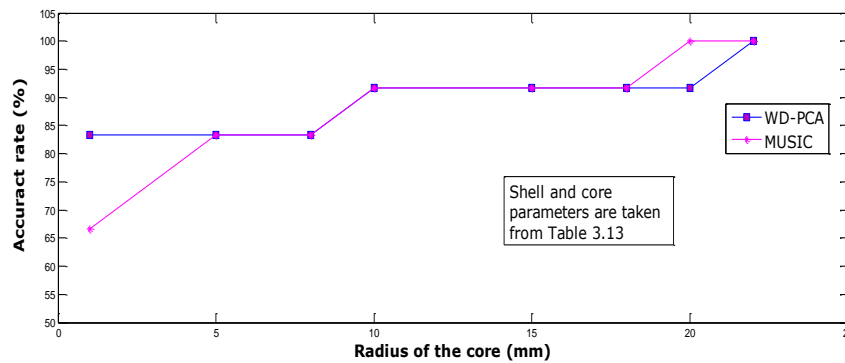
Figure 3.50 Tumor detection rates of the classifier C10 for different core sizes using MUSIC and WD-PCA methods.

As shown in Figure 3.50, both classification methods produce similar tumor detection performance when the breast tissue has a high-fat (adipose) content. They differ only at extreme ends. MUSIC algorithm method looks better for tumor radius larger than 20mm and the WD-PCA method performs better at very small tumor sizes with a radius less than 5mm.

When the breast tissue has moderate-adipose levels, the MUSIC based classifier performs much better for tumor radii from 8mm to 15mm. The WD-PCA method, on the other hand, looks better detecting tumors for radii larger than 15mm as seen in Figure 3.51.
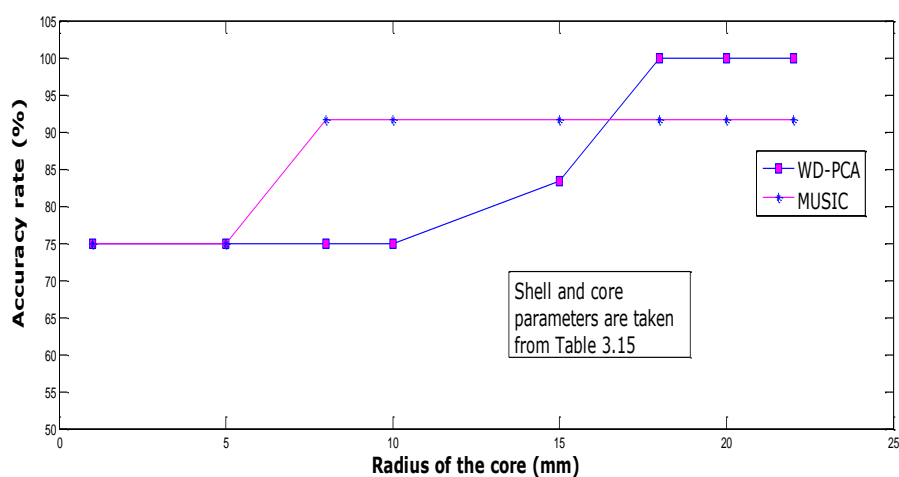
Figure 3.51 The change of accuracy rates for C11 with different core sizes for MUSIC and WD-PCA algorithm
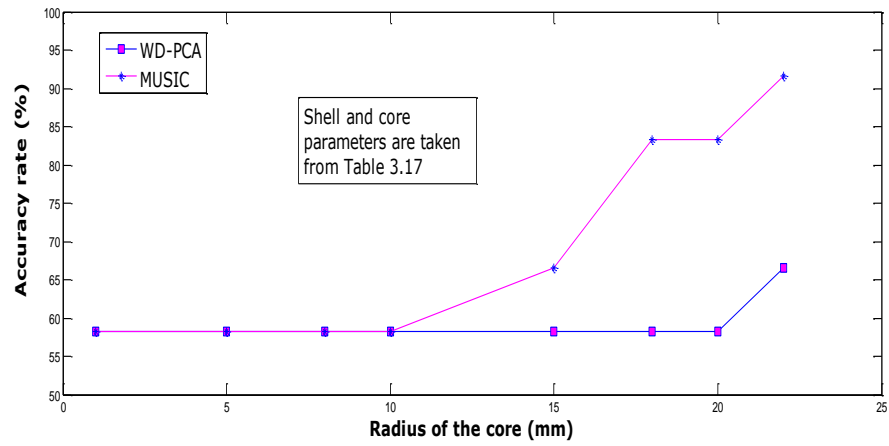
Figure 3.52   The change of accuracy rates for C12 with different core sizes for MUSIC and WD-PCA algorithm

