T.C.

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF SCIENCE AND TECHNOLOGY

GPR SIGNAL PROCESSING METHODS FOR WATER DETECTION

M.Sc. THESIS

Sarah Ayad TABANNA

Department of Electrical and Electronics Engineering
Electrical and Electronics Engineering Program

July, 2019

T.C.

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF SCIENCE AND TECHNOLOGY

GPR SIGNAL PROCESSING METHODS FOR WATER DETECTION

M.Sc. THESIS

Sarah Ayad TABANNA
(Y1713.300004)

Department of Electrical and Electronics Engineering
Electrical and Electronics Engineering Program

Advisor: Assoc. Prof. Dr. Saeid KARAMZADEH

July, 2019

T.C.
İSTANBUL AYDIN ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜ

Yüksek Lisans Tez Onay Belgesi

Enstitümüz Elektrik- Elektronik Mühendisliği Ana Bilim Dalı Elektrik- Elektronik Mühendisliği (İngilizce) Tezli Yüksek Lisans Programı **Y1713.300004** numaralı öğrencisi **SARAH AYAD IBRAHEEM TABANNA'** nın **"GPR SIGNAL PROCESSING METHODS FOR WATER DETECTION''** adlı tez çalışması Enstitümüz Yönetim Kurulunun 12.06.2019 tarih ve 2019/12 sayılı kararıyla oluşturulan jüri tarafından *oy.birliği* ile Tezli Yüksek Lisans tezi olarak *kabul* edilmiştir.

Öğretim Üyesi Adı Soyadı	İmzası

Tez Savunma Tarihi : 11-07-2019

1)Tez Danışmanı: Doç. Dr. Saeid KARAMZADEH   ....

2) Jüri Üyesi   :   Dr. Öğr. Üyesi Evrim TETİK   ....

3) Jüri Üyesi   :   Dr. Öğr. Üyesi Oğuz ATA   ....

Not: Öğrencinin Tez savunmasında **Başarılı** olması halinde bu form **imzalanacaktır**. Aksi halde geçersizdir.

DECLARATION

I hereby declare that all information in this thesis 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 and referenced all material and results, which are not original to this thesis.

Sarah Ayad TABANNA

FOREWORD

In the beginning I would like to thank my mother and father, My Ideals who raised me to be the good person I'm today, for being patient, loving and supporting through it all. I hope one day I can return some of what they gave me, everything I have accomplished is because of them.

I would like to express my gratitude to my advisor Dr. Saeid karamzadeh for his guidance and encouragement, the suggestions and comments during the making of this thesis. I'm truly grateful for working with him all this time.

Finally, I would like to thank my friends for always believing in me and encouraging me from day one.

TABLE OF CONTENT

ABBREVIATIONS

GPR   : Ground Penetration Radar
EM   : Electro Magnetic
PCA   : Principle Component Analysis
LNA   : Low Noise Amplifier
MSE   : Mean Square Error
PSNR   : Peak Signal to Noise Ratio
SNR   : Signal to Noise Ratio
SVD   : Singular Value Decomposition
MS   : Mean Square
ICA   : Independent Component Analysis
PICA   : Probabilistic Independent Component Analysis
SAR   : Synthetic Aperture Radar
GLM   : Generalized Linear Model
CDF   : Cumulative Distribution Function
IE   : Image Enhancement
FFT   : Fast Fourier Transform
IFFT   : Inverse Fast Fourier Transform

LIST OF FIGURES

LIST OF TABLES

GPR SIGNAL PROCESSING METHODS FOR WATER DETECTION

ABSTRACT

Recently various developments have been introduced in GPR Signal Processing Technology. Multispectral sensors have been used for years and provide images with multi bands up to 10-20 bands. The information extracted from the multispectral images are useful and helped in many different applications in the real world.

While signal processing is very important technique on the other hand its challenging.

In underground water detection a lot of laboratory experiments were used to investigate the potential of ground penetration radar (GPR) efficient and accurate detection. GPR uses a high frequency electromagnetic wave reflection technique to obtain subsurface information.

In this study we will review basic processing techniques to remove cluttering and comparing their performance like PCA, ICA, SVD, PICA and other filters including Median, Gaussian, Histogram, etc.

In order to find the best method to process the raw data that is collected from different environments, we designed A box that was filled with sand with a sphere with radius 10 mm buried inside then water was injected inside the sphere after that our antenna scanned the surface approximately 56 times (5mm at a time) B-scan is formed from these collected multiple A-scans.

After collecting that data, we add to the box more roughness to the sand and grass on top sequentially and scanning the box each time the same way and getting different raw data to work with.

Detailed simulation is presented analyzing SNR (Signal-to-Noise Ratio) and PSNR (Peak-Signal-to-Noise Ratio) Values for our obtained results is compared, interpreted and discussed in this Research.

Keywords: *GPR, Signal Processing, Clutter Removal, PCA, ICA, Filtering, Water Detection.*

YER ALTI GÖRNTÜLEME RADARLARINDA SU TESPITI IÇIN SINYAL İŞLEME METOTLARI

ÖZET

Günümüzde yeraltı görüntüleme radarları (YAG) bir çok askeri ve sivil uygulamalarda kullanılmaktadır. Askeri alanda, mayın tespiti, sivil alanda ise arkeoloji çalışmaları, kazı yapmadan su sızıntısının bulunması, petrol veya yer altı su kaynaklarının bulma çalışmaları, YAG uygulamalarına örneklerindedir.

YAG sistemlerinde akademik ve mühendislik çalışmalar iki ana alana ayrılır. Birincisi, sistemin fiziksel tasarımı. Burada en önemli çalışmalar anten tasarımı için yapılmaktadır. Tasarlana uygun bir anten, daha derinlerde aranan hedefin daha iyi çözünürlükle görüntülenmesini sağlayacaktır. Doğru frekans seçimi, bant genişliği ve kazanç tasarımı bu alandaki meydan okumalardandır. Bunun yansıra, anten yapısı radar sistemine monte edilmesi için, oldukça kompakt yapıya sahi olması gerekmektedir.

YAG sistemlerinde ikinci akademik çalışma ise, işaret ve görüntü işlemedir. Radar sisteminden alanın bilgiden, hedef le ilgili tüm detayları görüntülemek, gürültü ve istenmeyen nesneler ve hedefler le ilgili bilgilerin minimize edilmesi gerekmektedir.

Bu tez çalışmasında ise, arka plan (toprak veya hedefin gömülü olduğu herhangi bir zemin) etkisini azaltmak ve olası gürültüleri minimize ederek su tespitinin yapılması planlanmaktadır.

Farklı toprak çeşitleri ve farklı alanlar; sadece toprak, engebeli toprak ve çimle kaplı toprak engebeli incelemeler yapılacaktır.

Farklı sinyal işleme metotları uygulanıp sonuçları karşılaştırılacaktır.

Anahtar Kelimeler: *Günümüzde yeraltı görüntüleme radarları (YAG), ICA, PCA ,PICA.*

1. INTRODUCTION

1.1 Purpose of This Thesis

GPR technology witnessed great development in the past few years it became one of the most important techniques for underground detection because of its non-destructive nature, there is wide diversity for what GPR can find such as liquids (water, fuel, gas, etc..), metal and non-metallic objects.

The main purpose of this research is comparing multiple processing methods for clutter removal and filters and obtaining their PSNR, SNR values and find the best algorithm for water detection using different data collected in different environment or obstacles in life, finding which method is best in every situation is our goal.

1.2 Literature Review

As known GPR technology has been widely used as a method of non-destructive testing of different subsurface observation. The vertical cross-section images obtained allow the identification of thickness and lithological horizons of different media. And over these years, GPR has been able to adjust to new areas [1].

Recently, it has been successfully used to characterize and map subaqueous and anthropogenic soils. It has also been extensively used in hydro-pedological and hydro-geophysical investigations, finding landmines, water, oil, pipes and fused wires. During all GPR surveys, noise in general (either environmental or systematic) with other radio frequency signals interference can blur and damage the desired signal.

Meanwhile, clutter which can be something such as a strong reflection from the antenna direct coupling and air-ground subsurface, the signature signals of the subsurface son be masked [3].

A considerable number of researches in literature on GPR signal processing multiple signal enhancement and clutter removal methods for underground water detection has been under studying for many years a lot of fresh water is wasted (leaked) throughout the years.

Moreover, a recorded response of the radar at a certain location is called an A-Scan waveform, which is explained as a measure of the amplitude in a reflected signal with respect to time. After combining all the collected A-scans by moving the radar antenna in a distinct direction the combined data forms what is called B-scan. In a B-scan image, EM wave transmission time (or penetration depth) is represented by the vertical axis and the horizontal axis represents the GPR spatial location [2].

1.2.1 Electromagnetic propagation of GPR

The electromagnetic wave in materials, its propagation velocity 'v' and ε is dielectric permittivity and μ magnetic permeability [6]:

$$v = \frac{1}{\sqrt{\varepsilon\mu}} = \frac{1}{\sqrt{\varepsilon_\circ \varepsilon_r \mu_\circ \mu_r}} \qquad (1.1)$$

Where the free space permittivity is $\varepsilon_\circ = 8.854 \times 10^{-12} F/m$

And the medium relative permittivity (dielectric constant) is donated by: $\varepsilon_r = \varepsilon/\varepsilon_\circ$ While the free space the magnetic permeability is explained as: $\mu_\circ = 4\pi \times 10^{-7} H/m$,

And the relative magnetic permeability is $\mu_r = \mu/\mu_\circ$

In several soils type, the magnetic properties are so small that its neglected, which leads to $\mu = \mu_\circ$ that makes the equation above as follow:

$$v = \frac{c}{\sqrt{\varepsilon_r}} \qquad (1.2)$$

Where the velocity of light is $c = 3 \times 10^8 m/s$

And the wavelength λ could defined as:

$$\lambda = \frac{v}{f} = \frac{2\pi}{\omega\sqrt{\varepsilon\mu}} \qquad (1.3)$$

Where $\omega = 2\pi f$ is the angular frequency and the frequency is donated by f.

In general, ε and the electric conductivity σ are complex content and it can be expressed as:

$$\varepsilon = \varepsilon' - j\varepsilon'' \tag{1.4}$$

$$\sigma = \sigma' - j\sigma'' \tag{1.5}$$

Where the dielectric polarization term is ε', and ε'' represents the energy loss due to the polarization lag, while the ohmic conduction is referred to as σ', and the faradic diffusion is related to σ'' [32].

Meanwhile A complex effective can expresses the total loss and storage effects of the material [7]:

$$\varepsilon = \left(\varepsilon' + \frac{\sigma''}{\omega}\right) - j\left(\varepsilon'' + \frac{\sigma'}{\omega}\right) \tag{1.6}$$

And the loss tangent $(tan\delta)$ could be define ass

$$tan\delta = \frac{\varepsilon''}{\varepsilon'} \equiv \frac{\sigma'}{\omega\varepsilon'} \tag{1.7}$$

Mostly when ε'' and σ'' are small, the right expression to say is its approximated.

In the planewave solution of the "Maxwell's Equations" [8], the electric field 'E' of an electromagnetic wave that it's travelling in Z- direction can be expressed as:

$$E(z,t) = E_0 e^{-i(\omega t - kz)} \tag{1.8}$$

Where the peak signal amplitude is E_0 and the wavenumber is presented by $= \omega\sqrt{\varepsilon\mu}$, which in the conductive it will be is complex;

$$k = \alpha + j\beta \tag{1.9}$$

The α is attenuation constant (Np/m) and β is the phase constant (rad/m), that could be shown as below:

$$\alpha = \omega\left[\frac{\mu\varepsilon'}{2}\left(\sqrt{1 + tan^2\delta} - 1\right)\right]^{\frac{1}{2}} \tag{1.10}$$

$$\beta = \omega\left[\frac{\mu\varepsilon'}{2}\left(\sqrt{1 + tan^2\delta} + 1\right)\right]^{\frac{1}{2}} \tag{1.11}$$

The attenuation constant is expressed in dB/m by $\alpha' = 8.686\alpha$.

$$\delta = \frac{1}{\alpha} \qquad\qquad\qquad (1.12)$$

δ is called the skin depth, it gives the depth that which the amplitude of the electric field decay $1/e$ $(\sim -8.7\ dB, \sim 37\%)$.

It is a useful parameter to describes how lossy the medium is. The table below provides the typical range of permittivity, conductivity and attenuation of various materials.

Table 1.1: ε, σ and α of various materials properties at 100 MHz [7][8].

Material	εr	σ [S/m]	α [dB/m]
Air	1	0	0
Fresh water	81	$10^{-6}\ to\ 10^{-2}$	0.01
Clay, dry	2 to 6	$10^{-3}\ to\ 10^{-1}$	10 to 50
Clay, wet	5 to 40	$10^{-1}\ to\ 10^{-0}$	20 to 100
Sand, dry	2 to 6	$10^{-7}\ to\ 10^{-3}$	0.01 to 1
Sand, wet	10 to 30	$10^{-3}\ to\ 10^{-2}$	0.5 to 5

1.2.2 Applications of GPR

- **Archaeology**

The GPR is used mainly in site investigation, assessment and virtual reconstruction**,** artefact location, grave location, structure location and mapping [20].

- **Civil Engineering**

A lot of area in civil engineering requires the GPR such as, concrete or structure investigation, so bridges and road investigations like the thickness of the asphalt, thickness and base layer profiling, some reinforcement evaluations, structure subsidence and voids, and other utilities locating [20].

- **Environmental Studies**

Waste disposal and ground contamination studies is one of the quite important areas that the GPR can be used in, to find ground water studies or saltwater intrusion, underground storage tank and drum investigations [20].

- **Geological Investigations**

In Geology research, GPR has been used widely such as in, in fresh water in rivers or lakes investigations, bedrocks identifying, some sinkholes mapping, karst investigations, and of course water table studies [20].

- **Law Enforcement and Military**

The GPR technology is widely used in locating some Explosives or weapons location like mine detection, in forensics like finding an evidence of buried objects or probably human remains, also it's a good method for searching and rescuing in different areas and can be used for locating bunkers and tunnels [20].

Figure 1.1: Applications of GPR.

1.3 Organization of This Thesis

From this point forward our research will be organized as following. In Chapter 2, general GPR system design and components are explained and how the GPR

system works. Chapter 3 gives information about data collection, the use of GPRMAX and SAR (Synthetic Aperture Radar) system is explained.

Pre-processing steps (like Background removal and filtering methods for raw data) and Signal processing algorithms for detection are analyzed will be described in the Chapter 4. While the results are demonstrated in Chapter5 and the thesis will be concluded in chapter 6.

2. GPR SYSTEM DESIGN

2.1 GPR System Component

The ground penetrating radar (GPR) as explained before is a type of radar which is built for the subsurface imaging. GPR is designed to see what is underground and behind walls in contrast to the usual radar that is used to identify airplanes, vehicles or ships through the medium of water and air [2]. GPR technology is used in several applications including archeological researches, forensics, earth sciences, finding liquids and its especially useful for locating buried metals and non-metallic materials underground [2].

Despite the variations in designing the GPR system, there is a common topology associated with GPR. All radar in some form have a signal generator, a power amplifier, transmit antenna, receive antenna, LNA, and a capture device. However, in some cases the receive and transmit antenna are the same. Although the details of which will vary from design to design [4].
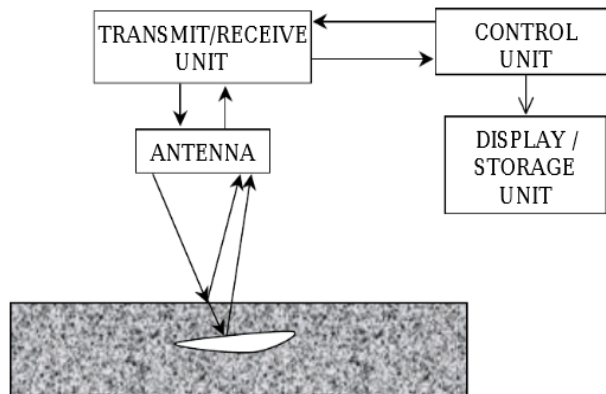
Figure 2.1: GPR System Component

There are many types in GPR systems, but the pulsed systems are mostly used, and they are the most commercially available. These pulsed systems have a simple principle based upon the transmission of EM pulses and analysis of its reflected pulse from interfaces where might be contrast in the dielectric

properties of underground structure. Figure 2.1 shows the major components of a pulsed radar system [2].

- **Signal Generator**

The signal generator is part of the control unit which is responsible for building up the appropriate waveform for the GPR. The waveform in question depends on the type of radar and can vary from an impulse to more advanced digital patterns.

In modern Digital Signal Processing (DSP) chip or a software defined radio can produce most signals in a flexible and dynamic way. However, at the power levels required by radar it may be difficult to construct a linear amplifier to bring the relatively small signal up in power to transmission strength. Therefore, it is useful to choose signals which are more easily matched to the amplifier in use, or to integrate both the generator and amplifier into one concrete device [31].

Figure 2.2: Example of Signal Generator

- **Amplifier**

The power amplifier portion of the radar is responsible for driving the signal produced by the signal generator to sufficient levels for transmission. The type of amplifier needed depends on the signal that is output and the application. Obviously, a hand-held radar does not need 10-kilowatt amplifiers; however, an aircraft radar might. When it comes to the pulse variety of radar, the signal generator and amplifier are often same unit [31].
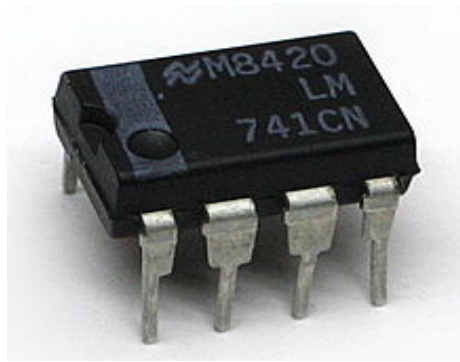
Figure 2.3: Example of Amplifier

- **Antenna**

The main purpose of any system antenna is to provide the coupling mechanism between the free propagating wave in the earth and the wave propagating within the feed transmission line. It's also responsible for spreading the signal out over a large area and going so increases the directionality of the coupling. There are so many antenna types according to the frequency and the application to use, such as bow-tie antenna, spiral, horn, Vivaldi antenna, etc.

