8-2010

An Automated Framework for Defect Detection in Concrete Bridge Decks Using Fractals and Independent Component Analysis

Fadi Abu-Amara
Western Michigan University

AN AUTOMATED FRAMEWORK FOR DEFECT DETECTION IN CONCRETE BRIDGE DECKS USING FRACTALS AND INDEPENDENT COMPONENT ANALYSIS

by

Fadi Abu-Amara

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Electrical and Computer Engineering
Advisor: Ikhlas Abdel-Qader, Ph.D.

Western Michigan University
Kalamazoo, Michigan
August 2010

AN AUTOMATED FRAMEWORK FOR DEFECT DETECTION IN CONCRETE BRIDGE DECKS USING FRACTALS AND INDEPENDENT COMPONENT ANALYSIS

Fadi Abu-Amara, Ph.D.

Western Michigan University, 2010

Bridge decks deteriorate over time as a result of deicing salts, freezing-and-thawing, and heavy use, resulting in internal defects. According to a 2006 study by the American Society of Civil Engineers, 29% of bridges in the United States are considered structurally deficient or functionally obsolete. Ground penetrating radar (GPR) is a promising non-destructive evaluation technique for assessing subsurface conditions of bridge decks. However, the analysis of GPR scans is typically done manually, where the accuracy of the detection process depends on the technician's trained eye. In this work, a framework is developed to automate the detection, localization, and characterization of subsurface defects inside bridge decks. This framework is composed of a fractal-based feature extraction algorithm to detect defective regions, a deconvolution algorithm using banded-ICA to reduce overlapping between reflections and to estimate the depth of defects, and a classification algorithm using principal component analysis to identify main features in defective regions. This framework is implemented and simulated using MATLAB and GPR real scans of simulated concrete bridge decks.

This framework, as demonstrated by the experimental results, has the

following contributions to the current body of knowledge in ground penetrating radar detection and analysis techniques, and in concrete bridge deck condition assessment: 1) developed a framework that integrated detection, localization, and classification of subsurface defects inside concrete bridge decks, 2) presented a comparison between the most common fractal methods to determine the most suitable one for bridge deck condition assessment, 3) introduced a fractal-based feature extraction algorithm that is capable of detecting and horizontally labeling defective regions using only the underlying GPR B-scan without the need for a training dataset, 3) developed a deconvolution algorithms using EFICA to detect embedded defects in bridge decks, 4) introduced an automated identification methodology of defective regions which can be integrated into a CAD system that allows for better visual assessment by the maintenance engineer and has the potential to eliminate human interpretation errors and reduce condition assessment time and cost, and 6) presented an investigation and a successful attempt to classify some of the common defects in bridge decks.

ACKNOWLEDGMENTS

First and foremost I would like to thank God for his blessings and guidance to complete this work. Secondly, I owe my deep gratitude to my dissertation advisors: Dr. Ikhlas Abdel-Qader and Dr. Osama Abudayyeh for their support, encouragement, and guidance throughout this research work. I have truly gained invaluable knowledge from working with them.

I would also like to extend my appreciation to my committee member Dr. Massood Atashbar for serving on my committee and many thanks to my sponsor Al Hussein Bin Talal University.

I would like to thank my wife Omaima Al-Matari for her love, patience, and pushing me when I am ready to give up. I can barely find the words to express my heartily gratitude and respect to my parents for their immense love, dedication, and encouragement throughout my life and this work. Finally, to my daughters Misk, Yasameen, and Rayhan: you are the best thing I have ever done in this life. I would like to thank and list all my family and friends, but it would end up being longer than the dissertation.

Fadi Abu-Amara

TABLE OF CONTENTS

CHAPTER

Table of Contents—Continued

CHAPTER

Table of Contents—Continued

LIST OF TABLES

LIST OF FIGURES

List of Figures—Continued

CHAPTER 1

INTRODUCTION

Bridge decks deteriorate over time as a result of deicing salts, water penetration, freezing-and-thawing, and heavy use, resulting in internal defects. The national bridge inventory includes more than 600,000 bridges in the United States [1]. According to a 2006 study by the American Society of Civil Engineers, 29% of bridges in the United States are considered structurally deficient or functionally obsolete due to overdue maintenance [1]. Also, $2.2 trillion dollars over a five year period are required to bring the U. S. roads, highways, and bridges back to reasonable conditions [2].

Bridge Deck Condition Assessment

Bridge deck condition assessment can be used to determine the necessity for maintenance, predict associated costs, and to determine safety and serviceability of the bridge [3]. The main challenge to bridge deck condition assessment is detecting subsurface defects before they develop into severe damage that would require a costly rehabilitation. Therefore, subsurface nondestructive techniques are needed to identify and diagnose embedded defects at their early stages.

Different nondestructive techniques are proposed for subsurface defects detection in bridge decks such as Ground penetrating radar, Impact Echo, Infrared

Thermography, Acoustic Emission, and Ultrasonic Pulse Velocity [3, 4, 5, 6]. A technique that deals with inhomogeneous, unknown, and difficult to access materials and structures is needed. This technique should not be time consuming or labor intensive and must provide reliable results.

In [3], ground penetrating radar, impact echo, and infrared thermography were evaluated as non-destructive evaluation techniques for subsurface defects detection in simulated bridge decks. Infrared thermography (IRT) provides good diagnostic information in real-time for near-surface targets but not suitable for targets deeper than 2 inches. Also, performance of IRT depends on the environmental conditions such as solar loading from direct sunlight and wind speed which indicates that it can only provide reliable results during specific times of day and year. Finally and most importantly, depth of the defects cannot be estimated from IRT scans which make it unsuitable for the problem at hand. Impact Echo (IE) was able to detect the embedded defects deeper than 2 inches. However, IE requires many testing points making it a tedious and time consuming technique. The ability of IE to detect shallow defects depends on surface roughness of the scanned concrete slab. Ground penetrating radar (GPR) is a more sophisticated technique that was able to detect depth of the embedded defects deeper than 1 inch. However, GPR reflections from deep objects are less visible than reflections from shallow objects due to the scattering and absorption of GPR waves at the medium boundaries and thus raw GPR data needs post-processing to enhance it. GPR offers rapid data collection and can be used anytime of the day or the year. Finally, radar waves are less sensitive to surface

roughness.

In [5], Infrared thermography and ground penetrating radar were evaluated as non-destructive techniques for bridge deck condition assessment. Results of [5] indicate that the smallest air-void defect that can be detected by IRT is of size $2" \times 2"$. However, IRT is not an effective tool to detect water-filled voids. Also, IRT can only provide reliable results during specific times of day and year under low wind velocity. The 1.5 GHz antenna outperformed the 2 GHz antenna in detecting the embedded defects. Finally, the 1.5 GHz antenna was able to detect air-filled and water-filled voids with sizes of $a1" \times 1"$.

Bridge Deck Condition Assessment via GPR

Ground penetrating radar (GPR) is a promising non-destructive evaluation technique for assessing subsurface conditions of bridge decks. The reported work of using GPR in bridge deck condition assessment can be categorized into three groups: manual analysis, signal and image processing, and inverse scattering approaches.

Manual Analysis of GPR Scans

The first group involves manual analysis of GPR scans, requires using post-processing such as RADAN (RAdar Data ANalyzer) to enhance raw GPR data for better visual inspection, is time consuming, and the accuracy of the detection process depends on the technician's trained eye [3, 5, 6, 7, 8, 9].

Signal and Image Processing Approaches

The second group involves using signal and image processing methods to detect and/or characterize subsurface defects. In [10], an algorithm was developed to detect and characterize subsurface defects in bridge decks where principal component analysis algorithm is used to identify main features from each block extracted from the GPR scan after removing rebar reflections combined with an Euclidean distance as a dissimilarity measure for classification into normal, air-void, or water-void. Unfortunately, the proposed work by [10] did not include delaminations, a common defect in bridges, in the detection and characterization process. Also, defect coordinates were not provided.

In [11], an algorithm was developed based on the learning vector quantization-based neural network to classify subsurface defects in simulated bridge deck slabs into cracks, delaminations, and voids. Raw scans from healthy and defective traces were used to train the algorithm after subtracting a reference noise signal from them. The proposed algorithm in [11] utilized a single a-scan (trace) to characterize the corresponding embedded defect. Their detection process was manual and was able to characterize delaminations with flat reflections and air-voids with arch-shape reflections. They also concluded that even when a single trace was enough to detect the existence of a defect; it was not enough to characterize its type.

In [12], a statistical framework was proposed to detect subsurface defects from GPR scans of bridge decks. Their framework consisted of receiver characteristic curve, control chart analysis, and repeatability analysis. Using this framework, they

were able to find an optimal threshold value at which all reflections are thresholded and eliminated except for those reflections associated with the embedded defects. Then, the contour maps were used to quantify the deterioration percentage. However, some defects with weaker reflections than the threshold were not detected nor they attempted to compute the defect location or determine the defect type.