Finally, as presented in Figure 3.52, both classification methods detects a tumor by an accuracy rate of about 59 percent for tumor radii equal to or less than 10mm in the case of having a breast tissue with low-fat content. But for larger tumor sizes, the MUSIC algorithm performs much better.

CHAPTER 4

CONCLUSION

In this thesis, two different classifier design methods, the WD-PCA method and the MUSIC algorithm based method, are utilized to classify healthy and malignant breast tissues and to detect the presence of a tumor in a given breast tissue using wideband electromagnetic scattered signals.

While both of these classification methods have been successfully applied in literature to conducting and/or lossless dielectric targets so far [25, 29-38], their use is extended to the recognition of lossy and dispersive objects for the first time in this study. The object geometries used in this thesis are chosen to be spherical due to the possibility of computing their scattered signals analytically [59]. The breast tissue samples as well as the tumor spheres are all assumed to be spherical shape. The tumor embedded in a breast tissue is approximately modeled as a dielectric sphere (core) coated by a dielectric shell. Both the core and shell materials are characterized by complex-valued and frequency-dependent permittivity functions [55]. These permittivity functions are modeled by Cole-Cole parameters. Realistic values of those parameters were provided in literature based on experimental studies. Healthy tissue samples were obtained from the breast reduction and breast cancer surgeries.

Normal (healthy) breast tissues can be classified according to their adipose (fat) content. Detection of a tumor becomes easier or more challenging depending upon the adipose content of the breast tissue covering the tumor. Furthermore, the size of the tumor is another important factor affecting the correct detection rate. Effects of all these factors in tumor detection problem are investigated in this thesis by

designing and testing various classifiers using both WD-PCA and MUSIC algorithm based methods.

Also, as a part of the scattered field computation task, a MATLAB code is generated to compute the multi-aspect scattered fields of a dielectric shell coated dielectric sphere when both core and shell are made of lossy and dispersive dielectric materials.

In the first part of the thesis, 9 different classifiers (C1, C2, ..., C9) are designed (using scattered data at 4 different aspect angles) and tested (at 12 different aspects). Each classifier is designed separately using both WD-PCA and MUSIC algorithm methods. Purpose of these classifiers is to discriminate different breast tissues (spherical breast tissue samples having different adipose contents) from each other. Tumor samples are also included in the target library of some of these classifiers.

Correct classification rates of all these classifiers are summarized in Table 4.1.

Table 4.1 Accuracy rates of all classifiers designed and tested in section 3.1 based on both MUSIC algorithm and WD-PCA methods

Classifier	WD-PCA method accuracy rate	MUSIC method accuracy rate
C1	100%	91.66%
C2	91.66%	100%
C3	97.22%	81.94%
C4	Distinguishing tumor: 100% Overall classification: 95.23%	Distinguishing tumor: 94.34% Overall classification: 77.38%
C5	79.16%	88.88%

Table 4.1 (cont'd)