The transmitting antenna (Tx) and the receiving antenna (Rx) are identical and in the same place despite the situations that requires it to be separated, and the antennas will be selected to meet the characteristics of the generated waveform [31].

d) horn antenna

a) Vivaldi   b) Bowtie   c) Loaded Bowtie
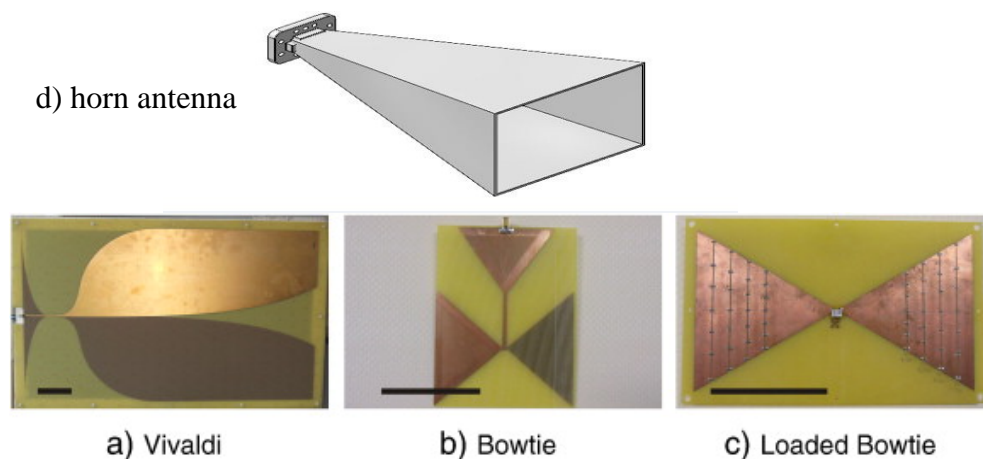
Figure 2.4: Example of Antenna Different Shapes.

- **Pulse Extender (LNA)**

Low-Noise Amplifiers are one of the significant parts of the GPR receiver circuit (in the control unit) whereby the received signal is processed and

converted into information. The Low-noise amplifiers (LNA) are designed to be close to the receiving devices to minimize the losses caused by interferences.

As the name suggests, LNA adds a minimum amount of noise (useless data) in the received signal because any more would highly corrupts the already weak signal.

LNAs are used when the SNR is too high and in need to be degraded by around 50 percent and the power needs to be boosted [31].

Figure 2.5: Example of LNA

- **A/D converter**

In the control unit of the GPR systems the signal is received in the frequency range of some GHz. So, the first step here is to sample and hold the signal which is used to hold sample value until the nest pulse occurs. In basic principle this process is fulfilled by what is called 'dwell time' which can be explained as connecting the input to the capacitor.

The dwell time can be very short time ranging from nanoseconds to microseconds. As typically known the analog to digital converter (A/D converter) is an amplifier with an amount of gain that can be programmed digitally [31].

This conversion involves multiple steps, including sampling, quantization and coding. This conversion process can be done in one stage (the correlator), and hence, no intermediate -frequency – conversion stage is needed, this reduces the system complexity greatly. As the correlator is the heart of the receivers, its consisted of a multiplier, integrator and sample and hold (S/H) circuit. As simple as it can be explained the multiplier mixes or to be precise, it multiplies the received pulse with the pulse signal from the signal generator to gut the output result [31].

- **Display Unit**

The main objective of the collected GPR data to show as an image of the subsurface. There are three display types of the subsurface data [31]:

- A -Scan

- B -Scan
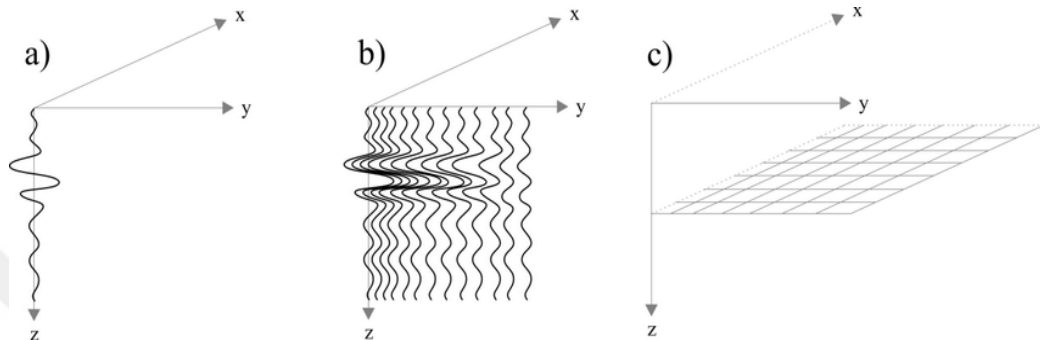
- C -Scan

Figure 2.6: Type of scans: a) shows A-scan, b) shows B-scan, c) shows C-scan.

2.2 How Does GPR Work

As mentioned, and previously known in basic radar systems, it has one or multiple antennas. Those antennas will transmit and receive the radio frequency waves, in each process those transmitted waves penetrate the subsurface of ground or wall (whatever surface the user wants to investigate). While a large amount of this signal is dissipated throw the layers, a small amount of this transmitted wave returns or reflects be received by the receiving antenna. The returned or reflected wave bring back data, this recorded data is elucidated and processed [21].

This data collection is shown on the computer as an image, an information about the object underground or behind walls such as the size of the target, position of target, and its depth [21].

while the radar systems are pretty sensitive to the changes of material characteristics, to detect these changes with the radar system movement is required. For example, in the air traffic radar system, targets such as airplanes are the one moving while the radar transmitter is stationary.

Meanwhile, in the matter of ground penetrating radar, we are searching for stationary targets so moving the GPR antenna and scan the concerned area to detect the underground target [21].
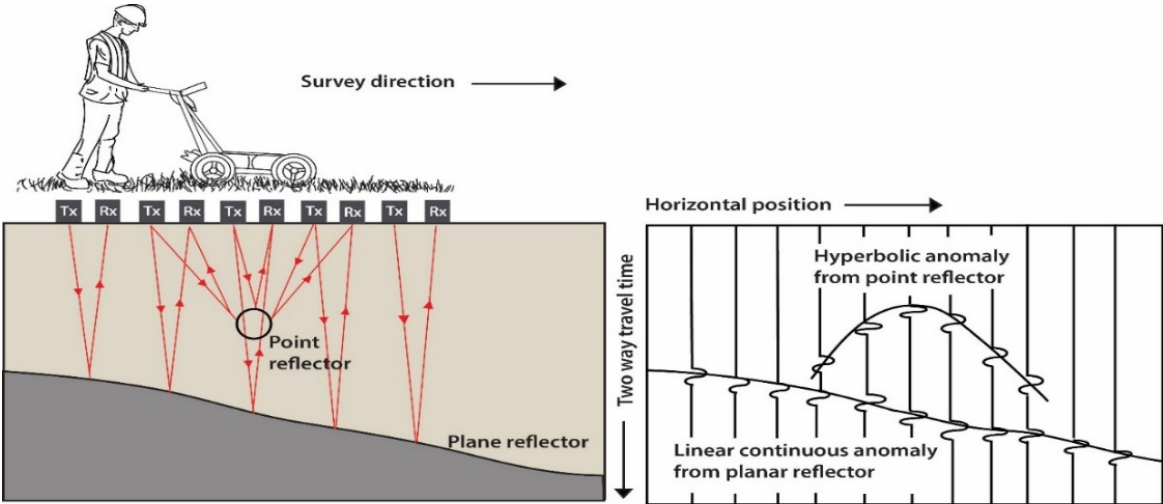
Figure 2.7: Explains How GPR Works.

2.3 GPR Depth Potentials

Considering depth in the ground penetrating radar, how much the wave can reach beneath the surface is mainly dependent on several factors including:

- **Soil type**

Whether its soil or rocks in the survey area, GPR can reach depth of up to 100 feet (30 meters) in some conducting materials like granites or day sand. While moist clay, and some relatively high conducting materials, will attenuate or absorb the GPR signal, causing penetrating depth to decrease greatly to 3 feet (1 meter) or less.

- **Antenna frequency**

Another factor effecting the penetrating depth is the antenna frequencies used. While the use of the low frequencies gives a great depth from 100-200MHz obtains subsurface reflections from deeper depth (about 25 to 100 feet or more) depending on the conductivity of the soil or rocks, the main disadvantage in using low frequencies is having a low resolution normally these low frequencies

are used to locate large and deep targets such as locating sinkholes or trenches [23].

- Meanwhile higher frequencies give shallow (limited) penetration depth but much higher resolution. In order to have better detection deeper utilities must have larger diameters than shallow utilities. Locating rebar or conduits in concrete requires higher resolution, such as (1,000 MHz) GPR system is used. This will give detailed high resolution approximately up to 24 inches in depth [23].
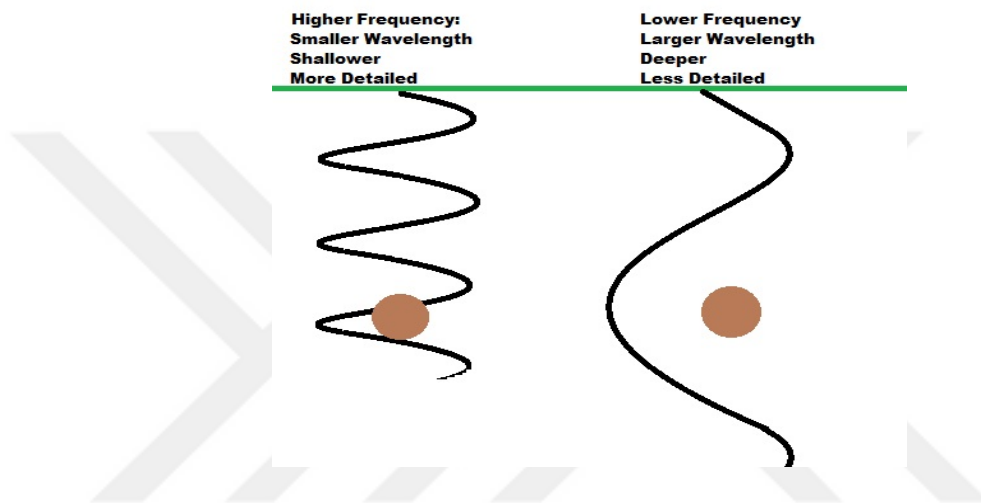
Figure 2.8: Antenna Frequencies Capabilities

Depending on the materials the depth can range from few inches to thousands of feet, figure below shows some depth potentials.
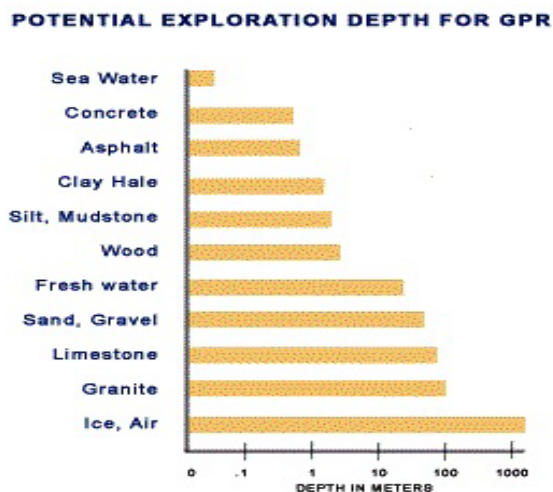
Figure 2.9: Potential Depth For GPR.

2.4 Why Can't GPR Find Everything

Mostly GPR technology can find metallic and non-metallic objects, so as a utility locater there are two main reasons that the GPR has to consider when scanning an area; first one is soils conductivity and the second is the dielectric constant [21].

- **The Conductivity**

Conductivity can be defined as the ability or can be expressed as a power used to conduct or to transmit the heat or maybe electricity, or some sound of a substance. So, as known the units of the conductivity is siemens per meter[S/m].

The clay is a one example for conductive soils. Within clay structure there is an amount metallic material. Therefore, clay can be evaluated as conductive soil type. GPR signals will get scattered as a result of the conductive nature of the soil before it can return to the receiving antenna. Another example would be wet soils or damp that have a high salt content; normally near sources of salt or brackish water. Salty soils have more ions in them which creates a much conductive environment, and due to that it limits the GPR wave resolution [21].

Table 2.10 : Examples of Soil Material and their Conductivity, Permeability values.

Material	Intrinsic permeability (arcy)	Hydraulic σ (cm/s)
Clay	$10^{-6}\ to\ 10^{-3}$	$10^{-9}\ to\ 10^{-6}$
Silt, sandy silts, clayey sand, fine sands	$10^{-3}\ to\ 10^{-1}$	$10^{-6}\ to\ 10^{-4}$
Silty sands, fine sands	$10^{-2}\ to\ 1$	$10^{-5}\ to\ 10^{-3}$
sands, glacial outwash	$1\ to\ 10^{2}$	$10^{-3}\ to\ 10^{-1}$
gravel	$10\ to\ 10^{3}$	$10^{-2}\ to\ 1$

Meanwhile pure-water is considered relatively bad conductors. But oceans water, rivers, and lakes contain some dissolved salts remains. The conductivity of water will increase as the ions concentricity increases [21].

14

Table 2.2: Water conductivity

c	σ
Pure water	$5.5 \cdot 10\text{-}6$ S/m
Drinking water	$0.005 - 0.05$ S/m
Sea water	5 S/m

- **Dielectric Constant**

The dielectric constant can be explained as a measure of how permissive the soil and the target in it are to the radar ground penetrating wave. As easy as it sounds, but it is more complicated than that , consider you have a sandy-mixed soil with a concrete pipe and metal pipe buried in it, imagine the GPR wave will enter the soil at a certain rate and as it goes through different soil layers or different utility structures that rate will change as it passes theses layers. Most utilities are actually noticeable in sandier soils because the wave penetrating through the soil will meet a little resistance unlit it hits the utility. However, some utilities have a dielectric constant value close to the surrounding soil. This might result that the utility structure may fade into the background of the GPR image and make it harder to see it. This concludes that highly conductive soils and similar dielectric constant are two factors that can make a GPR technician have a difficult time in locating utilities that might be in the ground. That's why we can't rely on one piece of equipment technology when it concerns underground utility locating, although GPR is a great piece of technology that can be used [21].

Table 2.3: Dielectric Constant of Some Material.

ϵ_r of some materials	
Material	ϵ_r
Vacuum	1.0
Air	1.0006
Oil	2.2
Polyethylene	2.26
Beeswax	2.8
Fused quartz	3.78
Water	80
Calcium titanate	168
Barium titanate	1,250

2.5 Antenna Capability

Table 2.4: Antenna Capabilities According to Frequency in Wet clay and Dry sand.

Antenna MHz	Penetration in Dense wet clay	Penetration in Clean Dry Sand	Smallest objects visibility
100	20 ft (6m)	60 ft+(18m+)	Tunnel @60ft (18m) depth 2 ft(60cm) Pipe@ 20ft (6m) depth
250	13 ft (4m)	40 ft (12m)	3 ft (90cm) Pipes @ 12m
			6 in (15cm) Pipe@13ft (4m)
500	6 ft (1.8m)	14.5 ft (4.4m)	4 in (10cm) Pipe@4cm
			3/16 in (0.5cm) Hose 1.8m and less
1000	3 ft (90 cm)	6 ft (1.8m)	3/16in (0.5cm) Hose@ 3ft (90cm) wire mesh, shallow.
2000	0.5 ft (15 cm)	2 ft (60cm)	Monofilament fishing line

As has been discussed before when it comes the choosing depth capabilities, we must decide what frequency range we will use according to the application and the target. For example, for locating utilities 500MHz antenna can be selected, while for rebar locating the 1000MHZ antenna is preferable in walls and under the ground detection [21].

2.5.1 The relation between the resolution, signal penetration and antenna frequency

The profiling image resolution and the penetrating depth are dependent on the frequency that is selected to use and that they are inversely proportional. As seen, there is one relevant feature that will rise from this interaction between frequency and penetration depth. In way or another in here it is quite necessary to delineate the size of targeted objects in use, in order to select (or decide) what is the most suitable frequency for the given survey [23].

The possibility of obtaining high resolution will be high, when in use of high frequencies however, its only shallow penetration, meanwhile the use of low frequencies antenna the obtained result will be opposite of the high frequencies as explained before. The environment depolarization and saturation pore could also interfere with signal penetration [23].

Table 2.5: The relation between depth and antenna frequency.

fc (MHz)	Depth(m)
1000	0.5
500	1
200	2
100	7
50	10
25	30
10	50

Table 2.6: the signal attenuation and the speed relative to the changes in the saturation pores and lithology for antenna with frequency in the range of 80 to 120 MHz**.**

Lithology	Velocity(m/ns)	Attenuation (dB/m)
Unsaturated sand	0.1to 0.2	0.01to 0.014
Saturated sand	0.05 to 0.08	0.03 to 0.05
Unsaturated clay	0.09 to 0.12	0.028 to 300
Saturated clay	0.05 to 0.07	0.028 to 300

3. DATA COLLECTION

As known Ground Penetration Radar or any normal radar, collecting the measurement regarding the two-way travel time which is defined as, a time that each pulse spends to be transmitted into the ground and then be reflected off the subsurface after hitting the target and be received by the antenna.