In [13], a framework was proposed to estimate thickness of the layers and depth of the embedded objects in a pavement. The framework consists of, 1) incident pulse removal stage, 2) noise reduction stage using an elliptic filter, 3) deconvolution stage using the Homomorphic deconvolution, 4) estimation of round-trip travel time of layers and objects, 5) dielectric constant estimation stage using the Reflections Amplitude method, and 6) layer thicknesses and embedded objects depth estimation stage. The proposed framework by [13] provided depth of all the embedded objects and defects without detecting and characterizing the subsurface defects.

Inverse Scattering Approaches

Inverse scattering approaches (ISA) use the Finite Difference Time Domain (FDTD) method as a subsurface modeling tool through approximating the physical geometry, material properties, and embedded targets. Thus, ISA are considered physics-based inversion techniques [14]. In case of bridge deck condition assessment, strong scatterers such as rebar may mask reflections from underneath objects which complicate the inversion approach and thus require new techniques to fit the traditional ISA to the problem at hand.

In [14], the FDTD method based on iterative computational modeling was used as an electromagnetic modeling to obtain simulated 2D and 3D GPR scans of a healthy bridge deck. By comparing the simulated healthy scans with real scans from the field, defective regions can be identified. In order to minimize reflections from the rebar that may mask underneath reflections, an excitation source was used to model the rebar. Characteristics of the GPR antenna were used to develop a virtual sensor implemented into the FDTD model to simulate the GPR antenna. Results of [14] indicate that the 2D FDTD modeling is more suitable than the 3D FDTD modeling for bridge deck condition assessment. Unfortunately, the developed model was not validated with real data from bridge decks to evaluate its performance in subsurface defects detection. Also, the developed model did not localize and characterize the subsurface defects.

Research Objectives

The goal of this research project is to automate the detection, localization, and characterization of bridge deck defects to provide a more accurate condition assessment method in a timely and cost effective manner to improve the inspection process. To achieve this goal, the following questions will need to be addressed:

- Can the defect detection process in bridge decks be automated?
- Can any type of defects be detected and characterized?
- Can coordinates of detected defects be estimated accurately?

To answer these questions, the following objectives are to be accomplished:

- Develop a feature extraction algorithm that can detect defective regions and localize them horizontally with reasonable computational complexity,
- Develop a deconvolution algorithm to reduce overlapping between reflections from closely spaced objects,
- Develop a mechanism to estimate the depth of defects, and
- Develop a classification algorithm to characterize and identify detected defects.

Overview of the Dissertation

In this work, a novel framework is developed to automate the detection, localization, and characterization of subsurface defects in bridge decks. Background on ground penetrating radar, principal component analysis, independent component analysis, and fractal analysis is presented in chapter 2. Chapter 3 explores the pertinent literature. Chapter 4 presents the proposed defect detection algorithm while chapter 5 discusses and analyzes experimental results followed by conclusions, contributions, and future work in chapter 6.

CHAPTER 2

BACKGORUND

Defects in Bridge Decks

Bridge decks are vulnerable to different forms of deterioration ranging from surface cracks to large holes. The most common defects in bridge decks are cracks, voids, delaminations, spalling, rebar corrosion, and accidental damage [3]. Rebar corrosion results from the infiltration of water and chlorides (from the deicing salt) through surface cracks. The corroded rebar expand causing cracks within the surrounding concrete. Also, the corroded rebar debones from the surrounding concrete reducing the structural integrity and resulting in delaminations and/or spalling. Cracks are caused by shrinkage tensile stresses, temperature changes due to freezing-and-thawing, or rebar corrosion. Detecting cracks is important as they are considered early signs of a physical damage. Unfortunately, longitudinal surface cracks cannot be detected using the GPR antenna [3, 7]. Another type of defects is spalling which occurs due to the freezing-and-thawing process where the corroded part of the rebar becomes heavier than before, causing a distress to the concrete. As a result, some parts of the concrete falls leaving discontinuities in the concrete.

The next type of defects is delaminations which mainly result from rebar corrosion that lead to the development of fracture planes. Typically, delaminations

will form right above or under the corroded rebar. Figure 1 shows a delamination defect in a core taken from a deteriorated bridge deck.

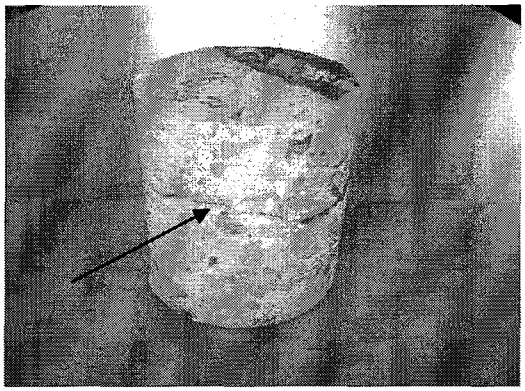

Figure 1. A core from concrete deck showing a delamination.

Voids can be air-filled or water-filled. Air-filled voids have a black-white-black reflection due the phase inversion since the electromagnetic waves propagate from a higher dielectric constant material to a lower dielectric constant material (concrete-air interface) while water-filled voids have a white-black-white reflection. Halabe and Bhandarkar [15] found that the process of detecting voids and delaminations in bridge decks becomes difficult in the presence of asphalt overlays.

Ground Penetrating Radar (GPR)

Ground penetrating radar (GPR) is a nondestructive technique that has been successfully used in bridge deck condition assessment [3, 5, 6, 7, 8]. The GPR antenna transmits polarized pulses of electromagnetic waves through the scanned

medium where portion of these radiations get attenuated due to natural absorption. At the boundary between two electrically different materials (i.e., different dielectric constants or electrical conductivities), some radiations reflect back while the rest refract and continue their penetration. Refractive indices of the two materials affect amplitude, reflection angles, and refraction angles of the reflected and refracted signals.

Higher contrast in the electrical properties between a target and the surrounding materials results in a stronger (brighter) reflection and consequently more visible target [16]. Higher electrical conductivity indicates higher water content and consequently more attenuation due to the natural absorption. Roughness of the scanned surface determines direction of the scattered reflections while the dielectric constant (real part of the dielectric permittivity normalized to air) determines propagation speed (and consequently penetration depth) within layers. The scattered reflections are detected by an antenna, recorded by the control unit against the two-way travel time, and then the signal is amplified. Due to the attenuation resulting from scattering and absorption, the penetration depth is limited in GPR systems.

Antennae with different frequencies can be used in GPR systems. If the antenna has a low frequency range (high wavelength), radio waves can reach a depth ranging from 30 to 40 feet in sandy soils [17]. In this case, the recorded scans will have a low resolution which can be used to locate fractures and deeply buried large objects. At a high frequency range, shallow surfaces can be inspected that range up to 10 feet. The generated scans can be used to detect and locate shallowly buried objects

such as reinforcing steel bars (rebar) and defects in bridge decks. The GPR antennae are either ground- or air-coupled. Air-coupled antennae are held at a close distance above surface of the scanned medium while ground-coupled antennae are in full contact with surface of the scanned medium. Air-coupled antennae provide rapid surveys (about 55 mph on Highways), record clean signals, but they offer limited penetration capabilities. On the other hand, ground coupled antennae suffer from slow surveys but they provide higher penetration capabilities [17].

Figure 2 shows the configuration and orientation of the 1.5 GHz (GSSI model 5100) bistatic antenna used in this project [16, 18]. The GPR system consists of an antenna, a system cart, a data acquisition system, and a post processing software called RADAN. The 1.5 GHz antenna provides a penetration depth up to 1.5 feet with a range of 10-15 ns.

Bridge decks consist of concrete slabs of varying thicknesses and may sometimes be covered by an asphalt overlay. Generally speaking, a bridge deck may consist of the following interfaces:

- Air-bridge deck surface interface (concrete or asphalt).
- Asphalt-water-concrete interface.
- Concrete-rebar interface.
- Rebar-deterioration interface.
- Concrete-void interface.
- Bottom of the structure interface (substrate).

