ADVANCED FEATURE BASED TECHNIQUES FOR LANDMINE DETECTION USING GROUND PENETRATING RADAR

A Thesis Presented to the Faculty of Graduate School

University of Missouri – Columbia

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science

---

by

ZHENHUA MA

Dr. Dominic K.C. Ho, Thesis Supervisor

DECEMBER 2007

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

ADVANCED FEATURE BASED TECHNIQUES FOR LANDMINE DETECTION USING GROUND PENETRATING RADAR

presented by Zhenhua Ma

a candidate for the degree of Master of Science

and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Dominic K.C. Ho

Dr. Justin Legarsky

Dr. Michael Jurczyk

I would like to express my sincere thankfulness to my father, Gengli Ma, my mother, Manna Ling, and all my other family members. Without your endless love, support and encouragement, I would not be able to accomplish my goal. Thank you for being always there for me!

I also want to say thank you to all my friends, especially Le Yang. Thanks for the help from all of you. I really enjoy the time being with all of you!

ACKNOWLEDGEMENTS

I would like to take this opportunity to gratefully acknowledge everyone who made this research possible. I want to express my sincere appreciation to my advisor Dr. K.C. Ho for his continued patience, support, inspiration and valuable guidance. Without his help and guidance, I would not be able to complete my research.

I would also like to express my sincere gratitude to my committee members Dr. Justin Legarsky and Dr. Michael Jurczyk for spending their valuable time and giving helpful suggestion to improve my research.

I thank University of Missouri – Columbia, for giving me the opportunity to pursue my Master's degree and obtain high quality education and wonderful experience that added a lot in my personality and made me a better person to face my future life.

ADVANCED FEATURE BASED TECHNIQUES FOR LANDMINE DETECTION USING GROUND PENETRATING RADAR

ZHENHUA MA

Dr. K.C. HO, Thesis Supervisor

ABSTRACT

Landmine detection is an important and yet challenging problem that remains to be solved. It is not only a problem for military, but also for humanitarian concern. The goal of this research is to propose some techniques for landmine detection. Two advanced feature based techniques are developed. One algorithm applies the clustering method based on the spectral feature vectors formed by the energy density spectra of return sensor signals, the idea behind is to find out whether there are some "hidden patterns" among the spectral feature vectors. The other one is the subspace detector technique that utilizes the energy density spectra of return signals directly. These techniques are tested in various testing data sets collected from the vehicle mounted ground penetrating radar to evaluate their ability to improve the detection result and reduce the false alarm rates. Both of them are proved to be useful in improving the detection of landmines.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

Chapter 1

Introduction

1.1 Forward

The detection of land mines has been a very important and popular research area in recent years. Nowadays the unstable world also highlights the importance of land mine detection technique. Widely published reports estimate that there are 110 million land mines in 70 countries [1]. The land mine detection is a problem of significant military, economic, and humanitarian concern. Not only do land mines pose a deterrent to military activity during conflict, they also remain lethal long after that conflict is over [2]. So the research is driven not only by the need in military operations, but also for humanitarian purposes to clear up mine fields left after wars that claim more than 30,000 deaths and injuries every year [3]. The statistics are staggering: 80 percent of the victims of antipersonnel mines are civilians, many of whom are children [4]. Moreover, the number of anti-personnel land mines is increasing at the rate of one half to one million mines per year [2]. In fact, current antipersonnel land mines cost between $3 - $25 per mine and clearance methods cost between $300 - $1000 per mine removed [5]. So the motivation of developing new counter-mine systems which will operate with a high degree of accuracy while being simple enough to deploy at low cost is great.

The detection of land mines is not an easy but difficult and challengeable work. Land mines can be buried in almost every type of terrain and environmental condition, so to locate a them accurately is very hard. The detection has to rely on the signal responses from land mines captured by the sensors. However, very little technology is currently employed in the real world for the detection of land mines [1]. Many types of detection sensors are developed or currently being developed as well as many signal processing techniques which aims to improve the detection accuracy in both academic and industrial fields. Among them, ground penetrating radar (GPR) has been gaining momentum recently as a sensor for land mine detection [3]. It is very promising in detecting the low-metal and plastic land mines but in the meanwhile some signal processing techniques are needed to improve the detection result. The research involved in this thesis will introduce a signal-processing technique for ground penetrating radar and investigates two advanced feature based detection techniques using the signal collected by the ground penetrating radar.

1.2 Problems and Techniques for Land Mine Detection

Land mines are different from unexploded ordnance (UXO). UXO are bombs that were drop from planes, or fired rocket launchers, whose fuses did not detonate when they hit the ground. They are not designed to blow up when stepped on or driven over [1]. For land mines, they are present intentionally and are designed to kill or maim the people step on them or demolish the vehicle drive over them, so it makes the detection of land mines much more dangerous than the detection of UXO. Also, metal detector will be very efficient to detect the UXO, but nowadays the newer land mines contain very little metal so to detect them with a metal detector becomes not feasible.

Land mines are usually embedded in metal, plastic or even wooden case and they can be in different shapes and sizes. Figure 1.1 shows four kinds of anti-tank (AT) and anti-personnel (AP) land mines.

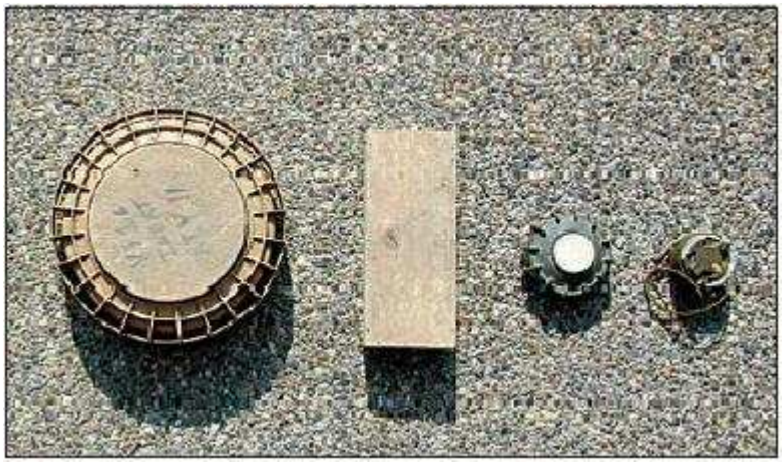

Figure 1.1: Examples of land mines. Left to right, VS1.6 plastic AT mine, PMD6 wood AP mine, VS50 plastic AP mine, and M14 plastic AP mine [1].

These land mines can be placed just on the surface of earth or buried underground at different depth, generally shallow for anti-personnel land mines and can be up to 16 inches for anti-tank land mines. Also the climate and geographic condition of the fields on where land mines are placed could be very

complicated. All these factors can limit the efficiency of the detection and are the challenges that have to be conquered.

So an efficient land mine detection system should have the capability to detect varying kinds of land mines no matter what size, shape, burial depth, casing type and explosives signature they might have, in a good standoff distance for the safety consideration. The system should be able to distinguish the land mines from the background clutter and provide a highly precise detection result with very low false alarm rate, acceptable operational speed and imaging capability in a reasonable cost. It is really a challenge to meet all these demanding requirements. However, so far many technologies and detection sensors have been developed to solve these problems. Forward-Looking Infrared (FLIR), Thermal Neutron Activation (TNA), Quadruple Resonance (QR), Visible Wavelength Cameras, Acoustics detection system, Ground Penetrating Radar (GPR) [1], laser detection and penetrating radiation are some of the new land mine detection technologies used in various applications, some traditional sensors such as metal detector can be very useful for UXO and metal-cased land mine detection. However, the resulting false alarm rate, particularly in metal-cluttered areas such as war zones, can be extremely high, also it has difficulty in urban environments where there is a lot of background metal [1]. Moreover, unfortunately not all land mines are metal-cased, so other detection techniques are needed.

Forward-looking infrared (FLIR) systems attempt to take advantage of the circumstance where the soil directly above a land mine heats and cools at a slightly different rate than surrounding soil [6]. And as the name implies, instead of detecting the land mines directly beneath the system, it is able to detect the land mines which are ahead of the sensor. The FLIR can form

an image of the land mines, or the soil on top of the them, so it can cover a large area in a short time. When the purpose is to identify minefield, this technology is the best suit one. But the disadvantage of this system is that it cannot be applied when the soil and mine are in thermal equilibrium condition.

Thermal neutron activation (TNA) and quadruple resonance (QR) sensors can directly detect the explosive signature of land mins. TNA works on the principle of the detection of neutron-activated gamma-rays from nitrogen, which is a constituent of all explosives. Nitrogen, following thermal neutron capture, emits a characteristic gamma-ray which is a higher energy than almost any other gamma-ray likely to be encountered in the field. QR utilizes a magnetic field pulse at a radio frequency specifically targeted to explosive compounds (e.g., RDX and TNT) [7]. So far, TNA and QR are effectively used as confirmation sensors to confirms or denies the presence of a land mine when other sensor such as GPR indicates there may be one in some location.

Visible wavelength cameras and acoustics detection systems are another two land mine detection technologies. The autonomous system in the visible wavelength cameras can detect the anti-tank land mines which are placed just on the surface of earth, when the human eyeball can barely see them. Or even when the mines are buried, the fresh dig and weathered dig location of the mines can be noticed by the camera system. Acoustic detection methods typically employ loudspeakers to vibrate the ground and laser Doppler vibrometers to scan an area and detect the acoustically induced motions [8]. Detection rests on the principle that motion of the land mine is different from that of surrounding soil because the land mine contains rigid structure that resonates in a particular way [1]. It is a more recently out-of-gate technique than GPR or metal detector.

Laser detection utilizes the difference in the reflectance and polarization of soil when disturbed by laser energy. This requires however a large laser power and a complex data interpretation process [9], which makes this system not easy to implement. Penetrating radiation (neutrons and photons) offers another probe for standoff land mine detectors. It does not rely on radiation transmission since it requires access to two opposing sides of the object, a situation not attainable with land mines. Therefore, one has to rely on radiation scattering or activation. Photons, in the form of x- or gamma-rays, incoherently collide with atomic electrons with a probability that is dependant on the electron density, and consequently the mass density, of the medium. As the scattered photons travel back towards the detector, they are removed by further scattering or absorption. With the photo-absorption probability being strongly dependant on the atomic number. The difference between the atomic number and density of mine and soil allows the identification of the former. This is the essence of the x-ray backscattering system of Campbell and Jacobs [10].

Ground-penetrating radar (GPR) is a geophysical method that uses radar pulses to image the subsurface. This non-destructive method uses electromagnetic radiation in the microwave band (UHF/VHF frequencies) of the radio spectrum, and detects the reflected signals from subsurface structures. It has been actively applied to variety of geophysical applications for nearly 70 years and is recently gaining momentum as a sensor for land mine detection. When parameters such as frequency range, antenna size, antenna separation, and system timing are optimized for detection of mine-sized objects in the near subsurface, GPR is quite effective in detecting both metal and plastic land mines in a variety of soils [1]. The radar signals are reflected by denser, buried objects. Due to its higher density, a mine has a different radar signature than the soil that surrounds it. Also, ground penetrating radar can detect the

presence of subsurface targets by distinguish them as land mines and clutter objects based on the level of symmetry exhibited by the targets. GPR is able to identify the dielectric discontinuities below the surface and any discontinuities appear as a signal, instead of detecting the metallic content of an object, so it is capable detect the low-metal cased or even plastic and wooden cased land mines. Some signal processing techniques are typically applied to calculate and subtract background in order to enhance signal. However, GPR's ability to detect non-metallic targets can be offset by the false alarms caused by some materials such as roots of tree and rocks. Besides, GPR is technologically complex, and the weight and power requirement of the system make GPR most easily deployed on a vehicular platform [1], i.e., the vehicle mounted GPR system.

There are certain challenges pertaining to the GPR system, the major one is the visualization of the signal data. Also, the data collected by the radar has to be presented in a comprehensible format in order to attain a comprehensive conclusion. This data may be a color-mapped image of the energy of a block of data, or a time domain representation of the measured signal. Image processing sometimes becomes equally important to the signal detection. Besides, pre-processing the data is also an important factor in order to achieve better detection of the land mines. This can be done in both the time and frequency domains.

1.3 Previous Works on Land Mine Detection

Since land mines are buried deeply under the ground, land mine detection has to rely mostly on sensors that can capture some types of signal responses from land mines. The initial standard land mine detector is metal detector (MD), also known as Electro Magnetic Induction (EMI) sensors. EMI sensors can easily detect metal land mines. Unfortunately, newly produced land mines nowadays contain very low metal content and can be plastic or even wooden cased. As a result, metal detectors will have problems in detecting this kind of land mines. Ground penetrating radar (GPR) is the emerging technique used in mine detection that can detect plastic or low metal content land mines [11]. GPR system can be operated in both the time domain and the frequency domain. In this system, the lower frequency components can penetrate the ground further, while it is the higher-frequency components that are necessary to image and resolve smaller targets below the ground. GPR systems are usually of two types: handheld system and vehicle mounted system. Extensive research has been done for these two GPR systems.

Wide varieties of signal processing techniques have been used in the past to process GPR received signals. Some works done Gader et al. [12]-[13] has suggested a gradient-based method for land mine detection. This method efficiently used the Fuzzy clustering technique by generating multiple prototypes of training data from a fuzzy clustering of gradient features. In [14]-[17], Hidden Markov Models (HMM) have been used to process GPR received signals. Though the technique is still under investigation, it has been proved that HMM can be used to efficiently detect the presence and location of land mines. In [18]-[19], a least-squares method has been proposed to remove the ground bounce obtained on a rough surface. However, its performance depends heavily on the reference ground bounce and no method has been given

in [18]-[19]. A system identification based clutter removal algorithm is given in [20]-[21], which uses ARMA model to describe clutter and abrupt change detection technique to classify clutter and target signal. The technique focuses on pre-processing the GPR data to reduce the influence of near-surface clutter. Once the clutter is satisfactorily known, this technique easily identifies the target as a small anomaly within the known clutter background. However, the dependency on the reference clutter data is a major drawback of this technique. A linear prediction approach for the detection of land mines has been proposed by Dr. Ho, which uses a linear prediction model to generate a mine detection alarm if the current and past few signal samples do not fit the linear prediction model. Apart from the above methods, wavelet analysis, principal Component analysis (PCA), Independent component analysis (ICA) have been proposed for clutter suppression in [22]-[23]. Wavelet decomposition and reconstruction will lose some useful target signal component because the ground bounce is usually very strong and would overlap the target signal in the time-frequency plane. Deming [24] proposed Maximum Likelihood Adaptive Neural System (MLANS) for land mine detection, which is a model based neural network that combines the adaptability of a neural network with the priori knowledge of signal models. The MLANS technique is designed to adapt to unknown and changing soil conditions, while incorporating signal models based on the physics of electromagnetic scattering for a specific object or clutter type. However, for accurate feature extraction, the technique requires well defined clutter and mine signature models. The handheld GPR system has time varying clutter and does not provide clear signature models. Hence, MLANS technique would be more appropriate for vehicle-based systems.

1.4 Thesis Contribution

In the research of this thesis, we propose and investigate two techniques trying to improve the detection result and reduce the false alarm rates for the land mine detection. We realize that the GPR signal spectrum for certain types of land mine sometimes can reveal important features of these mines even when the signals themselves are quite weak, so both techniques we investigate in this thesis are based on the spectral features.

The first technique is the clustering method based on the spectral feature vectors which is formed by the energy density spectrum of the GPR signals. The purpose of this technique is to find out whether there are some "hidden patterns" among the GPR signals from the training data field. If some patterns are found through clustering, then we want to see whether these "hidden patterns" can be used to improve the detection performance. After the energy density spectrum of the GPR signals are obtained, they will be used as the elements of the spectral feature vectors. Then we will study and define the proximity measure of these vectors. Proximity here is defined as the "distance" between one spectral feature vector and other vectors. Later according to the proximity measure, some studies will be done trying to decide which clustering method we are going to implement to these spectral feature vectors. As a result, the K-Means algorithm, which is one of the partitioning algorithms, is implemented to cluster these vectors. When the clustering process is done, the quality of the clusters will be studied, then according to the experiment results we will decide how many clusters the whole data set should be partitioned, and the K-Means algorithm will be implement to cluster the vectors to obtain the final clusters. Eventually the centroids of the clusters will be taken out and combined with the spectral feature vectors to verify the effect of the clustering method in the testing field.

The subspace detector is the second technique we have investigated. Not like the clustering method based on the spectral feature vectors formed by the energy density spectrum of the GPR signals, the subspace detector technique is based on the energy density spectrum directly. In this technique two hypotheses for detection are set to decide whether the received GPR signal contains the response signal from target plus the noise or just has the noise. The energy density spectra of a GPR signal can be considered to lie in the subspace spanned by some basis vectors $\mathbf{s}_i$ and some coefficients a_i. According to maximum likelihood estimation, the value of the coefficient vector can be estimated using the subspace basis matrix formed by the basis vectors and the energy density spectra. Then the two hypotheses will be related to the value of the estimated coefficient vector: if its value bigger than certain threshold value, then the received GPR signal contains the response signal of target. In order to obtain the value of the coefficient vector, we have to obtain the subspace basis matrix, so later independent component analysis technique and eigenspace separation transform technique are investigated. Finally eigenspace separation transform and eigenvalue decomposition are used to obtain the value of the subspace basis matrix. Also the effects of varying the total number of basis vectors used and the length of energy density spectrum used in the subspace detector on the detection performance will be evaluated. Eventually the subspace detector technique is implemented on the testing field to verify its ability to improve the detection result and reduce the false alarm rates.

1.5 Thesis Organization

This thesis consists of six chapters. The introduction is covered in this chapter. Chapter 2 will covers the background of this land mine detection project, some basis knowledge of the GPR system as well as the description of the mathematics used in this thesis will be introduced, and the training and testing data set will be describe. In chapter 3, we will introduce the vehicle mounted GPR system and discuss how to use the frequency domain features from the GPR in the land mine detection. Chapter 4 will be the investigation of the clustering method based on the spectral feature vectors. In chapter 5 we will study the subspace detector based on the energy density spectrum of the GPR signals, and also evaluate the effects of varying the total number of basis vectors used and the length of energy density spectrum used in the subspace detector on improving the detection result. Finally in chapter 6, conclusions and possible future works will be included.

Chapter 2

Background

2.1 Basic Knowledge of Land Mine Detection

Metal Detector, also known as Electromagnetic Induction (EMI) sensor, is the most widely used sensor in land mine detection. Since in metal case land mines, the metal content is typically ferrous, so this kind of land mines can be easily detected by EMI sensors. In EMI sensors, there is usually a pair of coils, one of them will be used to transmit an electromagnetic pulse and the other one will be used to receive and detect any current induced by the pulse in subsurface metallic objects. EMI sensors can be of two types, Time domain or Frequency domain. The response from a time domain sensor can be modeled mathematically as a sum of decaying exponentials, and numerical methods of frequency domain EMI sensors have been developed. The major disadvantage with the EMI sensors is that they cannot detect low metal or plastic-encased land mines, and the false alarm rates in urban environments where lots of background metal exist will be very high. Nonetheless, it remains as the front-line tool in the detection of land mines [1].

GPR sensor has been used for land mine detection in recent years since it has the advantage of having the capability to detect plastic case mines [16]-[18]. GPR sensor is typically moved forward along a test track by hand or mounted

on a vehicle. GPR sensor operates by transmitting high frequency pulse down into the ground through an antenna, the transmitted pulse enters the ground, contacts objects or subsurface with different electrical conductivities and dielectric constants, and then is reflected by them. Ultimately, GPR images the dielectric properties of the soil, and any discontinuities appear as a signal [1]. Another antenna of GPR sensor will receive the reflected pulse and store it in a digital control unit. The control unit present in the antenna system registers the reflections against the ground surface and then amplifies these signals. GPR pulse can reach depth up to 100 feet in low conductivity materials and decreases to around 30 feet in high conductive materials which may attenuate or absorb the GPR signals. The depth of penetration is also determined by the types of radar antenna used in the sensor. Antennas with low frequencies from 25 to 200 MHz can obtain 30-100 feet or more penetration depth. When antennas with high frequencies from 300-1000 MHz just have shallow penetration depth, the high frequency components can achieve high resolution for small targets. One problem of the GPR sensor is that the presence of clutter will has a holdback in the detection accuracy of land mines. The clutter in various soil and environmental conditions can lead to false alarms in land mine detection. Another problem is the soil itself, even under near-ideal test track conditions, soil itself is a remarkably inhomogeneous medium, and false alarms are easily generated from the background itself [1]. Therefore it is necessary to use signal processing techniques on the GPR data to overcome the these negative effects, especially for detecting small, shallow objects.

Another radar approach is synthetic aperture radar (SAR), which utilizes multiple antenna locations to improve resolution of the resulting image, essentially creating a larger antenna from multiple smaller antennas [28]. Lately, Energy Focusing Ground Penetrating Radar (EFGPR) has been proposed, this sys-

tem incorporates both bistatic impulse radar and synthetic aperture (SAR) principles. EFGPR is designed specifically for the detection of land mines, it is unique in its ability to focus in hardware and is designed by taking a wide variety of environments into consideration.

Though several sensors are developed in the sensor industry, ultra wideband GPR sensor have been accepted as popular sensor for the detection of land mines. GPR systems are usually divided into two board types: vehicle-mounted and hand-held GPR systems, we will focus on the vehicle-mounted GPR system in this thesis. Vehicle-mounted GPR is a rapid, continuous data collection system that contains highly accurate and repeatable data as it is moved over a pre-defined path. There are several constraints that need to be taken care of before designing any vehicle-mounted system which makes the system expensive. The system has the radar antenna mounted on a vehicle which is remotely controlled from an armor-protected vehicle. The vehicle's ground pressure should make it possible to drive over mines without detonating them. The antenna has to be very stable with respect to speed, vibration and other movements. All these factors have strong influences in the design of the vehicle-mounted GPR system. More contents concerning the vehicle-mounted GPR system will be introduced in chapter 3.

2.2 Ground Penetrating Radar Data Format

GPR data can be represented in three different forms, A, B, and C scans, according to the scanning dimension [29]. Figure 2.1 shows the 3D coordinate system defined on a section of earth, where the XY plane represents the ground surface and the Z-axis represents the direction into the ground.