By moving the GPR antenna over a certain search area in one line (one dimension) there by placing it over an area each time something called an A-scan is collected and by repeating this scan process these multiple A-scans now forms what is called a RADAGRAMS of as B-scan (a two-dimensional scan) [4].
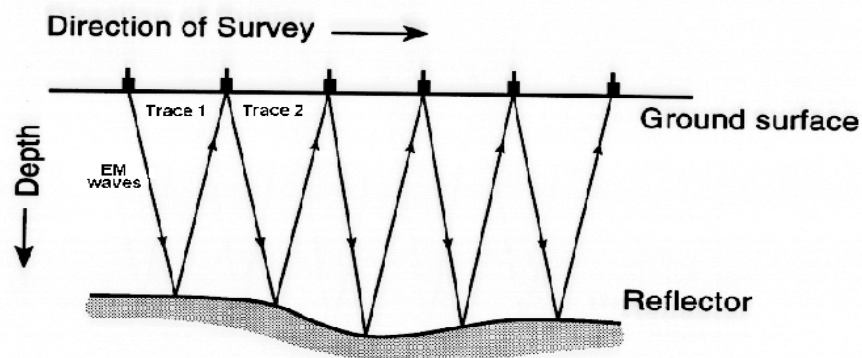
Figure 3.1: How Data is Collected.

So as explained the data recorded by the GPR are usually represented as one, two, or three-dimensional dataset [4]:

- **A-Scan**

An A-scan can be explained as a single GPR recorded waveform, with the antennas at fixed location, the changing variable is time.

- **B-Scan**

B-scan is explained as gathering multiple recorded A-scans but to be clearer, in this B-scan process our antenna is a changing variable also. By moving our

antenna over this certain subsurface, the GPR is recording a two-dimensional dataset.

- **C-Scan**

In here the idea of collection the GPR dataset is becoming more obvious each step, the C-scan can be formed after collecting or recording multiple parallel B-scans or in other words, the GPR antenna is moving on a grid in the xy-plane.

3.1 Synthetic Aperture Radar

As explaining the Synthetic Aperture Radar (SAR) system, we know that these systems are taking great advantages of the long-range propagation characteristics of the radar signals and providing high-resolution images by modern digital processing of the complex information. For us here SAR radar system is used as a data collection method, a technique that allows us to have a high-resolution image. Simply the radar images are produced by an active system that sends the specific signal to the ground and then detects the backscattered wave that is reflected from the ground searching area [5].

These synthetic aperture radar (SAR) systems, produces a two-dimensional (2D) images, one of the dimensions in the image is called range (or cross track) and the other is called azimuth (along track) and these dimensions are defined in details bellow [5].

3.1.1 Range

In all Synthetic Aperture Radar (SAR) systems, good resolution and range calculations are achieved in the same manners in other radar systems. The range component can be defined as measurement of time, from the transmitting the pulse to receiving an echo from the underground target in question, as simply as it can be explained; range resolution is determined by the width of the transmitted pulse, for example the narrow pulses give a good range resolution [5].

3.1.2 Azimuth

The other dimension regarding synthetic radar (SAR) is referred to as the azimuth (or along track). SAR systems have the ability to assembles fine

20

azimuth resolution. For obtaining good resolution, a large antenna can be used to stress the energy of this transmission and reception into a relatively shape beam, this beam sharpness is defined as azimuth resolution. Similar to SAR systems, optical systems like a telescope, it requires a large aperture (a lens or a mirror similar to antenna) to obtain good resolution images [5].

Even though, synthetic aperture radar (SAR) uses pretty much lower frequencies than optical systems. the moderated synthetic aperture radar (SAR) resolution requires a very large antenna to be carried by an airborne system; the antenna length of several hundred meters long are sometimes needed [5].

Using a narrow beam width will lead to having a finer data resolution with the use of relatively large antennas than possibility of having smaller antennas, that the airborne radar collects while flying some distance [5].
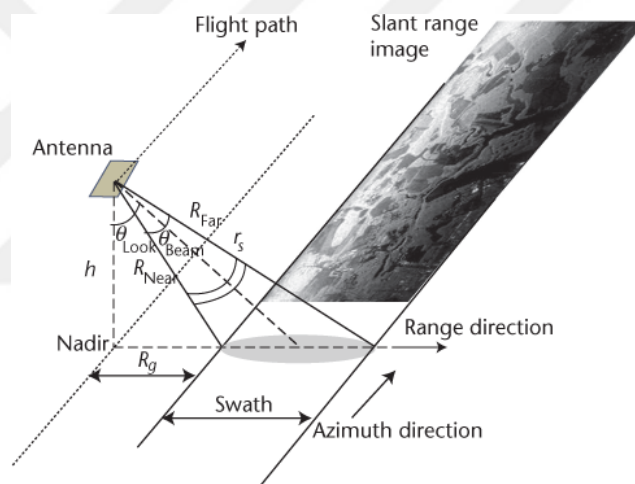
Figure 3.2: How Data is Collected by SAR System

3.2 Applications of SAR System

As shown with the use of the SAR system, it became possible to have a quick detailed information for monitoring the environment, mapping earth sources, and some military applications that requires wide area imaging with high resolution.

Usually, collected images are obtained in night or during bad weather conditions. As we previously mentioned, SAR systems are used in variety

applications in life because of its ability to collect data bad conditions so it is used in many applications such as [19]:

- Satellite surveillance in military.

- Mapping earth-resources.

- Environmental changes (climate changes, disasters, hazards).

- Traffic monitoring (mobility).
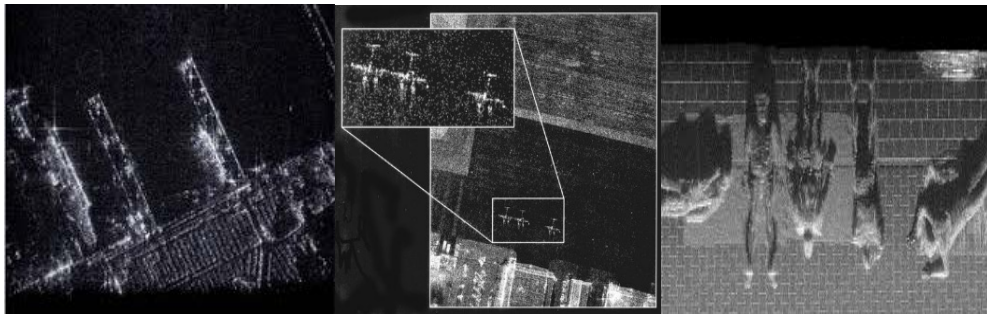
Figure 3.3: SAR Images for Different Applications.

3.3 GPRMax

The GPRMax was developed by DR Antonis Giannopoulos, A GPRMax is an open source simulation, a software that simulates the EM wave propagation. It can solve the Maxwell's equations using Yee's algorithm in the 3D shape with the use the FDTD (finite difference time domain) methods among a lot of things. It is modeled to be used in different fields such as engineering, geology, medicine, etc.

GPRMax in command-line principal is written in Python (hhttps://www.python.org). all software features along with examples explained in the user guide on the GPRMax website above. GPRMax has so many features that are powerful such as:

- Modeling of soil with realistic properties.

- Building objects with rough surfaces.

- Built-in libraries of antenna models.
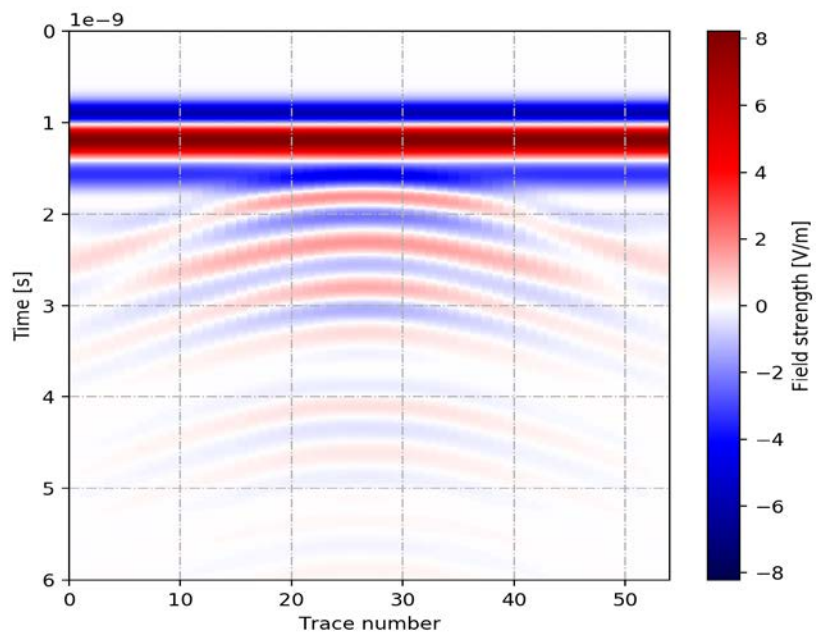
Figure 3.4: 2D Section, B-scan generated by GPRMax.

4. GPR DATA PROCESSING

4.1 Preprocessing and Processing Steps

After collecting the GPR data the main advantage of using the GPR technology, is that raw data that was collected in a way that it can be viewed easily an image using a computer (or as explained before any display unit).

Usually, the amount of processing depends on the data conditions and the application, mostly few processing steps is required for the representation of the collected data. On the other hand, more data processing is needed depending on what application and the object in place. A lot of techniques, algorithms, and filtering can be applied according to the condition of the data or the specific application in use, the figure 4.1 shows the basic step that can be used to process GPR data [22][25].
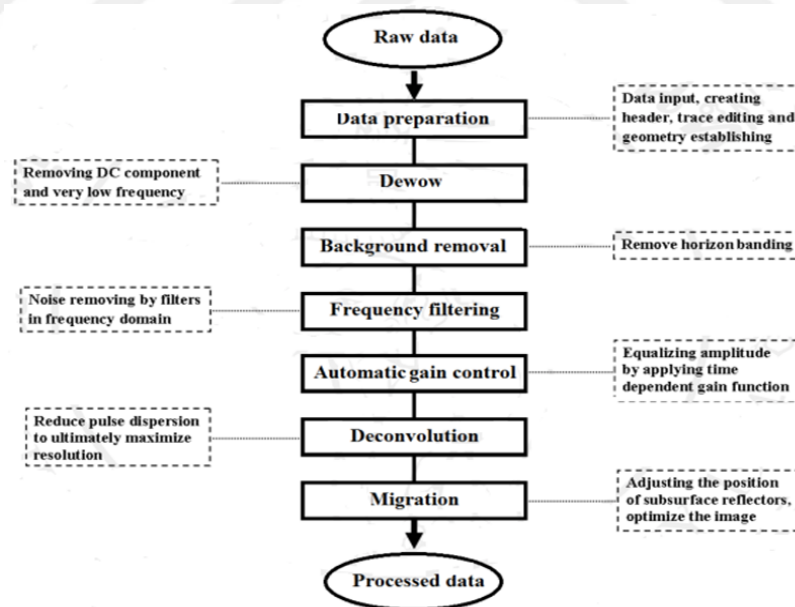
Figure 4.1: Some basic GPR processing step**s**.

As it is going now you can see that the main three steps in image processing can be listed as, collecting the data (obtaining the image), analyzing the image and manipulation, and obviously the last one is displaying the image [22].

In the first step as its named collecting an object in the image is recognized and converted to digitized format with an equipment like digital camera. Followed by improving the quality of the image is considered to be the analysis step, with some additional objects being removed from our image. And finally displaying it of the computer device to see what's under the ground [22]. Meanwhile the main goals of the image processing can be categorized as:

- Visualization which is the ability to detect invisible objects.

- Retrieval which is obtaining or detecting the image of the surface we are interested in.

- Sharpening and restoring by creating a high-quality image.

- Recognition where we can classify a correct object we want within the collected image.

- Measurement of the pattern for determining the object of interest within the image.

One of the significant steps in explaining GPR image processing is the preprocessing step, it is applied to remove noises and increasing the image quality. However, as we explained not all the steps are applied it all depends on the condition of the collected images from GPR scans, preprocessing includes data editing, dewow, zero- time adjustment, and filtering [22].

Figure 4.2: GPR Raw Data (Collected by GPRMAX).

4.1.1 Data editing

The first step can be represented by data editing which is the first activity to applied in any data processing. Mostly, the collected data needs to be sorted and arranged to insure high reliability in data interpretation, especially in the case of road investigations because of its considerable large amount of data.

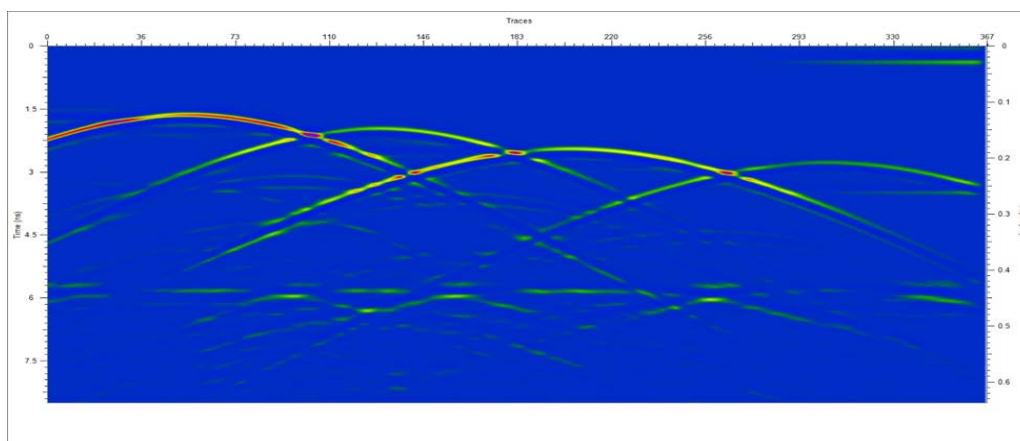There two kinds of faults levels that can happen generally in this step. The first level can be that the data is related to incorrect or inaccurate parameters setting during the survey. While the second level error can happen during the collection of the data.

Obviously, the quality of an image in the survey section could be affected somehow be incoherent, fuzzy or clipped traces [25][8].

4.1.2 Zero time Adjust

So, in any GPR inquiry, a need to zero-time adjustment mostly is required because that the first wave to arrive is the air wave which is delayed in time, the arriving time of the first break of the radar wave due to the length of cables connecting the antennas and the control unit.

For this reason, zero depth is needed associated with the zero-time, an offset caused by the instruments recording the data is removed before the analyzation of the GPR image [25].

Mostly zero-time adjustment (or correction) is used to correct or fixing all the traces to the zero-time common position before image processing [7].

Mainly this problem is solved by cutting the air layer and fixing a threshold, the proper threshold setting depends on both of central frequency and the antenna type in investigation [25].

4.1.3 Dewow

The wow elimination – (Dewow) or what we can refer to as signal saturation correction. The Dewow applies an averaging filter to each trace to remove the DC signal component or the low frequency from the GPR collected data [7]. This is wow- is caused by the DC Component that's is added to the data by early arrivals or the effects of inductive coupling.

The Dewow is preformed to reduce the collected data to a mean zero level, but if the Dewow is applied incorrectly this can result the data to have decaying, low frequency which will ruin the whole trace. Luckily, nowadays most GPR system is applying the Dewow automatically to each trace while setting the filter parameters to the optimal conditions [4][8].
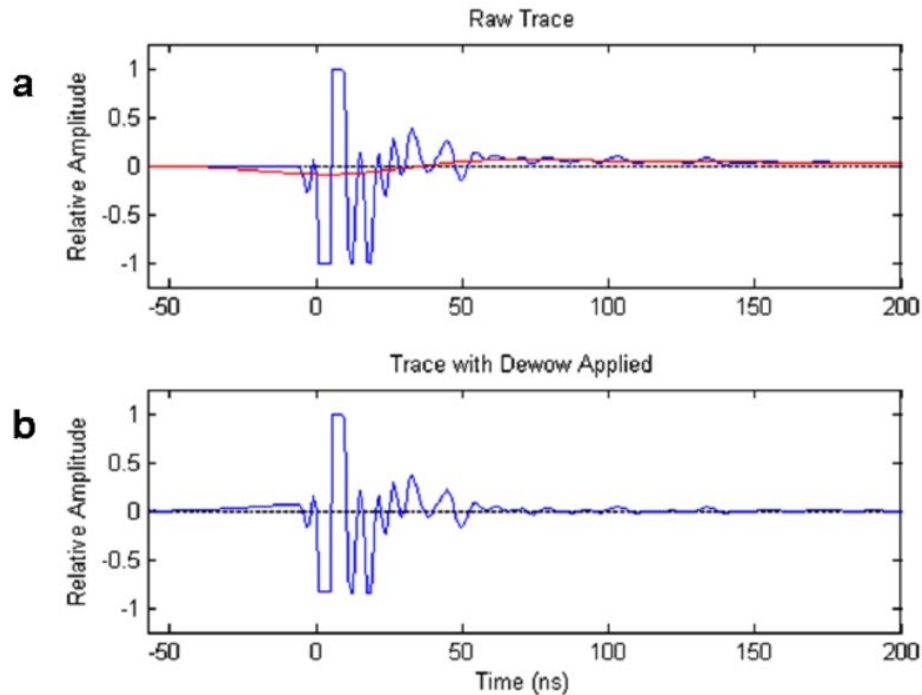
Figure 4.3: Dewow filtering: a) Raw Trace. b) Trace with Dewow Applied [8].