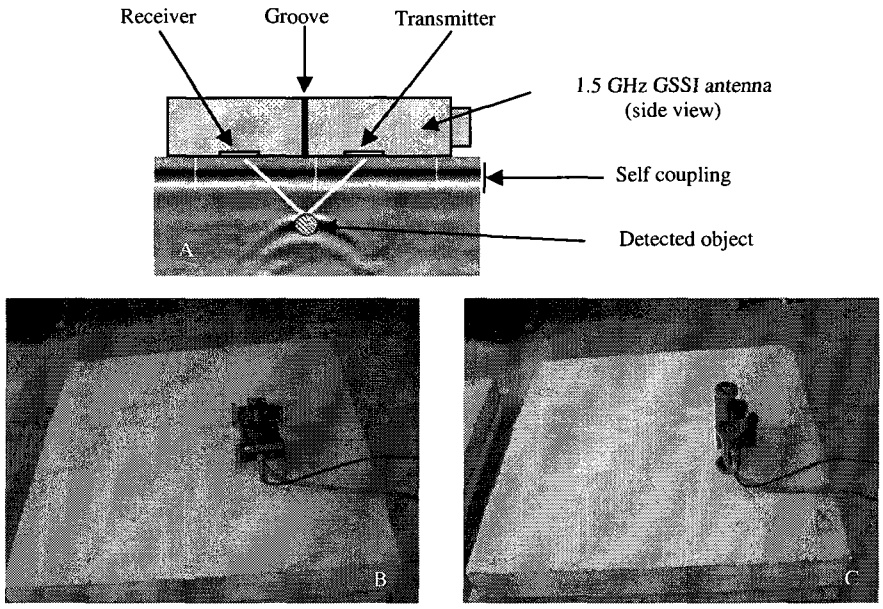

Figure 2. (A) 1.5 GHz antenna configuration, (B) cross-polarized orientation, and (C) normal orientation.

By moving the GPR antenna over bridge decks, scans can be obtained. Next, the three commonly used formats for raw GPR data presentation will be investigated. They are A-scan, B-scan, and C-scan.

A-Scan

The A-scan, also known as a trace, is obtained by placing the GPR antenna above the target surface and recording reflected signals. The A-scan is displayed as graph of amplitude versus round-trip travel time. Eq. 1 can be used to model raw GPR data if prior information about the structure of the scanned medium is not available. This equation indicates that a detected GPR pulse consists of a number of

reflected signals with each signal being a delayed and attenuated version of the initial transmitted signal:

$$g(t) = \sum_{i=1}^{n} a_i f(t - d_i) \qquad (1)$$

where $f(t)$ represents the transmitted pulse, n represents the number of detected pulses, and a_i *and* d_i represent the attenuation and delay time associated with the i^{th} received reflection respectively.

B-Scan

The B-scan, also known as a line-scan, is obtained by moving the GPR antenna over the target surface and recording the reflected signals at regular intervals. Usually, the recorded data is presented as a gray scale image of size x by y where x is the scan horizontal locations (distances) and y is the round trip travel time. Each column of the B-scan is a single A-scan taken at the x^{th} location.

The GSSI antenna transmits a pulse with a positive peak followed by a negative peak and then a small positive peak due to overshoot. Therefore, a detected object appears as a band of white-black-white or black-white-black (due to phase inversion at the boundary between two layers) lines/arcs as shown in Figure 3. Since the GPR antenna transmits a cone-shaped beam, a target can be detected when it is before, under, or after the antenna's vertical position.

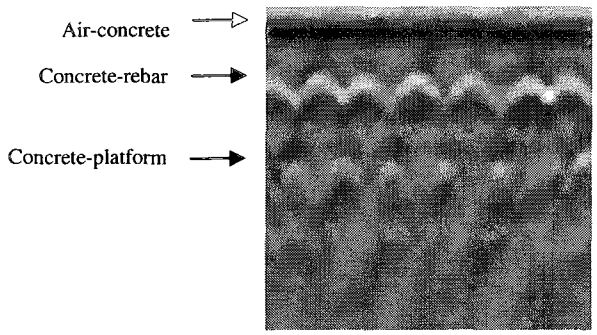

Figure 3. Raw GPR scan data from a 6-inch concrete slab.

The detected pulses by the GPR antenna will follow a travel path with either a single-reflection or with multiple-reflections. In case of single-reflection travel paths, the transmitted radar pulse reflects once of a target and travels back to the receiving antenna. In case of multiple-reflection travel paths, as the transmitted signal enters a specific layer, multiple reflections from top and bottom of that layer occur and then return back to the receiving antenna. It is also possible for a pulse to reflect off a target, reflect again off a second target, and then return back to the receiving antenna. Due to the attenuation of each reflection, single-reflection paths tend to be the strongest signals detected in a raw scan. Ringing emerge as repetitive reflection patterns throughout the GPR scan which may obscure the visual appearance of weaker reflection such as those from targets at lower depths.

Once the transmitted radar pulse penetrates the target, it is reflected and refracted at every boundary between any two materials with different dielectric constants in its path. As Figure 3 shows, the self-coupling bands at the top of the

image represent air-concrete interface. The hyperbolic arcs represent the concrete-rebar interface. We expect to see horizontal lines representing reflections from the concrete-base interface. These reflections show non-continuous segments since reflections from targets under rebar peaks are masked by their strong reflections. The interference between the hyperbolic tails causes the appearance of spurious hyperbolic arcs which may obscure our ability to observe the reflections from underneath targets. The migration method can be used to eliminate such artifacts [16].

C-Scan

The C-scan can be constructed from a collection of B-scans. It is a 3D image used to investigate diagnostically buildings, walls, and bridges. The main advantage is to locate accurately fine details inside structural objects. In order to construct a C-scan image, two points should be achieved. First, a dense grid of antennae is required to perform a successful 3D survey where antenna orientations and positions should be setup accurately. Secondly, 3D migration software is required. The slow development in the 3D GPR equipments is the main reason for the limited use of C-scans [16].

GPR Data Preprocessing

Raw GPR data is highly subjective to degradation for two reasons. First, the internal structures of the scanned medium are unknown to the GPR antenna. Secondly, hardware limitations result in a low resolution scan (low contrast) [19]. In general, GPR data processing can be categorized into basic and advanced data

processing. The most commonly used basic data preprocessing methods are time gaining and temporal and spatial filtering of the data [20]. The most commonly used advanced data processing methods are deconvolution, background removal, and velocity analysis. It is worth noting that choosing the appropriate method(s) is an application dependent. Background removal methods include average trace removal and orthogonal trace decomposition [20]. In the next subsections, time gain and migration will be discussed.

Time Gain (Gain Boosting)

As radar signals spread into the scanned medium, their energies are quickly attenuated due to scattering and absorption. As a result, reflected signals from deep objects are barely visible. Time gaining is the process of equalizing signal amplitudes where a time dependent gain function is used in order to enhance signature of weak signals reflected from deep objects. The non-uniform variation of attenuation with depth is the main difficulty of this method since some layers have low attenuation while others have high attenuation [20]. A simple equation that can be used to solve this problem is shown in Eq. 2.

$$g(t) = \frac{A.e^{-\alpha.v.t}}{v.t} \tag{2}$$

where A represents signal amplitude, α represents attenuation factor, v represents velocity of radar waves, and t represents time. RADAN can provide the following

gain boosting methods: automatic gain, linear gain, and exponential gain. On the other hand, in [15] the following gain boosting equation is used.

$$g(t) = 2.27e^{-.02705t} \tag{3}$$

Unfortunately, the enhanced image may not be appropriate for mathematical analysis since the time gaining process may modify signals' amplitude and consequently their shape.

Migration

Reflections from deep objects may be obscured by diagonal components such as side reflections from boundaries in the scanned medium, as shown in Figure 13, which should be detected and eliminated. Migration can be used to reduce the diagonal components and diffractions by returning the detected features back into their accurate spatial location [21]. A general migration algorithm has been formulated [21, 22] to follow the following steps:

1. An opaque filter is applied to the diagonal regions to set their pixels to zero where shape, angle, and size of the filter are user defined,
2. Fourier transform is used in order to eliminate the remaining diagonal components while preserving the horizontal components.

Migration can be implemented using the frequency-wavenumber (F-K) filtering method which can be done by cascading filters in the frequency and wavenumber domains where long wavelength components and high frequency

components are filtered out based on the assumption that most of the energy is concentrated in regions with small wavenumber and low frequency [22].

Principal Component Analysis (PCA)

The covariance can be used to measure the correlation between elements of a centered vector X as shown in Eq. 4.

$$C = E[(x-\mu)(x-\mu)^T] \tag{4}$$

where E is the expected value and μ is mean of the vector X. The diagonal elements of the covariance matrix contain variances of components of X. For two vectors X and Y, their covariance matrix is shown below.

$$C_{xy} = E[(x-\mu)(y-v)^T] = E[xy^T] - \mu v^T \tag{5}$$

When the two vectors are uncorrelated (they have a zero correlation coefficient), the following equation holds: $E[xy] = E[x]E[y]$. This implies that their covariance matrix is equal to the identity matrix.