C6	Distinguishing tumor: %100 Overall classification: 71.42%	Distinguishing tumor: %100 Overall classification: 80.95%
C7	84.72%	84.72%
C8	Distinguishing tumor: 95% Overall classification: 86.66%	Distinguishing tumor: 99.16% Overall classification: 77.5%
C9	Distinguishing tumor: 97.3% Overall classification: 67.38%	Distinguishing tumor: 97.3% Overall classification: 60.5%

From the results of classifiers C1, C2, C3 and C5, it is seen that MUSIC algorithm based method gives better accuracy rates as compared to the WD-PCA based method for distinguishing adipose-dominated tissue samples and WD-PCA algorithm based method gives better accuracy rates as compared to the MUSIC based method for distinguishing low-adipose content tissue samples while both methods give the same accuracy rate results for mid-adipose content tissue samples (see the results of classifier C7).

From the results of the classifiers C4 and C6 obtained in section 3.1 we can say that, distinguishing the tumor tissue sample among high-adipose content healthy breast tissue samples much easier as compared to recognize the malignant sample among the low-adipose content breast tissue samples. However, high correct classification rates are found for recognizing the tumor tissue among low or high adipose content healthy breast tissue samples by using both WD-PCA method and MUSIC algorithm based method.

When we examine the results obtained for the classifiers C8 and C9 designed by using either the WD-PCA method or MUSIC method, it is seen that the tumor tissue sample can be recognized among all other possible healthy breast tissue samples with high accuracy rates even if the number of targets increased up to 19.

In the second part of the thesis, 3 more classifiers (C10, C11 and C12) are designed and tested to detect the presence of a tumor embedded into a breast tissue sample. Each classifier used a different adipose content (low, moderate and high) for the breast tissue shell. For each shell type, different sizes of tumor are simulated in the raw database to figure out the smallest tumor size that the classifier can detect. The resulting of these classifier design and test simulations are summarized in Table 4.2 for WD-PCA based classifiers, and in Table 4.3 for MUSIC algorithm based classifiers.

Table 4.2 Tumor detection rates obtained for the classifiers (based on the WD-PCA method) C10, C11 and C12

Radius of the core	Tumor detection rates for C10	Tumor detection rates for C11	Tumor detection rates for C12
2.2cm	100%	100%	66.6%
2cm	91.66%	100%	58.3%
1.8cm	91.66%	100%	58.3%
1.5cm	91.66%	83.33%	58.3%
1cm	83.33%	75%	58.3%
8mm	83.33%	75%	58.3%
5mm	83.33%	75%	58.3%
1mm	83.33%	75%	58.3%

Table 4.3 Tumor detection rates obtained for the classifiers (based on the MUSIC method) C10, C11 and C12

Radius of the core	Tumor detection rates for C10	Tumor detection rates for C11	Tumor detection rates for C12
2.2cm	100%	91.66%	91.66%
2cm	100%	91.66%	83.3%
1.8cm	91.66%	91.66%	83.3%
1.5cm	91.66%	91.66%	66.6%
1cm	91.66%	91.66%	58.8%
8mm	83.3%	91.66%	58.8%
5mm	83.3%	75%	58.8%
1mm	66.6%	75%	58.8%

When we examine the results listed in Table 4.2 and Table 4.3, we can come into the following conclusion: When the breast tissue has a high fat content; MUSIC algorithm looks better for tumor radius larger than 20mm, the WD-PCA method performs better at tumor sizes less tan 5mm and both methods produce similar tumor detection performance for other tumor sizes. When the breast tissue has a moderate-fat content; the WD-PCA method looks better detecting tumors smaller than 15mm radii and the MUSIC based classifier performs better for tumor radii 8mm to 15mm. When the breast tissue has a low- fat content; both methods give similar results for tumor radii equal or less than 10mm but for larger tumor sizes MUSIC algorithm performs much better.

In this thesis, we have shown the usefulness of the WD-PCA and MUSIC algorithm based methods in breast tissue (healthy and/or malignant) and tumor detection problems. However, all the investigations carried on so far have been completed for noise-free electromagnetic scattered data. As a feature work, the performance of the target classification problems can be extended to include the noise effects on arbitrary shape targets.

