

Nonlinear Principal Component Analysis for Clutter Reduction in GPR Data for Landmine Detection

Fawzy Abujarad

Engineering Department, Faculty of Engineering & Information Technology, Al Azhar University Gaza
Al-Azhar University - Gaza Palestine

ABSTRACT: This paper illustrates clutter reduction in stepped-frequency ground penetrating radar (SFGPR) data for anti-personnel landmines detection. Ground reflected clutter is often a performance limiting factor in ground-penetrating radar (GPR) detection of near surface targets including antipersonnel mines. In this paper the Non-linear principal component analysis (NLPCA), as a method for blind source separation is used to separate the clutter and the target signals. The algorithm has been experimentally evaluated and compared with principal component analysis (PCA) algorithm using different non-metallic AP landmines with different depths. The experimental data have been collected by using frequency stepped GPR system covering two ranges from 1 GHz to 4GHz and from range 1.5 GHz to 20GHz, from laboratory measurements. The test gives good result, in agreement with the theory. The surface reflection and the reflection within the earth have been removed approximately for both data. Based on these results the NLPCA showed significant improvement compared to the PCA.

Keywords: GPR, NLPCA, clutter, signal processing.

I. INTRODUCTION

The United Nations estimates that tens of millions of mines lie buried around the world. Anti-personnel mines and anti-tank mines have not been cleared out in our living earth and can be found in more than 70 countries. The detection of landmines is an important mission to save lives for many innocent victims. Detection of anti-personnel non-metallic landmines using Ground Penetrating Radar (GPR) has shown promising results in recent years [1].

A GPR senses electrical in-homogeneities caused by a dielectric contrast [2]. A related key problem is how to extract the mines scattered signals from the received data when the contrast is very weak (e.g., when the GPR used for the detection of non-metallic anti-personnel (plastic) landmine in the presence of soil) implying that the landmine GPR signal is very weak. [3].

The largest contrast typically exists between the air and the soil and therefore GPR is typically characterized by a very large ground bounce. If a landmine is buried at a shallow depth, then the GPR, which is used to detect and localize AP mines, will lead to several problems especially if the surface of the soil is rough or the soil is inhomogeneous. Unfortunately, this technology can suffer false alarm rates as high as that of metal detectors. Thus, it is clear that signal processing of the GPR data is needed in order to extract useful information for the target. Signal-processing algorithms, which filter out clutter signals and select objects to be declared as mines, is considered the most critical part of the GPR system.

There are two distinct types of GPR: Time-domain and frequency domain. Time domain or impulse GPR transmits discrete pulses of nanosecond duration and digitizes the returns at GHz sample rates. Frequency domain GPR systems transmit single frequencies uniquely, either as a series of frequency steps, or as a chirp.

The amplitude and the phase of the returned signal are measured, and the resulting data is converted back to the time domain [4]. Clutter signal sources are the cross-talk between the transmitter and the receiver antennas as well as scattering from the ground surface and within the soil. In addition other objects are generally buried under the ground, which appear as targets.

The clutter signal due to the reflection from the ground surface is the main problem. In case, the target is buried deeply below the surface, the backscattered signal from the surface can easily be distinguished from that of the target. The signals can be separated, in this situation, by a time gating technique. However, if the target is shallowly buried near the surface interface, gating is not a proper solution, as the backscattered signal from both the target and the surface will arrive almost simultaneously. The detection problem can thus be described as how to separate the target signal from the ground surface reflection.

Several signal processing approaches have been suggested to improve the performance of GPR systems. These include: simple mean scan subtraction [3], two-dimensional digital filtering [5], wavelet packet decomposition [3], likelihood ratio test [3], [6], parametric system identification [7], and Kalman filter [4], [6]. Most of these methods depend on the background signal estimation by taking the mean value of the unprocessed ensemble collected GPR data, followed by employing the simple mean scan subtraction. Although these methods have been used widely in GPR applications.