PCA is an orthogonal transform and a decorrelation method that projects the high dimensional data into a lower dimensional space. Since data redundancy can be measured based on the correlation between components of a vector, most of the information contained in the original vector can be represented by a much smaller vector after the PCA stage. PCA algorithm can be implemented by variance maximization, by minimum mean-square error compression, by stochastic gradient ascent algorithm, by subspace learning algorithm, by recursive least-squares method,

or by back propagation learning algorithm, [23]. In variance maximization, PCA captures maximum variance of the data components in a finite number of orthogonal (uncorrelated) principal components. In minimum mean-square error compression, a set of orthogonal basis vectors is found that minimizes the error between the original vector and its projections. In general, the PCA algorithm can be summarized as follows:

1. Remove sample mean of each signal vector.
2. A linear transformation is applied in order to rotate the coordinate system where the first axis aims toward the maximum variance and the second axis contains the largest portion of the remaining variance and is orthogonal to the first axis.

After the PCA stage, v (out of n) principal components (largest Eigenvalues) are retained. Eq. 6 can be used to measure the retained variance [24].

$$\rho_m = \frac{\sum_{i=1}^{v} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{6}$$

where λ_i represents an eigenvalue and $\lambda_i \geq \lambda_{i+1}$.

Independent Component Analysis (ICA)

Linear ICA Model

Second-order statistical methods such as PCA assume a Gaussian distribution for the data. Unfortunately, in real life the data have non-Gaussian distribution and hence these methods fail to separate components of a multivariate data. In other

words, uncorrelatedness is not enough to separate components with non-Gaussian distribution while independence is enough. Thus independence implies nonlinear uncorrelatedness (decorrelation).

If the joint density function of two random variables equals to the multiplication of their marginal density functions, they are independent. Many methods have been proposed in the literature to decompose source signals (impulse response signals) within observed signals (mixture). The most used classical analytical methods are Wavelets and Fourier methods. On the other hand, the most used statistical methods are PCA and ICA algorithms. Transforms such as Wavelets estimate basis vectors (mixing matrix coefficients) independently from the data while ICA estimates basis vectors from the data under consideration. This makes ICA a data-driven technique and consequently adaptable to any kind of data. The linear ICA model can be formulated as

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = \sum_{j=1}^{n} a_{ij}s_j, \qquad for all i = 1,2,\cdots,n \tag{7}$$

Eq. 7 can be written in matrix form as follows

$$x = As \tag{8}$$

where $x \in R^n$ is the observed vector with its elements are in fact the mixture $x_1,\cdots,x_n$, $s \in R^n$ is a zero mean vector contains statistically independent and stationary impulse response signals $s_1,\cdots,s_n$, and $A \in R^{nxn}$ is a full rank (i.e., nonsingular matrix) matrix called the mixing matrix. In other words, x contains a linear mixture of a number of impulse response signals (original signals to be recovered).

Let y represents an estimate of s (with scaling and permutation ambiguity of its components) that is related to x according to Eq. 9 where W is the separating (demixing) matrix defined as the (pseudo) inverse of A.

$$y = Wx = WAs \tag{9}$$

Since both A and s are unknown, variances of the independent components cannot be found since scaling s by a factor may be cancelled by dividing A over the same factor. In the same way, there is a sign ambiguity. This means that the order, sign, and variances of the impulse response signals cannot be determined. For this reason, sample variance of y is assumed (without loss in generality) equal to one. Also since arithmetic mean of the mixture is irrelevant to its mutual information, it can be removed [25]. Usually, a whitening step achieves the previous two points.

The ICA algorithm projects the data into a subspace of statistically independent components. Statistically independent components means that value of any one of the components gives no information about value of the other components. In other words, ICA algorithm finds a linear transformation W that maximizes the non-Gaussianity (super-or-sub Gaussianity) of the impulse response signals s so that they are as statistically independent as possible which results in obtaining the independent components as shown in Eq. 9.

A super-Gaussian probability density function has a sharp peak with longer tails in comparison with the Gaussian density function. On the other hand, sub-Gaussian density is flat or multimodal in comparison with the Gaussian density. The

uniform distribution is an example of a sub-Gaussian distribution while the Laplacian distribution is an example of a super-Gaussian distribution.

The ICA decomposition holds under three assumptions. First, the impulse response signals must be as statistically independent as possible. Second, at most one impulse response signal may have a Gaussian distribution while the others must have non-Gaussian distribution. Third, number of observed signals should be greater than or equal to number of impulse response signals. On the other hand when the impulse response signals have a Gaussian distribution, the use of second-order statistics is enough to find their mean and covariance matrix which are enough to uncorrelate and separate the impulse response signals. This implies that uncorrelated Gaussian data is also independent data.

The central limit theorem states that if a random variable consists of a sum of independent random variables ($s = s_1 + s_2 + ...$), its distribution will be closer to a Gaussian distribution in comparison with distribution of the random variables since the convolution of their densities is a smoothing operation ($p(s) = p(s_1) * p(s_2) * ...$). In other words, the mixing process results in more Gaussian distribution while the separation process increases non-Gaussianity of the mixed signals. In case of ICA, a mixture of the measured signals with minimal Gaussian properties has to be found.

ICA can be implemented either on-line or off-line (batch). Batch algorithms assume the whole data is available for the ICA estimation phase. On the other hand, on-line algorithms assume that the mixing matrix may change during the ICA

estimation phase. In this case, stochastic gradient algorithms are the most useful methods [23].

Optimization Methods

Estimation of the W matrix cannot be done in a closed form (as a function of the training data) since we have one equation (Eq. 9) with two unknowns. Fortunately, it can be done by optimizing (maximizing or minimizing) an objective function with respect to W. In this section, the most common optimization methods in the ICA field will be explored.

Gradient Methods

Gradient methods represent the basic optimization methods [23]. Let's assume g a function defined as follows

$$g(w) = g(w_1, w_2, ..., w_n) \tag{10}$$

where $w = [w_1, w_2, \cdots, w_n]$. Gradient of g can be found using

$$\nabla g = \frac{\partial g}{\partial w} = \left(\frac{\partial g}{\partial w_1} \quad \cdots \frac{\partial g}{\partial w_n} \right)^T \tag{11}$$

The second-order gradient (the Hessian matrix) can be found using

$$\nabla^2 g = \frac{\partial^2 g}{\partial w^2} = \begin{pmatrix} \frac{\partial^2 g}{\partial w_1^2} \cdots \frac{\partial^2 g}{\partial w_1 w_n} \\ \vdots \\ \frac{\partial^2 g}{\partial w_n^2} \cdots \frac{\partial^2 g}{\partial w_n^2} \end{pmatrix} \tag{12}$$

Taylor series expansion can also be used to derive gradient-type learning algorithms as shown below

$$g(w') = g(w) + (\frac{\partial g}{\partial w})^T (w' - w) + \frac{1}{2}(w' - w)^T \frac{\partial^2 g}{\partial w^2} + \ldots \quad (13)$$

The optimization step consists of step size (length) and its direction. Gradient methods can be used to estimate the w vector as follows [23]. First, we start from an initial value of w and compute its gradient at this point. Then, we move to the (or opposite in case of minimizing the function) direction of its gradient by a small value (step size). Then, the gradient is calculated at the new value of w and the parameter vector (w) is updated, and so on until w converges to a stable point (minimum or maximum point). In case the function to be optimized involves random variables, stochastic gradient algorithms can be used. Gradient descent methods can be used to minimize a function while gradient ascent methods can be used to maximize a function. In case of maximizing $g(w)$, the positive sign is used while the negative sign is used to minimize it as shown in Eq. 14.

$$w_{k+1} = w_k \pm \eta \frac{\partial g(w_k)}{\partial w_k} \quad (14)$$

where η represents the learning rate (step size) and g represents the objective function to be optimized. If the objective function is not simple and non-smooth, a local minimum (maximum) point will be reached before a global minimum (maximum) point. This means that the initial value of w is very important in gradient methods. Also, the selection of the corresponding learning rate affects the estimation process. In case of a small value, the convergence speed will be slow. In case of a

large value, an overshooting may occur that prevents the convergence to a stable point [23].