REFERENCES

[1] Lord Rayleigh, "On the scattering of light by small particles", Phil. Mag., vol. 41, pp. 447-454, 1871.

[2] E.M. Kennaugh and R.L. Cosgriff, "The use of impulse response in electromagnetic scattering problems," in IRE Nat. Conv. Rec., Part I, pp. 72-77, 1958.

[3] E.M. Kennaugh and D.L. Moffatt, "Transient and impulse response approximations," Proc. IEEE, vol.53, pp. 893-901, August 1965.

[4] C. E. Baum, "Toward an engineering theory of electromagnetic scattering: the singularity and eigenmode expansion methods", in P.L.E. Uslenghi (Ed.), Electromagnetic Scattering, Academic Press, New York, 1978.

[5] Jouny.I, Garber.F.D, Ahalt,S.C "Classification of radar targets using synthetic neural networks", Aerospace and Electronic Systems, IEEE, vol. 29 , no. , pp. 336-344, April 1993.

[6] C. E. Baum, "Emerging technology for transient and broad-band analysis and synthesis of antennas and scatterers", Proc. IEEE, vol. 64, no. 11, pp. 1598-1616, Nov. 1976.

[7] C. E. Baum, E. J. Rothwell, K.-M. Chen, D. P. Nyquist, "The singularity expansion method and its application to target identification", Proc. IEEE, vol. 79, no. 10, pp. 1481-1492, Oct. 1991.

[8] C. E. Baum, "The singularity expansion method", in L. B. Felsen (Ed.), Transient electromagnetic fields, Springer-Verlag, Berlin, 1976.

[9] L. B. Felsen, "Comments on early-time SEM", IEEE Trans. Antennas and Propagation, vol. 33, no. 1, pp. 118-119, Jan. 1985.

[10] F. M. Tesche, "On the analysis of scattering and antenna problems using the singularity expansion method", IEEE Trans. Antennas and Propagation, vol. 21, no. 1, pp. 53-62, Jan. 1973.

[11] M. L. Van Blaricum, R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response", IEEE Trans. Antennas and Propagation, vol. 23, no. 6, pp. 777-781, Nov. 1975.

[12] B. Drachman and E. Rothwell, "A continuation method for identification of the natural frequencies of an object using a measured response", IEEE Trans. Antennas and Propagation, vol. 33, no. 4, pp. 445-450, April 1985.

[13] E. M. Kennaugh, "The K-pulse concept", IEEE Trans. Antennas and Propagation, vol. 29, no. 2, pp. 327-331, Mar. 1981.

[14] E.M. Kennaugh, D.L. Moffatt and N. Wang, "The K-pulse and response waveforms for non-uniform transmission lines", IEEE Trans. Antennas Propagation, Vol.AP-34, No.1, pp. 78-83, January 1986.

[15] F.Y.S. Fok, D.L. Moffatt and N. Wang, "K-pulse estimation from the impulse response of a target", IEEE Trans. Antennas Propagat., Vol.AP-35, No.8, pp. 926-933, August 1987.

[16] H. T. Kim, N. Wang, D. L. Moffatt, "K-pulse for a thin circular loop", IEEE Trans. Antennas and Propagation, vol. 33, no. 12, pp. 1403-1407, Dec. 1985.

[17] G. Turhan-Sayan and D. L. Moffatt, "K-Pulse Estimation and Target Identification of Low-Q Radar Targets", Wave Motion, No. 11, pp. 453- 461, 1989.

[18] G. Turhan-Sayan and D. L. Moffatt, "K-Pulse Estimation Using Legendre Polynomial Expansions and Target Discrimination", Journal of Electromagnetic Waves and Applications, Vol.4, No. 2, pp. 113-128, 1990.