Various researchers have shown interest in subspace techniques including Singular Value Decomposition (SVD) [3], [8], Linear Discriminant Analysis (LDA) [9], Principal Component Analysis (PCA) [10], [11], and Independent Component Analysis (ICA) [10], [12]–[14].

Because of its computational and conceptual simplicity, a linear transformation is often assumed for such a representation. As such a transformation linearly mixes the target and clutter signals, it is usually expressed as a so-called mixing matrix.

The main goal of the detection here is to estimate the target and clutter signals (source signals) without any knowledge of the mixing matrix. Such a problem is usually called blind source separation.

Within the class of linear methods, the optimal information preserving transformation is given by principal component analysis (PCA) [15]. Reduction of dimensionality by PCA has been shown to facilitate many types of multivariate analysis, including data validation and fault detection quality, correlation and prediction, and data visualization. The feature variables in PCA, also called factors, are linear combinations of the original problem variables. The coefficients of the linear transformation are such that if the feature transformation is applied to the data set and then reversed, there will be a minimum sum of squares difference between the original and reconstructed data.

However, PCA suffers from two important limitations. First, it assumes that the relationships between variables are linear, and second, its interpretation is only sensible if all of the variables are assumed to be scaled at the numeric level (interval or ratio level of measurement). These assumptions are frequently not justified, and therefore, PCA as linear transformation may not always be the most appropriate method of analysis. To circumvent these limitations, an alternative, referred to as nonlinear principal components analysis (NLPCA), [16].

The aim of this article is to review and examine nonlinear principal component analysis (PCA) criterion in blind source separation (BSS). The numerical studies on NLPCA reported in the literature are only limited to some simple toy problems. The potentials and the drawbacks of the technique are not well understood yet. Our aim is thus to analyze the NLPCA features in the case of real data.

II. BLIND SOURCE SEPARATION

Recovering signals from only the mixed observations without knowing the priori information of both the source signals and the mixing process is often referred to as Blind Source Separation (BSS). The most elementary class of such methods is based on linear mixture models that combine source signals or mixture components as weighted linear sum. Such linear blind source separation methods have a wide spectrum of applications.

We consider the standard signal model used in BSS which is given by Equation (1)

$$x_i = \sum_j^n a_{ij} s_j \quad (1)$$

or in the matrix notation as given by Equation (2)

$$x = A s \quad (2)$$

Here $s(t) = [s_1(t), \dots, s_n(t)]^T$ is the source vector at time t . $x(t) = [x_1(t), \dots, x_m(t)]^T$ is a vector of sensor signals (observed variables). The components of the m dimensional data vector $x(t)$ are some linear mixtures of the source signals. The $m \times n$ mixing matrix A is an unknown full rank constant matrix.

BSS consists of identifying mixing matrix A and/or retrieving the source signals without resorting to any a priori information about A ; it uses only the information carried by the received signals themselves, hence, the term blind. Therefore, the objective of BSS is to find an inverse (un-mixing) system, sometimes termed a reconstruction system. In order to estimate the all primary source signals $s(t)$ or only some of them with specific properties. This estimation is performed on the basis of only the output signals $y(t) = [y_1(t), \dots, y_n(t)]^T$ and

the sensor signals. Then the BSS problem is to determine a constant un-mixing (weight, separating) matrix W (where W is $n \times m$) so that the solution is sought in the form as given in Equation (3)

$$\hat{s}(t) = y(t) = Wx(t) \quad (3)$$

Different methods have been developed to find such a linear representation, including singular value decomposition [17], factor analysis [18], principal components analysis [19], independent component analysis [20], dependent component analysis [21], non-negative matrix factorization [22], projection pursuit [23], etc.