The convergence speed of gradient methods can be boosted using the momentum method, adaptive learning rate, or choosing the initial value of *w*. For example, Amari proposed the following adaptive learning rate [26].

$$\eta_{k+1} = \eta_k \exp(\alpha(\beta g(w_k) - \eta_k)) \tag{15}$$

where η is the learning rate and α and β are constants. On the other hand, the momentum method is defined according to Eq. 16 [27]. In case there is a narrow and long valley in the *w* surface, the gradient is roughly perpendicular to the long axis of that valley and moves slowly. Then, it starts to oscillate along the short axis. The momentum term increases convergence speed along the long axis and average out the oscillations along the short axis.

$$w_{k+1} = w_k \pm \eta \Delta w_k + \alpha \Delta w_{k-1} \tag{16}$$

The numerical analysis field reported different methods that offer fast convergence speed in comparison with the gradient methods but they are computationally demanding. For example, the Newton method (Eq. 17) is derived from the first three terms of the Taylor series (Eq. 13) [28]. This methods requires that the Hessian matrix to be positive definite to attain fast convergence speed. Unfortunately, inverse of the Hessian matrix results in an ill-conditioned or close to a singular matrix. To solve this problem, Marquardt-Levenberg algorithm adds a diagonal matrix αI to the Hessian matrix before inverting it where α should be small [29].

$$w_k^{'} = w_k - \eta[\frac{\partial^2 g(w_k)}{\partial w_k^2}]^{-1} \frac{\partial g(w_k)}{\partial w_k} \quad (17)$$

Other methods such as the Gauss-Newton method and the conjugate gradient method offers compromise between gradient methods and Newton method. On the other hand, the FastICA method (will be shown later) approximates the Newton method to get fast convergence speed with less computational complexity.

Natural Gradient Method

Conventional gradient methods compute the gradient of the vector w in the Euclidean orthogonal coordinate system [30]. Since the parameter space of the gradient is curved and distorted, the gradient in the Riemannian metric space (differential geometry) of the parameters can optimize the objective function in a better way than being in the Euclidian space [26, 31].

The natural gradient provides a better steepest direction for the nonlinear function $g(w)$ and consequently provides a higher performance than standard gradient methods. However, the gradient matrix must be nonsingular. The natural gradient learning rule is shown in Eq. 18 [23].

$$\Delta W_k = \pm\eta \frac{\partial g(w_k)}{\partial w_k} W_k^T W_k \quad (18)$$

Efficiency of gradient methods depends on structure of the optimized objective function. On the other hand, the natural gradient can handle large class of objective functions with high efficiency [31]. Comparing the natural gradient method

with the Newton method, the later one may converge to spurious local minima/maxima since its Hessian matrix may not be positive definite for all *W*. on the other hand, Riemannian metric space is always positive definite [31].

Stochastic Gradient Descent Methods

Stochastic gradient methods work with specific objective functions that are twice differentiable according to Eq. 19. Most stochastic gradient methods have slow convergence speed in comparison with the gradient methods but they have low computational complexity.

$$W_k = W_{k-1} \pm \alpha_k \frac{\partial}{\partial w} g(W_{k-1}, x) \tag{19}$$

Bell and Sejnowski derived the following online stochastic gradient ascent learning rule [32].

$$\Delta W = \left(\frac{\partial y}{\partial x}\right)^{-1} \frac{\partial}{\partial w}\left(\frac{\partial y}{\partial x}\right) \tag{20}$$

If the optimization step is constrained by some conditions, the Langrage method or the projection method can be used to meet these conditions. For example, if we have the constraint $\| w \|^2 = 1$, we normalize w after each step as shown in Eq. 21. This is equivalent to orthogonal projection of w onto the unit sphere to keep the variance of the independent components constant.

$$w_{k+1} = \frac{w_k}{\| w_k \|} \tag{21}$$

The literature reported other optimization methods such as relative gradient and exhaustive search by rotation [33, 34]. On the other hand, the literature reported different methods for ICA estimation such as non-Gaussianity maximization, mutual information minimization, maximum-likelihood estimation, tensorial methods, nonlinear decorrelation, and nonlinear PCA [23]. The next sections explore the most used ICA estimation methods.

ICA by Non-Gaussianity Maximization

As mentioned before, maximizing non-Gaussianity of the independent components increases their statistical independence. In other words, densities of a mixture with dependent components are more Gaussian than densities of a mixture with independent components. Typically, non-Gaussianity is measured using kurtosis and negentropy.

Measuring Non-Gaussianity using Kurtosis

Kurtosis is the fourth-order cumulant which can be estimated using Eq. 22. Usually, fast gradient methods are used to maximize Kurtosis [23].

$$k(s) = E[s^4] - 3(E[s^2])^2 = \mu_4 - 3(\mu_2)^2 \tag{22}$$

where μ_4 represents the 4th moment and μ_2 represents the second moment for a centered vector. If the data is whitened, a unity variance is obtained. Thus, the fourth moment can be used to characterize the whitened data as shown in Eq. 23.

$$k(s) = E[s^4] - 3 \tag{23}$$

On the other hand, Eq. 24 shows the normalized kurtosis.

$$\tilde{k}(s) = \frac{E[s^4]}{(E[s^2])^2} - 3 \tag{24}$$

In case of a normalized random variable with Gaussian distribution, its variance is equal to one making its kurtosis equal to zero as shown below.

$$k(s) = E[s^4] - 3(E[s^2])^2 = (E[s^2])^2 - 3(0) = 0$$

This implies that kurtosis can be used to measure non-Gaussianity. Sub-Gaussian distributions have negative kurtosis while super-Gaussian distributions have positive kurtosis. In other words, non-Gaussianity can be measured using the absolute value of kurtosis. The gradient method can be used to derive a general learning rule by maximizing non-Gaussianity of the impulse response signals as shown below [23].

$$\Delta W = \frac{\partial |k(w^T x)|}{\partial w} = 4 sign(k(w^T x))(E[x(w^T x)^3] - 3w \| w \|^2) \tag{25}$$

Since the variance of $w^T x$ must be unity, data must be whitened prior to ICA and then w must be normalized. Since we normalize w, its direction is the goal and not its magnitude. Therefore, Eq. 25 can be simplified into [23]

$$\begin{aligned} w_{k+1} &= w_k + \eta sign(k(w_k^T x))(E[x(w_k^T x)^3]) \\ w_{k+1} &= \frac{w_k}{\| w_k \|} \end{aligned} \tag{26}$$

Unfortunately, the previous rule has a slow convergence speed and totally dependent on value of the learning rate η. A more efficient fixed-point algorithm can be derived where effect of the learning rate is eliminated as shown in Eq. 27 which represents one version of the FastICA algorithm [23].

$$w_{k+1} = w_k + [x(w_k^T x)^3] - 3w_k$$
$$w_{k+1} = \frac{w_k}{\| w_k \|} \tag{27}$$

Measuring Non-Gaussianity using Negentropy

Although kurtosis offer simple computational complexity but they are very sensitive to outliers (few large values in the mixture affects the kurtosis dramatically). Also, they measure tails of a distribution but not its center.

Entropy is defined as the amount of uncertainty (randomness) a random variable has. For a random variable x, its entropy is defined as

$$H(x) = -\sum_i P(x = \alpha_i) \log P(x = \alpha_i) \tag{28}$$

where α_i are the possible values of x. on the other hand, entropy of the transformation $y = Wx$ is defined as

$$H(y) = H(x) + \log |\det(W)| \tag{29}$$

Since the Gaussian distribution is the least structured distribution, it has more randomness than a non-Gaussian distribution and thus it has the maximum entropy. This implies that entropy can be used to measure non-Gaussianity. On the other hand, differential entropy (negentropy) will be zero for a Gaussian distribution and positive

for a non-Gaussian distribution as defined by Eq. 30. Therefore, non-Gaussianity of the independent components can be maximized by maximizing their negentropy.

$$J(x) = H(x_g) - H(x) \tag{30}$$

where x_g is a random vector with Gaussian distribution that has the same mean and variance as x and a covariance matrix C. Its entropy is defined as [23]

$$H(x_g) = \frac{1}{2}\log|\det C| + \frac{n}{2}(1 + \log 2\pi) \tag{31}$$

where n is the dimension of the vector x. Estimating non-Gaussianity based on equations 30-31 is computationally demanding. Fortunately, negentropy can be approximated by estimating its probability density function. For example, the Gram-Charlier expansion can be used to approximate negentropy as shown by Eq. 32 where x must have a zero mean and unit variance (standardized vector) [23].

$$J(x) \approx \frac{1}{12}E[s^3]^2 + \frac{1}{48}k(x)^2 \tag{32}$$

Unfortunately, Eq. 32 is totally dependent on the kurtosis which has drawbacks as mentioned before. On the other hand, Hyvarinen et al. proposed Eq. 33 to approximate the negentropy [23].

$$J(x) \approx \frac{1}{2}\sum_i c_i^2 \tag{33}$$

where $c_i = E[G(y_i)]$ and $G(y)$ is a nonlinear and moderately growing function. On the other hand, Equation 34 shows a simple method to approximate negentropy where y is a standardized vector with Gaussian distribution and x is a standardized vector. Eq. 34 offers fast, robust, and simple way of measuring non-Gaussianity [23].

$$J(x) \propto (E[G(x)] - E[G(y)])^2 \tag{34}$$

The derivative of G (i.e., g) must be growing (or decreasing) in a non fast way (i.e., has a unique inverse) and sufficiently smooth even/odd function. The literature reported the following choices of G with better results than results of equation 32 [23].