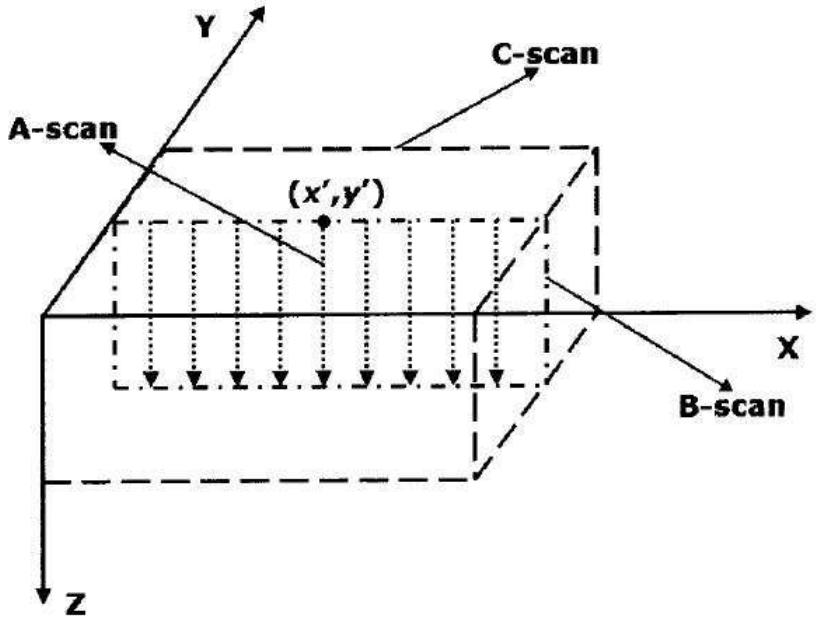

Figure 2.1: The 3D coordinate system defined on a section of ground [29].

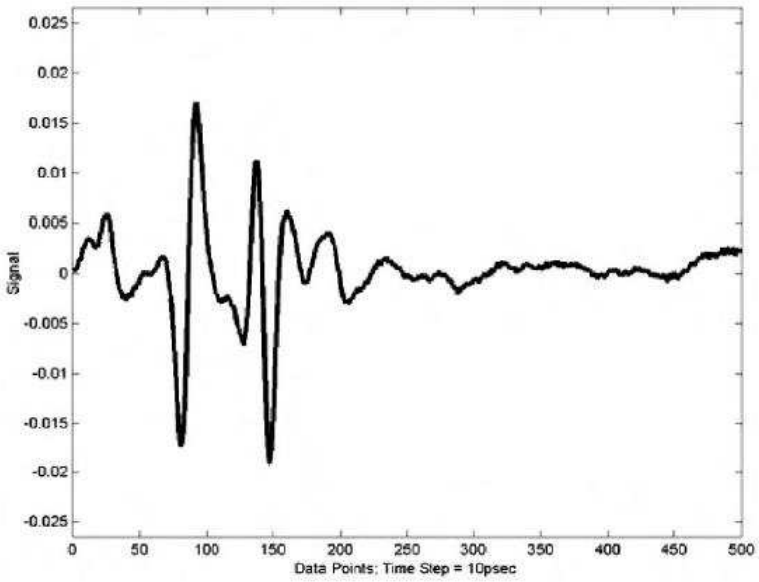

Figure 2.2: An example of an A-scan signal (1×500) [30].

The A-scan signal is obtained by a stationary measurement after placing an antenna above a specific position, such as (x, y) in Figure 2.1. The collected signal is presented in the form of a group of signal strength versus time delay. Figure 2.2 shows an example of an A-scanned signal obtained using an ultra-wide band (UWB) GPR. The positions of these peaks correspond to the distance between the antenna and various reflecting surfaces. The first peak is the air-ground reflection and the second is the mine target.

B-scan signal is obtained as the horizontal collection from the ensemble of A-scans. The collected signal is presented as intensity on the plane of scanned width versus time delay. Therefore, the B-scanned signal measured at $y = y'$ can be considered as a 2D signal, Figure 2.3 shows an example of B-scan.

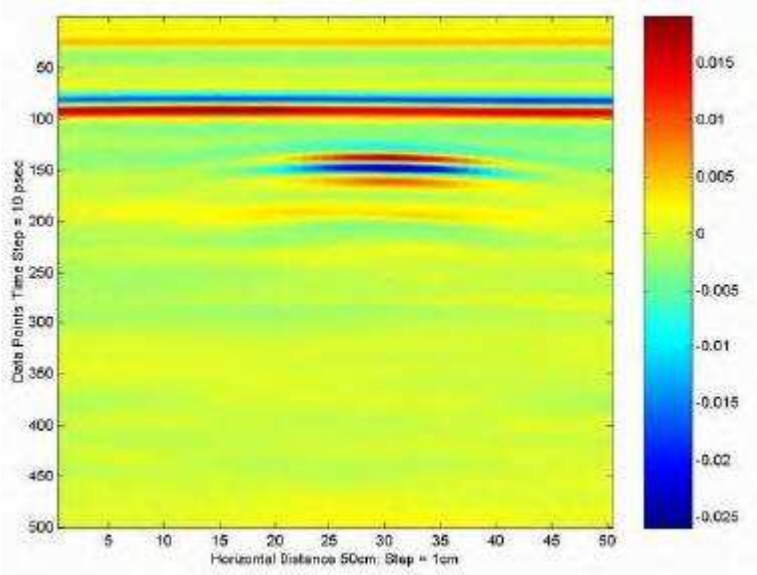

Figure 2.3: An example of a B-scan signal (500×50) [30].

In Figure 2.3, the vertical axis corresponds to the horizontal axis of Figure 2.2 and the horizontal axis represents the scanned width, which is the number of A-scans. The intensity or color of each pixel indicates the signal strength, and corresponds to the vertical axis of Figure 2.2. We can see from these two figures that the A-scan could only show some sign of object reflection in Figure 2.2, whereas a B-scan can distinguish a mine-like target from the air ground

surface and also can give information about the position of the object.

C-scan signal is obtained from the ensemble of B-scans, measured by repeated line scans along the plane. The collected C-scanned signal forms a 3D signal, which is depicted by the hexahedron shown in Figure 2.1. Since visualization of a three-dimensional data is not easy, a C-scan is usually represented by a collection of horizontal slices for a specific data point, that is XY-planes at each specific position on the Z-axis. Each slice corresponds to a certain depth level, which is equivalent to the vertical axis of the B-scan. Figure 2.4 shows an example of a C-scan signal.

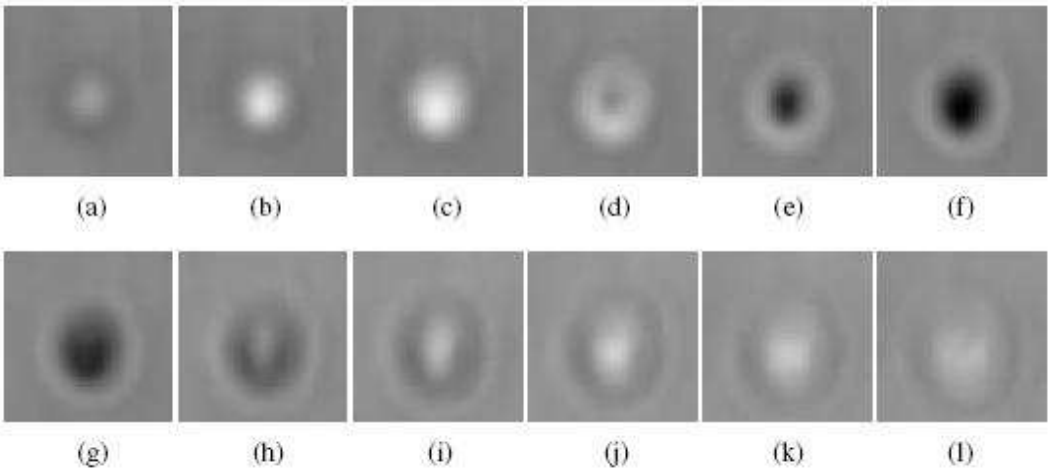

Figure 2.4: An example of a C-scan (50×50) [11]; consisting of horizontal slices at data points 133 to 166 at intervals of 3 depth points.

In Figure 2.4, both horizontal and vertical axes correspond to the horizontal axes of a B-scan, and the depth level corresponds to the vertical axis of B-scan. Although we cannot clearly define the air-to-ground boundary from these images, we can distinguish the target mine from its background, and can roughly figure out the shape of the target.

2.3 Testing and Evaluation

In the testing facilities, there are carefully constructed mine lanes. On these mine lanes, the land mines are real but unfused and thus do not detonate if they are run over by people or a detection system. All testing and evaluation environment is made purposely to be as close to real-world environment as possible.

These mine lanes will be divided into two parts: one part is called the training data set, the other one is testing data set. Training data set is used to obtain the GPR signals of land mines, and then the energy density spectrum of these signals are generated. Later, the clustering method and the subspace detector technique based on these energy density spectra are performed. The testing data set is used to evaluate whether the performance of the clustering method and the subspace detector technique can improve the detection result and reduce the false alarm rate. Also, the training and testing data set are not fixed, it means in one experiment, some mine lanes are belonging to the training data set, and then in next experiment, these mine lanes will belong to the testing data set, as shown in Figure 2.5.

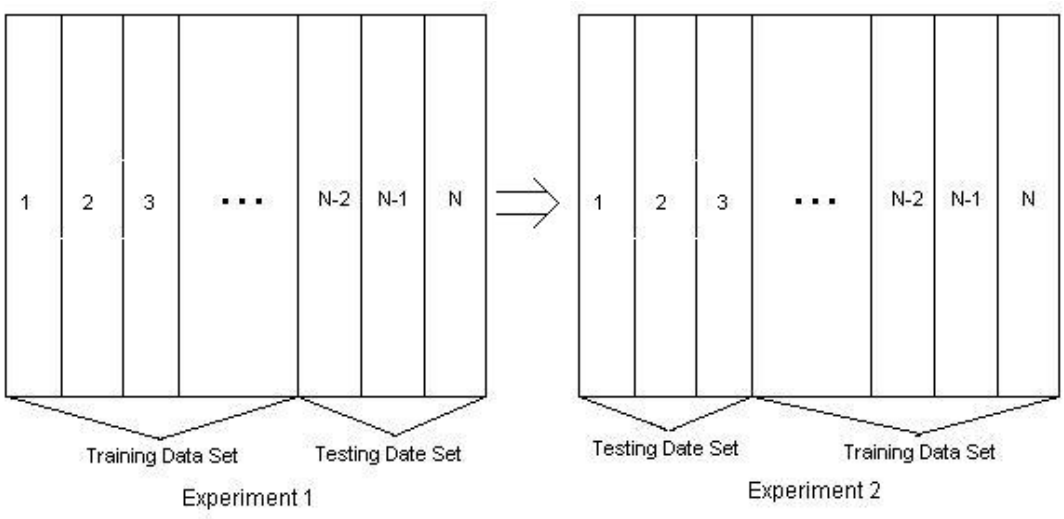

Figure 2.5: Mine lanes can belongs to different data set in different experiments.

2.4 Mathematics used in the Thesis

Maximum likelihood estimation and eigen decomposition are involved in the subspace detector technique which will be introduced and discussed in chapter 5. Here the basis introductions about these two mathematic methods will be given.

2.4.1 Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) is a popular statistical method used to calculate the best way of fitting a mathematical model to some data. Modeling real world data by estimating maximum likelihood offers a way of tuning the free parameters of the model to provide an optimum fit. The method was pioneered by geneticist and statistician Sir R. A. Fisher between 1912 and 1922.

In mathematical modeling, once a model is specified with its parameters, and data have been collected, the next step is to evaluate its goodness of fit, that is, how well it fits the observed data. Goodness of fit is assessed by finding parameter values of a model that best fits the data, a procedure called parameter estimation. There are two general methods of parameter estimation: least-squares estimation (LSE) and maximum likelihood estimation (MLE). Between these two methods, MLE has many optimal properties in estimation [31]: sufficiency (complete information about the parameter of interest contained in its MLE estimator); consistency (true parameter value that generated the data recovered asymptotically, i.e. for data of sufficiently large samples); efficiency (lowest-possible variance of parameter estimates achieved asymptotically); and parameterizations invariance (same MLE solution obtained independent of the parametrization used). From a statistical point of view, the method of maximum likelihood is considered to be more robust (with some exceptions) and yield estimators with good statistical properties.

In other words, MLE methods are versatile and apply to most models and to different types of data. In addition, they provide efficient methods for quantifying uncertainty through confidence bounds.

Consider a family D_θ of probability distributions parameterized by unknown parameters $\theta_1, \theta_2, \cdots, \theta_k$, which we need to estimate, associated with either a known probability density function (for continuous distribution) or a known probability mass function (for discrete distribution), denoted as $f(x; \theta_1, \theta_2, \cdots, \theta_k)$. We draw a sample $x_1, x_2, \cdots, x_n$ of n values from this distribution, and then using $f(x; \theta_1, \theta_2, \cdots, \theta_k)$ we compute the probability density associated with our observed data, which is

$$f(x_i; \theta_1, \theta_2, \cdots, \theta_k), \;\; i = 1, 2, \cdots, n \tag{2.1}$$

Commonly, one assumes that the data drawn from a particular distribution are independent, identically distributed (iid) with unknown parameters. This considerably simplifies the problem because the likelihood can then be written as a product of n univariate probability densities:

$$\begin{aligned} L(\theta_1, \theta_2, \cdots, \theta_k | x_1, x_2, \cdots, x_n) &= \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \cdots, \theta_k) \\ &= L \end{aligned} \tag{2.2}$$

Then the maximum likelihood estimator of $\theta_1, \theta_2, \cdots, \theta_k$ are obtained by maximizing L. Since L is not affected by monotone transformations, the logarithmic likelihood function is given by:

$$\Lambda = ln\ L = \sum_{i=1}^{n} ln\ f(x_i; \theta_1, \theta_2, \cdots, \theta_k) \tag{2.3}$$

Right now the maximum likelihood estimator of $\theta_1, \theta_2, \cdots, \theta_k$ can be obtained by maximizing Λ, which is much easier to work with than L. The maximum likelihood estimators of $\theta_1, \theta_2, \cdots, \theta_k$ are the simultaneous solution of k equations:

$$\frac{\partial \Lambda}{\partial \theta_j} = 0, \;\; j = 1, 2, \cdots, k \tag{2.4}$$

One thing needs to know is that the maximum likelihood estimator may not be unique, or indeed sometimes it may not even exist.

2.4.2 Eigenvalue, Eigenvector and Eigen Decomposition

Eigenvalues are a special set of scalars associated with a linear system of equations, i.e. a matrix equations, which are sometimes also known as characteristic values. The determination of the eigenvalues and eigenvectors of a system is extremely important in physics and engineering, where it is equivalent to matrix diagonalization. Each eigenvalue is paired with a corresponding so-called eigenvector. The decomposition of a square matrix $\mathbf{A}$ into eigenvalues and eigenvectors is known as eigen decomposition.

Let the matrix $\mathbf{A}$ be a linear transformation. If there is a vector $\mathbf{x} \in \mathbf{R}^n \neq 0$ such that

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x} \tag{2.5}$$

for some scalar λ, then λ is called the eigenvalue of the matrix $\mathbf{A}$ with corresponding eigenvector $\mathbf{x}$.

Let $\mathbf{A}$ be a $k \times k$ square matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{bmatrix} \tag{2.6}$$

with eigenvalue λ, then the corresponding eigenvectors satisfy

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \tag{2.7}$$

which is equivalent to the homogeneous system

$$\begin{bmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} - \lambda & \dots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2.8}$$

Equation (2.8) can be written as

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0, \tag{2.9}$$

where $\mathbf{I}$ is the identity matrix. According to Cramer's rule, a linear system of equations has nontrivial solutions if and only if the determinant vanishes, so the solutions of equation (2.9) are given by

$$det(\mathbf{A} - \lambda \mathbf{I}) = 0. \tag{2.10}$$

This equation is known as the characteristic equation of $\mathbf{A}$, and the left-hand side is known as the characteristic polynomial. If all k eigenvalues are different,

then plugging these value back in can give us $k-1$ independent equations for the k components of each corresponding eigenvector, and the system is said to be nondegenerate.

Assume that A has k nondegenerate eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ and their corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are linear independent. Define the matrices composed of eigenvectors

$$\mathbf{Q} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_k] = \begin{bmatrix} x_{11} & x_{21} & \ldots & x_{k1} \\ x_{12} & x_{22} & \ldots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1k} & x_{2k} & \ldots & x_{kk} \end{bmatrix}, \tag{2.11}$$

and eigenvalues

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_k \end{bmatrix}, \tag{2.12}$$

where $\mathbf{\Lambda}$ is a diagonal matrix. Then

$$\begin{aligned} \mathbf{AQ} &= \mathbf{A}[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_k], \\ &= [\mathbf{A}\mathbf{x}_1 \quad \mathbf{A}\mathbf{x}_2 \quad \ldots \quad \mathbf{A}\mathbf{x}_k], \\ &= [\lambda_1\mathbf{x}_1 \quad \lambda_2\mathbf{x}_2 \quad \ldots \quad \lambda_k\mathbf{x}_k], \end{aligned} \tag{2.13}$$

which can be rewritten as

$$
\begin{aligned}
\mathbf{AQ} &= \begin{bmatrix} \lambda_1 x_{11} & \lambda_2 x_{21} & \dots & \lambda_k x_{k1} \\ \lambda_1 x_{12} & \lambda_2 x_{22} & \dots & \lambda_k x_{k2} \\ \vdots & \vdots & & \vdots \\ \lambda_k x_{1k} & \lambda_2 x_{2k} & \dots & \lambda_k x_{kk} \end{bmatrix}, \\
&= \begin{bmatrix} x_{11} & x_{21} & \dots & x_{k1} \\ x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1k} & x_{2k} & \dots & x_{kk} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix}, \\
&= \mathbf{Q}\mathbf{\Lambda}.
\end{aligned}
\tag{2.14}
$$

According to equation (2.13) and equation (2.14), $\mathbf{A}$ can be expressed in terms of $\mathbf{Q}$ and $\mathbf{\Lambda}$ as

$$
\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}. \tag{2.15}
$$

The whole process above is called eigen decomposition. According to the eigen decomposition theorem, as long as $\mathbf{Q}$ is a square matrix, this decomposition is always possible for a square matrix $\mathbf{A}$.

Chapter 3

Land Mine Detection Using Frequency Domain Features from Vehicle Mounted Ground Penetrating Radar

3.1 Introduction

Metal detector can detect some small and shallow-buried metal antipersonnel mines. However, when it comes to the deeply buried land mines, a metal detector will have a difficulty to detect them. Also, nowadays more and more land mines are plastic-encased, these mines contain varying degree of metal, in the low degree of metal case, the metal detector may be unable to detect these land mines. Furthermore, in urban environment where there is a lot of background metal, the detection result of a metal detector will have a very high false alarm rate which would make the result useless. Because of all these disadvantages, detection technology development has been invested and implemented in other areas. Ground Penetrating Radar(GPR) is one of these technologies. Basically there are two types of GPR systems: handheld GPR systems and vehicle mounted GPR systems. In this research, we focus on the use of the vehicle mounted GPR systems. In the case of vehicle mounted GPR systems, the signal returns are usually corrupted by background noise and clutter. Also the pattern of signal return from a mine could be distinctive,

even if the energy of the signal return from a mine was relatively low. Hence, feature based processing techniques are usually employed in this case.

3.2 Ground Penetrating Radar and Vehicle Mounted GPR Systems

Ground penetrating radar is a geophysical method that has been developed for shallow, high-resolution, subsurface investigations of the earth. It has been applied to the problem of land mine detection for nearly 20 years [1] and gaining momentum in recent years as a sensor for land mine detection [3]. GPR is quite effective in detecting low metal or plastic land mines in different types of soils and it can be excited by pulse or frequency swept. During the detection process, high frequency signals, which are in the order of GHz, are sent out to penetrate the ground and the returned echo signals are then used for the detection of land mines. The penetration depth is related to both the frequency range produced and the soil attenuation: lower frequency components can penetrate further, but it is the higher-frequency components that can image and resolve smaller targets. Eventually, GPR gives out the image of the dielectric properties of the penetrating ground, any discontinuities occurs will be considered as a signal.

Vehicle mounted GPR systems, as the name implies, is a land mine detection systems that with the ground penetrating radar mounted on a vehicle. It is a concept relative to the handhold GPR systems. Comparing with the handhold GPR systems, this kind of land mine detection equipment should have the capability of detecting land mines in large and complicated condition surfaces such as roads, urban environment ground or even the battle front, at a acceptable speed.

Nowadays, many different types of vehicle mounted GPR systems have been developed and implemented in the land mine detection projects. All of them have the capability of detecting the low-metal or plastic-cased land mines. A conventional vehicle mounted GPR system will use one sensor or many sensors to proceed the detection, i.e, an array of sensors will be employed. Improved Landmine Detection System(ILDS) [32] is an instance: it is a multisensory teleported system which uses 24 metal detector coils to cover a 3 m swath, also 3 GPR modules are used, each consisting of a number of monostatic antennas, to cover the required detection field. The output of this system will be a combination of the outputs from all these detectors with data fusion technique in order to reduce false alarm rate from each individual detector. Another example for vehicle mounted GPR systems is the articulated robotic sensor system [33], a generic robotic device is used to move a land mine detection sensor over detection field in this system. The robotic device can be operated remotely so the dangers for the operator can be reduced. In this system, instead of an array of sensors, just one single sensor is employed to provide similar coverage so that the cost, size and overall complexity of the system is reduced with minor increase in mechanical complexity.

Wichmann/NIITEK system is the system which is being used recently to collect data in the field for detecting antitank land mines. In this system, the Wichmann/NIITEK radar mounted on a vehicle with antitank land mine overpass abilities. A marking system aiming to mark the locations where targets are determined is located at the rear part of the vehicle system. The Wichmann/NIITEK radar has very wide bandwidth (around 200 MHz to 7 GHz) and extremely low radar cross section, so even the inner structure of the buried objects could be easily determined by the radar. Base on this advantage, the identification and discrimination of targets are then possible with the use of

returned echo signals measured by the Wichmann/NIITEK system. The frequency domain feature based detection technique [3] is using the data collected from the Wichmann/NIITEK radar.

3.2.1 Wichmann/NIITEK GPR Data

The Wichmann/NIITEK to radar vehicular system uses an array of 24 antennas(channels), each antenna is approximately 5 cm apart from each other, and the whole array board is 1.2m wide. Since there are 24 channels, so the data consists of 24 cross-track channels for each scan and is sampled every 5 cm down-track as the vehicle moves in the down track direction. Each individual channel records 415 point vector of time samples and totally there are 24,415 point time-domain vectors for every 5 cm down track. Therefore pre-processing is needed on the data to avoid computational complexity in locating the position of interest. The data after the pre-processing is as shown in figure(3.1) below.