III. PRINCIPAL COMPONENT ANALYSIS

A central problem in signal processing is to find a suitable representation of the data, using a suitable transformation. In most cases, the representation is sought as a linear transform of the observed variables as given in Equation (4)

$$Y = W X \quad (4)$$

where X and Y are $m \times n$ matrices related by linear transformation A . The purpose of principal component analysis is to derive a relatively small number of decorrelated linear combination (principal component) of a set of random zero-mean variables while retaining as much of the information from the original variables as possible [24]. The basic idea in PCA is to find the rows of the $[y_1(t), \dots, y_n(t)]^T$ that explain the maximum amount of variance possible by N linearly transformed components.

Although the basic goal in PCA is to decorrelate the data by performing an orthogonal projection, we often reduce the dimension of the data from m to n ($n < m$) to remove unwanted components in the signal. In practice, the computation of A can be simply accomplished using the covariance matrix. The covariance matrix (C_x) reads then Equation (5)

$$C_x = \frac{1}{n} \sum_{i=1}^n X_i X_i^T = \frac{1}{n} X X^T \quad (5)$$

The eigenvector and eigenvalue matrices of (C_x) Φ and Λ , respectively, are computed as:

$$C_x \Phi = \Phi \Lambda$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, and $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues. If one assumes that eigenvalues are sorted in a decreasing order, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, then the n leading eigenvectors matrix A is given by Equation (6)

$$A = [\Phi_1, \Phi_2, \dots, \Phi_n] \quad (6)$$

In the application of landmine detection in GPR signals, the PCA can be used to detect landmines and reduce the noise [25]. By selecting some components which mainly carry mine information, say A_k we can remove the clutter. The reconstructed signal space in the original GPR signal space is then We select the matrix A to be a matrix where each row p_k is an eigenvector of XX^T (the principal components of X). In the standard numerical approach for extracting the principal components, first the covariance matrix $C_x = 1/n (XX^T)$ is computed and then its eigenvectors and (corresponding) associated eigenvalues are determined by one of the known numerical algorithms.

However, if the input data vectors have a large dimension (GPR data), then the covariance matrix C_x becomes very large and it may be difficult to compute the required eigenvectors. A neural network approach with adaptive learning algorithms enables us to find the eigenvectors and the associated eigenvalues directly from the input vectors $x(t)$ without a need to compute or estimate the very large covariance matrix C_x . Computing the sample covariance matrix itself is very costly. Furthermore, the direct diagonalization of a matrix or eigenvalue decomposition can be extremely costly since this operation is of complexity $O(m^3)$.

IV. NONLINEAR PRINCIPAL COMPONENTS ANALYSIS:

In the field of GPR signal processing, it is important to reduce the dimensionality of data. As a method of dimensionality reduction, principal component analysis (PCA) is considered to be an appropriate way to perform such data reduction.

PCA is an orthogonal transformation of a coordinate system in which we describe data. A basis of the objective coordinate system efficiently represents data distributed on a linear hyper plane as the coordinate value that is called a principal component.

PCA suffers from two important limitations. First, it assumes that the relationships between variables are linear, and second, its interpretation is only sensible if all of the variables are assumed to be scaled at the numeric level (interval or ratio level of measurement).

PCA may not always be the most appropriate method of analysis. To circumvent these limitations, an alternative, referred to as nonlinear principal components analysis.

Some methods of nonlinear principal component analysis (NLPCA) have been developed [26, 27]. These methods can be classified into three categories. One is the application of a sandglass-type multi-layered perceptron (MLP). The other method is based of a partial linear approximation. In these methods, the number of principal components has to be specified in advance. The third method is Kernel PCA recently proposed by Scholköpfung [28]. Kernel PCA is a method that executes the linear PCA algorithm for the image of input data mapped by a nonlinear mapping function. Consequently, the method constructs the ordered principal components. However, the adequate way to determine the nonlinear mapping function for a given data set is not known. Moreover, the method has to solve an eigen- equation for a covariance matrix of the image of data in order to obtain an eigen base. The method also requires calculations of kernel functions between an objective data and all training data to calculate a principal component score.