[19] K. M. Chen, D. P. Nyquist, E. J. Rothwell, L. L. Webb, B. Drachman, "Radar target discrimination by convolution of radar returns with extinction-pulses and single-mode extraction signals", IEEE Trans. Antennas and Propagation, vol. 34, no. 7, pp. 896-904, July 1986.

[20] D. P. Nyquist, K. M. Chen, E. Rothwell and B. Drachman, "Radar Target Discrimination using the Extinction-Pulse Technique", IEEE Transaction on Antennas and Propagation, vol. AP-33, pp. 929-937, Sept. 1985.

[21] E. J. Rothwell, K.-M. Chen, D. P. Nyquist, "Extraction of the natural frequencies of a radar target from a measured response using E-pulse techniques", IEEE Trans. Antennas and Propagation, vol. 35, no. 6, pp. 715 720, June 1987.

[22] E. J. Rothwell, K. M. Chen, D. P. Nyquist, J. E. Ross and R. Bebermeyer, "A radar target discrimination scheme using discrete wavelet transform for reduced data storage", IEEE Trans. Antennas and Propagation, vol. 42, no. 7, pp. 1033-1037, July 1994.

[23] K. T. Kim, I. S. Choi, and H. T. Kim, "Efficient radar target classification using adaptive joint time-frequency processing", IEEE Trans. Antennas Propagation, vol. 48, no. 12, pp. 1789-1801, Dec. 2000.

[24] Y. Shi and X.-D. Zhang, "A Gabor atom network for signal classification with application in radar target recognition", IEEE Trans. Signal Processing, vol. 49, no. 12, pp. 2994-3004, Dec. 2001.

[25] G. Turhan-Sayan, "Natural resonance-based feature extraction with reduced aspect sensitivity for electromagnetic target classification", Pattern Recognition, Vol.36, No. 7, pp. 1449-1466, July 2003.

[26] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations", IEEE Signal Process. Mag., vol. 9, no. 2, pp. 21–67, Apr. 1992.

[27] L. Cohen, "Time–Frequency Analysis, Prentice-Hall", Englewood CliBs, NJ, 1995.

[28] J. E. Jackson, "A User's Guide to Principal Components", New York: John Wiley & Sons, Inc., 1991.

[29] G. Turhan-Sayan, "Real Time Electromagnetic Target Classification Using a Novel Feature Extraction Technique with PCA-Based Fusion", IEEE Transactions on Antennas and Propagation, vol. 53, no.2, pp. 766-776, February 2005.

[30] M. Secmen and G. Turhan-Sayan, "Radar target classification method with reduced aspect dependency and improved noise performance using multiple signal classification algorithm", IET Radar, Sonar and Navigation, pp. 583-595, December 2009.

[31] M. Seçmen, G. Turhan-Sayan, "MUSIC algoritması ile yeni bir elektromanyetik hedef sınıflanndırma yöntemi", 14. Sinyal İşleme ve Uygulamaları Kurultayı (SIU), pp. 1-4, Nisan 2006, Antalya, Türkiye.

[32] M. Seçmen, E. Ekmekçi, G Turhan-Sayan, "A resonance region method for recognition of multiple targets using MUSIC algorithm and time correlation technique", IEEE Antennas-and-Propagation-Society International Symposium, pp. 1-4, 2008.

[33] M. Ayar, "Design of an electromagnetic classifier for spherical targets", M.Sc. Thesis, METU, Ankara, Turkey, April 2005.

[34] E. Ergin, "Investigation of MUSIC algorithm based and WD-PCA method based electromagnetic target classification techniques for their noise performances", M.Sc. Thesis, METU, Ankara, Turkey, October 2009.