$$G(x) = \frac{1}{\alpha_1} \log(\cosh(\alpha_1 x)) \tag{35-a}$$

$$G(x) = -e^{-\frac{x^2}{2}} \tag{35-b}$$

$$G(x) = \frac{1}{4} x^4 \tag{35-c}$$

where α is in the range [1, 2]. In [35], the following nonlinear function is proposed which is robust against outliers and works with different density functions.

$$G(x) = \log(x + 0.1) \tag{36}$$

Based on Eq. 34, a fixed-point algorithm can be derived which represents another version of the FastICA algorithm [23].

$$\begin{aligned} w_{k+1} &= w_k + E[z g\ (w_k^T z)] - E[g'(w_k^T z)] w_k \\ w_{k+1} &= \frac{w_k}{\| w_k \|} \end{aligned} \tag{37}$$

where g(.) and g'(.) represents the first and second derivatives of the nonlinear function G(.) respectively. All the previously discussed ICA estimation algorithms estimate only one independent component. In order to estimate more than one independent component, w_1 is found as mentioned before. Then, w_2 has to be found

such it is orthogonal to w_1. In the same way, w_3 must be orthogonal to w_1 and w_2. The Gram-Schmidt orthogonalization or the symmetric orthogonalization can be used to ensure the orthogonality [23].

ICA by Maximum Likelihood Estimation

It is the most commonly used approach in ICA estimation. In this approach, the parameters that result in the highest probability for the observations of a vector are kept. The probability density function of the mixture ($x = As$) is defined as shown below where p_i represent probability densities of the independent components.

$$p_x(x) = |\det W| \prod_i p_i(w_i^T x) \tag{38}$$

If we have T observations of x, its likelihood can be written as

$$L(W) = \prod_{t=1}^{T} \prod_{i=1}^{n} p_i(w_i^T x(t)) \, |\det W| \tag{39}$$

By taking logarithm of Eq. 39,

$$\log L(W) = \sum_{t=1}^{T} \sum_{i=1}^{n} \log p_i(w_i^T x(t)) + T \log |\det W| \tag{40}$$

Eq. 40 can be further reduced into [21]

$$\frac{1}{T} \log L(W) = E[\sum_{i=1}^{n} \log p_i(w_i^T x(t)) + \log |\det W|] \tag{41}$$

The process of estimating p_i is complex since a very large number of parameters have to be estimated (i.e., nonparametric estimation). The Bell-Sejnowski

algorithm (Infomax ICA) utilizes the gradient method to maximize the likelihood of Eq. 41 as follows where the function g is defined according to Eq. 43-B [32].

$$\Delta W \propto (W^T)^{-1} + g(Wx)x^T \tag{42-a}$$

$$g_i(s_i) = \frac{\partial}{\partial s_i}\log p_i(s_i) = \frac{p_i'(s_i)}{p_i(s_i)} \tag{42-b}$$

Unfortunately, Eq. 42 has a slow convergence speed and computationally demanding due to the matrix inversion. On the other hand, the natural gradient method can be used to simplify the previous learning rule. If we multiply the right hand side of Eq. 42 by $W^T W$, we obtain

$$\Delta W \propto (I + E[g(y)y^T])W = \eta(I + E[g(y)y^T])W \tag{43}$$

The previous algorithm converges when the actual response $E[g(y)y^T]$ is equal to the target response I. Regardless of the nature of the matrix W whether it is close to singular or ill-conditioned, the dynamic behavior of Eq. 43 still the same [36]. In order to obtain independent components with unit variances, the previous learning rule is modified into [37]

$$\Delta W \propto (I - diag\{E[y_1^2],...,E[y_n^2]\} + E[g(y)y^T] - diag\{E[g(y)y^T]\})W \tag{44}$$

The choice of the nonlinear function g depends on distribution of the independent components. As Eq. 45 indicates, g_1 can be used for independent components with super-Gaussian distribution while g_2 and g_3 can be used for sub-Gaussian independent components.

$$g_1(y) = -2\tanh(y) \tag{45-a}$$

$$g_2(y) = \tanh(y) - y \tag{45-b}$$

$$g_3(y) = -y^3 \tag{45-c}$$

A fixed-point algorithm can be derived utilizing the approximative Newton version of Eq. 42 which represents another version of the FastICA algorithm [23].

$$W_{k+1} = W_k + diag(\alpha_i)[diag(\beta_i) + E[g(y)y^T]]W_k \tag{46}$$

where $\beta = -E[y_i g(y_i)]$, $\alpha_i = -\frac{1}{E[g'(w_k^T z) + \beta_i]}$ and $y = Wz$, and z is the whitened mixture.

In general, FastICA algorithms have the following properties [38, 39]: Easy implementation since they do not depend on the learning rate, have high accuracy, and have fast convergence speed.

Eq. 47 can be used to select the best nonlinear function G for the ICA estimation among a set of nonlinear functions [23]. It states that a nonlinear function G that minimizes the trace of the asymptotic variance of W is chosen as the best candidate. This equation indicates that negentropy-based nonlinear functions are better than cumulant-based nonlinear functions.

$$V_G = \alpha \frac{E[g^2(y_i)] - (E[y_i g(y_i)])^2}{(E[y_i g(y_i) - g'(y_i)])^2} \tag{47}$$

where α is a constant. In summary, by comparing the nonlinear functions of Eq. 35, we conclude the following.

1. The hyperbolic function $g(s) = \tanh(s)$ can be used regardless whether the independent components have sub-or super-Gaussian distribution.

2. The following nonlinear function produces promising results [39].

$$g(y) = \frac{y}{1+y^2} \tag{48}$$

3. The exponential function $G(y) = -e^{-\frac{y^2}{2}}$ can be used with super-Gaussian independent components or in case the robustness against outliers is important.
4. Kurtosis can be used in case of sub-Gaussian independent components and the robustness against outliers is not an important issue.

Nonnegative ICA

Nonnegative ICA can be used in case the data has nonnegative properties as expressed in the following definition [40].

If $Pr(s<0) = 0$, then s is a nonnegative impulse response (49)

Yuan proposed the following nonnegative FastICA algorithm [41].

1. The input data is whitened. This results in the whitened vector z.
2. Initialize the weight vector, orthogonalize, and normalize it.

$$w_{k+1} = w_k - \sum_{j=1}^{n-1}(w_k^T w_j)w_j$$
$$w_{k+1} = \frac{w_k}{\| w_k \|} \tag{50}$$

   where n represents number of independent components.

3. Set $w_k = -w_k^T$ if $\max_{z \neq 0}(w_k^T z) \leq 0$.

4. Set $w_{k+1} = -w(r)(w(r)^T w_k)w_k$ if $\min_{z\neq 0}(w_k^T z) \geq 0$, where $w(r)$ represents a vector in the null space null (Z) with $Z =: \{z \neq 0 : w_k^T z = 0\}$.

5. The weight vector is updated and normalized.

$$w_{k+1} = w_k + E[(z - E[z])g(w_k^T z)] - \eta E[g'(w_k^T z)]w_k$$
$$w_{k+1} = \frac{w_k}{\| w_k \|} \tag{51}$$

where $g(y) = -\min(0, y)$

6. Stop the procedure when w converges or maximum number of iteration reached. Otherwise, go to step 3.

The proposed nonnegative FastICA is more computationally demanding than the traditional FastICA but requires less number of iterations [41].

Preprocessing Methods for ICA

In this section the most used preprocessing methods for ICA algorithm will be explored.

Data Whitening

The whitening process uncorrelates components of the centered vector resulting in an identity covariance matrix. The whitening process may also reduce a full-rank mixture into an orthogonal mixture (i.e., orthogonal columns) [42]. After the whitening process, the sample mean of the data is removed.

Any transformation that results in an identity covariance matrix can be used as a whitening method. A simple, fast, and efficient decorrelation transform is the Mahalanobis transformation as shown in Eq. 52 [43].

$$Z = C^{-0.5}(X - \overline{X}) \tag{52}$$

where C is the covariance matrix defined as

$$C = \frac{(X - \overline{X})(X - \overline{X})^T}{N} \tag{53}$$

where $\overline{X}$ represents sample mean of the mixture X and Z is the whitened data. To prove that X is decorrelated after applying this transformation, we can rewrite Z as

$$Z = HXC^{-0.5}$$

where H represents the centering matrix defined as

$$H = I - \frac{1}{n}J$$

where J is a matrix of all 1's. Utilizing the centering matrix equation, the covariance matrix of Z can be rewritten as [44]

$$\begin{aligned} C &= \frac{1}{n}Z^T HZ \\ &= \frac{1}{n}(C^{-.5}X^T H^T)H(HXC^{-.5}) \\ &= C^{-.5}(\frac{1}{n}X^T HX)C^{-.5} \\ &= C^{-.5}CC^{-.5} = C^{.5}C^{-.5} = I \end{aligned}$$

On the other hand, PCA algorithm can be used to decorrelate impulse response signals and to reduce their dimensionality. Reducing data dimensionality prevents over learning [23]. Over learning means that number of impulse response

signals to be estimated is larger than number of observed signals (mixture). In case of over learning, PCA is used to obtain a square mixing matrix which is a constraint for the ICA algorithm results to be valid. On the other hand, in case the mixtures are more than the available impulse response signals, no improvement in the signal separation results should occur [39].