Nonlinear principal component analysis (NLPCA) as a nonlinear generalisation of standard principal component analysis (PCA) means to generalize the principal components from straight lines to curves. Fig.1 presents the geometry principle curves. In such cases, it may be more appropriate to assume that the feature subspace is defined by nonlinear functions of the process variables.

It is more appropriate to assume that the hidden factors are nonlinear functions of the observed variables. Furthermore, the reconstruction of the variables from the factors may also be a nonlinear mapping. In general, we assume that the m-dimensional observation vector $x(t) = [x_1(t), \dots, x_m(t)]^T$ is generated by an underlying feature vector $y(t) = [y_1(t), \dots, y_n(t)]^T$ ($n < m$) via nonlinear continuous function as given by Equation (7)

$$y = \varphi(x), \quad \varphi \in S_e \tag{7}$$

and a nonlinear reconstruction function from y onto the reconstructed vector as \hat{x} as given by Equation (8)

$$\hat{x} = \psi(y) \quad \psi \in S_r \tag{8}$$

where S_e and S_r are the sets of nonlinear functions. The coding function φ from \mathcal{R}^m to \mathcal{R}^n and the decoding function ψ from \mathcal{R}^n to \mathcal{R}^m are members of some classes S_e and S_r of nonlinear functions. The data nonlinearly correspond to the principal components through the nonlinear extraction function and the nonlinear reconstruction function.

Our problem is to minimize the mean square reconstruction error is given by Equation (9)

$$E = E \left[\|x - \hat{x}\|^2 \right]$$

$$E = E \left[\|x - \psi(\varphi(x))\|^2 \right] \tag{9}$$

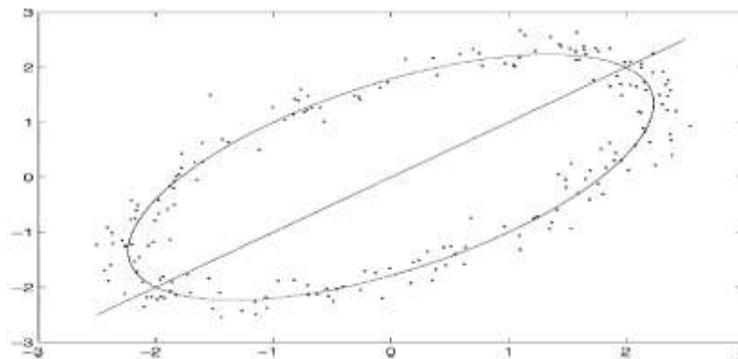


Figure 1 Set of ellipsoid data demonstrates the difference between nonlinear PCA and standard PCA. The data (dots) are much better represented by the monoparametric nonlinear principal curve (ellipse) than by the linear principal component curve (straight line).

When we find the optimal ψ and ϕ , we can efficiently describe the data with fewer principal components than that of PCA. Clearly, the solution to the nonlinear PCA problem depends on both the choice of the sets S_e and S_r and the distribution of x . Ordinary (linear) PCA is now a special case for S_e, S_r being the set of linear mappings. In this work, we introduce nonlinear reconstruction functions that map a vector in the principal component space into the original input space as given by Equation (10)

$$\begin{aligned} \hat{x}_1 &= \psi_1(y_1) \\ \hat{x}_2 &= \psi_2(y_1, y_2) \\ &\cdot \\ &\cdot \\ &\cdot \\ \hat{x}_n &= \psi_n(y_1, \dots, y_n) \end{aligned} \tag{10}$$

where y_i represents the score of the input vector with respect to the i th principal component, and \hat{x}_i is the n -dimensional vector reconstructed by the i th reconstruction function utilizing the representation of the original input vector x in the i -dimensional principal component space.