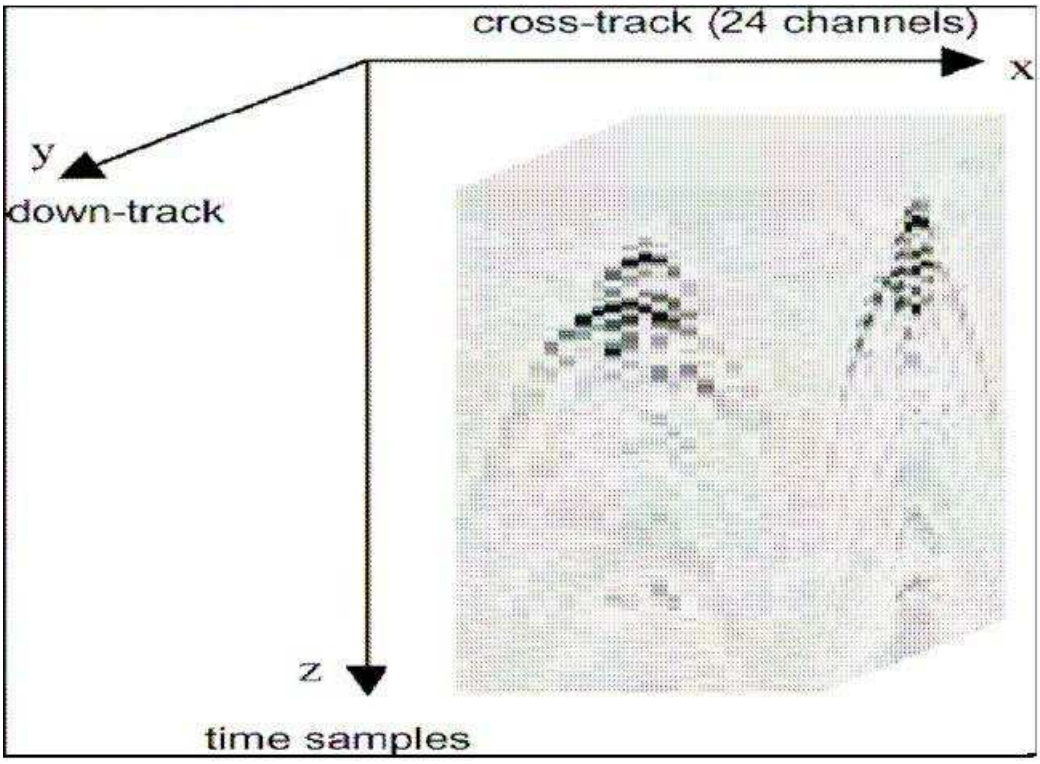

Figure 3.1: Wichmann/NIITEK GPR data after pre-processing [34]

3.3 Frequency Domain Feature Extraction Techniques

In most case, the collected GPR data from radar is very susceptible to ground bonce, which is an undesired signal generated due to high dielectric discontinuity present between the soil and air, and clutter response. Therefore signal processing is necessary to reduce false alarms caused by the ground bonce and clutter response, and improve the detection of land mines. Most GPR signal processing algorithm perform processing in the depth domain [35]-[38]. For certain specific mine types, the GPR signal return is very weak which presents difficulty for their detection [3]. In case where the time domain signature of GPR returned signal is week, sometimes the GPR returned signal spectrum can reveal important features for detection. The rest of this chapter study such a frequency domain technique using the spectral characteristics of land mines to improve the detection and reduce false alarms. Next section will summarize the procedure to generate the energy density spectrum of a suspected land mine location, and then use the energy density spectrum to extract the feature vectors and generate confidence values to detect the presence of land mine.

3.4 Energy Density Spectrum Generation

Ho et al. [3] proposed the frequency domain feature technique that uses the energy density spectrum produced by land mine targets to improve their detection. In this chapter, I will use the algorithm proposed by Ho at al. [3] as a reference to improve the detection result.

The GPR data in this study is a frequency sweep radar that collects data in the time domain. The bandwidth of the radar is up to 7 GHz and the sampling rate is about 62 GHz. A vector contains 415 data points is collected in each

physical location. Because of the wide bandwidth, this radar has a very high resolution in depth and the frequency resolution is:

$$Freq\ Resolution = \frac{62 \times 10^9}{415} = 150MHz \tag{3.1}$$

As mentioned in the section 3.2.1, the spatial sampling resolution in both cross-track and down-track is 5 cm. We use x to represent cross-track position and use y to represent down-track position, then the GPR data from certain location can be denoted as $g(x, y, z)$, where z is the index of depth that has values from 1 to 415. Then the data measurement can be modeled as:

$$d(x, y, z) = g(x, y, z) + v(x, y, z) + w(x, y, z) \tag{3.2}$$

where $g(x, y, z)$ represents the GPR signal, $v(x, y, z)$ represents the clutter responses and $w(x, y, z)$ denotes the noise component. After obtaining the data measurements, the generation of energy density spectrum contains 4 steps and will be studied in the following subsections.

3.4.1 Pre-processing

The ground level will be estimated and the data above and just below the ground will be removed in the pre-processing step. The depth of the ground level is estimated to be the mean of the maximum value and the minimum value of each vector sample averaging across down-track and cross-track. Then each vector sample will be shifted to perform data alignment. Eventually, we just use the data starting from 25 depth pixels below the ground surface to the $415th$ depth sample for further processing.

3.4.2 Non-linear Smoothing

In this step, 1-D median filter is applied to each depth bin separately to remove the noise component. The length of the median filter is 5.

3.4.3 Whitening

After median filtering, zeros are padded to each vector sample to make up the length of 512, then FFT is applied to each vector sample along depth. The FFT data before and after the current scan location of interest are used to compute the mean and standard deviation of the background for normalization.

Let (x_o, y_o) be the current position of interest, then the mean $m(x_o, k_z)$ and standard deviation $\sigma(x_o, k_z)$ of the background are:

$$m(x_o, k_z) = \frac{1}{2L}\Big(\sum_{i=y_o-G-L}^{y_o-G-1} D(x_o, i, k_z) + \sum_{i=y_o+G+1}^{y_o+G+L} D(x_o, i, k_z)\Big) \tag{3.3}$$

$$\begin{aligned} &\sigma^2(x_o, k_z) \\ &= \frac{1}{2L}\Big(\sum_{i=y_o-G-L}^{y_o-G-1} |D(x_o, i, k_z)|^2 + \sum_{i=y_o+G+1}^{y_o+G+L} |D(x_o, i, k_z)|^2\Big) - |m(x_o, k_z)|^2 \end{aligned} \tag{3.4}$$

where $D(x, y, k_z)$ represents the FFT data taken along depth at position (x, y), and k_z is the frequency domain index for the depth dimension. G is the number of guard samples that is set to be 6. L is the number of scans before or after the current location to perform averaging, which is also set to 6.

The normalization is then applied to the scans from $y_o - G$ to $y_o + G$, at each frequency bin k_z:

$$\begin{aligned} \widetilde{D}(x_o,y, k_z) &= \Big(\frac{D(x_o, y, k_z) - m(x_o, k_z)}{\sigma(x_o, k_z)}\Big), \\ y &= y_o - G, y_o - G + 1, \cdots, y_o + G \end{aligned} \tag{3.5}$$

After the normalization, the magnitude of $\widetilde{D}(x_o, y, k_z)$ is subtracted out the mean and clipped at the root-mean-square value, where the mean and root-mean-square value are computed at each frequency bin index k_z over $y =$

$y_o - G, y_o - G + 1, \cdots, y_o + G$. Finally the data is magnitude squared and denoted as $U(x_o, y, k_z)$.

3.4.4 Spectrum Generation

The spectrum is generated by averaging $U(x_o, y, k_z)$ over a square window of N samples cross-track times N samples down-track, where N is set to 5 in the current study:

$$P(x_o, y_o, k_z) = \frac{1}{N^2} \sum_{x=x_o-(N-1)/2}^{x_o+(N-1)/2} \sum_{y=y_o-(N-1)/2}^{y_o+(N-1)/2} U(x, y, k_z) \tag{3.6}$$

Figure 3.2 shows the energy density spectra of two different types of plastic anti-tank land mines which are known to be difficult to detect.

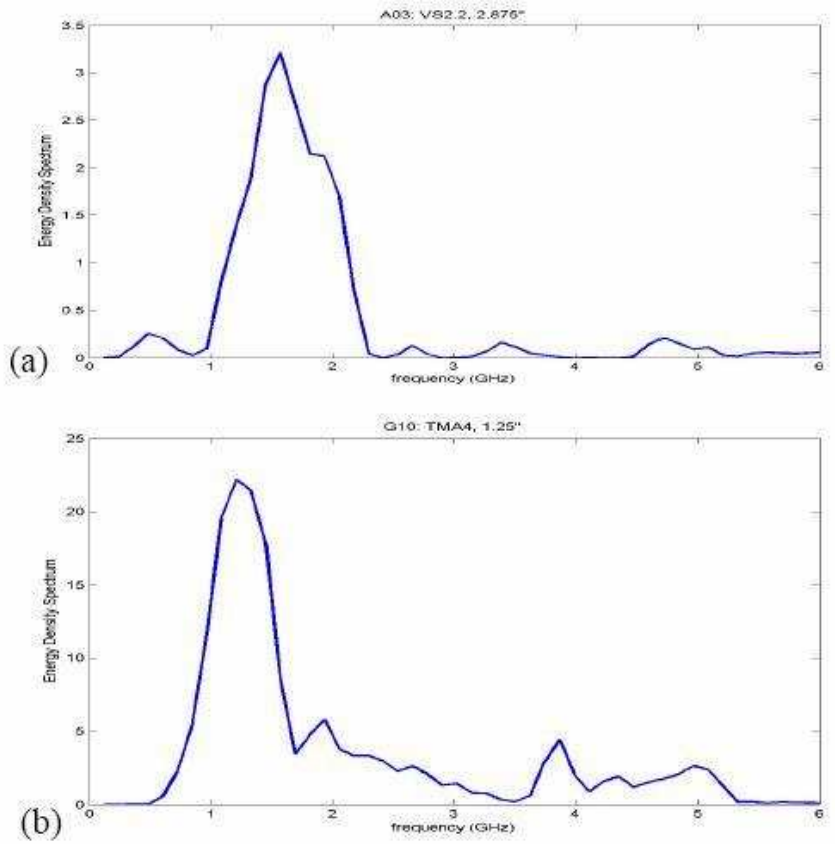

Figure 3.2: Energy density spectra. (a)Type-1 land mine, (b)Type-2 land mine. [3]

From the figure, we can observe that the energy density spectra for these two type of land mines have strong spectral peaks in certain point.

Figure 3.3 show the energy density spectra of two different types of clutter objects. One of them has metal content, the other is just a piece of irregular plastic.

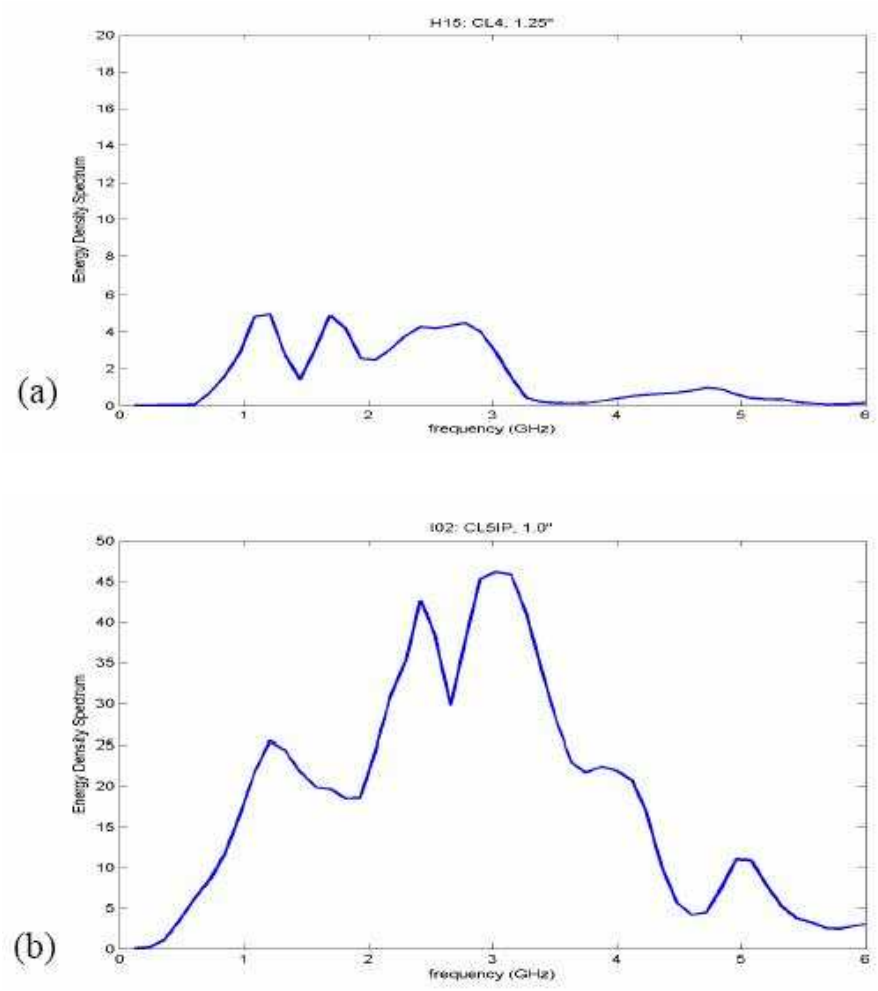

Figure 3.3: Energy density spectra. (a)Clutter object 1, (b)Clutter object 2. [3]

In the figure, both of the clutter objects have high GPR energy return, however, their spectra do not have any strong harmonic components.

As a result based on the observations from figure 3.2 and figure 3.3, it is expected that the energy density spectrum can be useful in improving the detection of some weak mines and reduce false alarms.

3.5 Spectral Feature Generation

To make use of the energy density spectrum to improve mine detection, a spectral feature vector shall be generated [3] in which the elements are the spectral energy at different frequency bands. Initially, the size of each frequency band is set to 600 MHz for the study, hence over the the total bandwidth which is 6 GHz, there is 10 spectral features. Since the sampling frequency is 62 GHz and the FFT size is 512 in this case, then the frequency bin size is computed as $62/512 = 120$ MHz. Note that each frequency band is 600 MHz, so each frequency band will covers $600/120 = 5$ frequency samples.

Because the GPR signal has a wide bandwidth, so sub-banding technique is used here in frequency domain to reduce the computational complexity. A cosine square window is used to decompose the frequency bands. The jth spectral feature is generated by:

$$Q(x_o, y_o, j) = \sum_{i=-(M-1)/2}^{(M-1)/2} P(x_o, y_o, Bj - \frac{B}{2} + i) cos^2(\frac{\pi}{M} i) \qquad (3.7)$$

where B is the frequency band size that is set to 5. M is the window size and is equal to $M = 2B - 1$. There is 50% overlap between two adjacent bands, and j takes values from 1 to 10.

In order to investigate the effectiveness of the spectral features in improving land mine detection result, a confidence value based on the spectral feature vectors should be generated. It is the weighted sum of the 10 spectral features we attain from (3.7). Use $\mathbf{Q}$ to represents the spectral feature vector and $\mathbf{W}$ to denote the weight vector which indicates the strength of the land mines, then the confidence value is

$$Conf = log\Big(\mathbf{W}^T\mathbf{Q} + 1\Big) \tag{3.8}$$

The weight vector $\mathbf{W}$ in this study is chosen to detect the weak land mines and it is found by experimentation as

$$\mathbf{W} = [0.2, 0.4, 1, 0.4, 0.2, 0, 0, 0, 0, 0]^T \tag{3.9}$$

Figure 3.4 below shows an example of the feature vectors obtained using 10 subbands.

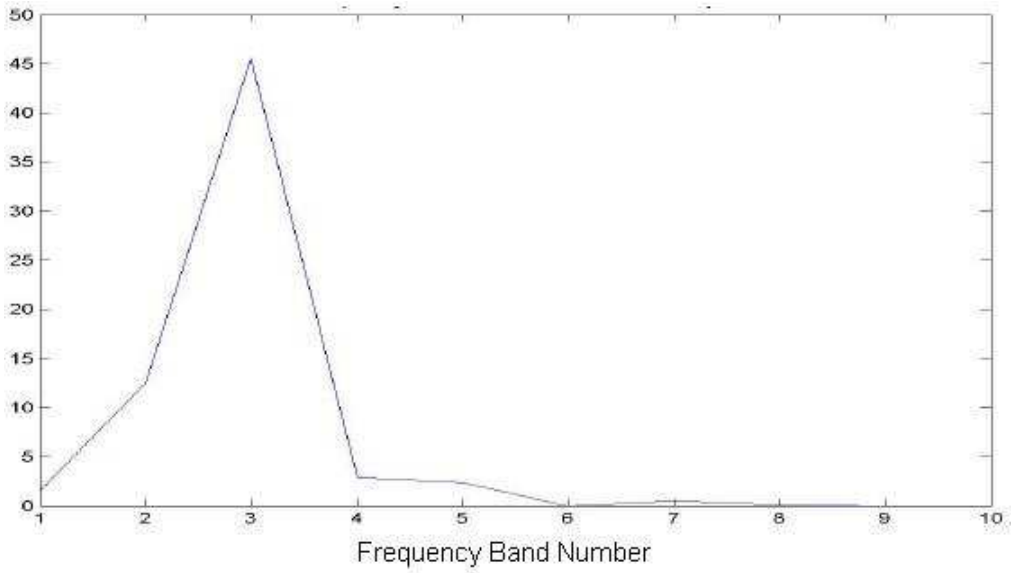

Figure 3.4: Spectral features for 10 subbands for plastic anti-tank land mine. [39]

3.6 Receiver Operating Characteristic

In this land mine detection project, special software developed by NIITEK, namely Counter Mine Test Measurement system (CMTS) is used to score the confidence values based on the data collected from various test sites. Later, the receiver operating characteristic (ROC) which is the plot of probability of detection versus probability of false alarms is generated. Then, the obtained ROC is compared with the FROSAW algorithm [38]. For the FROSAW algorithm, feature vectors are generated using feature based rules, order statistics and adaptive weighting. Figure 3.5 and figure 3.6 show the comparison result.

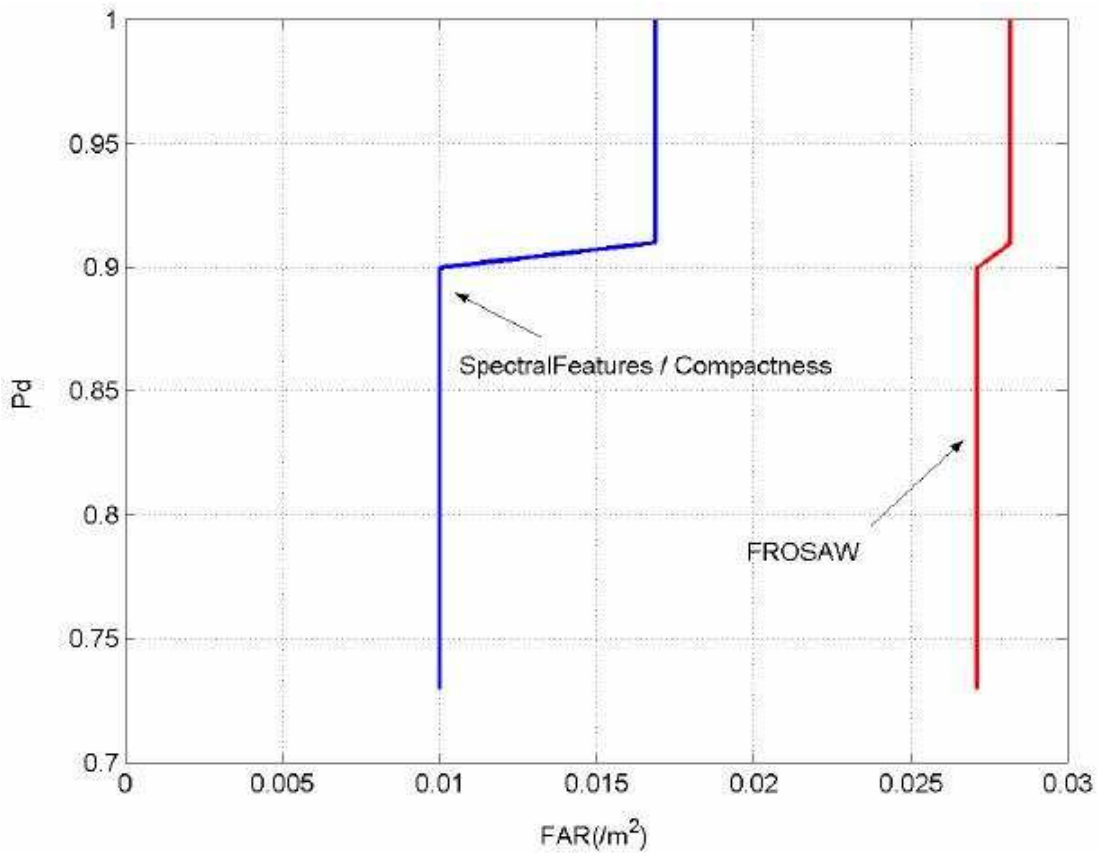

Figure 3.5: Performance of the spectral features method, dataset 1. [3]

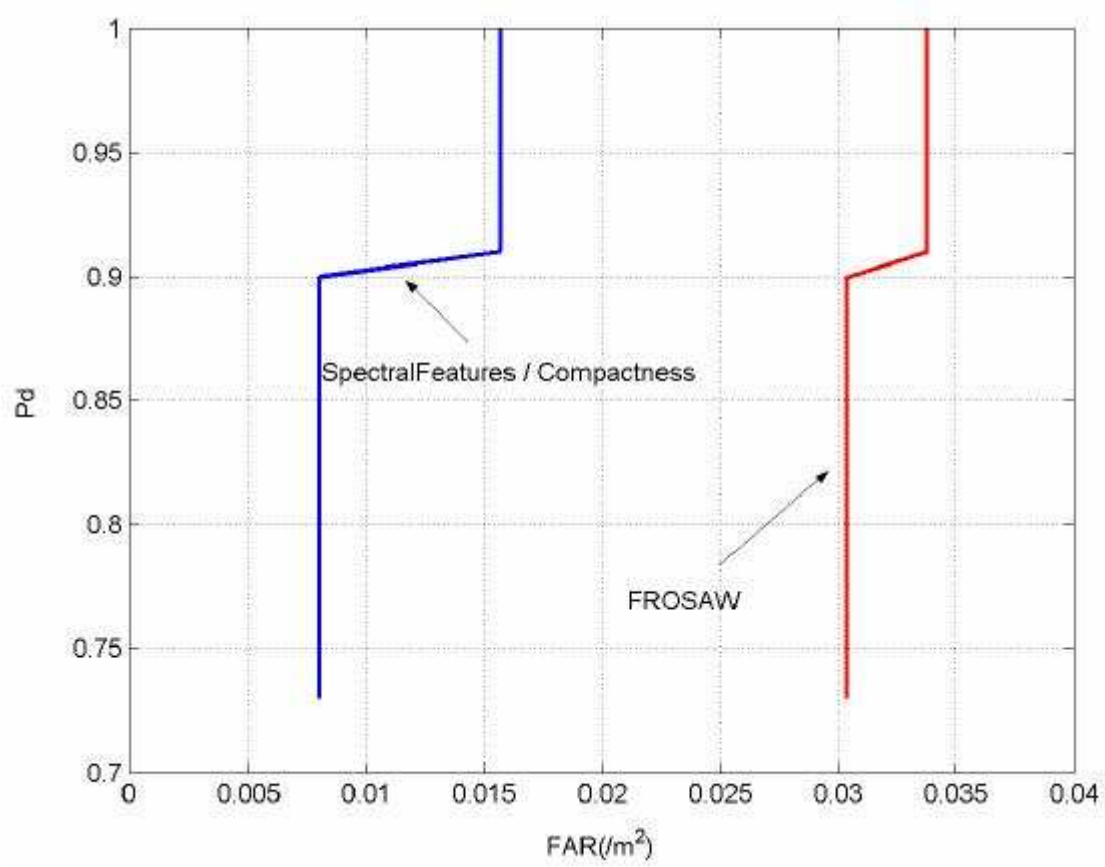

Figure 3.6: Performance of the spectral features method, dataset 2. [3]

In these two figures, "Pd" represents "Probability of Detection", and "FAR" denotes "Probability of False Alarms". It can be seen from these figures that the ROC curve from the spectral features has almost half of the false alarms from the FROSAW algorithm. We can have the conclusion that the spectral feature technique improves the probability of detection and reduces the false alarms, its performance is better than the FROSAW algorithm.