Since uncorrelation is a necessary condition for independence, a whitening procedure, as a preprocessing for ICA algorithm, has proven to be useful. It decorrelates the data and reduces the dimensionality leading to a reduction in the number of parameters to be estimated and in an increase in the convergence speed of the ICA algorithm [23].

Time Filtering

Since GPR scans are time signals, time filters can be used. These filters do not change coefficients of the mixing matrix but filter the independent components of the impulse response signals [23]. A high-pass filter can be used to sharpen and decrease the dependency between the impulse response signals in case the dependent components are located in the high frequency band [46]. This approach is computationally very efficient but it may introduce noise. On the other hand, a band-pass filter may be more suitable than the high-pass filter for certain signals.

Fractal Analysis

In this section, the concept of fractals and their use as a feature extraction tool will be explored.

Introduction to Fractals

In 1967, Mandelbrot used the word *fractal* to describe objects that are detailed at different scales. In other words, finer new features are revealed as the fractal object is magnified where shape (or statistics) of the smaller features is approximately similar to that of the larger features. This also implies that no new features will be revealed as a non-fractal object is magnified. Since fractals reveal more details at smaller scales, they are too complex to fit into a traditional geometrical model [47].

Euclidean geometry can be used to describe man-made shapes such as circles, cubes, etc. Unfortunately, some nature-made shapes such as defects in bridge decks are complex, non-uniform (irregular), and have rough edges. Therefore, a mathematical tool is required to describe heterogeneity of real objects that classical Euclidean geometry fails. At this point, fractal geometry begins as a complexity analysis tool where the fractal dimension is used to measure the scaling property of features of an object.

Fractal sets can be divided into self-similar sets and self-affine sets. Self-similar sets describe data that repeat themselves when different axes are magnified by the same factor while self-affine sets describe data that preserve their shape (or their statistics) only when different axes are scaled differently. Fractals have three kinds of

self-similarity. In exact self-similarity, a fractal set appears exactly the same at different scales. In quasi-self-similarity, a fractal set appears approximately identical at different scales. In this case, each scale contains small copies of the whole fractal set in a deteriorated form. Finally in statistical self-similarity, a fractal set has statistical properties which are preserved at all scales. This kind is regarded as the most realistic definition of self-similarity.

Mathematically if a signal $x(t)$ has a fractal feature structure, then it should satisfy the scaling law under the scale conversion ($t \rightarrow \lambda t$)

$$X(\lambda t) = \lambda^H x(t) \tag{54}$$

where H is the Hurst exponent. If H is equal to 1, then $x(t)$ is a self-similar function otherwise it will be a self-affine function. The general solution to the previous equation is [48]

$$x(t) = ct^H \tag{55}$$

where c is a random constant. In case of 2D images, fractal analysis offers a global description of the inhomogeneities in an image which means that complexity of images (more precisely their texture composition) can be characterized well using the fractal geometry.

Both roughness and topology of interfaces inside the scanned medium are imprinted in the recorded GPR traces where these traces are considered as self-affine functions of time [49].

Hausdorff-Besicovitch Dimension

The Hausdorff-Besicovitch dimension (HBD) of a fractal set is a fractional number greater than its topological dimension that can be used to measure irregularity of that set. Number of independent variables required to describe a point in a set is the topological dimension of that set. The Hausdorff dimension D of a set A can be defined as the infimum of $D \in [0,\infty)$ such that the D-dimensional Hausdorff measure of the set A is equal to 0. If we cover a set A with a number of closed sets each of diameter at most r where $r > 0$, Then, the D-dimensional Hausdorff measure will be the infimum of the summation of the D^{th} powers of the diameters, i.e., $HM := \inf\left\{\sum_i r_i^D \ , \ r_i > 0 \ \right\}$, [47].

The Hausdorff dimension can be defined in a simpler way. Number of the closed sets $N(r)$ to cover the set A increases as r decreases. Generally speaking, as r approaches the zero then $N(r)$ increases in the same rate as $1/r^D$ increases. In this case, the set A has a Hausdorff dimension equal to D. In practice, it is difficult to measure the HBD of a fractal set in its rigorous definition [47]. Therefore, several methods have been proposed in the literature to approximate it. Applying these methods to the same set does not necessarily result in the same estimated value of fractal dimension (FD). These differences are due to the estimation algorithm used by a particular method. Therefore, choice of the suitable method is an application dependent.

For a signal, the estimated fractal dimension will be in the range between 1 and 2. If the estimated FD is close to 1, a high neighbor-to-neighbor correlation between pixels will be present (smooth signal). If the estimated FD is close to 2, a high negative correlation will be present (non-smooth signal). Figure 4 shows healthy and defective traces. Both traces have a topological dimension of 1, meaning that classical Euclidean geometry is unable to differentiate between them. On the other hand, both traces have different fractal dimensions with higher value for the defective one. In other words, the estimated fractal dimension describes, in a compact way, the relation between signal variance and the time scale. Therefore, fractal techniques are purely a statistical tool [50]. In general, for an n-dimensional function, its estimated fractal dimension will be in the range [n, n+1].

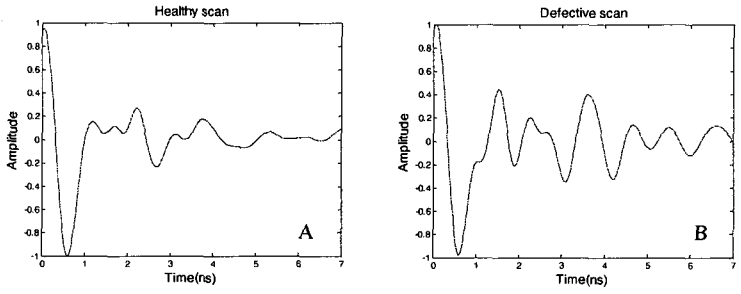

Figure 4. Healthy and defective scans with fractal dimensions of 1.56 and 1.65, respectively. Both signals have a topological dimension of 1.

In general, fractal dimension measuring methods are based on scale, measurement relationship, correlation function, distribution function, or the power spectrum [51]. Generally speaking, these methods follow the following procedure in approximating the Hausdorff-Besicovitch dimension of a fractal set.

1. For various step sizes, measure quantities of the object (such as length).
2. The logarithmic plot of the measured quantities versus step sizes is obtained.
3. A line is fitted through the data points using least-squares regression.
4. Slope of the fitted line is used to estimate the fractal dimension.

Next, the most commonly used methods to estimate the fractal dimension will be summarized.

The Divider Method

Sometimes it is called the structured walk method. This method is more suitable to signals that are considered as not perfect self-similar fractals (self-affine signals) such as GPR traces. In order to estimate a trace length, first it is approximated with several straight-line segments (steps) and it is total length estimated as the product of number of steps and the used step length. It is preferable to use small step length to get more accurate results. Klinkenberg [52] found that the minimum step size is equal to one-half the average distance between adjacent points. Eq. 56 can be used to check if the set follows a fractal model.

$$L(s) \alpha \delta^{1-D} \tag{56}$$

where L represents length of the trace, δ represents the step size, and D represents the fractal dimension. Next, Mandelbrot-Richardson logarithmic graph is formed by plotting logarithm of the corresponding signal's length versus logarithm of its used

step size. Then, a line is fitted into the graphed points and the fractal dimension is estimated using Eq. 57 where S is slope of the graphed line.

$$D = 1 - S \tag{57}$$

Box-Counting Method (BCM)

The box-counting is the most commonly used method since it is easy to implement and intuitive. In order to estimate the FD of an object, it is covered with boxes of different sizes. In case of a smooth object, it is enough to cover it with two boxes of different sizes r_1 and r_2 in order to estimate its FD according to Eq. 58 which indicates that the fractal dimension of a straight line and a square will be 1 and 2, respectively.

$$\frac{N_2}{N_1} = \left(\frac{r_1}{r_2}\right)^D \rightarrow D = \frac{\log \frac{N_2}{N_1}}{\log \frac{r_1}{r_2}} \tag{58}$$

In case of a non-smooth object, its FD can be found by averaging the estimated FD over different scales according to Eq. 59 [53].

$$D = -\lim_{r \to 0} \frac{\log[N(r)]}{\log[r]} \tag{59}$$

where $N(r)$ is the number of boxes required to completely cover a signal and D is slop of the logarithmic plot of $N(r)$ versus r. Unfortunately, this method is only valid for statistically self-similar binary signals. The main problem of box-counting methods is the determination of best number of boxes. A good approach is to let

number of boxes less than (but not much less than) number of available intensity levels.