[35] G. Turhan-Sayan, E. Ergin," Non-Destructive Recognition of Dielectric Coated Conducting Objects by Using WD Type Time–Frequency Transformation and PCA Based Fusion", International Journal of RF and Microwave Computer-Aided Engineering, Vol. 23, Issue 4, pp. 403-409, July 2013. (Published online in Wiley Online Library (wileyonlinelibrary.com) with DOI 10.1002/mmce.20726)

[36] M. Seçmen, G. Turhan-Sayan, "A novel method for electromagnetic target classification using the MUSIC algorithm: Applied to small-scale targets", EUropean Conference on Antennas and Propagation (EUCAP), November 2006, Nice, France.

[37] M. Okan Ersoy, "Application of a Natural-Resonance based feature extraction technique to small-scale aircraft modeled by conducting wires for electromagnetic target classification", M.Sc. Thesis, METU, Ankara, Turkey, September 2004.

[38] M. Seçmen, G. Turhan-Sayan, " The MUSIC Algorithm-Based Electromagnetic Target Classification for Isolated Targets from Incomplete Frequency Domain Data", IEEE Antennas-and-Propagation-Society International Symposium, pp. 4170-4173, 2007.

[39] Baum, C.E. "Signature-based target identification and pattern recognition", Antennas and Propagation Magazine, IEEE , vol 36, no 3 , pp. 44-51, Jun. 1994.

[40] J.-H Lee, I.-S. Choi, H.-T. Kim, "Natural frequency-based neural network approach to radar target recognition", IEEE Trans. on Signal Processing, vol. 51, no. 12, pp. 3191-3197, Dec. 2003.

[41] M. A. Morgan, "Singularity expansion representations of fields and currents in transient scattering", IEEE Trans. Antennas and Propagation, vol. 32, no. 5, pp. 466-473, May 1984.

[42] Azimi-Sadjadi.M.R, De Yao, Qiang Huang, Dobeck.G.J, "Underwater target classification using wavelet packets and neural networks", IEEE Trans. on Neural Networks, vol. 11 , no. 3 , pp. 784-794, May 2000.

[43] G.Turhan-Sayan, K. Leblebicioglu, and T. Ince, "Electromagnetic target classification using time-frequency analysis and neural networks", Microw. Opt. Technol. Lett., vol. 21, pp. 63–69, 1999.

[44] M. Seçmen, "A Novel Music Algorithm Based Electromagnetic Target Recognition Method in Resonance Region for the Classification of Single And Multiple Targets", Ph.D. Thesis, METU, Ankara, Turkey, Feb. 2008.

[45] P. Stoica, R. Moses, "Introduction to Spectral Analysis", Prentice Hall, Englewood Cliffs, New Jersey, 1997.

[46] Y. Li, J. Razavilar and K. J. Ray Liu, "DMUSIC algorithm for 2D NMRSignals", Engineering in Medicine and Biology Society, IEEE 17th Annual Conference, pp. 477 – 478, Sept. 1995.

[47] G. Turhan-Sayan and M. Kuzuoglu, "Pole Estimation for Arbitrarily- Shaped Dielectric Targets by a Genetic Algorithm-Based Resonance Annihilation Technique", IEE Electronics Letters, Vol. 37, No. 6, pp. 380-381, March 2001.

[48] G. Turhan-Sayan, K. Leblebicioğlu and T. İnan, "Input Signal Shaping for Target Identification Using Genetic Algorithms", Microwave and Optical Technology Letters, Vol.17, No. 2, pp. 128-132, February 1998.

[49] Y. Wang, N. Shuley, "Complex Resonant Frequencies for The Identification of Simple Objects in Free Space and Lossy Environments", PIER 27, pp. 1-18, 2000.

[50] S. Dewal and R. Bansal, "Complex resonant frequencies of biological targets for microwave imaging applications", Electronics Letters, vol.38, no.25, pp. 1633-1635 December 2002.

[51] Carl. E. Baum, "Identification of Hidden Mines", IEEE International Conference on Electromagnetics in Advanced Applications, Torino, pp. 692-695, 2007.