3.7 Effect of Increasing Subbands

The effect of increasing the number of subbands from 10 to 20 has been studied by Pisipati [40]. The study shows that when the number of subbands is increased to 20, the performance of the algorithm has some improvement in the detection result. In the case of 20 subbands, three different weighting mask are generated with three different pattern, and then the related maximum confidence value obtained from these three masks is used as the final confidence value. The three weighting masks are given below

$$\mathbf{W}_1 = [0.02, 0.1, 0.177, 0.558, 1, 0.243, 0.03, 0.01, 0.002, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$$

$$\mathbf{W}_2 = [0.014, 0.155, 0.638, 1, 0.042, 0.13, 0.539, 0.25118, 0.0137, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$$

$$\mathbf{W}_3 = [0.03, 0.1, 0.25, 0.57, 0.83, 1, 0.83, 0.57, 0.25, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$$

Figure 3.7 shows an example of the feature vectors with 20 subbands

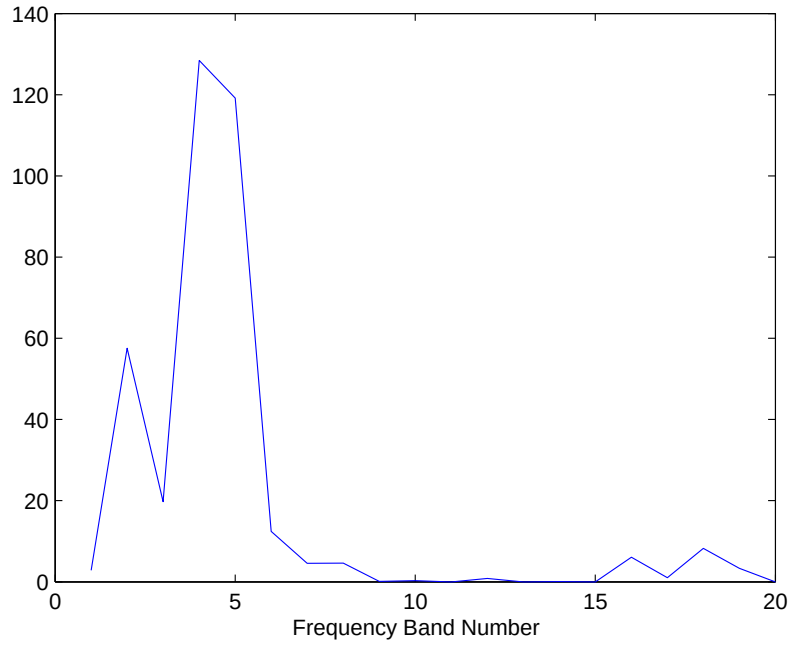

Figure 3.7: Spectral Features for 20 subbands

All my research and investigation of the clustering technique and subspace technique in the following chapters are based on the spectral feature vectors obtained using 20 subbands.

3.8 Conclusion

In this chapter, the knowledge of the vehicle mounted GPR radar system is reviewed. Its capability of performing detection in the complicated terrain such as urban environment and the ability to detect the deeply buried, low-metal or even plastic cased land mines make the vehicular system a much better detection tool over the metal detector and the handheld GPR radar system. Many kinds of vehicle mounted GPR radar system have been developed so far, and among them, the Wichmann/NIITEK system has been emphasized in this chapter since the data used in the land mine detection project at hand are collected from this system. Also, the data structure of the Wichmann/NIITEK system has been introduced.

Since the GPR data from the radar contains undesired signal which is known as ground bounce, signal processing technique is needed to improve the detection result and reduce the false alarms. The technique proposed by Ho and Gader [3] which is performed in the frequency domain is introduced as a performance reference in this chapter. This algorithm generates the energy density spectrum of the GPR data, and then some spectral feature vectors are produced to make use of the energy density spectrum. Later a confidence value is generated to investigate the effectiveness of the spectral features in improving land mine detection. Eventually through the investigation by Pisipati [40], when the number of subbands used is 20, the algorithm has the best performance in detection, and it becomes the basic setting of the clustering technique and subspace technique which will be introduced in the following chapters.

Chapter 4

Clustering Method Based on the Spectral Feature Vectors

4.1 Introduction

For the land mines that are buried under the ground, they could be divided into different groups according to the different "feature" they have. For example, according to the material, they can be divided into "Metal Mines", "Plastic Mines" and even "Wooden Mines". Or according to the shape, they can be divided into "Round Mines", "Rectangular Mines" and so on. The fact that the land mines in a certain group would share similar features motivated our research to investigate the possibility to improve the detection result using these features, to be more specific, the spectral features of land mines will be used. In this approach, the spectral feature vectors each has size 20×1 would be used as the database. All the feature vectors are generated by the energy density spectrum from each confirmed land mine in the training data set. The goal is try to find out if there are some "hidden pattern" within these land mines signals from the training data set, and then use these "hidden pattern" for the detection in the testing data set to see whether some improvement can be achieved in the detection result.

4.2 Classification and Clustering

Classification is produced by sorting similar objects into different groups, it is a basic ability of human beings.

> An intelligent being cannot treat every object it sees as a unique entity unlike anything else in the universe. It has to put objects in categories so that it may apply its hard-won knowledge about similar objects encountered in the past, to the object at hand.
>
> Steven Pinker, How the Mind Works, 1997.

After classification, a set of objects will be divided into different groups, and for each group there is a label which can concisely describes the similarities or differences for the objects. The concept of classification widely exists in our life without being noticed, for example, each noun in a language is essentially a label used to describe a class of things with common features [40], such as T-shirt, jeans and dress are called 'clothes'.

Nowadays, larger and larger databases are used in many research areas such as computer science, economics, geology science and so on, inside these databases there are many data that would shared certain properties, so with a classification scheme people can organize these databases in a better way so that they can access to these information more efficiently. Numerical techniques which aim is to provide objective and stable classifications have been applied in these area. Clustering is one of these numerical methods. It is an unsupervised classification, unsupervised means there is no predefined classes before the clustering, i.e the class labels and numbers of class are unknown.

The definition of clustering is: a processing of partitioning a set of data/objects into a set of meaningful sub-classes. This kind of sub-classes are called cluster which is a collection of data/objects that are similar to each other in some

way and can be treated collectively as one group. The complete set of clusters should contains all data/objects.

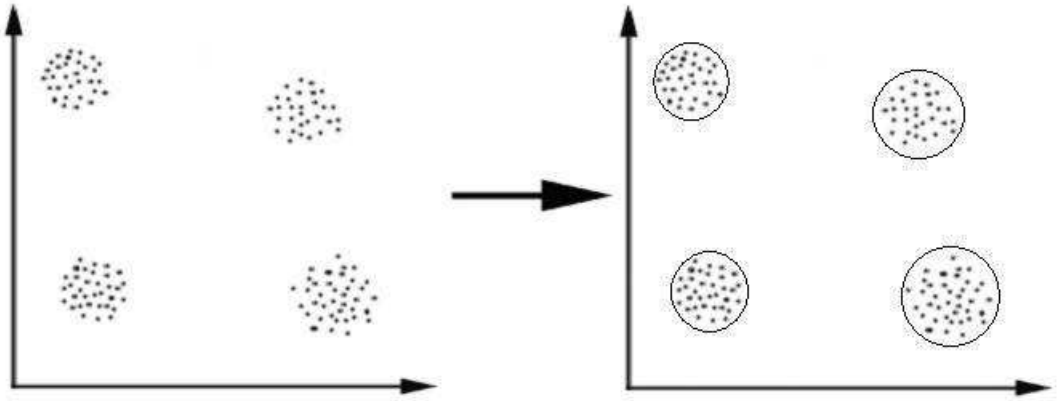

Figure 4.1: Principles of clustering. [41]

In the above figure, the similarity criterion is "distance": the elements in each cluster is close to other elements inside that cluster. As a good clustering method, it should produce high quality cluster in which the intra-class dissimilarity is low and the inter-class dissimilarity is high. Also, the quality of the cluster depends on its ability to discover the hidden patterns within the set of objects. The basic form of data that are used for analyzing after clustering, is usually a $m \times n$ multivariate matrix $\mathbf{X}$ whose elements are variables presenting the objects inside the cluster. The entry x_{ij} in $\mathbf{X}$ presents the jth variable on object i.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \tag{4.1}$$

In the land mine detection project, the data matrix is formed by putting all the spectral feature vectors together, each vector's size is 20×1, and there are totally 548 feature vectors, so the matrix dimension is 20×548.

4.3 Measurement of Proximity

In clustering, for the objects among all the clusters, it is very important to identify how 'close' or how 'far away' they are to others, since if two objects are from the same cluster but they are 'far away' from each other, then that cluster will not be a good cluster. Usually this kind of observation is not easy to do, so a quantitative measurement is necessary for this situation, commonly it is called *similarity, dissimilarity or distance*, and a general terminology is given to them as *proximity*. Two objects can be considered as 'close' if their similarity is high or their dissimilarity is low. Sometimes proximity can be obtained directly, such as in some experiments which involve a set of objects of interest and individuals are asked to determine the perceived similarity or dissimilarity among these objects. However, in most cases, proximities has to be determined indirectly.

There are almost endless number of similarity or dissimilarity measurements exists, which make it not so easy to choose a proximity measure method. However, there are some heuristics in selecting proximity measure: First of all, the nature of the data that needed to deal with will strongly influence the choice; Second, the choice should depend on the scale of data, for example, the choices of proximity measure should be different for binary data and continuous data; Last but not least, the clustering method to be used might also have some implications for the choice. In general the variables can be divided into three different types: categorical variables, continuous variables and data sets including both categorical and continuous variables. The spectral feature vectors in which we are interested in for the land mine detection belong to continuous variables, and we also want to know the distance between each feature vector and the centroids. So as a result here we will focus on the dissimilarity measures for continuous data. There are many measures for deriving the

dissimilarity matrix from a set of continuous variables, three commonly used dissimilarity measures are listed in Table 4.1.

Table 4.1 Dissimilarity measures for continuous data

Measure	Formula
Euclidean distance	$d_{ij} = \left\{\sum_{k=1}^{p}(x(k)_i - x(k)_j)^2\right\}^{1/2}$
City block distance	$d_{ij} = \sum_{k=1}^{p}\lvert x(k)_i - x(k)_j\rvert$
Minkowski distance	$d_{ij} = \left\{\lvert\sum_{k=1}^{p}\big(x(k)_i - x(k)_j\big)\rvert^r\right\}^{1/r} (r \geqslant 1)$
Canberra distance	$d_{ij} = \sum_{k=1}^{p}\lvert x(k)_i - x(k)_j\rvert/\big(\lvert x(k)_i\rvert + \lvert x(k)_j\rvert\big)$ for $x(k)_i \neq 0$ or $x(k)_j \neq 0$

In Table 4.1, $x(k)_i$ is the kth variable value of the p-dimensional observation for individual i and $x(k)_j$ is the kth variable value of the p-dimensional observation for individual j. The dissimilarity measure used in this clustering approach is very similar to the 'Euclidean distance'. Here the dissimilarities are the distances between the elements of different clusters and centroids of these clusters. The distance is defined as:

$$d_{k,l} = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{20}\Big\{x(k)_{ij} - c(l)_i\Big\}^2, \tag{4.2}$$

where $d_{k,l}$ represents the distance between the elements of cluster number k and the centroid of cluster number l. 20 and N are, respectively, the number of rows and columns of the spectrum feature matrix of cluster number k. $x(k)_{ij}$ represents the value of element in the ith row and jth columns of the spectral feature matrix of cluster k. Also it means that the spectral feature matrix is formed by N spectral feature vectors whose size is 20×1. The $c(l)_i$ represents the value of element in the ith row of the centroid vector of cluster number l, the size of the centroid vector is 20×1. After obtaining the proximity measure, we right now have a standard criterion for dividing the whole data set into different clusters.

4.4 Categories of Clustering Methods

After knowing the proximity measurement, we should move on to the step of deciding which clustering method to use for the land mine detection project. There are so many types of clustering methods, according to the algorithms implemented in these methods, they can be divided into five categories [42]: partitioning algorithms; hierarchy algorithms; Density-based algorithms which based on the connectivity and density functions of the data; Grid-based algorithms which based on a multiple-level granularity structure of the data; Model-based algorithms in which a model is hypothesized for each of the clusters and the idea is to find the best fit of that model to each other. Among them, partitioning algorithms and hierarchy algorithms are essentially two major clustering techniques, so more details about these two algorithms will be introduced here.

Hierarchical algorithms find successive clusters using preciously established clusters [43], which means it does not determine all the clusters at one step. Hierarchical algorithms can be subdivided into agglomerative ('bottom-up') methods or divisive ('top-down') methods. Agglomerative methods begin with each element as a separate cluster and fusion them into successively larger clusters. Divisive methods separate the whole set into successively better and smaller clusters. No matter what kind of hierarchical methods, i.e 'fusion' or 'division', once it is implemented, it is irrevocable [40], which means the clusters it generated can not be changed any more. The end result of the algorithms is a tree of clusters called a dendrogram, which shows how the clusters are related, with individual elements at one end and a single cluster containing every element at the other. By cutting the dendrogram at a desired level a clustering of the data items into disjoint groups is obtained. Agglomerative algorithms begin at the top of the tree, whereas divisive algorithms begin at

the bottom.

In contrary to hierarchical algorithms, partitioning algorithms will determine all clusters at once [43], it attempts to directly decompose the data set into a set of disjoint clusters. This clustering technique will produces various partitions of the whole data set, evaluate them by some criterion which may emphasize the local structure or the global structure of the data, and then divide them into a a specified number of clusters that optimize the chosen partitioning criterion. Typically the global criteria involve minimizing some measure of dissimilarity in the samples within each cluster, while maximizing the dissimilarity of different clusters. The number of clusters is fixed before hand and the clusters are changing during the process until the optimization is achieved.

For the land mine detection, we want to find out the hidden pattern among the collected data from the training data set according to the spectral feature vectors, recall that the proximity measure for the project here is the distance between each elements of a cluster to the centroids of all clusters, so we want the elements belongs to the same cluster to be very 'close' to each other and very 'far away' to the elements from other clusters, in other words, minimizing the dissimilarity within each cluster and maximizing the dissimilarity of different clusters. For these reasons, the partitioning algorithm which can generate the optimized disjoint clusters will be chosen to implement the clustering. To be more specific, the K-Means clustering algorithm will be implemented.

4.5 K-Means Algorithm used in Land Mine Detection

K-Means algorithm is one of the simplest unsupervised learning algorithms that can solve the well known clustering problem. As the name imply, given a k , the algorithm will find a partition of k clusters that optimizes the chosen partitioning criterion. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters(assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them far away as much as possible from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early status is done. At this point we need to re-calculate k new centroids as barycenters of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop we may notice that the k centroids change their location step by step until no more changes are done. In other words centroids do not move any more. Finally, this algorithm aims at minimizing an objective function, which is a squared error function [44]. The objective function is:

$$J = \sum_{i=1}^{k} \sum_{j=1}^{n} \left\| x(j)_i - c_i \right\|^2, \tag{4.3}$$

where $x(j)_i$ represents the jth object of cluster number i, c_i denotes the center of cluster number i and $\left\| x(j)_i - c_i \right\|^2$ is a chosen distance measure between the data point x_{ji} and the cluster center c_i, it is an indicator of the distance of the n data points from their respective cluster centers.

Given k, the K-Means algorithm can be implemented in 4 steps:

i. Partition objects into k nonempty subsets.

ii. Compute seed points as centroids of the clusters of the current partition, the centroid is the center(mean point) of the cluster.

iii. Assign each object to the cluster with the nearest centroid.

iv. Go back to step ii, stop when no more new assignment.

These steps can be illustrated by the figures below:

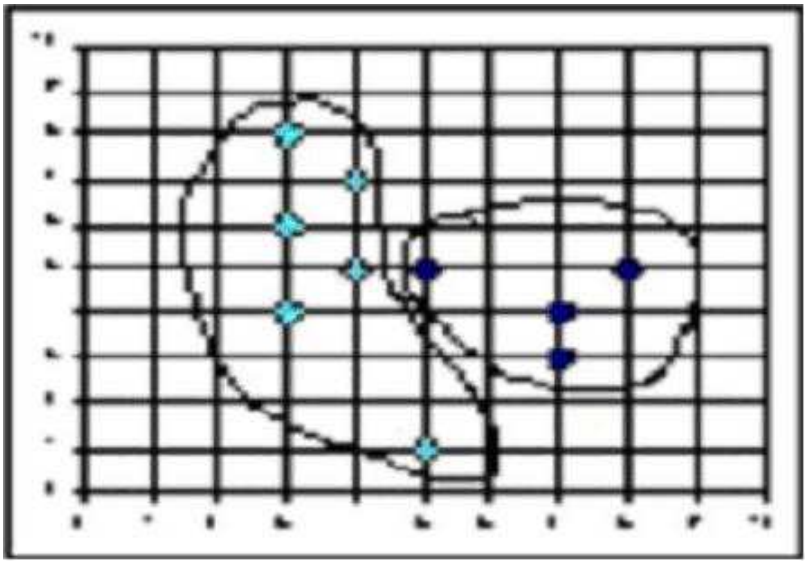

Figure 4.2: K-Means algorithm step 1. [42]

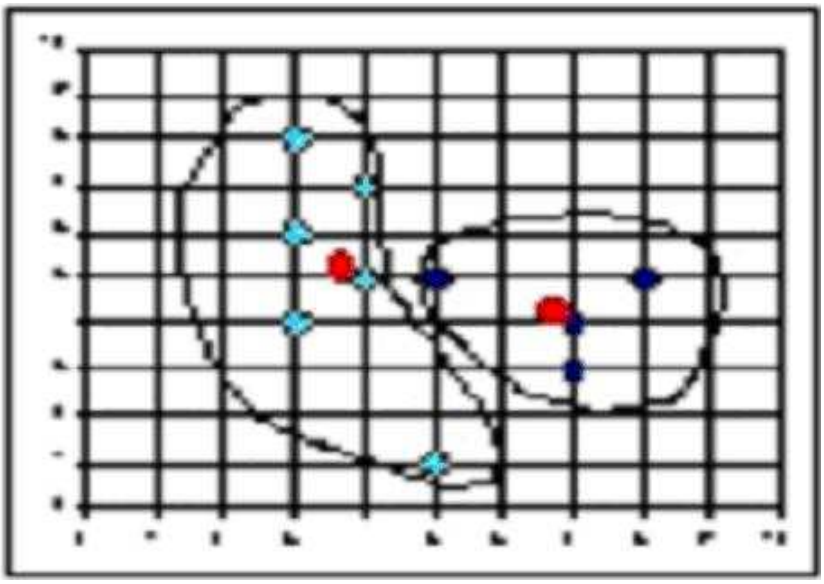

Figure 4.3: K-Means algorithm step 2. [42]

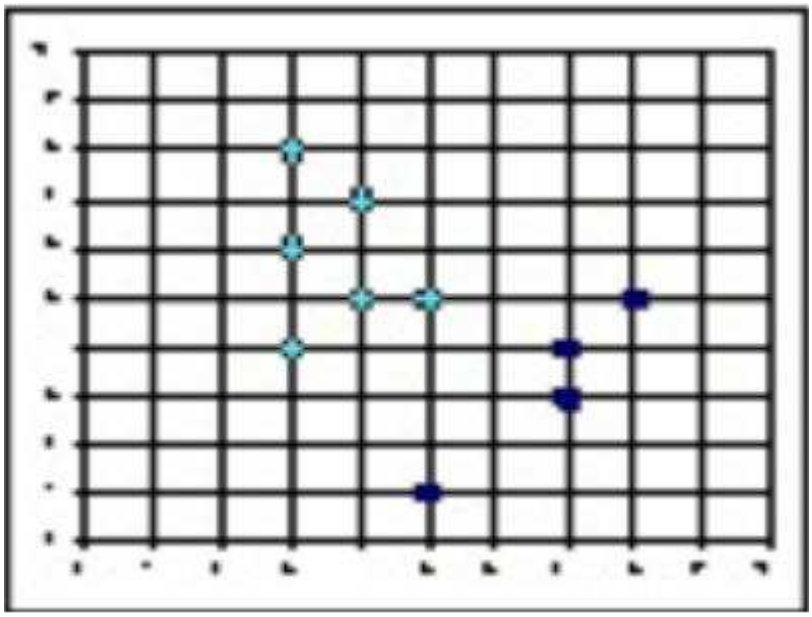

Figure 4.4: K-Means algorithm step 3. [42]

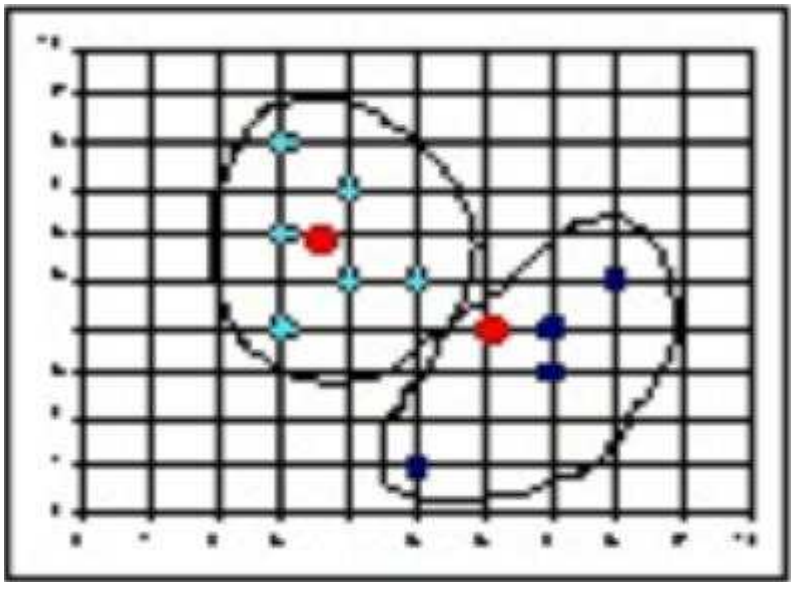

Figure 4.5: K-Means algorithm step 4. [42]

The strength of K-Means algorithm is its relatively efficient processing, which is on the order of $(t \times k \times n)$, where n represents the number of objects in the data set, k represents the number of clusters and t represents the number of iterations, normally, the value of k and t are much smaller than the value of n. Also, it often terminates at a local optimum, which is what we need here for the clustering results.