4.1.4 Background removal filter

The background noises in the signal are usually caused by the ringing in the antennas. This might produce coherent banding effects, parallel to the cross section and surface wave [26].

This is a simple filtering process that may be explained as it's the summation all the reflection amplitudes recorded at the same time in the process and dividing the number of summed traces, this resulting digital wave, which forms the average of the background noises, then it will be subtracted from the collected data. In this process the technician must be careful not to dismiss the real linear events in the data and must specify the operating window of the filter so that the filter is not implemented after the surface wave ends [8]. This filtering process as be explained in this equation:

$$y'(n) = y(n) - \frac{1}{k}\sum_{k=1}^{k} y_k(n) \qquad\qquad \textbf{(4.1.4)}$$

Where $y'(n)$ and $y(n)$ are the processed signal and the raw signal traces, in order, while the sample number is n and k are the number of traces within the selected data set of A-scans [8].
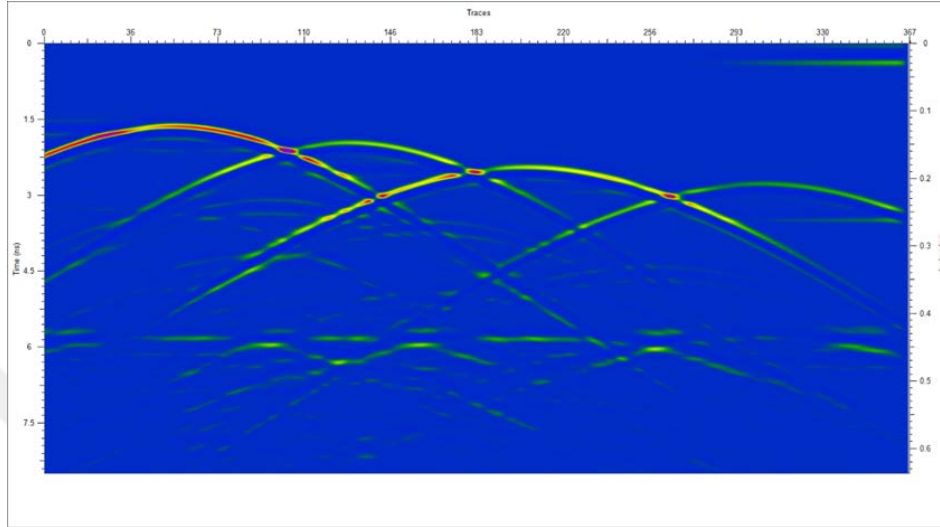
Figure 4.4: Background Removal 2D Section (GPRMAX).

4.1.5 Horizontal normalization

This horizontal stretch (or distance) is used to get the constant traces separating this correction requires removing non-constant motion effects all along the process. The correct representation of the image can be damaged if the required steps are not considered to correct the variable horizontal data coverage for the GPR collected data [8].

4.1.6 Gain

As it is known the gain part is used to indemnification (or make up) for the variations in the amplitudes of the GPR image. Some of the transmitted signals hits the ground and arrive faster (or to be more precise early arrivers), these early arrived signals have a much greater amplitude then the ones after it, this is caused by not traveling far like the rest of the signals. In a way or another the losses in the amplitudes of the signal are associated to its geometric spreading in the surface and their inherent attenuation. Normally different time variable gain functions are implemented for equalizing the recorded signals amplitudes.

One of the widely common methods used is the Automatic Gain Control (AGO), which is defined as a time varying gain of a running window in each trace for a specific length, for each and every place the average amplitude over the selected length of the window is found. So, the gain function is implemented so that the average at each specific point along the trace became a constant [8].

4.1.7 Topographic corrections

In some GPR survey profiles in the surveyed data elevation, topography is applied.

First of all, the trace window is implemented to the collected data to eliminate all these traces in surveys (as in early arrival signals) that will arrive before the time zero arrival. The real recorded data elevation in the GPR line are considered to enter the data processing step, by applying a time shifts for each individual trace these zero time arrivals are suspended from the topographic profiles [8][26].

4.1.8 Frequency filtering

Even though the GPR data is being collected with transmitting and receiving antennas with a certain known frequency domain, but the signals recorded will contain a band of frequencies around our frequency component. Simply frequency filtering is considered as a way to eliminate the undesired low and/or high frequency component to produce a much better GPR interpretable images [8].

In HPF, the high frequencies will be preserved in the signal and eliminate all unwanted low frequency component in that signal. For the low-pass filtering it is the other way around (just the opposite of high filtering process), it will remove all the unwanted high frequencies and keep the low frequency component of the signal [8]. While a band- pass filter is combining these previous effects, where this band-pass filter with reserve all the desired frequencies within the band and removes all frequencies outside of this band whether it was high or low [8].

4.1.9 Deconvolution

Deconvolution is defined as an inverse filtering operation that tries to eliminate the effect of the wavelet source in order to produce a better GPR images of the subsurface structure. So, when GPR pulsed signal propagates in the subsurface, what explains how the transmission and reflection responses of the subsurface (the physical process of the propagation wavelet interacts with the earth is called convolution [8].

The deconvolution operations can ruin (degrade) the GPR images when the signature of the source is not recognized, deconvolution works under the presumptions that propagation wavelet source is the minimum phase (i.e., the most of its energy is connected with early times in the wavelet operation). While this assumption will not necessarily work for the GPR signals. In the GPR technology, because it is so close to the ground, the ground itself will become a part of the antennas and the source pulse may vary from trace to trace and it is not necessarily the minimum phase. All filtering methods/operations must be implemented carefully in data processing [8][24].

4.1.10 Migration

In literature an important technique in processing is called migration, that aims to correct the fact that most energy in the GPR images are not necessarily correct and associated with the depths in the 2D survey line.

As previously explained with deconvolution process, migration can be seen also as an inverse processing step that effects the geometry of the GPR image from the subsurface with the respect to the geometry survey. As a simple explanation, in a subsurface scattering points would show up as a hyperbolic-shaped feature in the GPR image [24].

The migration process can connect all the energy in the wavelets making the hyperbolic feature with the point of diffraction and the imaging of the actual surface structure would be shown more clearly, the migration operation requires good evaluation of the surface EM wave speed in order to implement the right adjustment to the GPR images [24].

4.2 Data Processing Algorithms

4.2.1 Principle component analysis
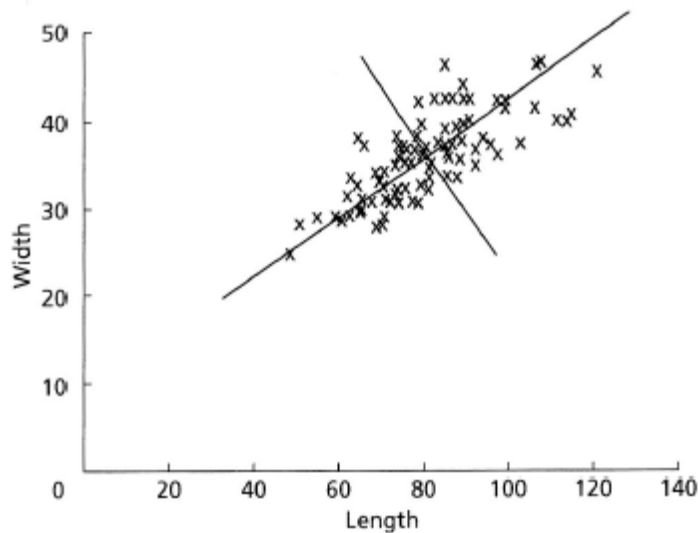
Figure 4.5: The Real Coordination Axis for Scattering the Length and the Width.
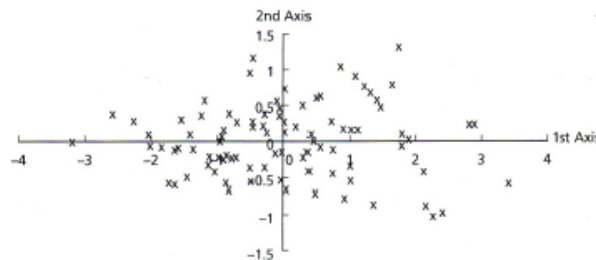
Figure 4.6: New Transformed Coordination Axis for Scattering the Length *and the Width.*

For explaining the principal component analysis (PCA) is way to identifying the patterns in a collected data set, and since patterns in some data might be little hard to establish such as, in high dimensional data. In that case the PCA technique is used to limit the number of dimensions, without ruining (or losing much of) the information.

Actually, a PCA technique is handful method for compression and classification of data sets. The main objective of this method is to minimize the dimensionality of the data set (samples) by obtaining a new set of variables,

much lower than the original data set, that small new set of data has most of the sample's information [14].

Here the general steps to perform a principle component analysis is explained:

1. Taking all the data set X that is consisted of D-dimensional data samples while disregarding the class labels.

2. Computing the mean of that d-dimensional data (the mean values for all the dimensions in the data set).

3. Computing the scattering matrix (or alternatively, the covariance matrix) of the data set.

- The following equation explains how the scattered matrix is computed:

$$S = \sum_{k=1}^{n}(x_k - \text{m})(x_k - \text{m})^T \tag{4.1}$$

Where the mean is referred to by m

$$\text{m} = \frac{1}{n}\sum_{k=1}^{n} x_k \tag{4.2}$$

Meanwhile, the covariance matrix can be explained as a dot product of the centered matrix $X_\circ$ that is divided by the number of the data samples:

$$C = \text{cov(X)} = \frac{1}{n-1}X_0\,X_0^T \tag{4.3}$$

4. Computing eigenvectors $(e_1, e_2, \ldots\ldots, e_d)$ and the corresponding eigenvalues $(\lambda_1, \lambda_2, \ldots\ldots, \lambda_d)$.

We can quickly check if the calculations of the eigenvector- eigenvalue is correct we apply it in this equation:

$$Ce = \lambda e \tag{4.4}$$

Where λ is it's the eigenvalue (a scalar) and e is its eigenvector. If all the vectors are independent there are M eigenvalues and eigenvectors, furthermore, for the correlation matrix which is real and symmetric with all of the eigenvalues are real and positive [12].

In general, once we find the eigenvectors and the C covariance matrix, the following move is to organize these vectors with the eigenvalues, from highest to the lowest, which shows us the components in the significant order [12].

Here we came to the last stage in the PCA method, and probably the simplest step, after we have picked up the components (eigenvectors) that we want to maintain in our data, we just take the transpose of that vector and multiply it on the left of the original data set. Now after transposing the matrix with the eigenvectors in the columns is now in rows (eigenvectors), with the most significant vector on top, with the mean-adjusted data transposed with every row containing a different dimension [12].

For additional reading regarding general information about PCA, we are referring to the reader to check section 7.5 of lay's textbook [16].

4.2.2 The singular value decomposition

Perhaps one of the widely used matrices is the singular value decomposition (SVD), this method is used to clutter reduction techniques that can be used to reduce noise (denoising), compression, and so on [17].

Mostly all the matrices have an SVD, which makes it more stable than the other methods, such as the eigen-decomposition. So, for that it is often used in wide area of applications including compression, data reduction, and denoising [18].

This SVD method of decomposing a matrix into three other matrices:

$$A = USV^T \tag{4.5}$$

Where: A is an $m \times n$

U is an orthogonal matrix $m \times n$

S is a diagonal matrix $m \times n$

V is an orthogonal matrix $n \times n$

In calculating the SDV method it is consisting of calculating the eigenvalues and eigenvectors AA^T and A^TA. The eigenvectors of A^TA will construct the columns of V, while the eigenvectors of AA^T will construct the columns of U [18].

besides, the singular values in S are the square roots of eigenvalues from AA^T or A^TA.

The singular values are the diagonal records of the matrix and are organized in a certain arrangement. These singular values are always real numbers, if the matrix A is a real matrix then U and V are real also.

We now know that for a matrix X with $n \times n$ dimension, then u is the nonzero vector which is the eigenvector of X if:

$$X u = \lambda u \quad \text{then calculate} \quad (X-\lambda I) x = 0 \tag{4.6}$$

So, for some scalar λ is what we can call eigenvalue of X, and u is called eigenvector of X, which is corresponding to λ. So, to calculate these values of the variables above, we evaluate these matrices AA^T and A^TA.

Just as we mentioned earlier, the eigenvectors of AA^T will form the columns of U matrix so we can carry out this analysis to calculate U [18].

4.2.3 Independent component analysis

One of the most important statistical and computing techniques is called Independent component analysis (ICA) method, this method is used for discovering and uncovering invisible elements that is situated under sets of random variables, measurements, or signals.

In ICA model, we assume that the data variables are linear mixtures of some unrevealed obscure (latent) variables, and the mixing system is also unidentified. These latent unknown variables are concluded to be non-gaussian and reciprocally independent, so obviously it was called the independent component of the recognized data.

These independent components are as well named sources or factors, can be calculated by the use of ICA [28].

Here, it comes to explain what makes ICA method different (or distinguish) from alternative methods, is that ICA searches for both statically independent, and non-gaussian. Now, we concisely explain the essential concepts, applications, and the principle calculations for the ICA model [27]:

ICA model is used in many applications with many different research areas due to its generality, some of them now are being mentioned [27]:

- Brain imaging technology.

- Econometrics.

- Image feature extraction.

The ICA model assumes that there are (n) independent signals, where $g_i(t) = 1,2,\dots,n$ and some mixing matrix B

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,n} \end{bmatrix}$$

Signals can be observed $x_i(t)$ that are represented by the linear superposition of $g_i(t)$:

$$x_i(t) = \sum_{i=0}^{n} b_{ij} g_j(t)$$

So as explained the main difficulty is in fact both b_{ij} and $g_i(t)$ are unknown to us, however when ICA solves a problem it uses assumptions, that the signals in the mixture are statistically independent and has a non-gaussian distribution. To make is simpler, the general idea is that if we have two random variables like y_i and y_j we can say that they are statistically independent if the information about one variable says nothing about the other variable [28].

To see it from the mathematical point of view, the word statistical independence means that the 2D probability density function $p(y_i, y_j)$ is the product of the 1D probability density function:

$$p(y_i, y_j) = p(y_i)p(y_j)$$

Meanwhile for the statistically independent signals, the covariance matrix of the odd functions $f(y_i)$ and $h(y_j)$ is a diagonal, all the mutual covariances are set zero:

For statistically independent signals the covariance matrix of odd functions $f(y_i)$ and $g(y_j)$ is diagonal: all mutual covariances are zeros:

$$E[f(y_i)h(y_j)] - E[f(y_i)]E[h(y_j)] = 0$$

and all self-covariances are non-zeros:

$$E[f(y_i)h(y_j)] - \mathrm{E}[f(y_i)]E[h(y_i)] \neq 0$$

4.2.4 Probabilistic independent component analysis

Probabilistic ICA (alternatively called noisy ICA, or independent factor analysis, although we will reserve the latter term to a more specific method in which factors are Gaussian mixtures) assumes a small number of independent components, with a residual term which is modeled as Gaussian noise. As we formalize the algorithm of probabilistic ICA as the generative of the linear latent variables in model, so similar to square noise-free case.

The PICA model is characterized by the assumption of that p variate vector of observation is calculated from the set of q statistically independent, non-gaussian sources via a linear instantaneous mixing process that is damaged by adding Gaussian noise $\eta(t)$ [29]:

$$x_i = As_i + \eta_i + \mu \qquad \forall i \in v$$

In here, we see that x_i shows the p dimensional column vector of individual measurements at voxel location i, and the s_i shows the q-dimensional column vector of the non-gaussian source signals that is contained in the data, and η_i is the gaussian noise , where $\eta_i \sim N(0, \sigma^2 \sum i)$.

We can assume that $< p$, i.e. which means that there are fewer source processes than the observations in time [30].

We can say that the covariance of the noise is authorized to be voxel dependent, so it allows the immensely different noise covariances that are recorded in different types. The mean of the recorded x_i is defined by vector μ, where i is the index over the set of all voxel locations V and $p \times q$ matrix A is supposed to be non-decadent, i.e. of the rank q. So, in order to solve the blind separation problem, we are required to find the linear transformation matrix W [30]:

$$s = Wx$$

This is fine estimation to the true source signals s.

To be clear PICA model is pretty equivalent to the standard GLM with small differentiations. As opposed to the designed matrix in the GLM model, the mixing matrix A is no longer pre-identified before model fitting, but it will be

calculated from the data as part of the fitting model. The spatial source signals are corresponded to the parameter estimations in the GLM model with the additional constraint of being independent statistically [29].

4.3 Image Processing Filters

In image processing the term filtering is known as technique used for adjustment or image enhancement. As a simple example, you can use a filter on an image to focus on some specific features or disregard other features. So many image processing operations is performed with filtering which includes smoothen, edge enhancement, and sharpening of an image [36].

As the image enhancement (IE) explained, it concerns to emphasizing, or sharpen of image features such as edges, boundaries or contrast to make the graphical display more handful for displaying and analyzing. Actually, the enhancement process will not increment the collected information content in the data. But in fact, it does increase the dynamic range of the selected features, so that they are smoothly observed.

So obviously, the main reason of the image enhancement is to get better the quality images, which is mainly removing noise, deblurring objects edges, and to highlight some specific features.

In general idea, image enhancement will improve humans viewing capabilities and it will increase the possibility to success in digital image processing to reach a specified goal [37].

Mainly image enhancement can categorize into two sections:

- **Spatial Domain Methods**

This method is consisted of the procedure that operates immediately on the pixels of the concerned image. For example, it can smooth the image by averaging, and sharpen by gradient type of operators, global enhancement by means of histogram modification techniques or contrast stretching [37].