The i th component of the extraction function ϕ_i and the i th reconstruction function ψ_i are paired in the following manner (Equation (11))

$$\begin{aligned} \psi_1(x) &= \psi_1(\phi_1(x)) \\ \psi_2(x) &= \psi_2(y_1, \phi_2(x)) \\ &\cdot \\ &\cdot \\ &\cdot \\ \psi_n(x) &= \psi_n(y_1, \dots, y_{n-1}, \phi_n(x)) \end{aligned} \tag{11}$$

The functions of each pair are adjusted to minimize the mean square reconstruction error in the above order. Therefore, y_i can be regarded as the i th significant nonlinear principal component.

The unique recovery of the hidden parameters is impossible in general because there are infinite solutions to the NLPCA minimization problem. Indeed, if a pair of functions $\phi_1(\cdot), \psi_1(\cdot)$ achieves the minimum error

$E \left[\|x - \psi_1(\phi_1(x))\|^2 \right]$ then so does any pair $\phi_1(q^{-1}(\cdot)), \psi_1(q(\cdot))$ for any invertible function $q(\cdot)$. Nevertheless,

the following sets are unique and can be considered as problem inherent [29]:

- 1) The set $\mathfrak{T} = \{I(y), \text{all } y \in \mathfrak{R}^n\}$ of contours $I(y) = \{x : \phi(x) = y\}$ for the function ϕ .
- 2) n -parametric surface $\Sigma = \{\psi(y), \text{all } y \in \mathfrak{R}^n\}$ generated by ψ .

Σ is called the n -parametric nonlinear principal component surface of (Fig. 1).

V. GPR MEASUREMENTS

The data have been acquired with a bistatic-stepped frequency GPR system at IESK, Magdeburg University, Germany. The system consists of a network analyzer (Rohde & Schwarz) and two ultra-wideband (UWB) transmitting and receiving antennas [30]. A wooden box with dimensions $1.1 \times 1.1 \times 1.1$ m whose internal sides are covered by absorption material and is filled by sand of 0.5 m height has been used. The transmitting and receiving antennas are mounted on a 2D scanning system and were placed above the ground surface at height 30 cm.

The measurement grid covers the area bounded by $x = 27 \rightarrow 76$ cm and $y = 39 \rightarrow 89$ cm with a distance between the measurements of 1 cm in both x and y directions. The measurements then form a two dimensional matrix, referred to as a B-scan. Column vector of the B-scan matrix (image) is called an A-scan and it represents the data, at each individual point on the surface of the soil. Using this experimental setup, two different measurements were made. In the first measurement the radar system was operated in the frequency range from 1 GHz to 4 GHz and the number of samples was 1024 for each A-scan.

In the second measurement the radar system was operated in the frequency range from 1.5 GHz to 20 GHz and the number of samples was 1601 for each A-scan. Examples of A-scans in the presence and absence of a landmine for both measurements in the frequency domain are displayed in Fig. 2 and Fig. 3. PMN anti-personal landmine was used at different depths.

Sample B-Scans showing PMN targets at depth 0, 3, 6, 9, 12 and 21 cm are displayed in Fig.4.

VI. EXPERIMENTAL RESULTS

To examine the ability of a non-linear principal component technique for the elimination or reduction of the clutter in raw GPR data, the non-linear principal component technique has been applied to both measurements. The algorithm has been applied to different landmines and different depths. Fig.5 show the data after applying the non-linear principal component technique to the data in Fig.4. In almost cases, the target signals have been extracted from the GPR data, and most of the clutter signal (cross-talk and air-ground interface) has been removed for the target. However for the small targets and for deep depth the clutter has not been removed completely and the result of applying the NLPCA is the same as applying linear transform. Non-linear principal component have been compared with principal component analysis. The signal-to-noise ratios (SNR) [31] have been used to compare between the two methods.

The results show that NLPCA have offers a significant improvement as compared to PCA.