When implement K-Means algorithm in the land mine detection project, the data set we used are the spectral feature vectors, these feature vectors are all generated by the energy density spectrum from each confirmed land mine signal. The aim of the clustering technique is try to find out if there are some 'hidden pattern' within these signals form the training data set and the proximity here is set to be the distance between objects within clusters and the centroids of these clusters.

In the initial stage, the number of clusters that the whole data set will be divided into is set to be 7. This number is decided by experimentation, after trying many value range from 3 to 10, we find out that when the whole data set is divided into 7 clusters, the best performance from the K-Means algorithm can be obtained. The quality of the performance is justified by whether or not it can produce "good quality" clusters. The quality of a cluster is then decided by the ratio of the distance between the objects within this cluster and the centroid of this cluster to the distance between these same objects within this cluster and the centroids of other clusters. If the ratio value is equals to or smaller than 0.5, which means the distance from the objects of this cluster to the centroids of other clusters is at least twice the range the distance between these objects and their centroid within this cluster, then it will be considered as a "good quality" cluster, otherwise, its quality is no so good.

The example figure for “good quality” cluster of cluster number 7:

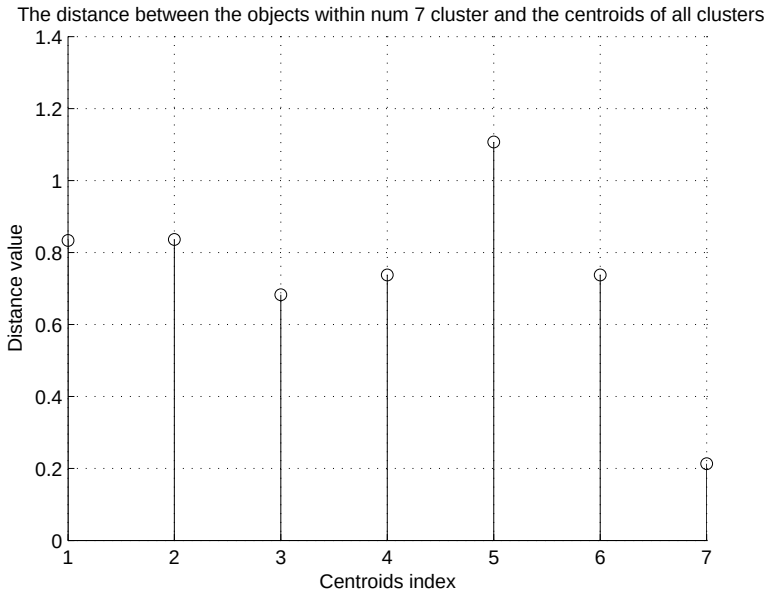

Figure 4.6: Example of good quality cluster.

The example figure for “not good quality” cluster of cluster number 2:

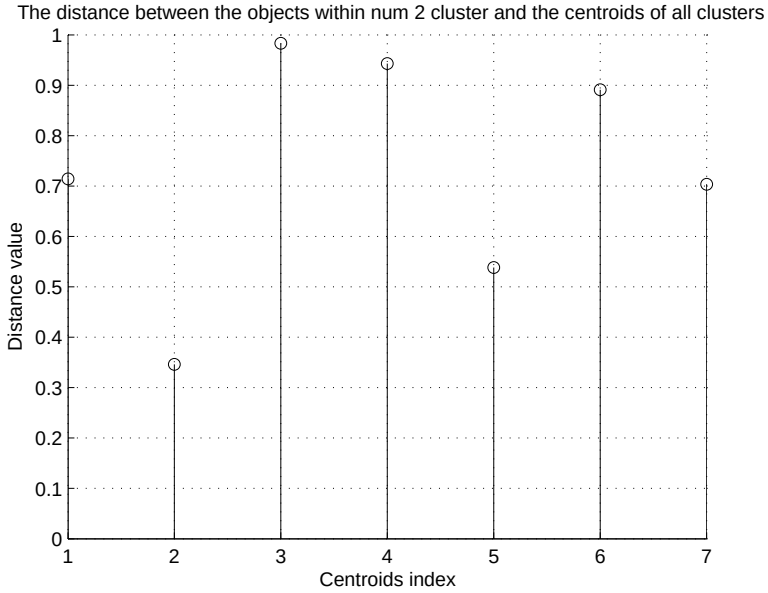

Figure 4.7: Example of not good quality cluster.

After the K-Means algorithm is implemented and its final step is completed, the value of the 7 centroids will not change anymore, and the whole data set will be divided into 7 clusters according to the locations of these 7 centroids obtained by the algorithm. Below the 7 centroids from all the clusters will be shown, as well as the standard deviation generated from all the objects within these clusters.

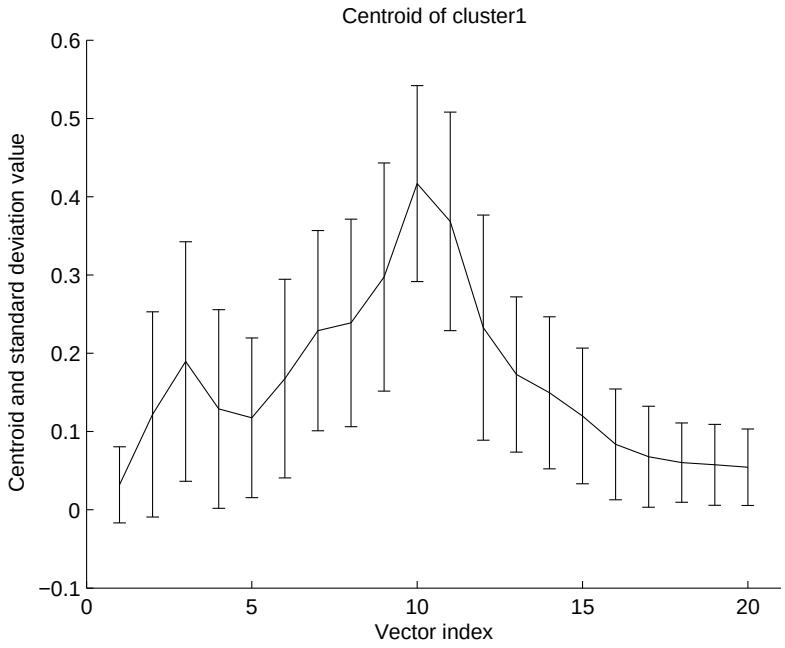

Figure 4.8: Centroid of cluster number 1.

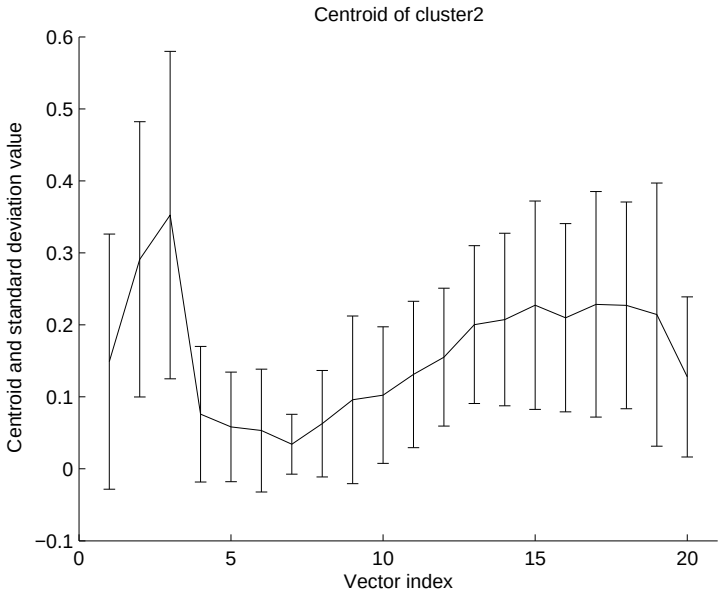

Figure 4.9: Centroid of cluster number 2.

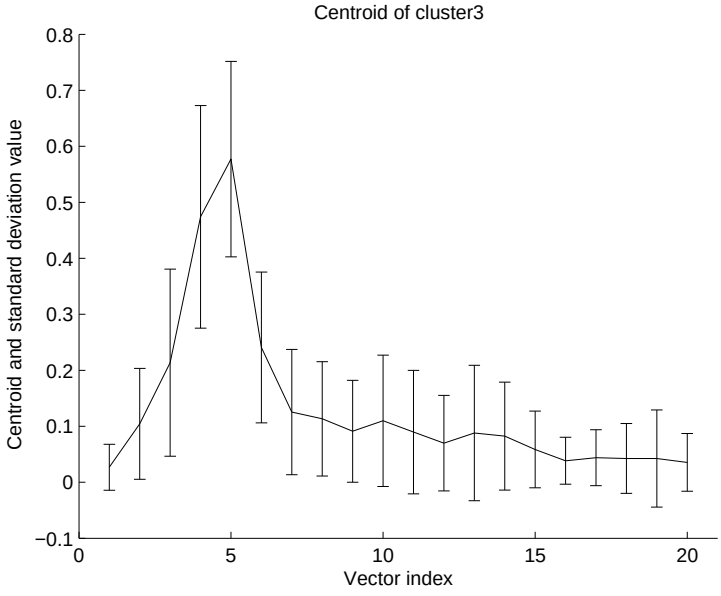

Figure 4.10: Centroid of cluster number 3.

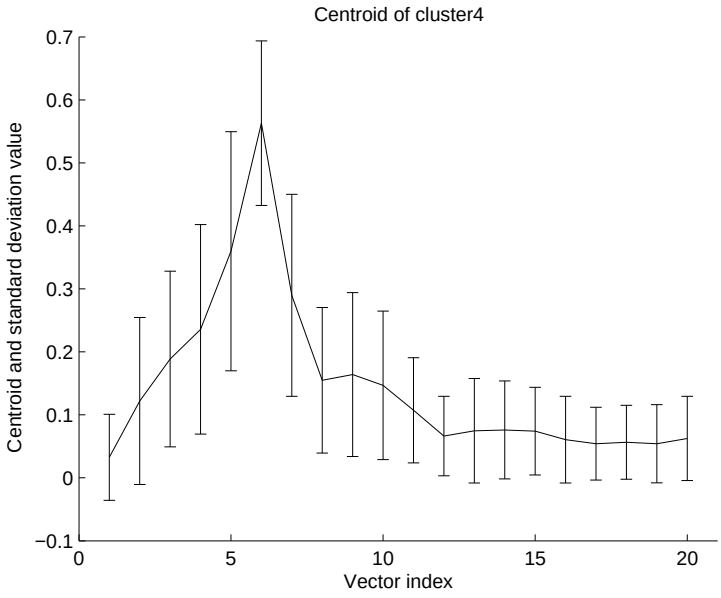

Figure 4.11: Centroid of Cluster Number 4.

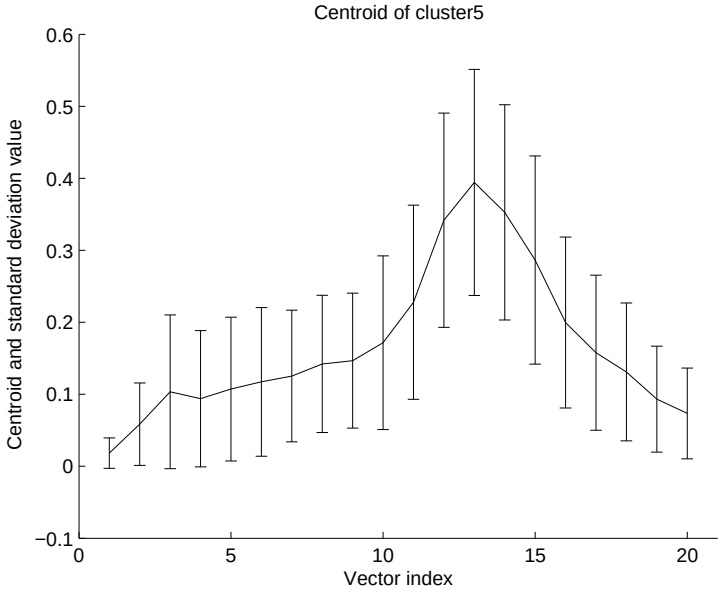

Figure 4.12: Centroid of cluster number 5.

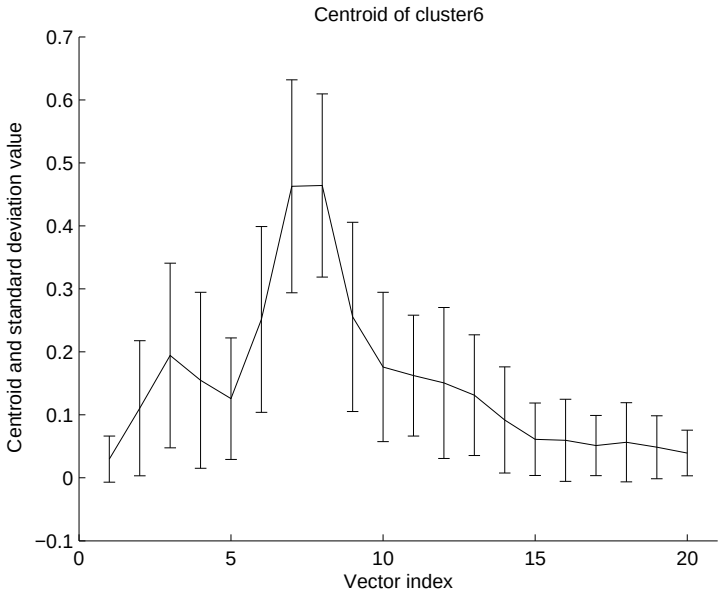

Figure 4.13: Centroid of cluster number 6.

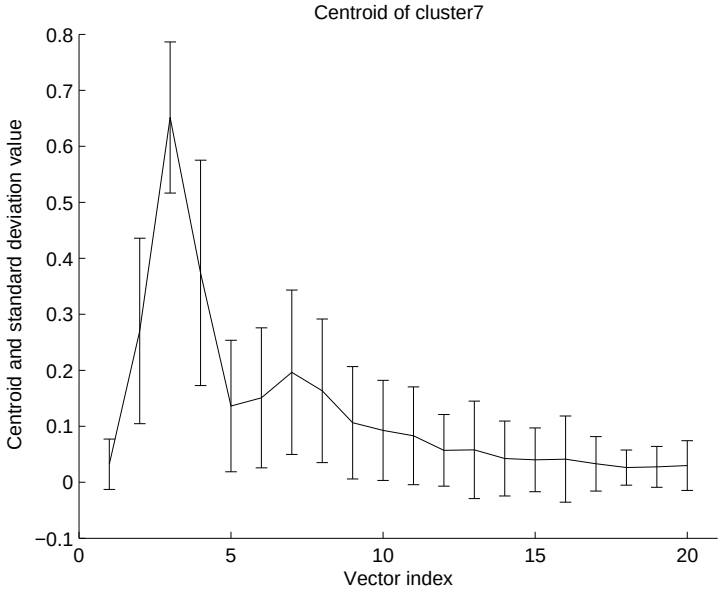

Figure 4.14: Centroid of cluster number 7.

The value of these 7 centroids, whose size is 20×1, will be taken out and used as the weight vector $\mathbf{W}$ to generate the confidence value. The concept of the confidence value is introduced in Chapter 3:

$$Conf = log\Big(\mathbf{W}^T\mathbf{Q} + 1\Big), \tag{4.4}$$

where $\mathbf{Q}$ represents the 20 spectral feature vector. In the original version, the elements in the weight vector $\mathbf{W}$ are found by experimentation to detect the weak mines. Right now, the elements are the value of the 7 centroids found by the clustering analysis. Since there are 7 centroids, so there would be 7 weight vectors $\mathbf{W}_{i=1,2...,7}$, respectively, there will be 7 confidence values $Conf_{i=1,2...,7}$, however, we will not use all of them. The operation is to sort all these confidence values and then take the biggest confidence value out for the detection use.

After taking out the biggest confidence value of the land mine signals from the training data set, detection will be performed in the testing data set using that confidence value. The aim is to see the effect of this confidence value obtained by the clustering technique, i.e, if there is any improvement on the detection results. Before analyzing the detection performance and showing the detection result figures, there are two fundamental parameters need to be explained: The first one is "PD", which means the probability of detection. The second one is "FAR", which means the probability of false alarm. Usually in the detection, increasing "PD" always requires accepting a higher "FAR" as well. So the performance of detection is decided by the combination of these two parameters. We will see in the detection performance figures that in the detection using the clustering method, when the "FAR" remains at the same point, the "PD" can reach a higher level than the detection without implementing the clustering method.

4.6 Results for Clustering Technique

First let us take a look at the confidence value generated by the K-Means clustering method:

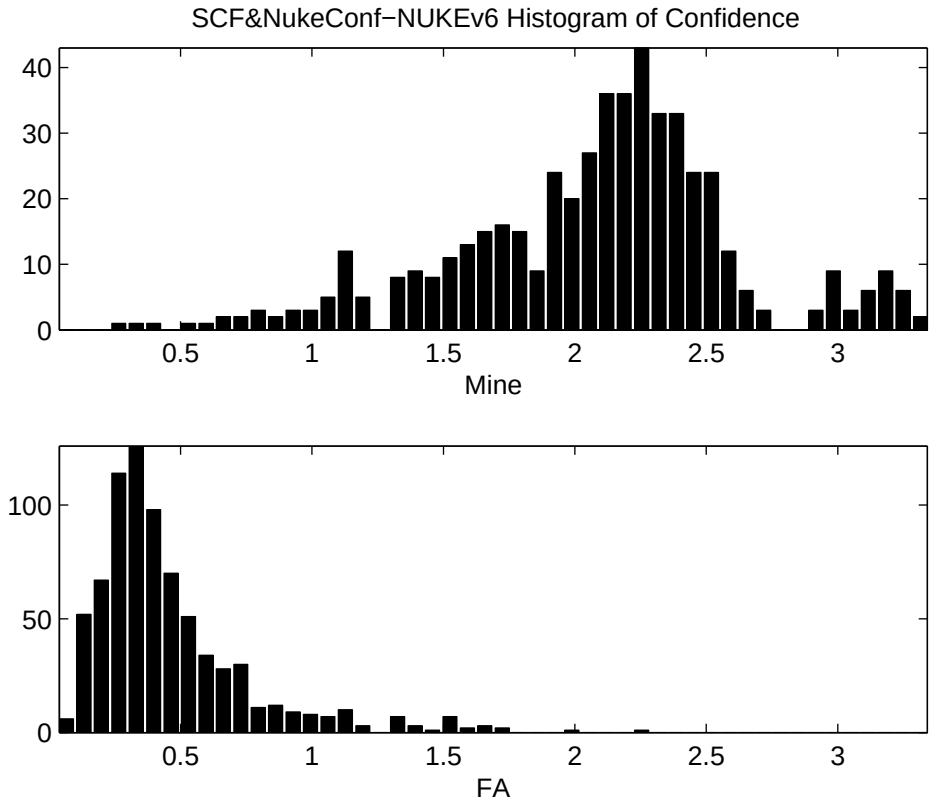

Figure 4.15: Confidence generated by K-Means clustering method.

The confidence value of mine and non-mine targets are displayed in the histogram. So here we shall explain what is histogram. A histogram is a graphical display of tabulated frequencies. It is the graphical version of a table that shows what proportion of cases fall into each of several or many specified categories. In the histogram it is the height of the bar that denotes the value, and the categories have uniform width, also the categories are usually specified as non-overlapping intervals of some variable. Take the bar above the number 2 in "mine" confidence part for example, it means that the confidence value of

'2' for mine appears around 20 times. Also from the histogram we notice that the confidence value of non-mine targets concentrate in the range from 0 to 1, when those of land mines are concentrated in the range bigger than 1.

Let us take a look at the detection performance which is a curve decided by the "PD" and "FAR", which is often called the receiver operating characteristics.