- **Transform domain methods**

The transform domain is consisting of computing the 2D frequency domain transform (Fourier/ Hadamard) of the image we want to enhance, changing the

transform, and calculate the invers to return the image that has been enhanced in a certain way. For example, high pass and low pass filters are used for smoothing and sharpen, and the homomorphic filtering is used to manipulating the illumination effects. In the following, spatial domain image enhancement methods which are widely implemented will be discussed [37].

- **The Spatial Filtering**

The number of changes in the brightness value per unit distance for any specific type of the image is defined as the spatial frequency. So, the spatial domain filtering is the context of depending operations (which they depend on the properties of the neighboring pixels) that alert the GL of the pixels corresponding to its relationship with the other pixels in that current surrounding area [35]. So, it was named a low frequency area, if we noticed that there were fewer changes in the GL per unit area. Meanwhile if there were a lot of changes in the GL per unit area, here we call it as the high frequency area [36].

Fourier analysis is defined as the mathematical techniques for separating within its various spatial frequency component. And it is possible to focus on a particular group of frequencies relatively close to others, and then combine these spatial frequencies to construct the enhanced image. An algorithm that preforms these enhancements are the filters, because as known the stop specific frequencies and let the others pass.

There are three general types of filtering that can be illustrated together to form more complexed filters [36]:

- Low Pass filtering (LPF)- is used to suppress high frequencies content and emphasize gradual changes, like smoothing details in the image and reducing the range.

- High Pass filtering (HPF)- is used to get fine details and edges, so it enhances the details.

- Band pass filtering (BPF)- as the name suggest, this filter does not have a general application in the image processing part, but it is primarily used in suppressing the periodic noises.

In this step we will be explaining some of the low-pass filters. A large variety of the low frequencies are implemented in image processing step, choosing a specific type of low pass filters, like the smoothing filters, as the name suggests to us, they smooth high frequencies noise/ information, depending on the image and the purpose of it. Some filters are explained below [36]:

- Mean filter.

- Weighted mean filter.

- Mode or majority filter.

- Median filter.

The mean filter replaces the pixel value by the median of its neighboring pixels. From the concept the median filter is very simple, but because of the sorting requirement, the computational part is expensive to implement. However, this filter is quite good for removing noises such as, salt and pepper noise, and impulse noises while maintaining the image details because they don't depend on values which are significantly different [36].

- **K-nearest Neighbor Filter**.

k- nearest neighbor filter is just another filter used in edge preserving smoothing, where the Central pixel in the image window is changed with the average of the K pixels, which are the closest to the central pixel in that window. A typical 3×3 square widow has a value of K=4.

- **Sigma Filter.**

This sigma filtering filters the sets of the central window pixels equal to the average of all the pixels in its neighborhood whose values are within the K counts of the central pixel value, where the adjustable parameter is K.

This process is called sigma filtering because the parameter K can be derived from sigma or the standard deviation of these pixel values in the window.

- **Gaussian Filter.**

The image in here is smoothed by the assumption that they are grey level values and they are distributed as gaussian function [36].

40

4.3.1 Gaussian filter

The gaussian (G) filters are one of the only filters which are separable and, circular symmetric (at least to the lattice approximation). The gaussian filters are linear filters with the chosen weights according to the gaussian function shape [33].

So, the gaussian smoothing filter is a quite good filter for removing noises drawn from normal distribution. And the zero-mean one-dimension gaussian function is:

$$g(x) = e^{\frac{-x^2}{2\sigma^2}} \tag{4.7}$$

Where the G spreads the parameter (J determines the width of the gaussian function). In image processing, the 2D zero-mean discrete Gaussian function is used as smoothen filters.

$$g[i,j] = e^{\frac{-(i^2+j^2)}{2\sigma^2}} \tag{4.8}$$

The G functions (F) has some properties that makes it very handful in early visions of processing. These properties show that the gaussian smoothing filters are an effective LPF from the both respective the spatial and the frequency domains.

They are efficient to implement and effectively used by engineers in wide area of applications. These five properties are listed below [33]:

1. For the GF in two-dimensions, they are rotationally symmetric this means that the smoothing amount will be performed equally in every direction.in general process, the edges in the images will not be oriented in any particular direction, which is known in advance.

2. These gaussian functions has a single lobe, which means that the gaussian filter is smoothing by changing the image pixels with a weighted average of the neighboring pixels such that the given weight decreases monotonically with the distance from the central pixel in the image. This property is very important since the edge is the local feature in the image, and smoothing operations that will give more significance to pixels farther away will distort the features.

3. The FT of the Gaussian has a single lobe in the frequency spectrum. This property is a direct consequence of the fact that the FT of the Gaussian function, is itself a gaussian. The images are sometimes damaged by unwanted signals with higher frequencies (noises). While the desired image features, like edges, will have the components at both high and low frequencies. The single lobe in the FT of the gaussian function means that these smoothed images will not be damaged by contributions from the undesired high frequency signals, and these desirable signals will be maintained.

4. In Gaussian filter, width (j) and smoothing degree are parameterized very simply. A large (j) signify a wider Gaussian filter and a better smoothen operation. The degree of smoothen in the Gaussian filter is adjustable, to perform a compromised between excessive blur of the features in the wanted image (caused by using smoothing too much) and the unreasonable variations in the smoothed image due to the noises and fine texture (caused by using small amount of smoothing).

5. The Gaussian filters can be used quite efficiently because these gaussian functions are separable. The convolution of the 2D Gaussian can be implemented by convolving the images with one- dimensional gaussian and then the convolving result with the same 1D filter to the Gaussian is performed in the earlier steps of the signal processing.

4.3.2 Median filter

As previously explained the median filter (MF) is non-linear digital filtering technique, that can be used to remove noises from images or signals. like a noise reduction method is a classical stage in preprocessing to amend the results of the processing such as, edge detection on an image [38].

The MF is a quite effective method that can be used, to some extent, to distinguish out of -range isolated noises from image features such as, edges and lines. To be very specific, the MF replaces the pixel by the median instead of the average of all pixels in the neighborhood W:

$$y[m,n] = median\{x[i,j], (i,j) \in W$$

Where W is representing the neighborhood, which is defined by the user, centered in the image around the location [m, n]. The non-liner function of the median filter is expressed as [37]:

$$y(n) = med[\, x(n-k), x(n-k+1), \ldots\ldots, x(n+K-1), x(n+k)] \qquad \textbf{(4.9)}$$

Where y(n) is our output and x(n) is the input signal. This filter collects a window containing N=2K+1 input samples, and then performs the median filter on these set of samples [39].

- **Threshold decomposition**

Analyzing the combinations of the linear filters using the principle of superposition is in so many cases much easier than analyzing the combinations of the non-linear devices like median filers. However, using what is called threshold decomposition method allow us to divide the analysis problem in the median filter into small parts.

This means we can decompose it into M-1 binary $x^1(n), x^2(n), \ldots\ldots, x^{M-1}(n)$

$$x_m(n) = \{\, 1 \; if \; x(n) \geq m_0 \; else \qquad\qquad \textbf{(4.10)}$$

$$x(n) = \sum_{m=1}^{M-1} x_m(n) \qquad\qquad \textbf{(4.11)}$$

Now, one of the very interesting properties of the median filter is that instead of filtering the original M-valued signal, we can decompose this into a M-1 channels (eq 2) each containing a binary median filter component [39].

The first use of the median was suggested for smoothing the statistical data. In this filter type, we have found most of its applications in the area of the digital image processing. An edge-preserving filter is like the median filter can removes noise and speckles without blurring the images. Removing artifacts from imperfect data collection, for example horizontal stripes sometimes produced optical scanners is done successfully using median filtering. Median filters are also has been used in radiographic systems. These types of filters are also likely to be found in digital commercials sets because of its ratio, very good cost-to-performance [39].

4.3.3 Wiener filter

In the term of signal processing, the wiener filter is the type of filter to use to produce an estimate of the desired or random process of the target by linear time-invariant (LTI) filtering of these observed noisy process, assuming the known stationary signal and noise spectra, and additive noise [40].

The wiener filtering is optimum in the manner of the mean square error. In other words, we can say that the wiener reduces the overall error in the process of the inverse filtering and smoothen noise. The wiener filter can be expressed in the Fourier domain as following [41]:

$$W(f_1, f_2) = \frac{H^*(f_1, f_2)S_{xx}(f_1, f_2)}{|H(f_1, f_2)|^2 S_{xx}(f_1, f_2) + S_{\eta\eta}(f_1, f_2)}$$

where $S_{xx}(f_1, f_2), S_{\eta\eta}(f_1, f_2)$ are respectively the power spectra of the original signal and the additive noises, and $H(f_1, f_2)$ is the blurring filters. It can be easy to see the wiener filter has two separated parts, the inverse filtering part and the noise smoothing part. The wiener filter does not only perform the deconvolution by inversing filtering (high pass filtering), but also it removes the noises with the compression operations (low pass filtering) [39].

- **Implementation of wiener filter**

For implementing the wiener filtering in practice, we need to estimate the power spectra of the original image and their additive noises. For the white additive noises of the power spectrum is equaled to the variance of that noise. In the term of estimating the power spectrum of the original images many methods can be used in it.

A direct estimation is the periodogram estimated of the power spectrum calculated from the observations [39].

$$S_{yy}^{per} = \frac{1}{N^2}[Y(k, l)Y(k, l)^*]$$

Where the observations of DFT is donated by Y (k, l). the main advantage of the estimation is that it can be implemented quite easily without concerning about the singularity of the inverse filtering. Another estimation leads to cascade implementations of the inverse filtering and noise smoothing is:

$$S_{xx} = \frac{S_{xx} - S_{\eta\eta}}{|H|^2}$$

This is a straightforward result of the fact: $S_{yy} = S_{\eta\eta} + |H|^2$, the power spectrum S_{yy} can be directly estimated from the observations using the periodogram estimation. Now this estimation results in the cascade implementation of the inverse filtering and the noise smoothing:

$$W = \frac{1}{H} \frac{S_{yy}^{per} - S_{\eta\eta}}{S_{yy}^{per}}$$

Meanwhile, the main advantage of this implementations is that when the inverse filter is set to be singular, then we have to use the generalized inverse filtering [39].

4.3.4 Histograms

Mainly a histogram of an image is a graph representation, a graph that displays the frequencies of any plotted data. Generally, histograms have bars that represents the plotted frequency of the appearing data in the all data set [44].

- **Image Histogram**

As explained the histogram of any image is just like other histograms it also displays the frequencies, but in the image histogram, they show frequency of pixels values. In the image of histogram, normally the x axis shows the gray level intensities and the frequencies of these intensities are shown on the y axis [44].

- **The Applications of Histograms**

The histogram filter has been implemented in many the image processing applications. Firstly, as it has been mentioned before the image analysis. We predict the image by checking its histogram. It is like watching an X-ray of a body bone.

Secondly, the histogram is implemented for the brightness purposes. So, histogram has a enormous applications in image brightness. Not just in brightness but the histograms also implemented in adjusting contrasts in the images. Another significant implementation of histogram filtering is equalizing

an image. So, they are also, utilized in thresholding techniques. This is mainly applied in computer vision [44].

4.3.5 Histogram equalization

The histogram equalization is one method used to process images by modifying the intensity distribution of histogram in order to adjust the contrast of these images. The main objective of this technique is to give a linear trend to the cumulative probability functions associated to the processed images [45].

As simply as it can be explained, the processing of the histogram equalization depends on the uses of the cumulative probability function (CDF). The CDF is a cumulative sum of all the probabilities in the domain is explained by:

$$cdf_x(i) = \sum_{j=0}^{i} P_x(j)$$

The main idea of this processing step is it to give to the resulting image in a linear cumulative distribution function. Indeed, as the CDF is clearly associated to the uniform histogram that we want the image to have in the result [45].

The histogram equalization usually provides unrealistic impact in the collected images. However, it is quite useful for scientific images such as, thermal, satellite or x-ray images, often the same classes of the images to which one can apply false-color.

Also, in histogram equalization would provide undesirable impact such as, visible images gradient. When its implemented to the images with low color depth. As example, if we implemented it to 8-bit images viewed with 8-bit gray-scale of these images. In the histogram equalization works better when it is implemented to images with relatively higher color depth than the palette size [46].

From this, two ways are presented to thing about and to implement   the histogram equalization, is you either implement as image changes or palette changes. This operation can be expressed as P(M(I)) where the I represent the original image; the M is the histogram equalization mapping operations and the palette is presented by P.

So, if we want to define a new palette such as $P' = P(M)$ and leave the image I the same (unchanged) then we can see that the histogram equalization here is implemented as a palette change. Meanwhile, if the P palette remains the same and the image is then one modified to $I' = M(I)$ we can say that the implementation is by image change. To conclude this in most cases the palette change is much better as it preserves the original collected data [46].

For example, here we examine a discrete grayscale image {x} and by letting n_i be the number of events of the level gray i. Then the probability of an event of a pixel of the level i in the image is here by:

$$p_x(i) = p(x = i) = \frac{n_i}{n}, 0 \leq i < L$$

By letting L be the total number of the gray levels in the image (typically 256), and the total number of pixels in the image donated by n, and $p_x(i)$ is as a matter of fact the images histogram a pixel value i, normalized to the values [0,1]. As way, considering the cumulative distribution functions corresponding to p_x as mentioned above which is also the image accumulated normalized histogram [46].

So, we would build a transformation of the form $y = T(x)$ to fabricate a whole new image {y}, with a flat histogram. Such image could have a linearized cumulative distribution function (CDF) across the range value, i.e. [46].

$$cdf_y = iK$$

For some values of constant K, the properties of the CDF allow us to implement such a transform (check the inverse distribution function), it can be explained as:

$$cdf_y(y') = cdf_y(T(k)) = cdf_x(k)$$

Where k is in the range of ([0, L]), notice above that T changes all the levels into the range of [0,1], and since we used a normalized histogram of {x}, in order to map all the values back to their original ranges, the simple transformation would be needed to be applied on the results [46]:

$$y' = y . (\max\{x\} - Min\{x\}) + \min\{x\}$$

47

4.4 Fast Fourier Transform

Actually, the fast Fourier transform (FFT) algorithm computes the discrete Fourier transform (DFT) of the sequence, or its inverse (IDFT). As known the Fourier analysis is converting a signal from its original domain (usually time or space) to its representation int the frequency domain or the other way around [42].

Usually the Fourier transform decomposes the images into its real and imaginary component which is the representation of the image in the frequency domain. If the input signal is an image, then the number of frequencies in the frequency domain is pretty much similar to the number of the pixels in the image or its spatial domain. The FFT and its inverse of the 2D image are given by these equations below:

$$F(x) = \sum_{n=0}^{N-1} f(n).e^{-j2\pi(x\frac{n}{N})}$$

$$f(n) = \frac{1}{N}\sum_{k=0}^{N-1} F(x).e^{-j2\pi(x\frac{n}{N})}$$

Where f(m, n) is the pixels at the coordinates (m, n), and F(x, y) represents the image in the frequency domain corresponding to the coordinates x and y. while M and N are the dimensions of that image. As it becomes obvious from seeing the equations, the implementation of this algorithm is quite expensive, but the beauty of using FFT is that it is separable, the 2D transform could be done as two 1D transforms as shown below (showing only in the horizontal direction) normally showing one in the horizontal direction then followed by the other vertical direction on the results of the transform.

This result is equivalent to preforming the 2D in the frequency space [43].

$$F(x.y) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)e^{-j2\pi(x\frac{m}{M}+y\frac{n}{N})}$$

$$f(m,n) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} F(x,y)e^{-j2\pi(x\frac{m}{M}+y\frac{n}{N})}$$

So, the FFT that was implemented required that the dimension of the image is power of two in this application. Anyway, the output of the Fourier transform is complex number and it has a much greater range than the spatial domain image. Therefore, to represent these values accurately, they are stored as floats. We can conclude that FFT is much simpler and faster than DFT as a method to implement, this is very useful when the value of N is quite large [43].

For the main and most important part of research is calculating the PSNR values for the methods, as we all know PSNR stands for the Peak Signal to Noise Ratio. Which is defined as the ratio maximum power and the power of the signal noise. So, PSNR is often explained in decibel which is a logarithmic scale [47]. So, the first step to calculate our PSNR, we must first calculate the mean square error, then sum these mean square errors and divide them by the number of matrix elements (rows* columns) [47].

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i,j) - K(i,j)\|_2^2$$

Then PSNR is expressed by:

$$PSNR = 10 \times \log_{10} \left(\frac{MAX_I^2}{MSE} \right)$$

In the next Chapter 5, we will be presenting our results regarding the methods or to be precise filtering and algorithms, that were defined in this chapter. The next chapter will only display the results we got from the data that we collected using GPRMax.

5. RESULTS

As we mentioned above in the previous chapter all the filtering methods and algorithms used in signal and image processing, in this research we are discussing underground water detection. We obtained 4 set of data for water detection to compare between our algorithms results to know which method is best in each situation by comparing their PSNR values from our MATLAB result by generating the formulas of these methods.

First, we brought a box of (480*148*170) mm filled with sand and scanned this box with our antenna on top of the box every 5mm.