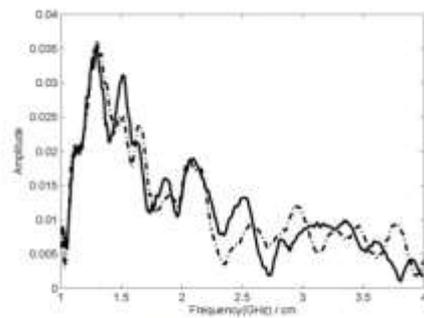


Figure2 A-scans in the presence (dashed) and absence (solid) of a mine for data with 2 GHz bandwidth in the frequency domain.

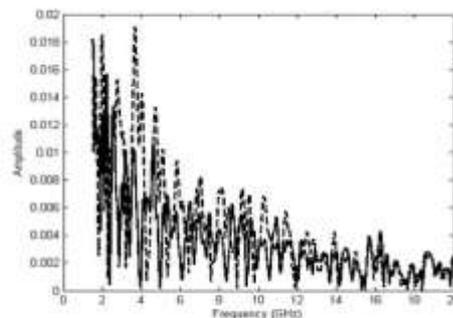


Figure3 A-scans in the presence (dashed) and absence (solid) of a mine for data with 2 GHz bandwidth in the time domain.

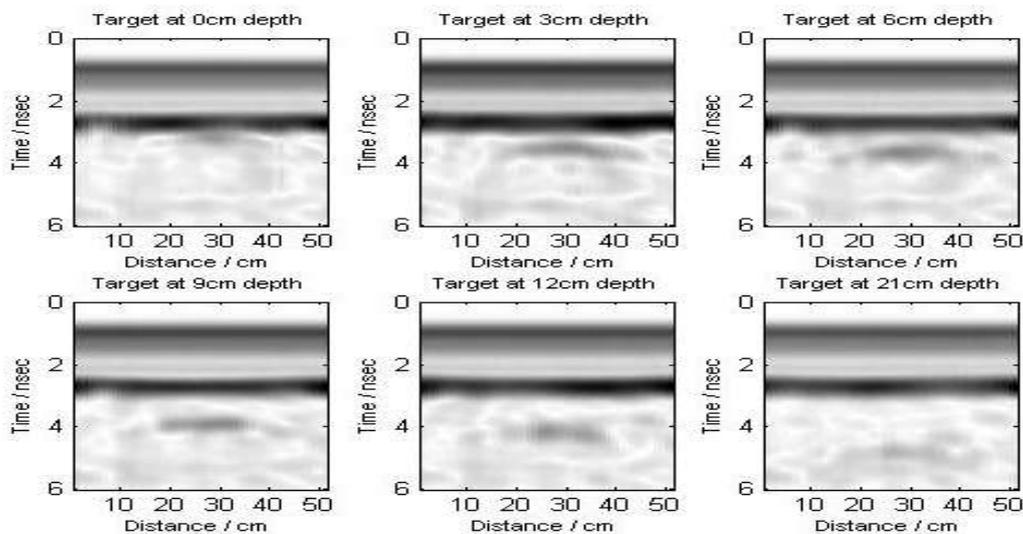


Figure 2 B-Scans of PMN anti-personal landmine at depth 0, 3, 6, 9, 12 and 21cm, respectively.