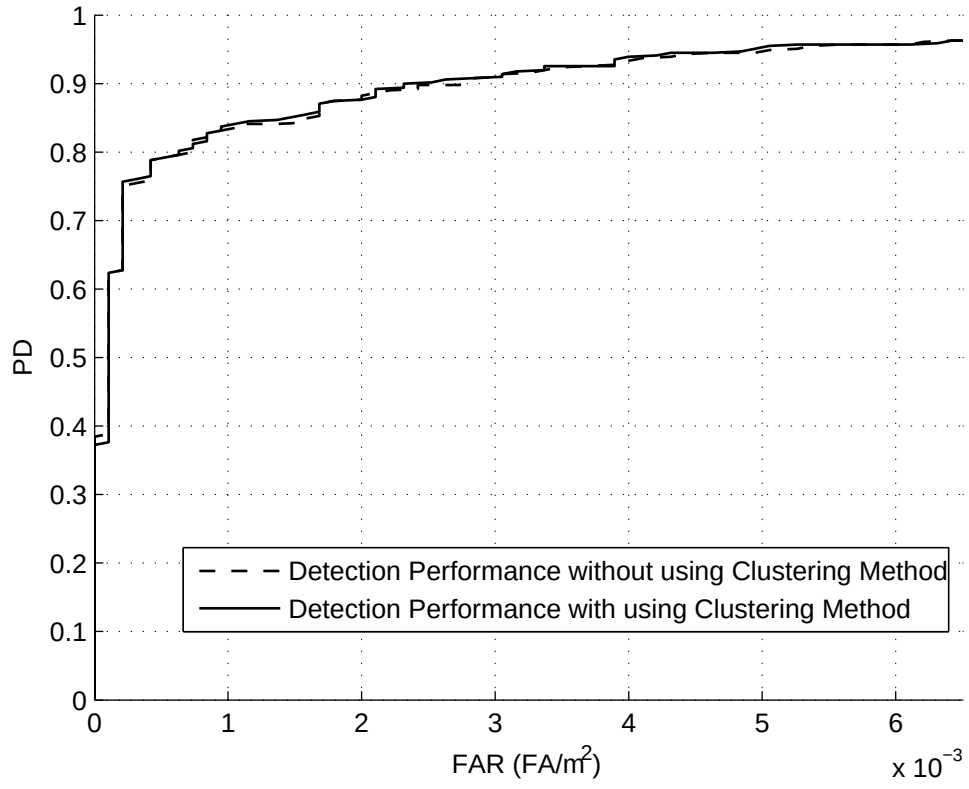

Figure 4.16: Comparison of detection performance w./w.o. clustering method.

From the figure above, even though we can tell there are some improvement, it is really hard to tell. So we need to "zoom in" to the figure. To be more specific, we will focus on the range where the "PD" are from 70% to 100%. In this range, we will pay attention to three points, which are "PD" at 85%, 90%, 95% respectively. The reason why this range is important is that only the 'high' probability of detection of land mines will make sense in the real world when doing the detection. There are no difference when you say this target has 1% probability to be land mine or 45% probability to be land mine, since once the detection is wrong, the loss is unbearable. So no matter how

low the "FAR" is, when the "PD" is low, the result will be useless. So the improvement is the range that "PD" from 70% to 100% is where we are really interested in.

The "zoom in" detection performance result is:

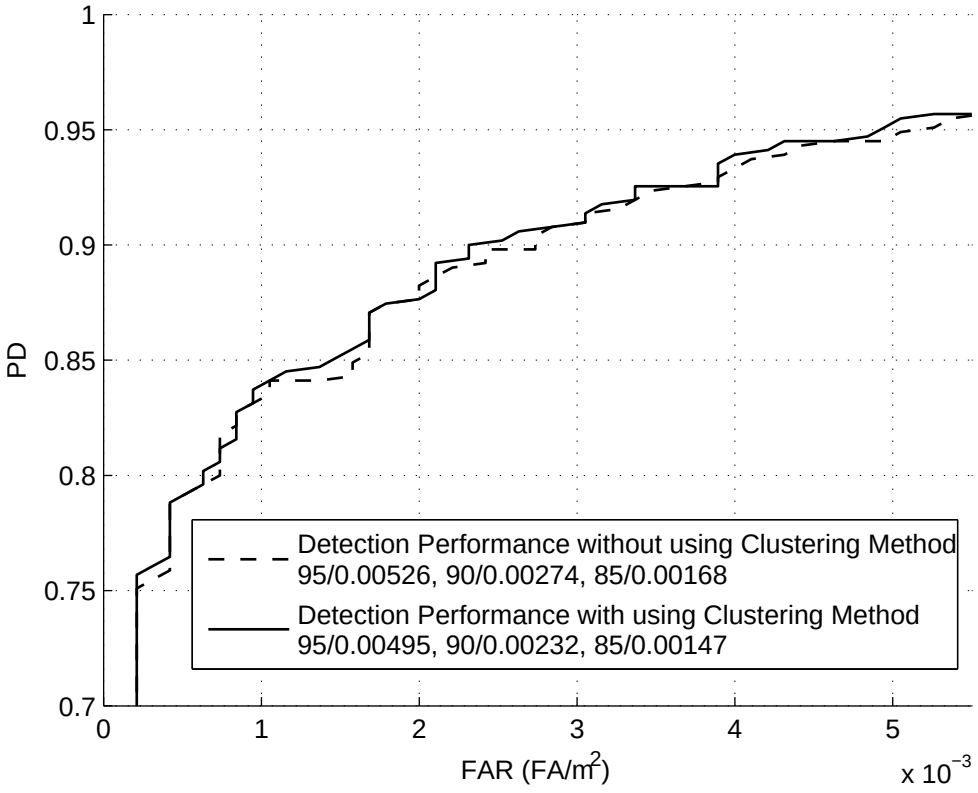

Figure 4.17: Comparison of "zoom in" detection performance w./w.o. clustering method.

From the figure above, the improvement in detection performance can be seen more easily. Also the probability of false alarm is given out when the probability of detection reaches 85%, 90%, 95% respectively.

4.7 Conclusion

After dealing with the energy density spectrum and the spectral feature vectors of the signals from the training data set, a approach aims to improve the detection performance of land mine based on the spectral feature vectors is proposed and implemented. The approach involve using clustering method whose goal is to find out if there is any hidden pattern within the data. The proximity measure of these data is defined and which kind of clustering methods should be used for clustering is investigated. Finally the K-Means algorithm is implemented to cluster the data. The quality of the clusters has been studied and later used to decide the number of clusters the whole data set should be divided into. Also the centroids of the clusters is taken out and combined with spectral feature vectors to generate the confidence value, which is used in the detection applied in the testing data set to verify the effect of the clustering method.

When the detection performance of the testing data set comes out, attention is paid to the range where the "PD" values are in a high level. Some improvement caused by the clustering method can be seen in the range between "PD at 80%" to "PD at 95%", however, the improvement is not remarkable, the reason might be that the pattern extracted from the training data set is unable to generalize very well, the centroids in the clusters can not represents the real world data of the land mines very well. Further investigation is necessary to find the better pattern of the data and make use of it, such as using some other clustering algorithm other than K-Means.

Chapter 5

Subspace Detector Based on the Energy Density Spectrum

5.1 Introduction

In chapter 3, an approach using clustering method based on the spectral feature vectors, which are generated by the energy density spectra of the signals from the training data set, is proposed and investigated. When the investigation of the clustering approach is completed, some questions remains in this research field: After the signal processing of the detected signals, we have the energy density spectrum and the spectral feature vectors. Between them, the spectral feature vectors are used directly in the clustering method to improve the detection performance. However, the energy density spectrum has not been used directly yet, so is that possible to use them to attain some positive effect on the detection result? if so, how is the performance of this method comparing with the results from the clustering method? To answer all these questions, an approach named subspace detector is implemented in this chapter, instead of the spectral feature vectors. This approach is based on the energy density spectra themselves. Some spectrum analysis techniques such as eigendecomposition, independent component analysis and eigenspace separation transform are studied and used for the subspace method.

5.2 Subspace Detector

In the land mine detection project, after the signal processing, the energy density spectrum of each signal data from the ground penetrating radar is a vector whose size is $M \times 1$. In most cases, the energy density spectra is corresponding to the target signal plus noise. However, sometimes it is nothing but noise. In order to solve this problem, we set up the detection hypothesis below.

5.2.1 Detection Hypothesis and Detection Rule

The detection hypotheses are set up as below:

$$H_0 \quad \text{(signal absent)} \qquad : \mathbf{x}_i = \mathbf{w}_i \tag{5.1}$$

$$H_1 \quad \text{(signal present)} \qquad : \mathbf{x}_i = \mathbf{t}_i + \mathbf{w}_i \tag{5.2}$$

where $M \times 1$ vector $\mathbf{x}_i$ represents the energy density spectrum, $M \times 1$ vector $\mathbf{w}_i$ represents the noise component, and $M \times 1$ vector $\mathbf{t}_i$ represents the component of the target signal. Let us take a look at the vector $\mathbf{t}_i$, which lies in the subspace spanned by the columns of matrix $\mathbf{S} = [\ \mathbf{s}_1\ \mathbf{s}_2\ \mathbf{s}_3\ \cdots\ \mathbf{s}_K\ ]$, where $\mathbf{s}_i$ are called the basis vectors, each has a size $M \times 1$. That is:

$$\mathbf{t}_i = a_{i,1}\mathbf{s}_1 + a_{i,2}\mathbf{s}_2 + a_{i,3}\mathbf{s}_3 + \cdots + a_{i,K}\mathbf{s}_K = \mathbf{S}\,\mathbf{a}_i, \tag{5.3}$$

where $\mathbf{a}_i = [\ a_{i,1}\ a_{i,2}\ a_{i,3}\ \cdots\ a_{i,K}\ ]^T$ represents a vector of independent random variables.

Under the hypothesis H_1, if we assume that the noise $\mathbf{w}_i$ is Gaussian distribution, then we can attain the maximum likelihood estimation of vector $\mathbf{a}_i$ as

$$\hat{\mathbf{a}}_i = \left(\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{S}^T\mathbf{x}_i, \tag{5.4}$$

After getting the value of $\hat{\mathbf{a}}_i$, the detection hypothesis can be translated to:

$$H_0 \quad \text{(signal absents)} \quad : \mid \hat{\mathbf{a}}_i \mid^2 = 0 \tag{5.5}$$

$$H_1 \quad \text{(signal presents)} \quad : \mid \hat{\mathbf{a}}_i \mid^2 > 0 \tag{5.6}$$

It means that right now we can determine the existence of the signal by the estimated value of vector $\mathbf{a}_i$. Hence a simple detection rule can be established as: select H_1 if $\mid \hat{\mathbf{a}}_i \mid^2 > 0$, i.e, $\sum \hat{a}_{i,j}^2 > 0$. Usually, a threshold will be set up in order to make the decision, so the exact detection rule we used is:
select H_1 if $\mid \hat{\mathbf{a}}_i \mid^2 > TH$, i.e, $\sum \hat{a}_{i,j}^2 > TH$, where TH represents "threshold".

Once the detection hypothesis and the detection rule are attained, the crucial part for the problem at hand is how to attain the matrix $\mathbf{S}$. There are two ways to get the elements of $\mathbf{S}$, one is the independent component analysis method, the other is use the eigenspace separation transform. Both methods are investigated here.

5.2.2 Independent Component Analysis

Since $\mathbf{x}_i$ is a linear combination of $\mathbf{a}_i$ and $\mathbf{S}$, so independent component analysis (ICA) can be considered here to get the value of $\mathbf{S}$ [45]. ICA is a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals. When the independence assumption is correct, ICA separation of a mixed signal gives very good results. The statistical method finds the independent components by maximizing the statistical independence of the estimated components. Typical algorithms for ICA use centering, whitening and dimensionality reduction as preprocessing steps in order to simplify

and reduce the complexity of the problem for the actual iterative algorithm. Whitening and dimension reduction can be achieved with singular value decomposition. Whitening also ensures that all dimensions are treated equally a priori before the algorithm is implemented.

When the data is represented by the observed $M \times 1$ vector $\mathbf{x}_i = (\ x_1\ \ x_2\ \ x_3\ \ \cdots\ \ x_L\ )^T$, and $\mathbf{x}_i$ are generated as a sum of the independent components $\mathbf{s}_i$, where $i = 1, 2, \ldots, K$:

$$\mathbf{x}_i = a_{i,1}\mathbf{s}_1 + a_{i,2}\mathbf{s}_2 + a_{i,3}\mathbf{s}_3 + \cdots + a_{i,K}\mathbf{s}_K = \mathbf{S}\,\mathbf{a}_i, \tag{5.7}$$

weighted by the $K \times 1$ weight vector $\mathbf{a}_i$. If we assume that $\mathbf{x}_i$ and $\mathbf{a}_i$ are zero-mean, and totally we have L $\mathbf{x}_i$ vectors, then the generative formula (5.7) can be written as:

$$\mathbf{X}_{M\times L} = \mathbf{S}_{M\times K}\,\mathbf{A}_{K\times L}, \tag{5.8}$$

where $\mathbf{X}_{M\times L}$ represents the measured energy density spectrum data, $\mathbf{S}_{M\times K}$ denotes the matrix spanned by the subspaces and $\mathbf{A}_{K\times L}$ represents the mixing coefficient matrix. Given the matrix $\mathbf{X}_{M\times L}$, the task right now is to estimate both the mixing matrix $\mathbf{A}_{K\times L}$ and the sources $\mathbf{S}_{M\times K}$. This is done by adaptively calculating some $\mathbf{w}_i$ vectors and setting up a cost function which either maximizes the nongaussianity of the calculated vector $\mathbf{s}_i = \mathbf{x}_i\mathbf{w}_i^T$ or minimizes the mutual information. In other words, we want to find a matrix $\mathbf{W} = \mathbf{A}^{-1}$ such that $\mathbf{S} = \mathbf{X}\mathbf{W}$.

However, the ICA approach has some disadvantages such as: the processing of the ICA is iterative, which means the result cannot be attained directly; The convergence is not guaranteed so we might not get the result we want eventually. Even if the convergence is generated, it might be the local conver-

gence, so the result might not be the optimized result. So instead of the ICA approach, the eigenspace separation transform method will be implemented here to obtain the subspace.

5.2.3 Eigenspace Separation Transform

The eigenspace separation transform projects each energy density spectrum data into an orthogonal subspace and thereby reduces the dimensionality of the input vector. Also the eigenspace separation transform preserves and makes more accessible the information that is most relevant to a classifier: information that allows separation of data clusters. To understand how this transform works, let us first take a look at the form of the energy density spectrum data:

$$\underset{(M\times 1)}{\mathbf{x}_i} = \underset{(M\times K)}{\mathbf{S}} \quad \underset{(K\times 1)}{\mathbf{a}_i} \quad + \quad \underset{(M\times 1)}{\mathbf{w}_i}\ , \qquad i = 1,\, 2,\, 3,\, \ldots,\, \mathrm{L} \tag{5.9}$$

where

$\mathbf{w}_i$ represents the M-length noise vector at time i whose elements are independent random variables with zero mean and variance σ^2, it is assumed uncorrelated with the signal;

$\mathbf{S} = [\ \mathbf{s}_1\ \ \mathbf{s}_2\ \ \mathbf{s}_3\ \ \cdots\ \ \mathbf{s}_K\ ]$ is the transformation matrix;

$\mathbf{a}_i = [\ a_{i,1}\ \ a_{i,2}\ \ a_{i,3}\ \ \cdots\ \ a_{i,K}\ ]^T$;

Hence, in (5.9) we obtain L vectors $\mathbf{x}_i$, i = 1, ..., L . The objective is to estimate the elements of matrix $\mathbf{S}$, by observing only the received signal. Note that $K < M$, let us form the covariance matrix $\mathbf{R}$ of the observing signal $\mathbf{x}$ [46]:

$$\begin{aligned} \mathbf{R} &= E(\mathbf{x}\mathbf{x}^H) = E[(\mathbf{S}\mathbf{a} + \mathbf{w})(\mathbf{a}^H\mathbf{S}^H + \mathbf{w}^H)] \\ &= \mathbf{S}E(\mathbf{a}\mathbf{a}^H)\mathbf{S}^H + \sigma^2\mathbf{I} \end{aligned} \tag{5.10}$$

The last line follows because the noise is uncorrelated with the signal, thus forcing the cross-term to zero. In the last line of (5.10) we also used the fact that the covariance matrix of the noise contribution (second term) is $\sigma^2\mathbf{I}$. This follows because the elements of the noise vector $\mathbf{w}$ are independent with equal power. The first term of (5.10) we call $\mathbf{R}_o$, which is the contribution to the covariance matrix due only to the signal component. Then let us look at the structure of $\mathbf{R}_o$:

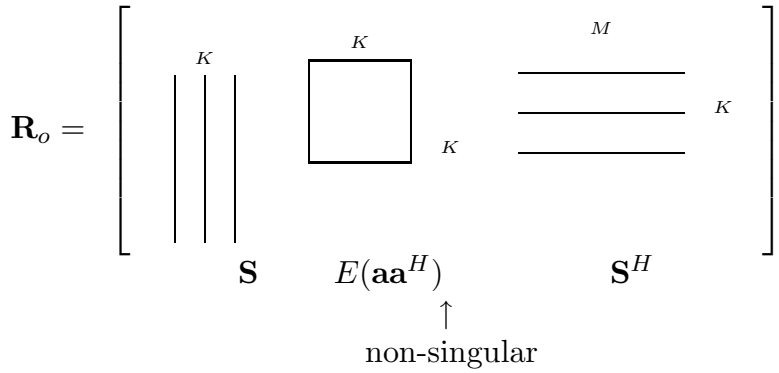

Figure 5.1: Structure of $\mathbf{R}_o$.

From the structure, we may conclude that $\mathbf{R}_o$ is rank K. This may be seen as follows. Let us define $\mathbf{A} \triangleq E(\mathbf{a}\mathbf{a}^H)$ and $\mathbf{B} \triangleq \mathbf{A}\mathbf{S}^H$. $\mathbf{S}$ is full rank (rank K), and since the $\mathbf{a}_i$ are independent, $\mathbf{A}$ is full rank (rank K). Therefore the matrix $\mathbf{B} \in \mathbf{R}^{K\times M}$ is of full rank K. Then, $\mathbf{R}_o = \mathbf{S}\mathbf{B}$. From this last relation, we can see that the ith, $i = 1, \ldots, M$ column of $\mathbf{R}_o$ is a linear combination of the K columns of $\mathbf{S}$, whose coefficients are the ith column of $\mathbf{B}$. Because $\mathbf{B}$ is full rank, K linearly independent linear combinations of the K columns of $\mathbf{S}$ are used to form $\mathbf{R}_o$. Thus $\mathbf{R}_o$ is rank K. Because $K < M$, $\mathbf{R}_o$ is rank deficient.

Let us now investigate the eigendecomposition on $\mathbf{R}_o$, where λ_k are the eigenvalues of $\mathbf{R}_o$:

$$\mathbf{R}_o = \mathbf{V}\Lambda\mathbf{V}^H \tag{5.11}$$

or

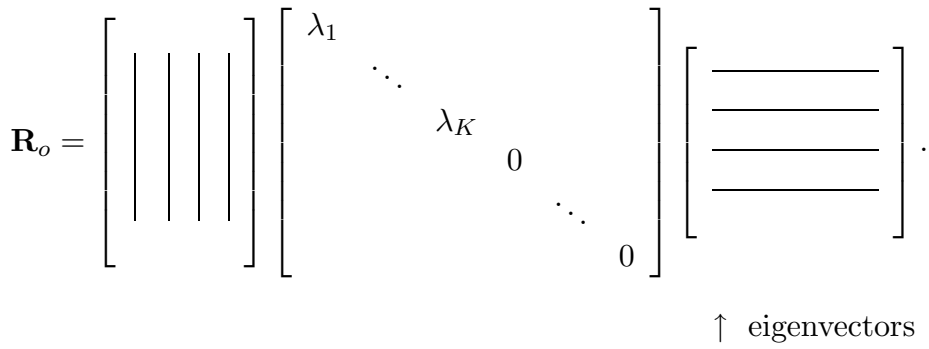

Figure 5.2: Eigendecomposition of $\mathbf{R}_o$.

Because $\mathbf{R}_o \in \mathbf{R}^{M \times M}$ is rank K, it has K non-zero eigenvalues and $M - K$ zero eigenvalues. We enumerate the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_K$ as those associated with the largest K eigenvalues, and $\mathbf{v}_{K+1}, \ldots, \mathbf{v}_M$ as those associated with the zero eigenvectors. From the definition of an eigenvector, we have:

$$\mathbf{R}_o \mathbf{v}_i = 0 \qquad i = K+1, \ldots, M \tag{5.12}$$

$$\text{or} \quad \mathbf{S}\mathbf{A}\mathbf{S}^H \mathbf{v}_i = 0, \qquad i = K+1, \ldots, M. \tag{5.13}$$

Now let us defined the so-called signal subspace $\mathbf{S}_S$ as:

$$\mathbf{S}_S = \text{span}[\mathbf{v}_1, \ldots, \mathbf{v}_K] \tag{5.14}$$

and the noise subspace $\mathbf{S}_N$ as:

$$\mathbf{S}_N = \text{span}[\mathbf{v}_K + 1, \ldots, \mathbf{v}_M] \tag{5.15}$$

From our discussion above, all columns of $\mathbf{R}_o$ are linear combination of the columns of $\mathbf{S}$. Therefore

$$\text{span}[\mathbf{R}_o] = \text{span}[\mathbf{S}]. \tag{5.16}$$

But also it is easy to verify that

$$\text{span}[\mathbf{R}_o] \in \mathbf{S}_S. \tag{5.17}$$

Comparing (5.16) and (5.17), we see that any received signal vector $\mathbf{x}$, in the absence of noise, is a linear combination of the columns of $\mathbf{S}$. Thus, any noise-free signal resides completely in $\mathbf{S}_S$. This is the origin of the term "signal subspace". Further, any component of the received signal residing in $\mathbf{S}_N$ must be entirely due to the noise. This is the origin of the term "noise subspace". Note that the signal and noise subspaces are orthogonal complement subspaces of each other.

In conclusion, the calculation of the transformation matrix [2] $\mathbf{S}$ can be proceed as follows:

1) Compute the $M \times M$ matrix:

$$\hat{\mathbf{M}} = \frac{1}{N_1}\sum_{i=1}^{N_1}\mathbf{X}_{1i}\mathbf{X}_{1i}^T - \frac{1}{N_2}\sum_{i=1}^{N_2}\mathbf{X}_{2i}\mathbf{X}_{2i}^T, \tag{5.18}$$

where $\mathbf{X}_{1i}$ is the energy density spectrum of the mine pattern i, N_1 is the number of mine patterns in the training set, $\mathbf{X}_{2i}$ is the energy density spectrum of the background pattern i and N_2 is the number of background patterns in the training set.

2) Calculate the eigenvalues of the matrix $\hat{\mathbf{M}}$, which are λ_i, $i = 1, 2, \cdots, M$.

3) Sort the eigenvalues λ_i, $i = 1, 2, \cdots, M$, then keep the biggest D values for use in next step.

4) Calculate all K orthonormal eigenvectors that are associated with the biggest D eigenvalues. Used these eigenvectors as the columns of the $M \times K$ transformation matrix $\mathbf{S}$.

After the transformation matrix $\mathbf{S}$ is constructed, $\hat{\mathbf{a}}_i = \left(\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{S}^T\mathbf{x}_i$ is then computed for each pattern $\mathbf{X}_i$ in the training and testing data sets. If the value of $\hat{\mathbf{a}}_i$ satisfy the hypothesis H_1, i.e, $|\,\hat{\mathbf{a}}_i|^2 > TH$, where TH means the threshold that decided by operation based on the pre-generated confidence value of land mine.