5.1 Detecting Fresh Water in Sand

For our first scenario using GSSI 1.5 GHz 'like' antenna with time window: 6e-9. We brought the box filled with sand with the dimensions (480*148*170) mm, a sphere (with 10 mm radius) is placed at (240*74*100) mm within the sand and filled with fresh water. Our antenna is placed 5mm above the top of the box and scanning every 5mm.
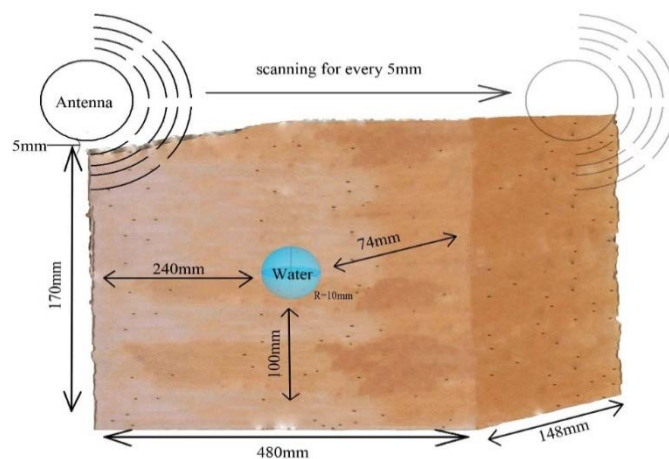
Figure 5.1: Shows the Box of Sand with a Sphere Filled with Water.

With Sand coefficients having these values for this experiment: Relative permittivity: 3, Conductivity :0.001, Relative permeability: 1.0, Magnetic loss:

51

0.0. And our Fresh Water coefficients are the Relative permittivity: 80.0, Conductivity :0.5, Relative permeability: 1.0, Magnetic loss: 0.0. The figure above shows our experiment box.
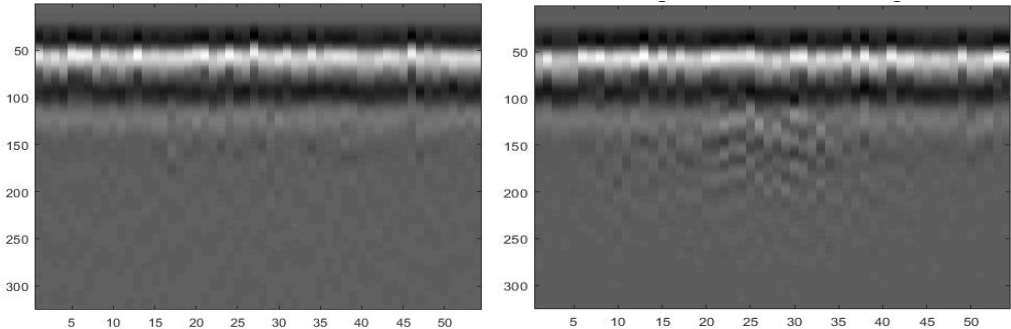
Figure 5.2: Shows First the Original Scanned Data for Soil and then the Soil and Fresh Water Data.

In these figures above you can see that while scanning these data from first and the second experience, that the image of the box having water in it there is a quite disturbance in the middle on the contrary of the other image that has only soil in it. This soil only image is our reference to show the great difference in the GPR image.

To go step by step firstly, in calculation we subtract the mean to simply decluttering and removing the noise from our data and its power spectrum using FFT with PSNR value 57.1904.
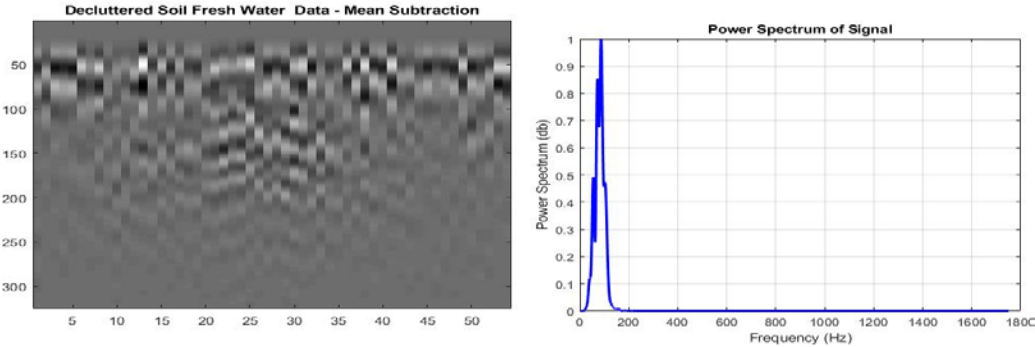
Figure 5.3: Shows the MS Method on the Data and its Power Spectrum.

The other method we used is SVD method which is as explained it is part of the other methods, calculating and transposing the matrix reducing the size and finding our new matrix, calculating its PSNR value 64.4371 and displaying its power spectrum.

52
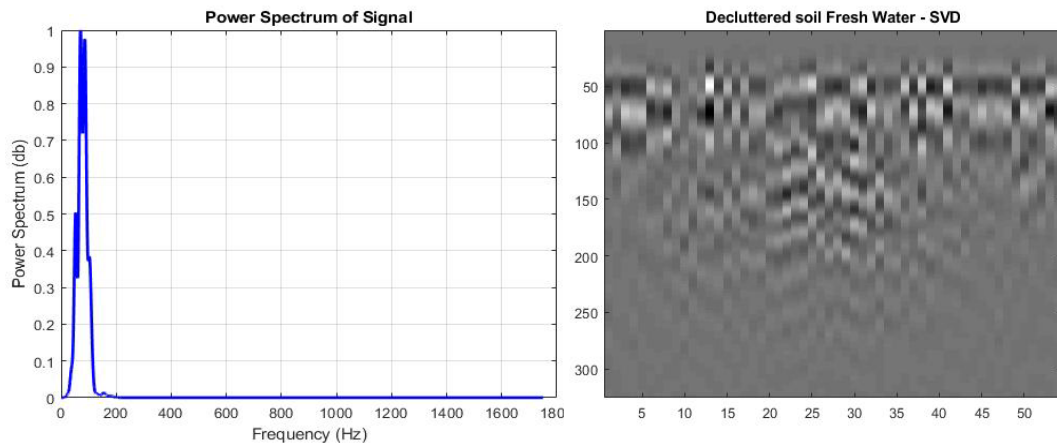
Figure 5.4: Shows SVD Method Effect on the Data and its Power Spectrum.

Then by using the algorithm explained, and by selecting the feature numbers for each of PCA, ICA, PICA we got these results:

As it was explained we calculated the PCA and selected the features and remove the mean value, estimating the covariance matrix, the eigenvectors and eigenvalues while sorting it in decreasing order, and finally constructing the image again then plotting its power spectrum using FFT, with calculating the PSNR value equals to 56.0816
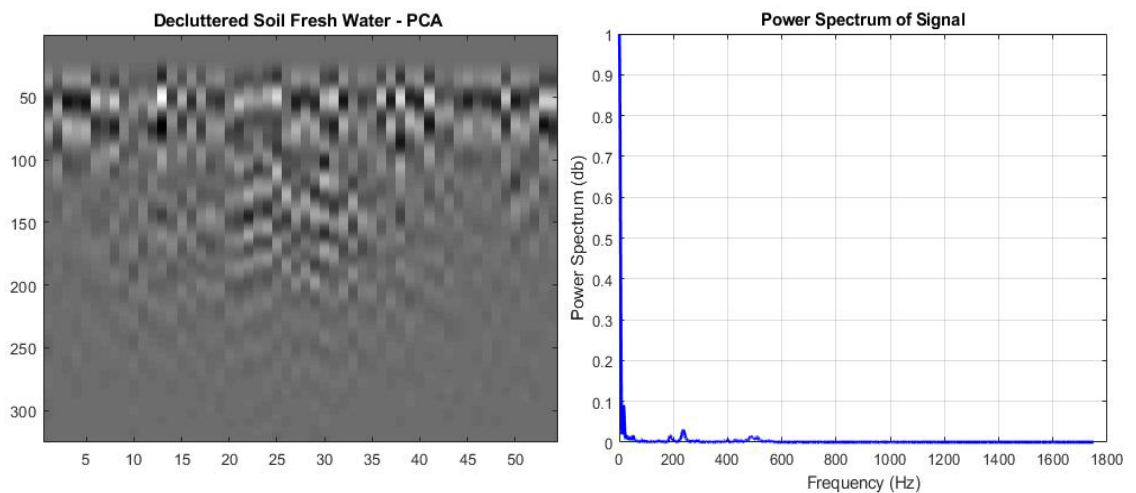
Figure 5.5: Shows PCA Method Effect on the Data and its Power Spectrum.

Even though these methods (PCA, ICA, PICA) are almost the same in process steps and there is much similarity in their coding algorithms but the efficiency of each method and how it effects the result is quite different.

As the PCA works, ICA method also requires finding the matrix and the independent values as previously discussed, and plotting its power spectrum, finally calculating its PSNR value 64.7930
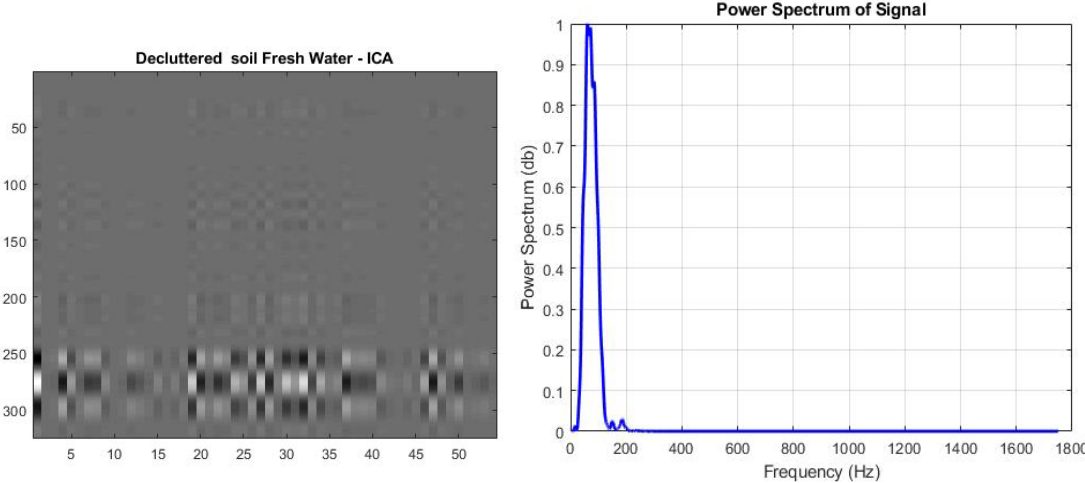
Figure 5.6: Shows ICA Method Effect on the Data and its Power Spectrum.

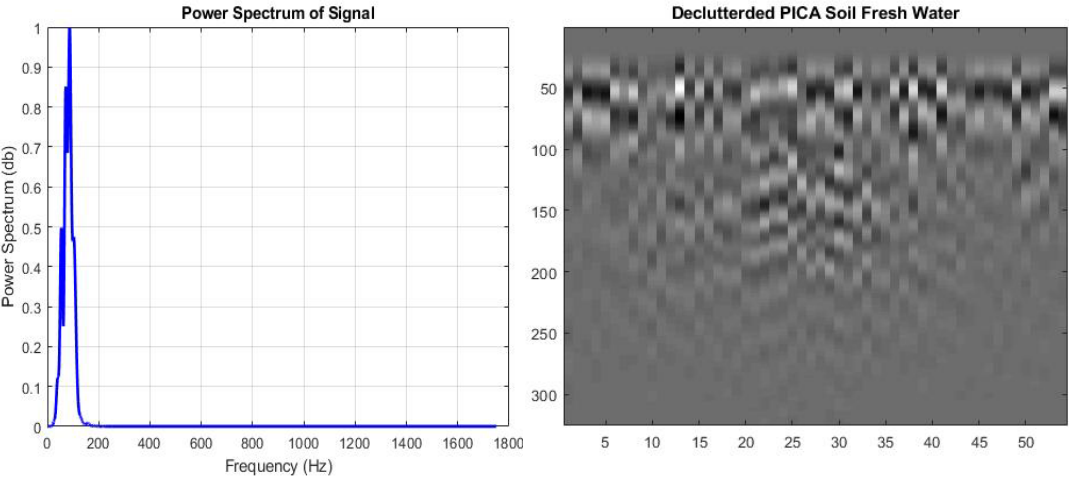In the same process we found the PICA effect of our data, we found that the PSNR value is equal to 65.2304

Figure 5.7: Shows the PICA Method on the Data and its Power Spectrum.

As for the filters we used the gaussian, histogram, median, and wiener filters in order are shown, we added some noises and removing it then adjusting these images using MATLAB codes to make our results enhanced.
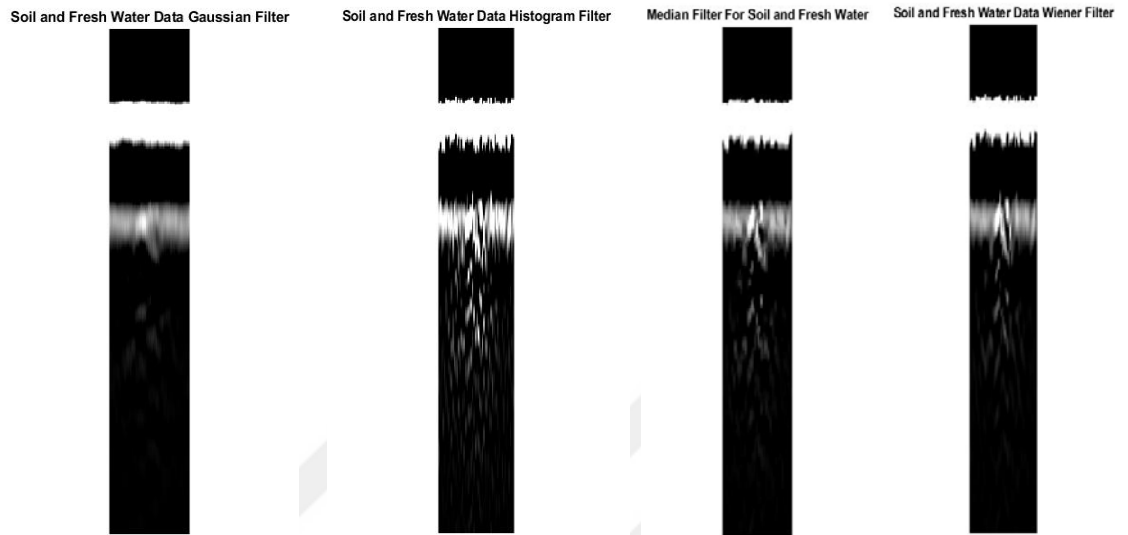
Figure 5.8: Shows the Filters Effect (Gaussian, Histogram, Median, and Wiener Filters) on Soil and Fresh Water Data in order.

And their PSNR values are like the following: 48.2118 for gaussian filter, 52.9641 for histogram, 48.0853 for median filter, and finally 48.0106 for the wiener filter.

5.2 Detecting Fresh Water in Soil with Some Roughness

For our second experiment, we have the same size of box, with same dimensions (480*148*170) mm, this box is filled with soil, the sphere (with 10mm diameter) is placed in that box with this location (240*74*100) mm, fresh water was injected into this sphere. Roughness is added now (valleys that they are now up to 5 mm deep and the peaks are up to 5 mm tall.)

Our antenna is on top of the box 5 mm high, and scanning the surface every 5mm at a time, with the use of GSSI 1.5 GHz 'like' antenna having a time window: (6e-9).
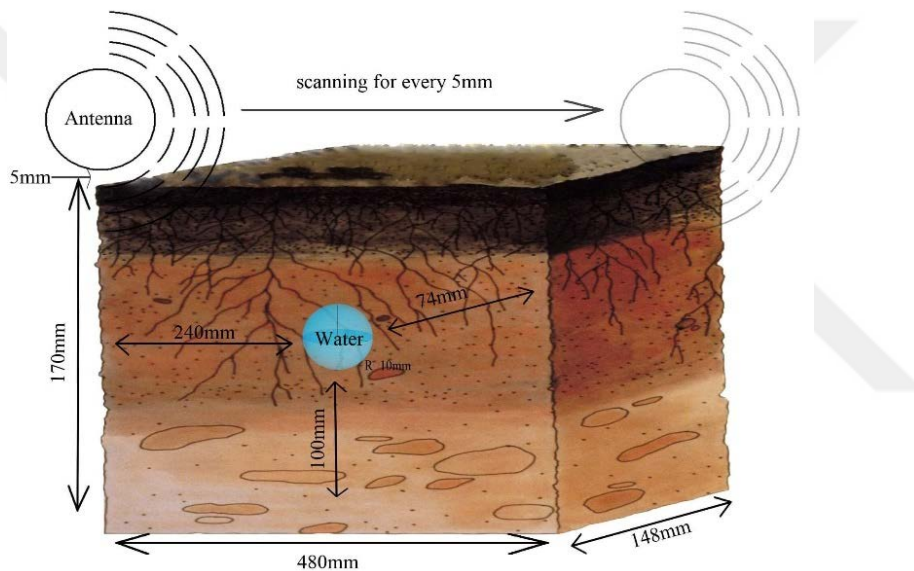
Figure 5.9: Shows the Box After Adding Some Roughness to the Soil.

Using soil Peplinski, a soil and sand fraction 0.5, clay fraction 0.5 also, bulk density 2g/cm^3, with the sand practical density of 2.66 g/cm^3 and a volumetric water fraction range of 0.001- 0.25.

As before our fresh water characteristics; the relative permittivity 80.0, the conductivity 0.5, relative permeability of 1.0, and with a magnetic loss = 0.0. As explained before in this experiment we are changing or to be more accurate we are testing different data from different soil situations, using the same methods that have been explained before.