[52] Motoyuki Sato, " Principles of Mine Detection by Ground Penetrating Radar", ISBN: 978-84882-345-7, pp. 19-26, Tohoku Univ., Japan, 2009.

[53] John H. Bradford," Frequency Dependent Attenuation Analysis of Ground Penetrating Radar Data", Center for Geophysical Investigation of the Sallow Subsurface (CGISS) Publications and Presentations, DOI: 10.1190/1.2710183, vol. 72, issue 3, pp. 1532-1544, Boise State University.

[54] Mariya Lazebnik, Cynthia B. Watkins, Susan C. Hagness, John H. Booske, Diana Popovic, Leah McCartney, Michal Okoniewski, Mary J. Lindstorm, Tara M. Breslin, Josephine Harter, Sarah Sewall, Walley Temple, Daphne Mew, Anthony Magliocco, Travis Ogilvie " The Dielectric Properties of Normal and Malignant Breast Tissue at Microwave Frequencies: Analysis, Conclusions and Implications from Wisconsin/Calgary Study", IEEE Antennas and Propagation Society International Symposium, pp. 2172-2175, 2007.

[55] Mariya Lazebnik, Dijana Popovic, Leah McCartney, Cynthia B Watkins, Mary J Lindstorm, Josephine Harter, Sarah Sewall, Travis Ogilvie, Daphne Mew, John H Booske, Michal Okoniewski, Susan C. Hagness, "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign and malignant breast tissues obtained from cancer surgeries", IOP Publishing Phys. Med. Biol., DOI: 10.1088/0031-9155/52/10/001, UK, pp. 2637-2656, 2007.

[56] Mariya Lazebnik, Michal Lazebnik, John H. Booske, Susan C. Hagness, "Highly Accurate Debye Models for Normal and Malignant Breast Tissue Dielectric Properties at Microwave Frequencies", IEEE Microwave and Wireless Components Letters, vol. 17, no.12, pp. 822-823, December 2007.

[57] Y. Huo, R. Bansal, Q.Zhu, "Breast Tumor Characterization Via Complex Natural Resonances", pp. 387,390, IEEE, MTT-S Digest, 2003.

[58] V. Komarov, S. Wang, J. Tang, "Permittivity and Measurement", DOI: 10.1002/0471654507 (Encyclopedia of RF and MW engineering), eme308, pp.1-20, 2005.

[59] Milton Kerker, "The Scattering of Light and Other Electromagnetic Radiation", Academic Press, New York, pp 39-50, pp 189-196, 1969.

[60] Clive M. Alabaster, "The microwave properties of tissue and other lossy dielectrics" Collage of defense technology, PhD Thesis, March 2004.

APPENDIX A

Sample MATLAB Program for Testing Lossy Dielectric Spheres with WD-PCA Algorithm Based Method

```matlab
%%
% Load all LTFVs created for all test targets and compare them with the FFVs at
database
load FFVT.dat
PCAA=FFVT;
load LTfeatvec1179020.dat
f1=LTfeatvec1179020;
...
load LTfeatvec315020.dat
f36=LTfeatvec315020;
FVC=[f1;f2;f3 ... f36];
FVCT=FVC';
 for i=1:36
    for j=1:3
      ac=corrcoef(FVCT(:,i),PCAA(:,j));
       corrcoeff3(i,j)= ac(1,2);
    end; end
 number2=0;number1=0;number3=0;
 for k=1:12
    if corrcoeff3(k,1)> max([corrcoeff3(k,2),corrcoeff3(k,3)])
       number1=number1+1;
```

```
        end
 end
for k=13:24
    if corrcoeff3(k,2)> max([corrcoeff3(k,1),corrcoeff3(k,3)])
        number2=number2+1;
    end
end
for k=25:36
    if corrcoeff3(k,3)> max([corrcoeff3(k,2),corrcoeff3(k,1)])
      number3=number3+1;
    end
end
```