To take advantage of the detection value generated previously, we shall combine the information from $|\hat{\mathbf{a}}_i|$ with the original confidence value as new confidence value

$$Conf_{new} = Conf_{original} log\Big(1 + \mathbf{b}_i\Big), \tag{5.19}$$

where

$$\mathbf{b}_i = |\ \hat{\mathbf{a}}_i|^{2wgh} \tag{5.20}$$

In equation (5.20), wgh is found through training to achieve better results.

5.3 Results for Subspace Detector

When the weight vector **W** is obtained using the training data, detection will be performed in the testing data set using the new confidence values generated. The purpose of the detection on testing data set is to verify the effect of the subspace detector technique, to see if there is any improvement on the detection results when this technique is implemented. The performance of the detection in the testing data set will be showed in next section.

First let us take a look at the confidence values attained without using the subspace detector:

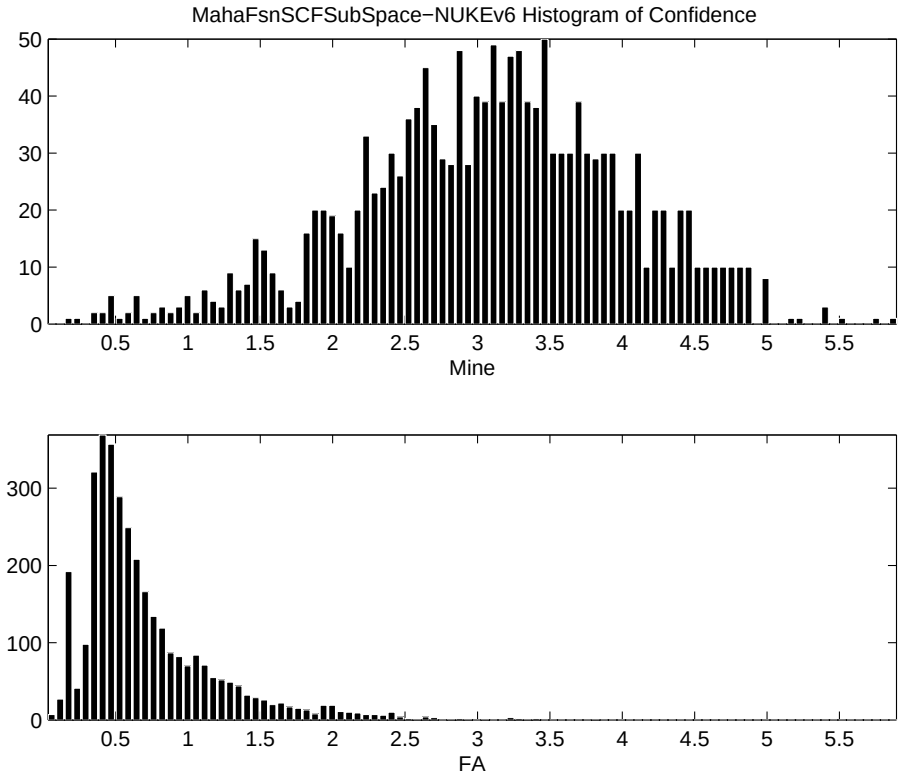

Figure 5.3: Confidence value obtained without subspace detector.

From the figure5.3, we can see that the confidence value of land mine spread out at the range from 0.2 to 5.5, and the confidence of false alarm lies in the range from 0.1 to 2.4. In the overlap area of these two categories of confidence value, it is hard to distinguish the difference between land mine and false alarm just according to the confidence values.

Then let us look at the figure of confidence value attained when the subspace detector is implemented:

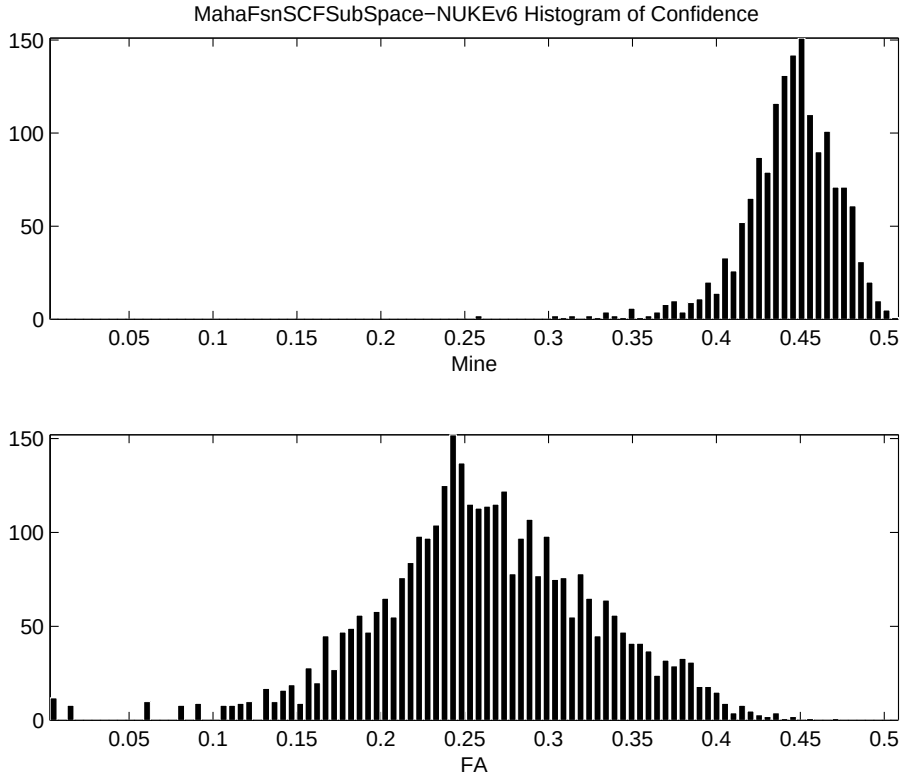

Figure 5.4: Confidence value obtained with subspace detector.

when the subspace detector is implemented, the confidence value of land mine are majorally in the range from 0.35 to 0.5, and the confidence value of false alarm spread out from 0.05 to 0.41. Also in the overlap area of these two categories of confidence value, now it is much easier to distinguish the difference between land mine and false alarm.

Besides the confidence value, the detection performance also will be compared with the ROC curve decided by the parameters "PD", which represents "Probability of Detection", and "FAR", which means "Probability of False Alarm", these two curves are showed in next figure:

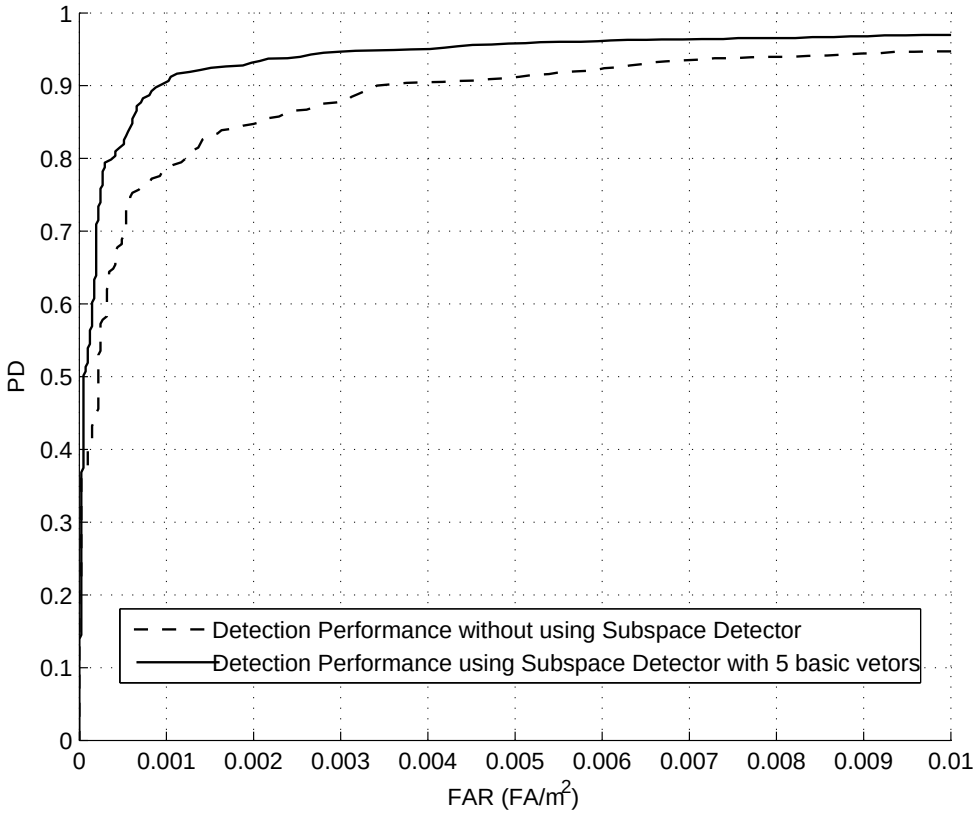

Figure 5.5: Comparison of detection performance w./w.o. subspace detector.

From the figure, we can easily tell that when the subspace detector is implemented, it does have very good effect on the detection performance. Especially at the range of "PD" from 0.85 to 0.95, which are the most important part to verify the detection performance. The improvement is very remarkable, and the "zoom in" results are showed in the figure(5.6) in next page.

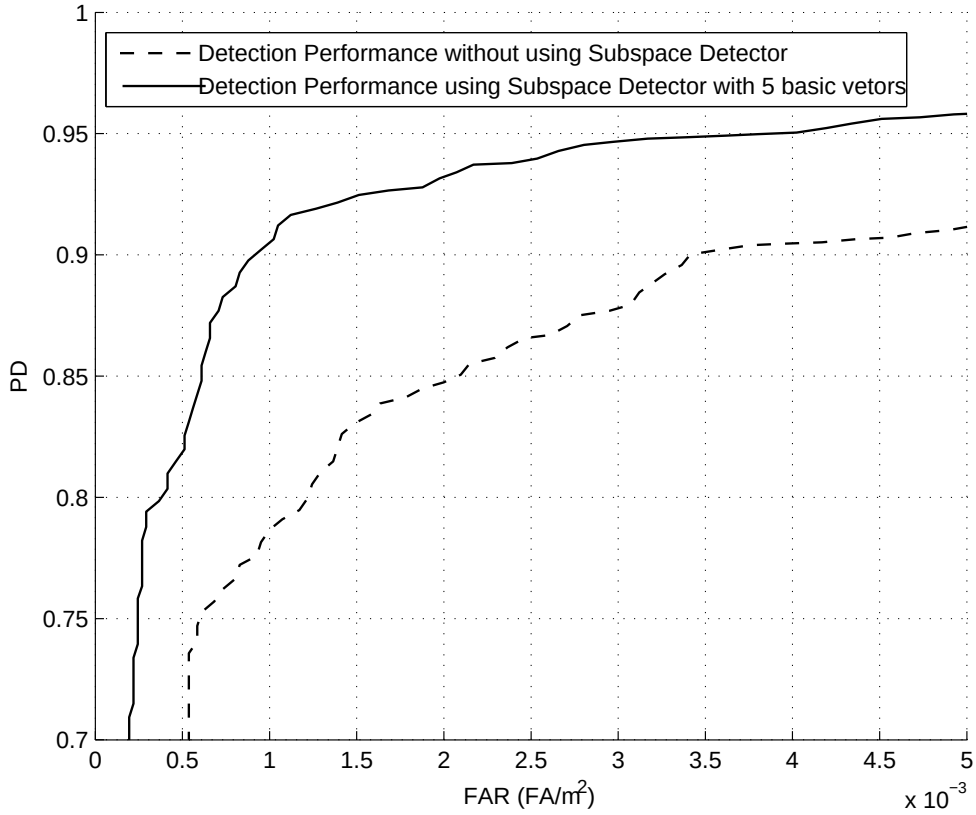

Figure 5.6: Comparison of "zoom in" detection performance w./w.o. subspace detector.

5.4 Effect of Varying the Parameters in Subspace Detector

Now we know that the subspace detector can improve the detection result remarkably. However, there is still one problem we can perform some further investigation: the setting of the parameters we used in the subspace detector.

In the subspace detector, there are two major parameters:

1. The number of basic vectors K, i.e, the number of vectors $\mathbf{s}_i$ that form the transformation matrix $\mathbf{S}$;

2. The length of energy density spectrum M, i.e, the size of vector $\mathbf{x}_i$ which characterizes the alarm i.

It is believed that when the parameters K and M vary, the performance of detection will be vary according to the change, so some investigations are performed to see what the value of parameters K and M should be in order to optimize the detection results.

5.4.1 Effect of Varying the Number of Basic Vectors Used in Subspace Detector

First of all, the parameter K, i.e the number of basic vector used in the subspace detector, will be changed to investigate the effect that it might have on the detection result. To investigate, we will focus on three important points at the ROC curve, they are: "FAR at PD = 85%", "FAR at PD = 90%", "FAR at PD = 95%". We will use the "FAR" values in these three points to figure out what is the number of basic vectors that should be used to optimize the detection result. When the number of basic vector K varies, the length of energy density spectrum M will be fixed to 120. The value of "FAR" at "PD = 85%", "PD = 90%", "PD = 95%" will be showed in the table below.

Number of Basic Vectors used	FAR at PD = 95% (FP95)	FAR at PD =90% (FP90)	FAR at PD = 85% (FP85)	PD at FAR 0 (%)	PD at FAR 0.00007 (%)	PD at FAR 0.0007 (%)
5	0.004023	0.000951	0.000610	14	50	87
6	0.003950	0.000927	0.000634	14	50	87
8	0.004096	0.000902	0.000585	13	47	88
10	0.003877	0.000902	0.000585	13	47	88
11	0.003852	0.000878	0.000585	13	47	88
12	0.003901	0.000853	0.000536	12	46	88
13	0.003877	0.000902	0.000585	13	47	88
14	0.003999	0.000902	0.000585	11	44	88
16	0.003999	0.000853	0.000585	16	44	88
17	0.003974	0.000878	0.000585	11	43	88
18	0.003974	0.000878	0.000585	11	43	88
20	0.003974	0.000878	0.000585	11	43	88

Figure 5.7: List of"FAR" values with different numbers of basic vector

From the table above, we can see that: when $K = 12$, "FAR" has the lowest value 0.000536 at "PD = 85%"; when $K = 12$ and $K = 14$, "FAR" has the lowest value 0.000853 at "PD = 90%"; when $K = 11$, "FAR" has the lowest value 0.003852 at "PD = 95%". In fact, the information from Table 5.7 just simply indicates the lowest points are in some local area, and it is not sufficient to draw a conclusion that which value of K will optimize the detection result. So another analysis method is need.

Since among the three points, "FAR at PD = 90%" is more important than the other two points "FAR at PD = 85%" and "FAR at PD = 95%", so a weighted sum of the values of "FAR" at these three points will be used here. A weight factor 0.5 will be assigned to "FAR at PD = 90%", and a weight factor 0.25 will be assigned to both "FAR at PD = 85%" and "FAR at PD = 95%". So the weighted sum of these three "FAR" values is:

$$FAR_f = 0.25 \times FAR_{PD85\%} + 0.5 \times FAR_{PD90\%} + 0.25 \times FAR_{PD95\%}, \quad (5.21)$$

Because this combined value of "FAR" has considered all the "FAR" values in these three points, it should give us better judgement on what number of basic vector should be used to obtain the best detection result. The fusion value is shown using the figure below:

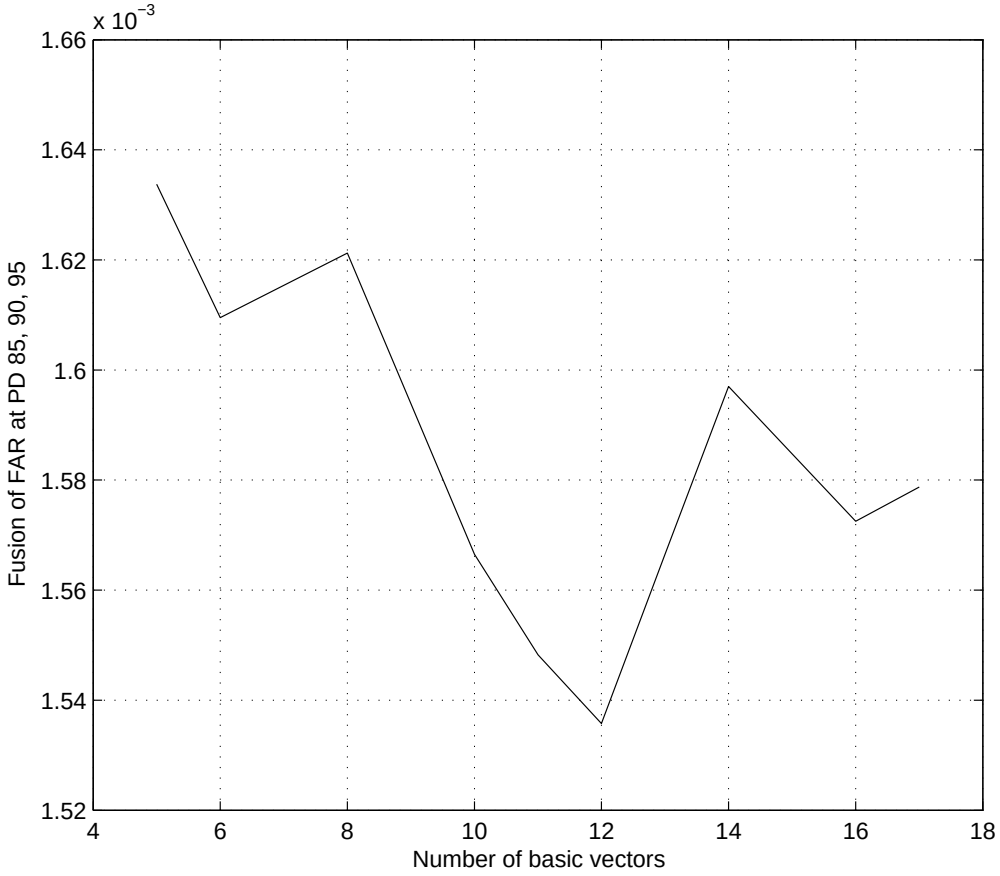

Figure 5.8: Fusion value of "FAR" at different numbers of basic vector

From figure 5.8, we can easily have the conclusion that when the number of basic vector used is 12, i.e $K = 12$, the combined value of "FAR" has the lowest value, which can lead to the conclusion that using 12 basic vectors can maximize the correct detection result.

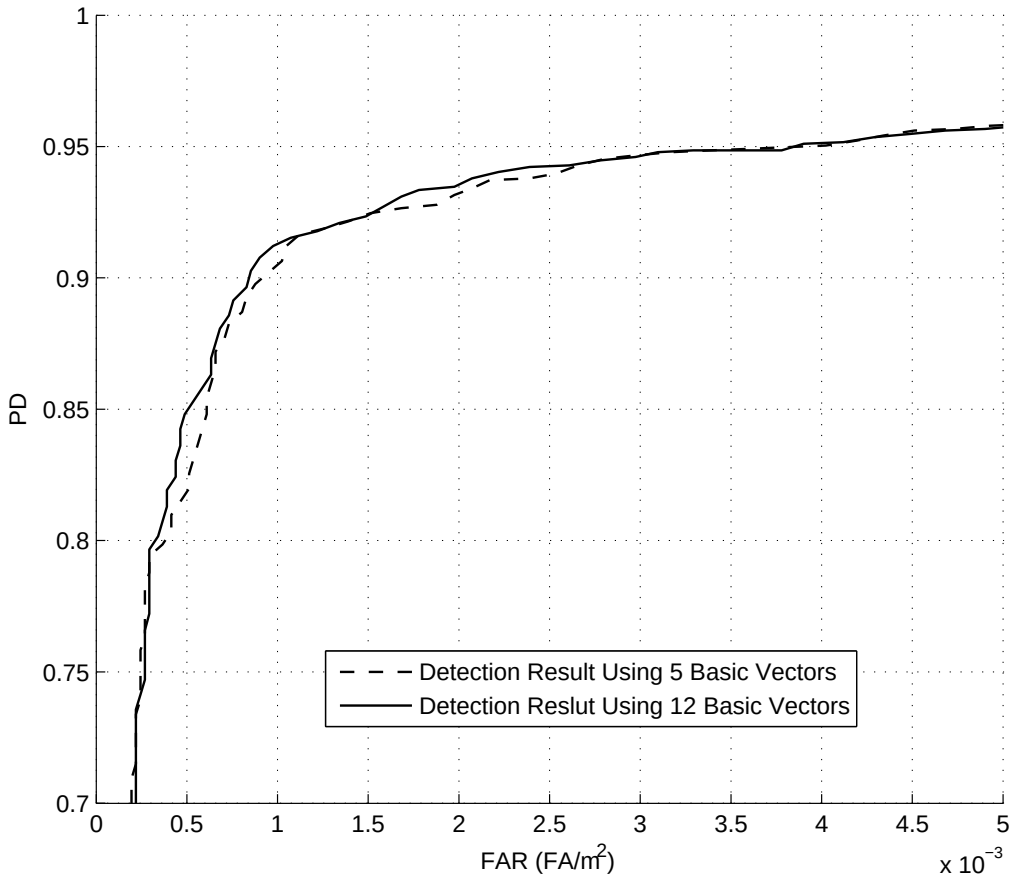

Figure 5.9: Comparison of detection result with different numbers of basic vector

5.4.2 Effect of Varying the Length of Energy Density Spectrum Used in Subspace Detector

After the number of basic vectors K is investigated and decided, we will move on to investigate the length of energy density spectrum M. Just as what we did in investigating K, we will vary the value of M, in the meanwhile set the value of K as 12, and then list the value of "FAR" at "PD = 85%", "PD = 90%", "PD = 95%" in the table below:

Length of EDS used	FAR at PD 95 (FP95)	FAR at PD 90 (FP90)	FAR at PD 85 (FP85)	PD at FAR 0 (%)	PD at FAR 0.00007 (%)	PD at FAR 0.0007 (%)
60	0.003682	0.000951	0.000610	17	47	86
70	0.003755	0.000951	0.000610	17	48	88
80	0.003804	0.000951	0.000610	16	47	88
100	0.004316	0.000975	0.000561	15	47	88
110	0.003950	0.000878	0.000585	14	46	88
120	0.003901	0.000853	0.000536	12	46	88
130	0.003999	0.000902	0.000634	13	43	87
140	0.003828	0.000927	0.000536	11	43	87
160	0.003950	0.000902	0.000561	12	44	87
170	0.003828	0.000902	0.000536	11	43	87
180	0.003901	0.000902	0.000561	11	43	87

Figure 5.10: List of "FAR" values with different length of energy density spectrum

From the table above, we can see that: when $M = 120$, $M = 140$, $M = 170$ "FAR" has the lowest value 0.000536 at "PD = 85%"; when $M = 120$, "FAR" has the lowest value 0.000853 at "PD = 90%"; when $M = 60$, "FAR" has the lowest value 0.003682 at "PD = 95%".