As has been mentioned previously, first, we are showing (or comparing) the soil only image with the one we collected after putting the sphere inside the soil and adding the roughness in this stage.
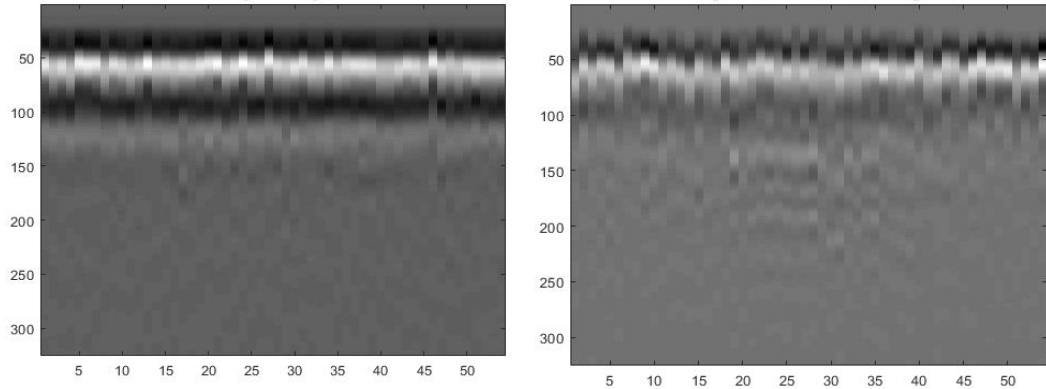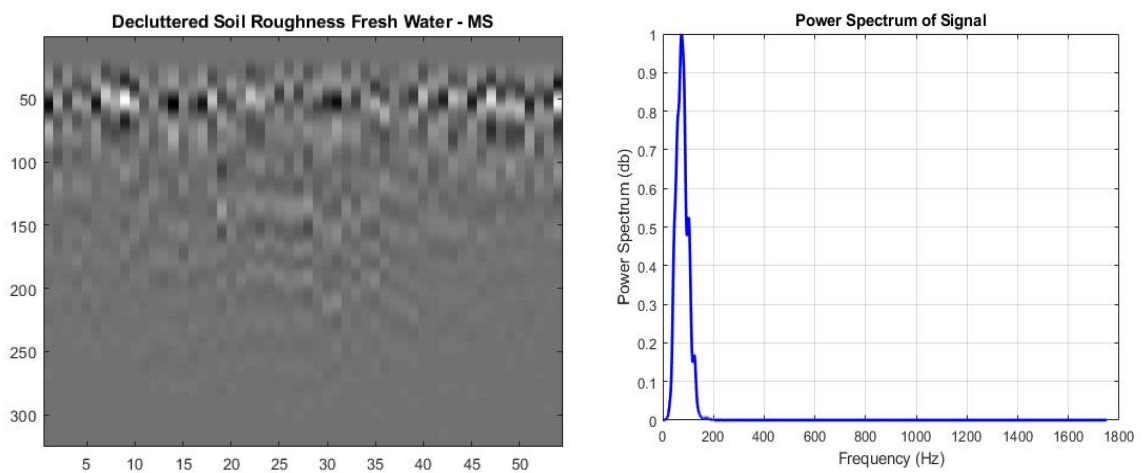

Figure 5.10: Compares Between Soil Only Image and the Soil Roughness with Water.

Here we calculated the Mean Subtraction to denoise the image, and displaying

its power spectrum using the FFT function, and it has a PSNR value 58.9175

Figure 5.11: Shows MS Method Effect on Soil Roughness Data and its Power Spectrum.

After calculating MS, we move on to calculate SVD and its PSNR value of this scenario, these figures below show the SVD image and its power spectrum using the FFT function, while its PSNR=69.3546.
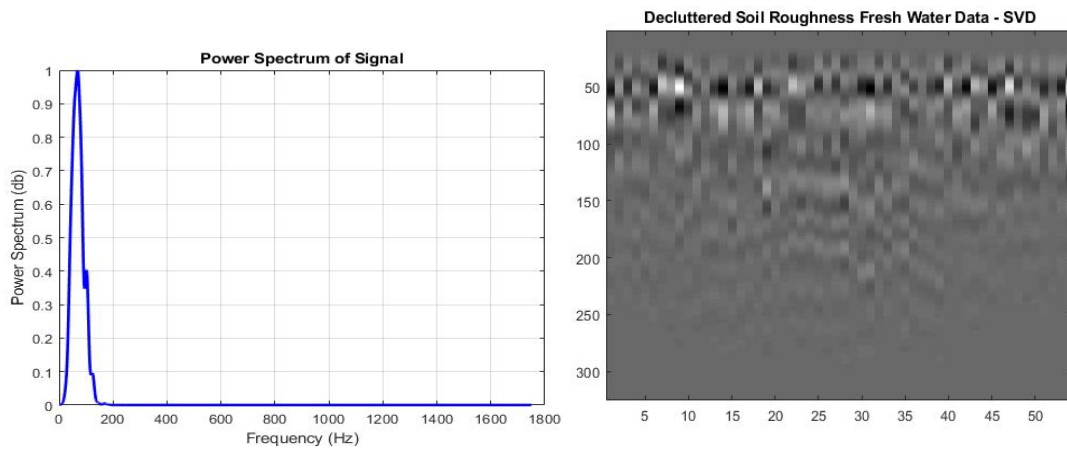
Figure 5.12: Shows SVD Method Effect on the Soil Roughness Data and its Power Spectrum.

Within the calculation of the SVD we calculated the covariance matrix and the eigenvectors and eigenvalues for the PCA method in the same way as before, its PSNR Value 70.5059
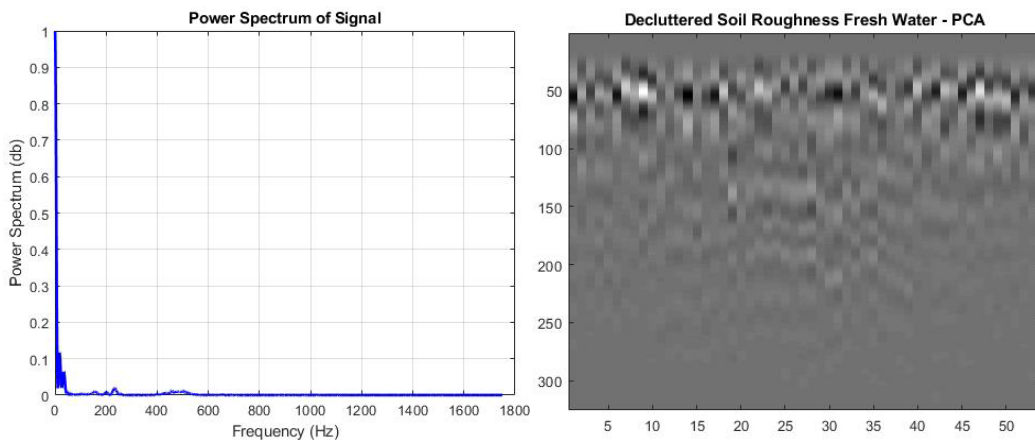

Figure 5.13: Shows the PCA Method of Effect on the Soil Roughness Data and its Power Spectrum.

Same way we calculate the ICA and the PICA algorithms to obtain our results and calculate their PSNR values, ICA 66.6977 and 70.4988 for the PICA.
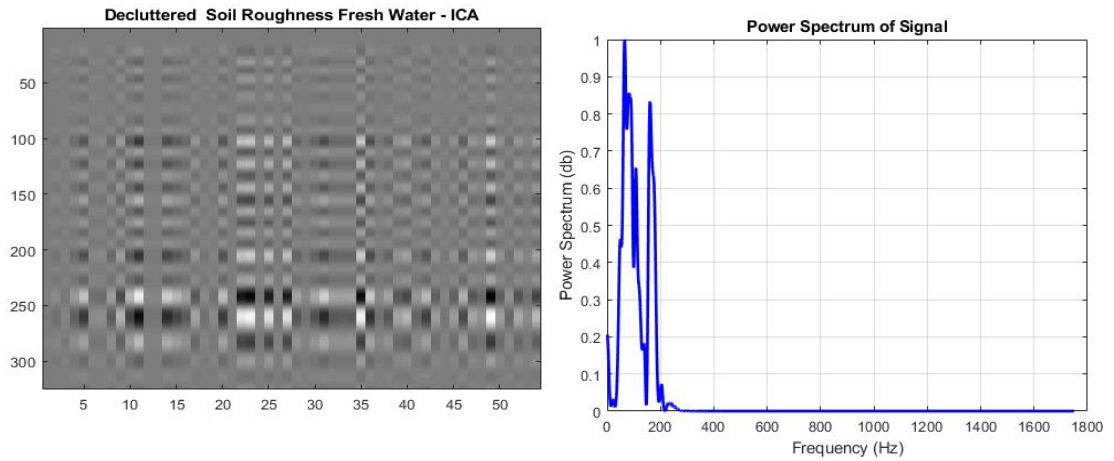
Figure 5.14: Shows the ICA Method Effect on the Soil Roughness Data and its Power Spectrum**.**
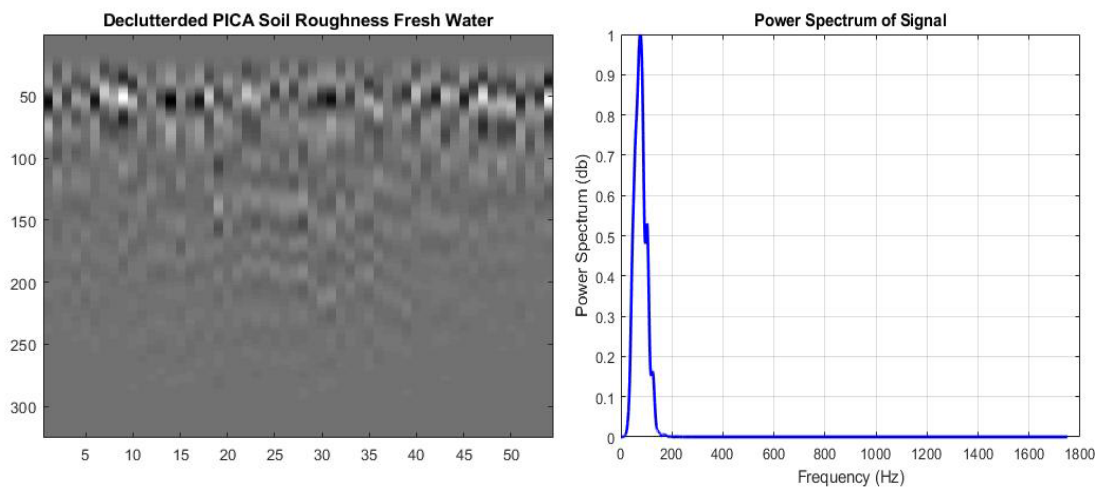

Figure 5.15: Shows the PICA Method Effect on the Soil Roughness Data and its Power Spectrum.

As we used the same filters on this data set also, using Gaussian, histogram, median, and wiener filter in same order showing in the following figure.
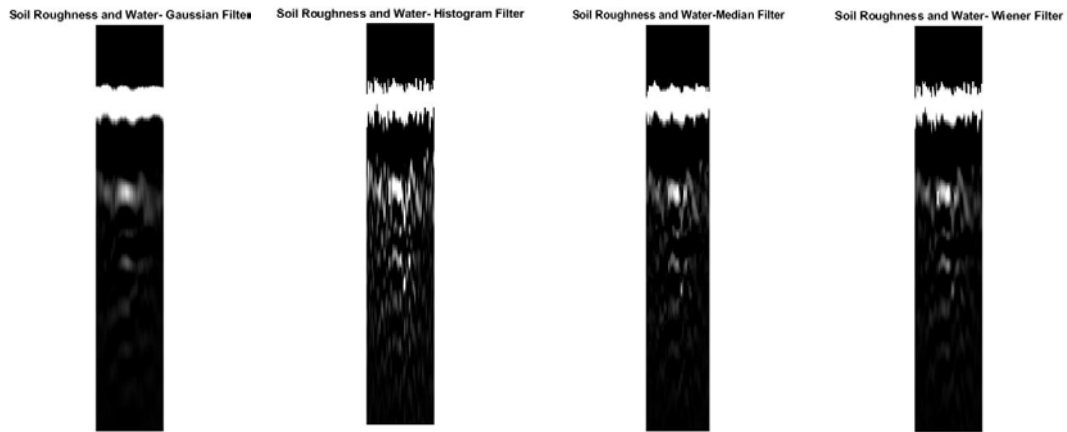
Figure 5.16: Shows the Filters (Gaussian, Histogram, Median, and Wiener) Effect on the Soil Roughness Data in Order.

Having these PSNR values: Gaussian with 48.7414, histogram 53.4258, Median 48.4714, and finally the wiener filter with the PSNR value 48.2815.

5.3 Detecting Fresh Water in the Same Soil with Some Grass

As the previous scenarios we discussed this one has the same box dimensions for the filled soil box (480*148*170) mm, with same sphere (10mm radius) was placed in the (240*74*100) mm location inside the box, and fresh water was injected in that sphere.

Adding 300 blades of grass on the top of the box with various height between 30 to 50 mm. With our antenna scanning the surface every 5mm, the antenna is placed on top of the box with the height 5mm.
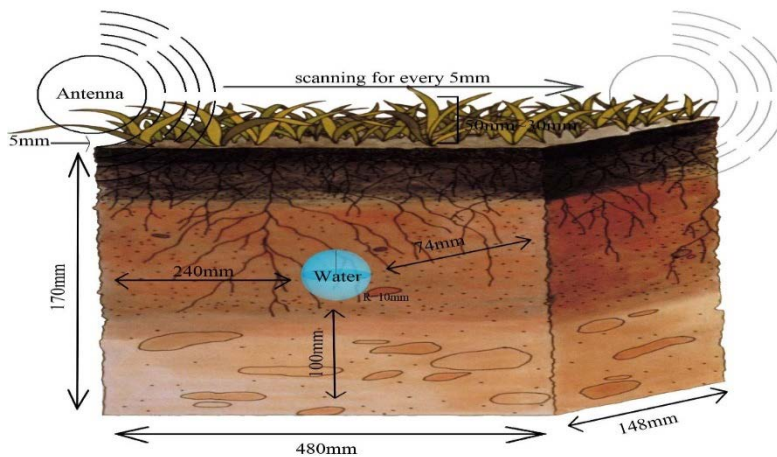

Figure 5.17: Shows the Box after Adding the Grass on Top.

60

With the same GSSI 1.5 GHz 'like' Antenna providing the same time window: (6e-9). And with the use of soil Peplinski, our soil is with sand fraction 0.5, and clay fraction up to 0.5, bulk density 2 g/cm^3 , with sand particle density 2.66 g/cm^3 and volumetric water fraction range of 0.001-0.25.

The fresh water has the same characteristics (relative permittivity: 80.0, conductivity: 0.5, relative permeability:1.0, with the magnetic loss:0.0).

For this data first, we also showed the difference in image collected if the area of search has water in it or not
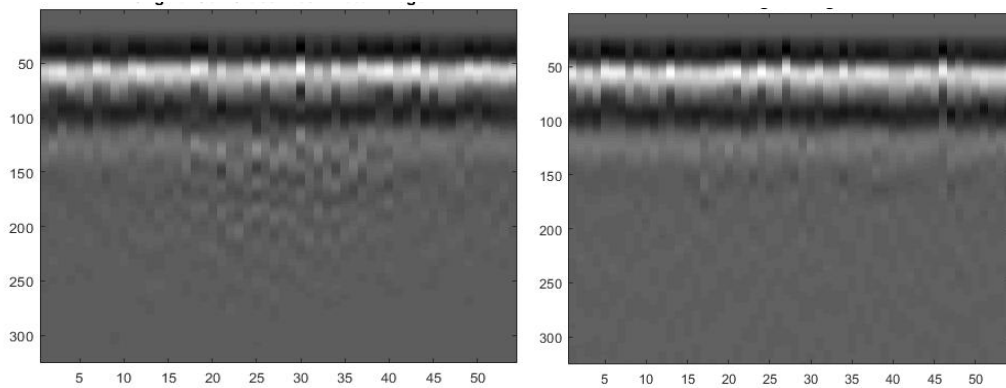

Figure 5.18: Compares Between Soil Only Image Collected and the Soil Grass with Fresh Water in it in order.

The same process is applied here, we are going to calculate the mean subtraction method and its PSNR value and display our processed image with its power spectrum. Also, we have computed the Mean Subtraction to remove the noise from our image with PSNR = 57.3919.
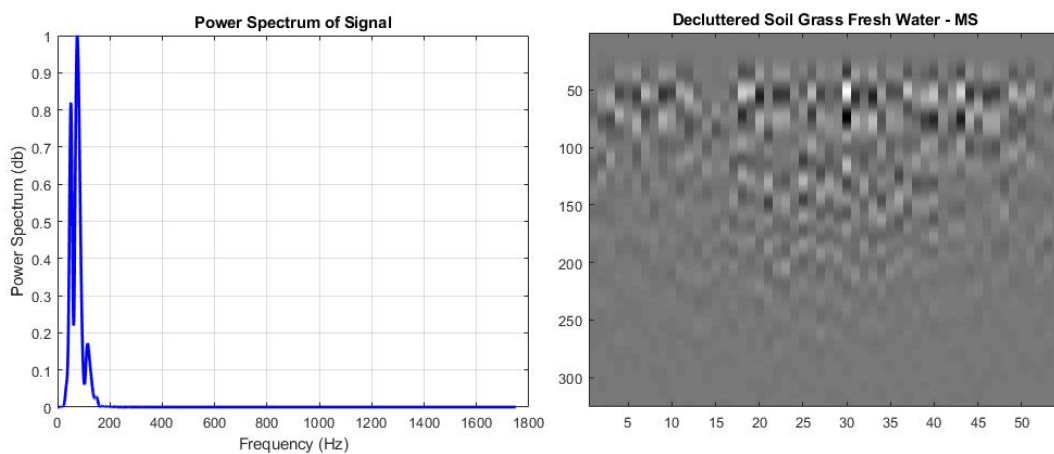

Figure 5.19: Shows the MS Method Effect on the Grass Data and its Power Spectrum.