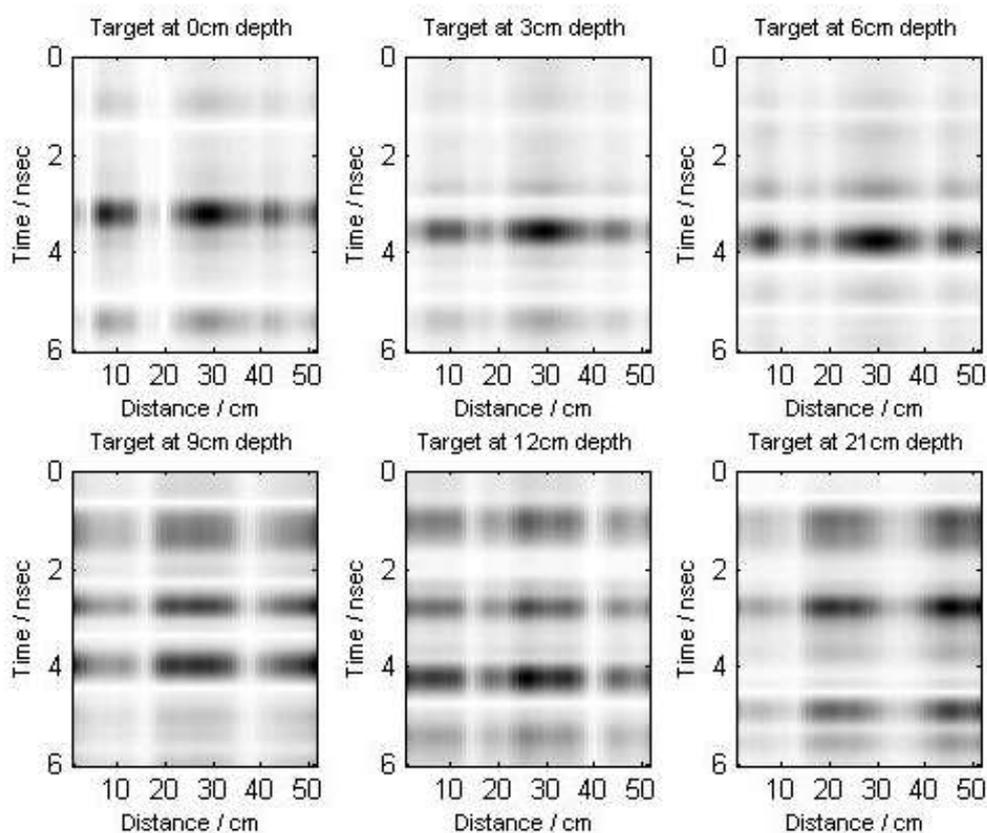


Figure 3 B- Scans of the data after applying the non-linear principal Component technique to the data in

VII. CONCLUSION

The main purpose of this communication has been the study of non-linear principal component test of GPR data. An algorithm for de-noising the GPR data has been described and has been explored for clutter reduction and non-metallic landmine detection from SF-GPR signals. The reflection from the surface has been removed or has been at least significantly reduced. From the results NLPCA showed significant improvement over PCA. On the other hand, the numerical implementation of PCA is an easier.

REFERENCES

- [1]. D. Daniels, Surface Penetrating Radar. IEE, 1996.
- [2]. X. Xu, E. Miller, C. Rappaport, and G. Sower, "Statistical method to detect subsurface objects using array ground penetrating radar data," IEEE Transactions on Geoscience and Remote Sensing, vol. 40, pp. 963–976, 2002.
- [3]. F. Abujarad, Jostingmeier, and A. S. Omar, "Clutter removal for landmine using different signal processing techniques," Proc. tenth international conference on ground penetrating radar, pp. 697–700, 2004.
- [4]. A. M. Zoubir, I. J. Chant, C. L. Brown, B. Barkat, and C. Abeynayake, "Signal processing techniques for landmine detection using impulse ground penetrating radar," IEEE Sensors Journal 2, pp. 41–51, 2002.
- [5]. D. Potin, E. Duflos, and P. Vanheeghe, "Landmines ground-penetrating radar signal enhancement by digital filtering," IEEE Transactions on Geoscience and Remote Sensing, vol. 44, pp. 2393–2406, 2006.
- [6]. van Kempen and H. Sahli, "Signal processing techniques for clutter parameters estimation and clutter removal in gpr data for landmine detectio," IEEE Workshop on Statistical Signal Processing Proceedings, pp. 158–161, 2001.
- [7]. van der Merwe and I. Gupta, "A novel signal processing technique for clutter reduction in gpr measurements of small, shallow land mines," IEEE Transactions on Geoscience and Remote Sensing, vol. 38(6), pp. 2627–2637, 2000.
- [8]. F. Abujarad, G. Nadim, and A. Omar, "Clutter reduction and detection of landmine objects in ground penetrating radar data using singular value decomposition (svd)," Proceedings of the 3rd International Workshop on Advanced Ground Penetrating Radar, pp. 37–41, 2005.
- [9]. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19(7), pp. 711–720, 1997.
- [10]. F. Abujarad and A. Omar, "Gpr data processing using the component-separation methods pca and ica," IEEE International Workshop on Imaging Systems and Techniques, pp. 59–63, 2006.