Also, a weighted sum of values of "FAR" at these three points using equation (5.23), will be used to decide what length of energy density spectrum should be used to attained the best detection result. The weighted "FAR" value is shown in the figure below:

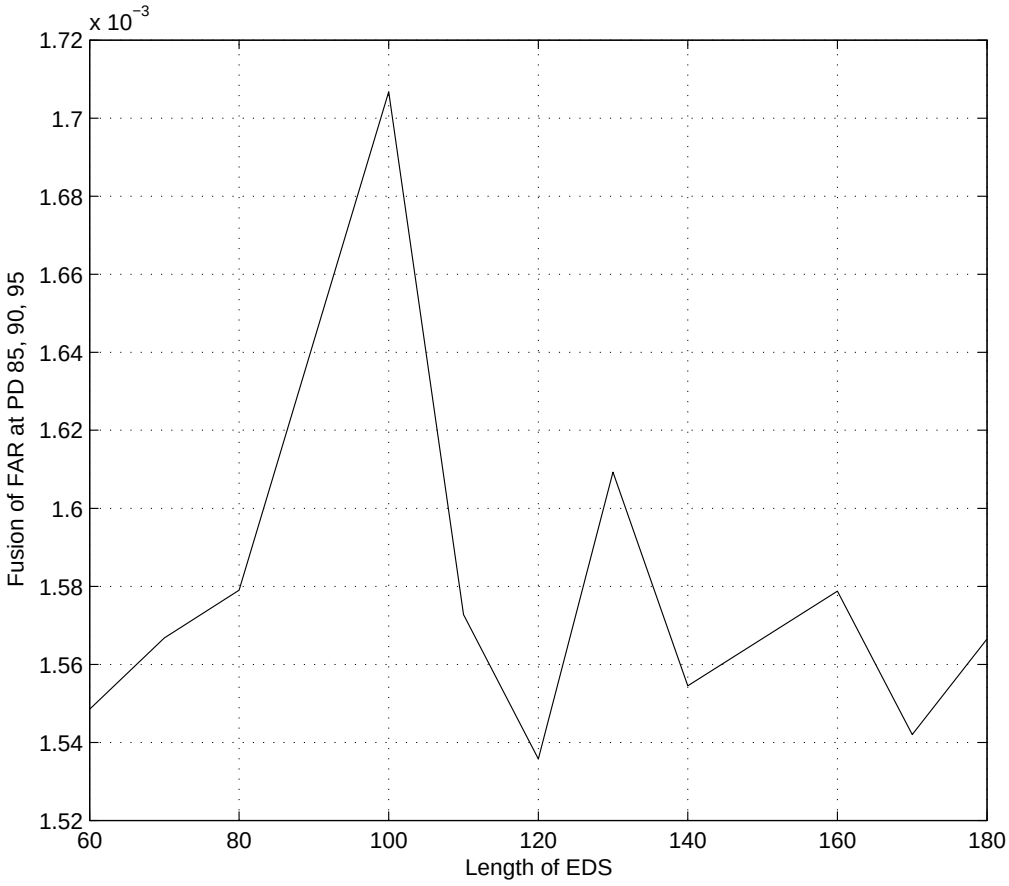

Figure 5.11: Fusion value of "FAR" at different length of energy density spectrum

From figure 5.11, we can easily conclude that when the length of energy density spectrum used is 120, i.e $M = 120$, the fusion value of "FAR" has the lowest value, which can lead to the conclusion that using 120 energy density spectrum can maximize the probability of detection result. The following figure shows the comparative results generated by setting $M = 120$ (solid line) and $M = 100$ (dashed line)

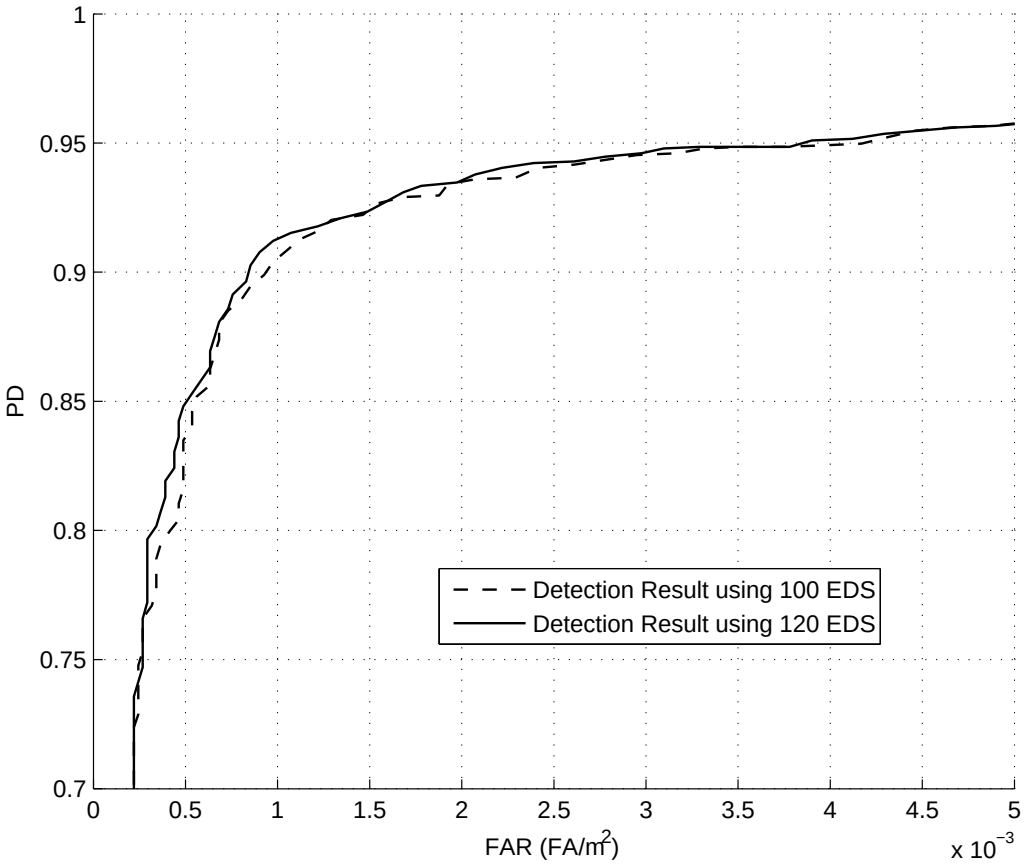

Figure 5.12: Comparison of detection result with different length of energy density spectrum

5.5 Conclusion

After using clustering technique to improve the detection performance, a problem remains for further research: since the clustering method are based on the spectral feature vectors, the energy density spectrum(EDS) has not been used to improve the detection performance directly. So the possibility of improving the detection result using these energy density spectrum are investigated. The subspace detector based on the EDS are then proposed to see if it can have good effect on the performance: Two hypotheses for detection are set. The maximum likelihood estimation of the subspace coefficient vector $\mathbf{a}$ are obtained. Independent component analysis technique and eigenspace separation transform technique are investigated. Finally eigenspace separation transform and eigenvalue decomposition are used on the training data to obtain the value of the subspace basis matrix $\mathbf{S}$. New confidence value are generated on the testing data set to verify the effect of the subspace detector.

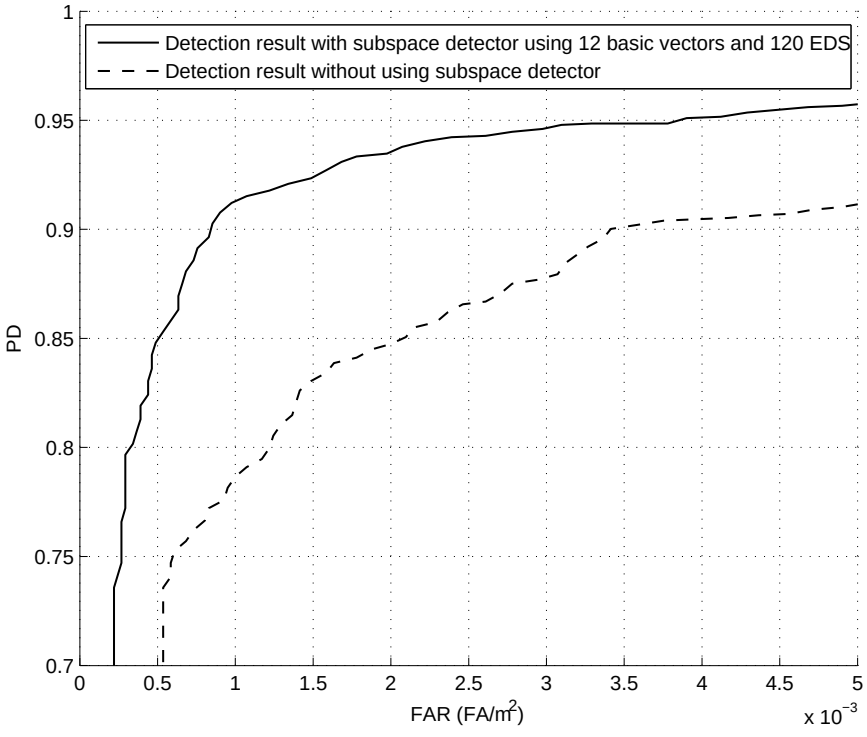

Figure 5.13: Improvement of detection result using subspace detector

From result of detection on the testing data set comes out, it is easy to draw the conclusion that the subspace detector has very good effect on the detection result, the improvement on the detection performance is very remarkable. Also, some investigation are done on the two parameters of the subspace detector: number of basis vectors and length of energy density spectrum. Eventually the investigations imply that with 12 basic vectors and 120 energy density spectrum length, we can achieved the best detection result.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, we have presented two advanced feature based techniques for land mine detection using vehicle-held GPR system. The contribution of this thesis involves investigating the clustering method based on the spectral feature vectors and the subspace detector technique based on the energy density spectrum, both techniques are aimed to improve the detection result and reduce the false alarm rates for the land mine detection. We summarize the contributions for each of the two techniques as follows:

In chapter 4, we investigate the clustering method. This method is based on the spectral feature vectors, which is formed by the energy density spectrum of GPR signals. The purpose of this method is trying to find the "hidden pattern" among the GPR signals, and use these patterns to improve the detection result. After the proximity measure is defined, we do some study to decide which clustering algorithms should be chosen to cluster the data set. As a result, K-Means algorithm is implemented to the data set, and the centroids of the resulting clusters are combined with the spectral feature vectors to generate the confidence value of land mines and objects that will cause the false alarm. Finally the confidence value is used in the detection applied in the

testing data set to verify the effect of the clustering method. From the evaluation result, we have a conclusion that the clustering method does improve the detection result and reduce the false alarm rates. However, the improvement is not remarkable. It might be because the centroids in the clusters cannot represent the real world data of the land mines very well so the patterns within the data set is not generalized well.

In chapter 5, the subspace detector technique is proposed and investigated. This technique is based on the energy density spectrum of the GPR signals. The energy density spectrum of a GPR signal can be considered to lie in the subspace spanned by some basis vectors $\mathbf{s}_i$ and some coefficients $a_i, i = 1, 2, \ldots, m$. Through the maximum likelihood estimation, the value of the coefficient vector can be obtained using the subspace basis matrix formed by the basis vectors and the energy density spectrum. Later, eigendecomposition and eigenspace separation transform are used to obtain the subspace basis matrix and eventually obtain the information of the coefficient vector. Then, the coefficient vector is combined with the previous detection confidence value to generate the new confidence value, and finally the new confidence value is used in the testing data set to verify the performance of the subspace detector technique. Also the effects of varying the total number of basis vectors used and the length of energy density spectrum used in the subspace detector on the detection performance are evaluated. From the testing result, we draw the conclusion that the subspace detector has very good effect on the detection result. It reduce the false alarm rates sharply, the improvement on the detection performance is very remarkable.

6.2 Future Work

Some directions of the future work are as follows: We might be able to improve the advanced feature based technique by considering the possibility of using the size and shape of the land mines. Also, other clustering algorithms rather than K-Means algorithm may be used in the clustering method to see if they can discover better patterns of the GPR signals of land mines in order to improve the detection results.

Reference

[1] Rob Siegel, "Land Mine Detection," *IEEE Instrumentation & Measurement Magazine,* Vol. 5, No. 4, Dec. 2002.

[2] Gregory L. Plett, Takeshi Doi, Don Torrieri, "Mine Detection Using Scattering Parameters and an Artificial Neural Network," *IEEE Transactions on Neural Networks,* Vol. 8, No. 6, Nov. 1997.

[3] K.C.Ho, P.D. Gader, J.N. Wilson, "Improving Landmine Detection Using Frequency Domain Features from Ground Penetrating Radar," *Proc. of IEEE International Symposium on Geoscience and Remote Sensing (IGARSS'04)*, 2004.

[4] D. A. Bloom, "Land Mines Keep Wars from Ever Coming To an End," *The Christian Science Monitor,* Vol. 87, No. 57, Feb. 1995.

[5] D. Webster, "One Leg, One Life at a Time," *The New York Times Magazine,* Jan. 23, 1994.

[6] "Current and Emerging Technologies for Use in a Hand-Held Mine Detector," Moody and LeVasseur, 19 June 2000, The Royal Military College of Canada, Dept. Appl. Military Science, p. 58.

[7] "Detection of Landmines by Quadrople Resonance at the Naval Research Laboratory (NRL)," press release, Feb. 22, 2000.

[8] J.M. Sabatier and N. Xiang, "Laser-Doppler Vibrometer-Based Anti-Personnel Mine Detection," in *Proc. Geoscience and Remote Sensing Symp.,* 2001.

[9] Ashley, S., "Searching for Landmines," *Mechanical Engineering,* Vol. 118, No. 4, p. 62, April 1956.

[10] Campbell, J.C. and Jacobs, A.M., "Detection of Buried Land Mines by Backscatter Imaging," *Nuclear Science and Engineering,* Vol.110, p. 417, 1992.

[11] K. C. Ho and Paul D. Gader, "A Linear Prediction Land Mine Detection Algorithm for Hand-Held Ground Penetrating Radar," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 40, No. 6, June 2002.

[12] P. D. Gader, B. N. Nelson, H. Frigui, G. Vaillette, and J. M. Keller, "Fuzzy Logic Detection of Land Mines with Ground Penetrating Radar," *Signal Processing, Spec. Issue Fuzzy Logic Signal Processing (Invited Paper)*, Vol. 80, No. 6, pp. 1069-1084, June 2000.

[13] P. D. Gader, B. N. Nelson, A. Koksal Hocaoglu, S. Auephanwiriyakul, and M. A. Khabou, "Neural versus Heuristic Development of Choquet Fuzzy Integral Fusion Algorithms for Land Mine Detection," in *NeuroFuzzy Pattern Recognition*, H. Bunke and A. Kandel, Eds, Singapore: World Scientific, 2000.

[14] P.D. Gader and M. Mystkowski, "Application of Hidden Markov Model to Detecting Land Mines with Ground Penetrating Radar," in *Proc. SPIE'99 Conf.*, Orlando, FL, 1999.

[15] P. D. Gader, M. Mystkowski, and Y. Zhao, "Application of Hidden Markov Model to Landmine Detection Using Ground Penetrating Radar," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 39, pp. 1231-1244, June 2001.

[16] K. C. Ho, P. D. Gader, "Fusion of Statistical Deviance and HMM GPR Algorithms for the Mine Hunter/Killer System," in *Proc. SPIE.01 Conf.,* Orlando, FL, 2001.

[17] P. D. Gader, M. Mystkowski, and Y. Zhao, "Adaptive Hidden Markov Models for Extended Land Mine Detection," in *Proc. SPIE.01 Conf.,* Orlando, FL, 2001.

[18] H. Brunzell, "Clutter Reduction and Object Detection in Surface Penetrating Radar," in *Proc. Radar.97* Conf., 1997 pp. 688-691.

[19] Dynatilaka, A. H. and B. A. Baertlein, "A Subspace Decomposition Technique to Improve GPR Imaging of Anti-Personnel Mines," *Proc. of SPIE, AeroSense 2000: Detection and Remediation Technologies for Mines and Minelike Targets,* Vol. 4038, 1008-1018, 2000.

[20] L. van Kempen and H. Sahli, "Signal Processing Techniques for Clutter Parameters Estimation and Clutter Removal in GPR data for Landmine Detection." *Statistical Signal Processing,* Aug 2001, pp. 158-161.

[21] L. van Kempen, H Sahli and J. Brooks and J. Cornelis, "New Results on Clutter Reduction and Parameter Estimation for Landmine Detection using GPR," *Eighth International Conference on Ground Penetrating Radar,* Gold Coast, Australia, pp 872-879, May 23-25 2000.

[22] Carevic, D., "Wavelet-Based Method for Detection of Shallowly Buried Objects from GPR data," *Proc. Of Information Decision and Control,* 201-206, 1991.

[23] Brian, Karlsen, Helge B.D. Sorensen, Jand Larsen and Kaj B. Jalobsen, "Independent Component Analysis for Clutter Reduction in Ground Penetrating Radar Data," *Proc. Of SPIE, Detection and Remediation Technologies for Mines and Minelike Targets,* Vol. 4742, 378-389, 2002.

[24] R. W. Deming, "Automatic Buried Mine Detection Using the Maximum Likelihood Adaptive Neural System (MLANS)," in P*roc. 1998 IEEE ISIC/CIRA/ISAS Joint Conf.*, pp. 428-433, Gaithersburg, MD, Sept. 1998.

[25] C. M. Rappaport and D. M. Reidy, “Focused Array Radar for Real Time Imaging and Detection,” in *Proc. SPIE.96 Conf.*, Orlando, FL, 1996.

[26] M. D. Patz and M. A. Belkerdid, “Evaluation of a Model-Based Inversion Algorithm for GPR Signal Processing Correlation for Target Classification,” in *Proc. SPIE.98 Conf.*, Orlando, FL, 1998.

[27] T. R. Witten, “Present State of the Art in Ground Penetrating Radars for Mine Detection,” in *Proc. SPIE.98 Conf.*, Orlando, FL, 1998.

[28] S.M. Shrestha and I Arai, “High Resolution Image Reconstruction by GPR Using MUSIC and SAR Processing Method for Landmine Detection,” *Geoscience and Remote sensing Symposium*, Volume 4, pp. 2921-2923, 2003.

[29] J. K. Paik, C. P. Lee, and M. A. Abidi, “Image Processing-Based Mine Detection Techniques Using Multiple Sensors: A Review.” *Subsurface Sensing Technologies and Applications: An International Journal*, Vol. 3, No. 3, pp. 153-202, July 2002.

[30] UWBGPR measurement at the Royal Military Academy, Belgium on May 31, 1999.

[31] I.J. Myung, “Tutorial on Maximum Likelihood Estimation,” *Journal of Mathematical Psychology* 47 (2003) 90–100.

[32] K. C. Ho and P. D. Gader, “Correlation Based Landmine Detection Using GPR,” in *Proc. SPIE .00 Conf.*, Orlando, FL, 2000.

[33] Y. Das, K. Russell, N.Kircanski and A.A.Goldenberg, “An Articulated Robotic Scanner for Mine Detection - a Novel Approach to Vehicle Mounted Systems,” *SPIE, 1999 Conference (Aerosense)*, Orlando, Florida, 5-9 April 1999.

[34] John McFee, Vic Aitken, Robert Chesney, Yogadhish Das, Kevin Russell, “A Multisensor, Vehicle-Mounted, Teleoperated Mine Detector with Data Fusion,” in *Detection and Remediation Technologies for Mines and Mine-like Targets III*, A.C.Dubey, J.F.Harvey, and J.T.Broach, eds., *Proc.SPIE* **Vol.3392**, pp. 1082–1093, (Orlando, FL, USA), 13–17 April 1998.

[35] L. M. Collins, P. Torrione, V. Munshi, C. S. Throckmorton, Q. Zhu, F. Clodfelter and S. Frasier, “Algorithms for Landmine Detection Using NIITEK Ground Penetrating Radar,” in *Proc. SPIE Detection and Remediation Technologies for Mines and Minelike Targets VII*, pp. 709 - 718, Orlando, Apr. 2002.

[36] P. Torrione, L. Collins, F. Clodfelter, S. Frasier, I. Starnes, “Application of the LMS Algorithm to Anomaly Detection Using the Wichmann/Niitek Ground Penetrating Radar,” in *Detection and Remediation Technologies for Mines and Minelike Targets VII Conf.—Int. Symp. Aerospace/Defense Sensing and Controls*, pp. 1127–1136, Orlando, FL, Apr. 2003.

[37] P. D. Gader, M. Mystkowski and Y. Zhao, "Landmine Detection with Ground Penetrating Radar Using Hidden Markov Model," *IEEE Trans. Geoscience Remote Sensing*, Vol. 39, pp. 1237-1244, June 2001.

[38] P. D. Gader, R. Grandhi, W. Lee, J. Wilson and K. C. Ho, "Feature Analysis for the NIITEK Ground-Penetrating Radar Using Order-Weighted Averaging Operators for Landmine Detection," in *Proc. SPIE Detection and Remediation Technologies for Mines and Minelike Targets IX*, Orlando, Apr. 2004.

[39] Udaynag Pisipati, "Techniques for Improving Landmine Detection Using Ground Penetrating Radar," Master Thesis, University of Missouri—Columbia, May 2006.

[40] Brian S. Everitt, Sabine Landau, Morven Leese, "Cluster Analysis," Oxford University Press Inc., New York.

[41] http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/

[42] http://www.cs.ualberta.ca/7Ezaiane/courses/cmput690/slides/Chapter8/

[43] http://en.wikipedia.org/wiki/Cluster_analysis

[44] Vance Faber, "Clustering and the continuous K-Means Algorithm," *Los Alamos Science*, Number 22, 1994.

[45] Seong-Wook Joo, "Subspace and Kernel Methods", Seminar slides. (Online) Available http://www.umiacs.umd.edu/~knkim/KG_VISA/Subspace/ SubspaceMethods.ppt

[46] Carl Meyer, "Matrix Analysis and Applied Linear Algebra," Society for Industrial and Applied Mathematics, Philadelphia.