And for the SVD method we also calculated the matrices and the eigenvectors and eigenvalues. The results obtained are represented below along with its FFT analysis, while it has a PSNR value of 62.7766.
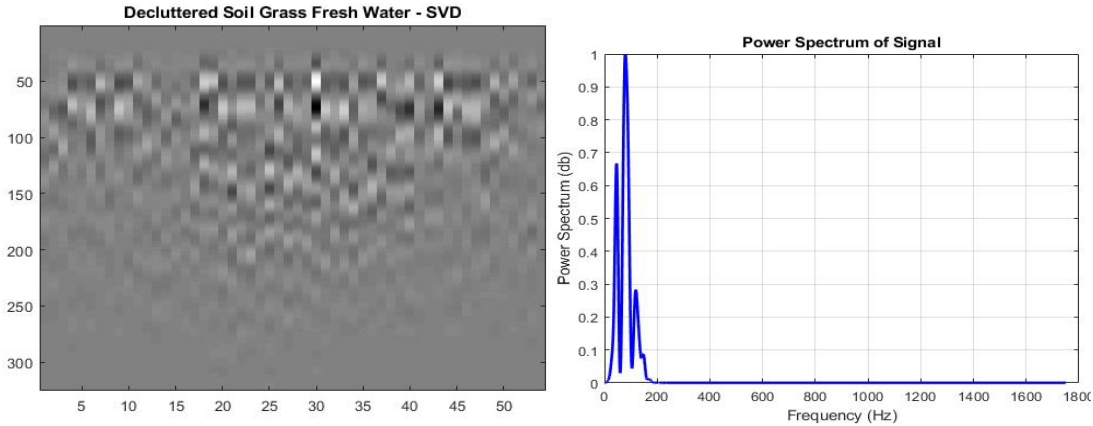

Figure 5.20: Shows the SVD Method Effect on the Grass Data and its Power Spectrum.

Meanwhile it is known that the SVD are main part of the other methods so it makes the calculation of the PCA easier, we calculated the covariance matrix and the eigenvectors and eigenvalues, ordering λ in a decreasing order, for this we selected the feature number to be 20 to estimate the matrix, below we will present the PCA effect on the data and its FFT analysis. With PSNR value 56.3077.
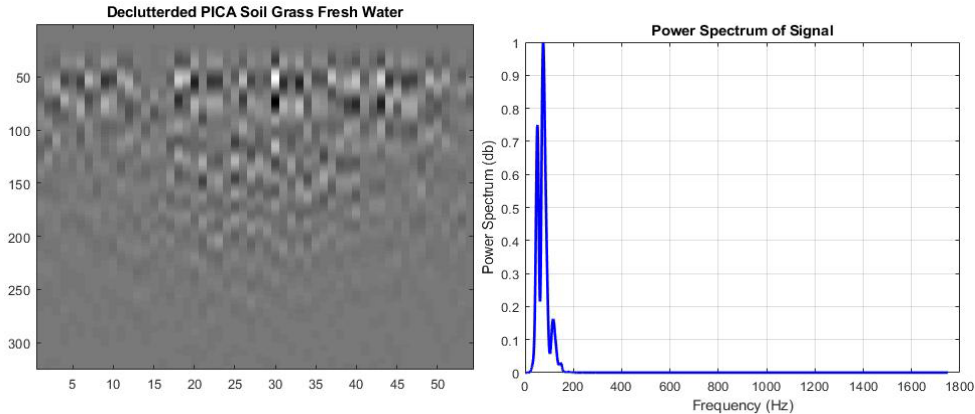

Figure 5.21: Shows the PCA Method Effect on the Grass data and its Power Spectrum.

As for the ICA Method all the parameters are calculated like before, with obtaining these independent images to reconstruct our ICA image the resulting effect is shown below along with the FFT analysis. The PSNR value =61.4543.
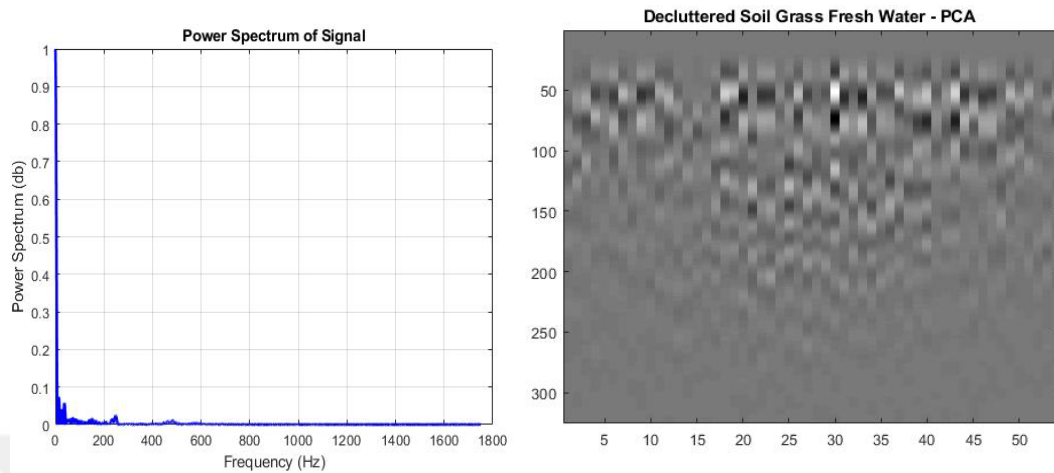
Figure 5.22: Shows the ICA Method Effect on the Soil Grass Data and its FFT Power Spectrum.

And the PICA method effect on the image is shown in the figure (5.23) with its FFT analysis, while its PSNR value =64.1537.
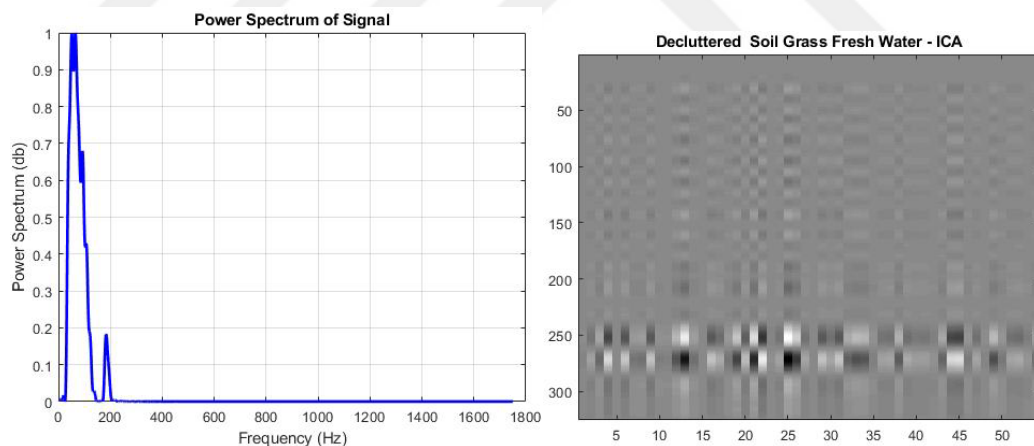
Figure 5.23: Shows the PICA Method Effect on the Grass Data and its FFT Power Spectrum.

The final method we applied is the filters and how it effects on our image results. These filters are applied to enhance our image for better detection or to be accurate for better visualization. Gaussian, Histogram, Median, and Wiener filters are applied and their PSNR results are like that: 48.3540, 53.0424, 48.2273, and 48.1700 sequentially.

Figure 5.24: Shows the Filters Effect (Gaussian, Histogram, Median, and Wiener Filter) on Grass Data in order.

6. CONCLUSION

In this thesis, we have discussed the signal processing algorithms for detecting underground water using the GPR technology. Briefly the GPR history, system design, the usage of GPR system, and its electromagnetic wave propa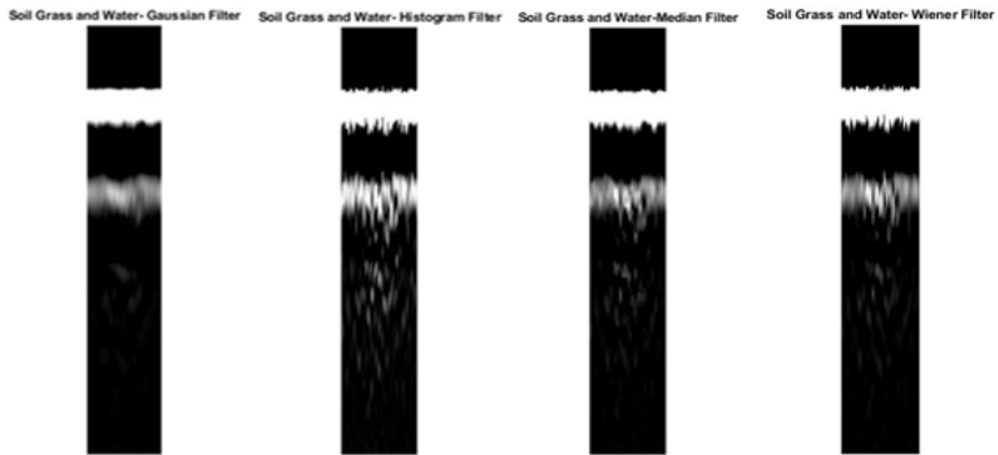gations were introduced. The collection method of SAR system and its applications with the use of GPRMax as a simulation software was explained. Both signal processing algorithms and image enhancement filters that we used in this research were fully discussed.

Water is very important element in this life, a lot of water is being wasted due to leaks or broken metal or PVC pipes or its just hidden underground resources. So, the GPR is a reliable technology that made it easier for us to find these sources, it can be used for soils with different electrical of properties to detect underground water. As mentioned, the GPR responds to the electric property changes, which can be caused by rocks or some roughness in the soil.

The signal processing is required to analyze the collected signals, here we used different methods to analyze and find water with the best method, to use in different soil situation some methods like MS, SVD, PCA, ICA and PICA algorithms were implemented to our collected data. For our first data the PICA proved to be better to detect water with in the soil while comparing all the methods PSNR values. While with the previously explained filtering method such as Gaussian, Histogram, Median, and Wiener, these obtained ratios were little bit less and as explained before filters are not preferred to be used alone but as a step to enhance the resulting image before we put it within one of the processing algorithms even though, the histogram equalization has shown to better than the rest of the filters applied on soil and fresh water data.

As for the second data for water detection with the data referred to as soil roughness with same methods applied, we found that the best algorithm is PCA while PICA showed really near results which is really good also depending on

the PSNR values that we calculated. In this data set the filters still weak compared to the algorithms nut histogram filtering results are still better than the rest.

And last data set we provided that for water detection is after adding the roughness to the soil we added grass on the top as shown before, this major change in the data we made to approximately simulate the reality that not all places for find water is as good as looking for it in a dry sand. The algorithms have been applied on this last data said we called Grass, the best algorithm in this situation to use is the PICA, this method that has proved itself the best in almost every situation so far. While in the filtering methods the histogram also is better for this Grass data set that we collected from the final experiment, and in all situations that we discussed earlier.

For future work, more complex environments can be built and more clutter effect in processing algorithm methods can be analyzed in detail, the application of more methods such as wavelet transformers, probabilistic principle analysis (PPCA) and support vector mechanism (SVM) methods, neural network, and automatic feature recognition among other different processing methods to consider

REFERENCES

[1] **Wikipedia,** Ground Penetrating Radar Introduction.

[2] **Daniels, D.J. (Ed.). (2004),** Ground Penetrating Radar (2nd Edition), IEE, ISBN 0-86341-360-9, London

[3] **Erica Carrick Utsi**, Ground Penetrating Radar: Theory and Practice.

[4] **Jeff Ambrose Mitchell,** Introduction to Locating Buried Utilities.

[5] **Curlander, J.C., McDonough, R.N,** Synthetic Aperture Radar: Systems and Signal Processing. Wiley, 1991.

[6] **Annan, A.P. (2009).,** Electromagnetic principles of ground penetrating radar, In: Ground Penetrating Radar: Theory and Applications, Jol, H.M., pp. 3-40, Elsevier, ISBN 978-0-444-53348-7, Amsterdam, The Netherlands

[7] **Cassidy, N.J. (2009),** Electrical and magnetic properties of rocks, soils, and fluids, In: Ground Penetrating Radar: Theory and Applications, Jol, H

[8] **H.M.Jol,** Ground Penetrating Radar: theory and applications,Elsevier,2009

[9] **Van der Merwe, J.Gupta,**"A Novel Signal Processing Technique for Clutter Reduction in GPR Measurements of Small, Shallow Land Mines", IEEE Transactions on Geoscience and Remote Sensing, Vol. 38, No. 6, November 2000.

[10] **SHLENS, J.,** A Tutorial on Principal Component Analysis. 2th ver. Institute for Nonlinear Science. University of California, 2005.

[11] **Lindsay I Smith**, A tutorial on Principal Components Analysis February 26, 2002.

[12] **Steven M. Holland, Univ. of Georgia,** Principal Components Analysis.

[13] **skymind.ai,** Eigenvectors, Eigenvalues, PCA, Covariance and Entropy.

[14] **J. P. Hoffbeck and D. A. Landgrebe,** "Covariance matrix estimation and classification with limited training data," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 7, pp. 763-767, 1996.

[15] **Gilbert Strang,** "Introduction to Linear Algebra, Fifth Edition'.

[16] **David C. Lay, Steven R. Lay, Judi J. McDonald,** Linear Algebra and Its Applications.

[17] **Aswini, N. Dyana, A. and Srinivas Rao, C. (2014),** Detection and classification of ground penetrating radar image using textrual features, International Conference on Advances in Electronics, Computers and Communications (ICAECC), Bangalore.

[18] **PO, Botstein D (2000)**, Singular value decomposition for genome-wide expression data processing and modeling. Proc Natl Acad Sci U S A, 97, 10101-6.

[19] http://saraero.com/sar-applications/.

[20] https://buanapt.wordpress.com/support/gpr-applications/.

[21] **http**://www.u-survey.com/blog/why-cant-ground-penetrating-radar-find-everything.

[22] https://www.engineersgarage.com/articles/image-processing-tutorial-applications.

[23] **Larissa Natsumi Tamura, Renato Paes de Almeida,,**Ground Penetrating Radar investigation of depositional architecture: the São Sebastião and Marizal formations in the Cretaceous Tucano Basin (Northeastern Brazil)**.**

[24] **Caner Özdemir, Şevket Demirci, Enes Yiğit, and Betül Yilmaz**, A Review on Migration Methods in B-Scan Ground Penetrating Radar Imaging.

[25] **L. Dojack**, Ground Penetrating Radar theory, data collection, processing and interpretation: A guide for archaeologist.

[26] **Conyers, L.B. and Goodman, D. (1997)**, Ground Penetrating Radar-An Introduction for Archaeologists. Altamira Press, London.

[27] **Aapo** Hyvärinen **Juha Karhunen Erkki Oja,** Independent component analysis (2001).

[28] http://shulgadim.blogspot.com/2014/02/independent-component-analysis-ica.html.

[29] **Christian F. Beckmann and Stephen M. Smith**, Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging.

[30] **M.W. Woolrich, B.D. Ripley, J.M. Brady, and S.M. Smith,** Temporal autocorrelation in univariate linear modelling of FMRI data. Neuro Image, 14(6):1370–1386, 2001.

[31] **Jonathan Fraser,** Design and Simulation of a Coded Sequence Ground Penetrating Radar, The University of British Columbia, 2009.

[32] **Knight, R., and A. L. Endres (2005),** An introduction to rock physics for near-surface applications, in Near-Surface Geophysics,vol. 1, Concepts and Fundamentals, edited by D. Butler, pp. 31–70, Soc. Of Explor. Geophys., Tulsa, Okla.

[33] **file:///E:/MachineVision_Chapter4.pdf.**

[34] **Tinku A. , Ajoy K. Ray ,** Image Processing principle and applications.

[35] **Schowengerdt, R.A., 1980**. Reconstruction of multi-spatial, multi-spectral image data using spatial frequency content. Photogrammetric Engineering & Remote Sensing, 45 (10).

[36] https://nptel.ac.in/courses/105104100/33.

[37] http://fourier.eng.hmc.edu/e161/lectures/smooth_sharpen/node2.html.

[38] https://en.wikipedia.org/wiki/Median_filter.

[39] **James D. Broesch**, in Digital Signal Processing, 2009.

[40] https://en.wikipedia.org/wiki/Wiener _filter.

[41] http://www.owlnet.rice.edu/~elec539/Projects99/BACH/proj2/wiener.html.

[42] https://en.wikipedia.org/wiki/Fast_Fourier_transform.

[43] **Raghu Muthyalampalli**, Implementation of Fast Fourier Transform for Image Processing in DirectX* 10, published on September 1, 2011.

[44] https://www.tutorialspoint.com/dip/histograms_introduction.htm.

[45] **Arthur COSTE,** Histograms University of Utah: CS6640 Image Processing.

[46]   https://en.wikipedia.org/wiki/Histogram_equalization.

[47]   **Bill Comstock**, Using PSNR thresholds to modulate the degree of lossy compression in JPEG2000 files, May 2011.

RESUME

Name : Sarah Ayad TABANNA

Place and Date of Birth : Baghdad, Iraq, 05-09-1994.

E-Mail: Sarahtabanna@yahoo.com.

Education :

- **Bachelor :** 2015-2016, AL-Mansour University College, Telecommunication Engineering .

- **Master   :** 2019, Istanbul Aydin University, Electric and Electronic Engineering.

-

 Certificates:

- IT Essentials

- CCNA

- Website Design using DREAMWAVER