- [11]. B. Karlsen, J. Larsen, H. B. D. Sørensen, and K. B. Jakobsen, "Comparison of pca and ica based clutter reduction in gpr systems for anti-personal landmine detection," IEEE Workshop on Statistical Signal Processing Proceedings, pp. 146–149, 2001.
- [12]. A. Zhao, Y. Jiang, and W. Wang, "Exploring independent component analysis for gpr signal processing," in Progress In Electromagnetics Research Symposium, 2005.
- [13]. B. Karlsen, J. Larsen, H. B. D. Sørensen, and K. B. Jakobsen, "Independent component analysis for clutter reduction in ground penetrating radar data," SPIE Aerosense, vol. 4742, pp. 378–389, 2002.
- [14]. Q. gao, T. Li, and R. Wu, "A novel kica method for ground bounce removal with gpr," Proceedings of 2006 CIE International Conference on Radar, vol. 1, pp. 1058–1061, 2006.
- [15]. Fukunaga, K., and W. Koontz, "Application of Karhunen-Loeve Expansion to Feature Selection and Ordering," IEEE Trans. Comput., C-19, 311 (1970).
- [16]. A. Gifi, Nonlinear multivariate analysis. Chichester, New York, Wiley, c1990.
- [17]. G. H. Golub and C. F. Van Loan. Matrix Computations. Baltimore, MD, USA, second edition, 1989.
- [18]. H. H. Harman. Modern Factor Analysis. University of Chicago Press, 2nd edition edition, 1967.
- [19]. E. Oja. Principal components, minor components, and linear neural networks. Neural Networks, 5:927{935, 1992.
- [20]. C. Jutten and J. Herault. Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. Signal Processing, 24:1{10, 1991.
- [21]. A. K. Barros. The independence assumption: Dependent component analysis. In M. Girolami, editor, in Advances in Independent Component Analysis. Springer, 2000.
- [22]. A. Cichocki, R. Zdunek, and S. Amari. New algorithms for non-negative matrix factorization in applications to blind source separation. 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP2006, pages 14{19, Toulouse, France, May 2006.
- [23]. P.J. Huber. Exploratory projection pursuit. J. of the American Statistical Association, 82(397):249{266, 1987.
- [24]. A. C. . S. ichi Amari, Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. New York, NY: John Wiley & Sons, Inc.,2002.
- [25]. F. Abujarad, A.S. Omar. Factor and Principle Component Analysis for Automatic Landmine Detection Based on Ground Penetrating Radar. German Microwave Conference (GeMIC) 2006, Karlsruhe, Germany, March 2006.
- [26]. K.I. Diamantaras, S.Y. Kung, Principal Component Neural Networks Theory andApplications, Wiley, New York, 1996.
- [27]. R. Gnanadesikan, Methods for Statistical Data Analysis of Multivariate Observations, Wiley, New York, 1977.
- [28]. B. Scholkopf, A. Smola, K. Muller, Nonlinear component analysis as a kernel eigenvalue problem, Neural Comput. 10 (5) (1998) 1299–1319.
- [29]. K. I. Diamantaras and S. Y. Kung,Principal Component Neural Net-works: Theory and Applications New York: Wiley, 1996
- [30]. F. Abujarad, "Ground penetrating radar signal processing for landmine detection," Ph.D. Thesis, 2007.
- [31]. H. Brunzell. "Detection of shallowly buried objects using impulse radar" Department of Applied Electronics, Chalmers University of Technology. May,13 1996.