An extended method is proposed to overcome the difficulties associated with the BCM. First, the fractal set is divided into subsets. Then, the BCM is used to estimate the fractal dimension of each set. Finally, the fractal dimension of the fractal set is equal to the maximum fractal dimension of the estimated ones.

On the other hand, a differential box-counting method that works with self-similar and self-affine signals is proposed in the literature [54]. First, a signal is covered with three boxes. The first box completely covers the signal; the second box covers the first half of the signal while the third box covers the second half of the signal. Second, the FD is estimated using

$$D = \frac{\log(N_1 + N_2) - \log N_3}{\log 2} \tag{60}$$

The Hurst Method

First, windows of different sizes are used where the maximum difference of data within each window is computed and plotted against its corresponding window size in the logarithmic space. Then, fractal dimension is estimated using

$$D = 2 - H \tag{61}$$

where H is the Hurst exponent that can be, obtained from the slope of the graphed line. An analysis of the effect of the used window sizes can be found in [49, 55]. In case small window sizes are used, the plotted data will be a straight line. However for large window sizes, the normalized maximum difference turns into a constant [56].

Fractional Brownian Motion (fBm)

Most fractals encountered in physical models are fractal Brownian motion (fBm) functions. According to Mandelbrot, fBm is a statistically self-affine function which can be regarded as a generalization of Brownian motion. According to the variance properties of fBm, the expected value of the intensity difference between two points is nonzero only when square of the difference is proportional to the distance between the points at a power of $2H$ as shown in Eq. 62.

$$E[|\Delta X(t,\Delta t)|^2] = |\Delta t|^{2H} \; E[X(t+1)-X(t)]^2 = |\Delta t|^2 \; H\sigma^2 \qquad (62)$$

The fBm algorithm can be summarized as follows [48]. The logarithmic plot of $E[x(t+\Delta t)-x(t)]^2$ versus step size, Δt is obtained for various values of the step size Δt and the fractal dimension, FD is estimated using Eq. 61.

The Information Dimension

As mentioned before, entropy can be used to measure the amount of uncertainty (randomness) a random variable r has. In fractal theory, entropy can be defined as [57]

$$I(r) = -\sum_{i=1}^{N(r)} P(r,i)\log P(r,i) \qquad (63)$$

where $P(r,i) = \dfrac{N(r,i)}{\| S \|}$, S is a fractal set, $N(r)$ is the minimum number of cells of size r to cover S, $\| S \|$ is cardinality of S (number of its elements), $N(r,i)$ is number of points in the i^{th} cell, and $P(r,i)$ indicates the probability that a point of S inside the

i^{th} cell. In case all the probabilities are equal, each $P(r,i)$ is equal to $1/N(r)$ and thus the above equation reduces to $I(r) = \log N(r)$. Therefore, the information dimension can be defined as

$$D_{\inf} = -\lim_{r \to 0} \frac{I(r)}{\log(r)} \tag{64}$$

The Correlation Dimension

The correlation dimension can be defined as [58]

$$d_{cor} = \lim_{r \to 0} \frac{C(r)}{\log(r)} \tag{65}$$

where the correlation is defines as

$$C(r) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=j+1}^{N} \theta(r - \| R_i - R_j \|) \tag{66}$$

where N represents number of points in the set S, θ is the Heaviside step function, and $\| R_i - R_j \|$ is the Euclidean norm.

CHAPTER 3

PERTINENT LITERATURE

ICA-Based Algorithms

In [59], a modified version of the FastICA algorithm of Eq. 37 is proposed where two iterations of the FastICA algorithm are merged into single iteration meaning that the Jacobean matrix is computed once per two iterations. This increases the convergence speed of Eq. 37 while preserving its performance. It can be summarized as follows.

1. The data is whitened, i.e. $E[zz^T] = I$ where z is the whitened data.
2. The separating matrix W is initialized randomly.
3. The following learning rule is used.

$$w_{k+1} = w_k - \frac{[F(w_k) + F(w_k - F(w_k)/JF(w_k))]}{JF(w_k)} \tag{67}$$

where $F(w_k)$ is the gradient of Eq. 37 and the Jacobean matrix is defined as

$JF(w) = E[zz^T g'(w_k^T z)] - E[g'(w_k^T z)]I$.

4. The separating matrix is normalized: $w_{k+1} = w_k / \| w_k \|$.
5. If $| w_{k+1} - w_k | < \varepsilon$, the algorithm is converged. Otherwise, go to step 3.

In [60], the FastICA algorithm of Eq. 37 is compared against CumICA for blind separation of non-destructive acoustic emission signals. The CumICA is based on computing the cross-cumulants (off-diagonal elements of the cumulant matrix) of

the mixtures. The experimental results indicate that the CumICA requires a high-pass filter to achieve the same SNR as FastICA.

In [39], a statistical efficient version of the FastICA algorithm is proposed (EFICA). EFICA is a slightly higher computationally demanding than FastICA algorithm but has superior separation performance in comparison with JADE, nonparametric ICA, and FastICA algorithms [39]. It can be summarized in the following steps.

1. The symmetric FastICA is executed using the hyperbolic nonlinear function until the weight matrix converges to a stable point.
2. The following nonlinear function is evaluated for the k[th] estimated independent component.

$$g_k(y) = \begin{cases} ye^{-\eta_1|y|}, & \mu_{4k} > 3 \\ sign(y).|y|^{\min\{\alpha_k - 1.14\}}, & 1.8 < \mu_{4k} \le 3 \\ sign(y).|y|^{14}, & \mu_{4k} \le 1.8 \end{cases} \tag{68}$$

where μ_{4k} represents the fourth-order moment of the k[th] independent component and the parameter α_k is defined as $\alpha_k = [0.29\sqrt{\mu_{4k} - 1.8} - 0.185(\mu_{4k} - 1.8)]^{-1}$.

3. The FastICA algorithm is used with the nonlinear function of the previous step in order to refine the estimated independent components.

The literature reported a number of nonlinear functions that can be used in addition to the functions of Eq. 48 and the first derivative of Eq. 35. Yang et al. [61]

derived the following function which is used to estimate the marginal probability density function of y using its central moments and cumulants.

$$g(y) = f_1(k_3, k_4) \circ y^2 + f_2(k_3, k_4) \circ y^3 \tag{69}$$

where k_3 and k_4 are the 3rd and 4th cumulants and ($\circ$) indicates Hadamard product of two matrices and

$$f_1(k_3, k_4) = 0.5 k_3 (4.5 k_4 - 1) \tag{70}$$

$$f_2(k_3, k_4) = 1.5 (k_3)^2 + \frac{1}{6} k_4 (4.5 k_4 - 1) \tag{71}$$

Singh and Rai [62] derived the following nonlinear function based on the Edgeworth expansion which is suitable for sub-Gaussian distributions only.

$$g(y_i) = 3y_i^3 - \frac{11}{4} y_i^5 - \frac{25}{12} y_i^7 - \frac{65}{48} y_i^9 - \frac{111}{24} y_i^{11} + \frac{1631}{288} y_i^{13} + \frac{47}{72} y_i^{15} - \frac{23}{12} y_i^{17} + \frac{505}{864} y_i^{19} - \frac{55}{864} y_i^{21} + \frac{1}{432} y_i^{23} \tag{72}$$

On the other hand, the following asymmetric generalized logistic function can be fitted to any kind of data [63].

$$g(y) = y^p (1 - y)^r \tag{73}$$

The numerical integration of this function results in a sigmoidal function. Based on the values of p and r, this function can be adapted to peaked, flat, and unit-like distributions.

On the other hand, the literature reported other learning rules for the ICA estimation. For example, Bell and Sejnowski [32] derived the following learning rule for the logistic function of Eq. 73.

$$\Delta w \propto \frac{1}{w} + x[p(1-y) - ry] \tag{74}$$

$$\Delta w_0 \propto p(1-y) - ry \tag{75}$$

where x is the input vector and y is the output vector. Park et al. [64] derived the following learning rule.

$$\Delta W \propto \left(\frac{1}{W^*} - fft[g]Z^* \right) |W|^2 \tag{76}$$

where W and Z are the discrete Fourier transforms of w and. Abu-Amara and Abdel-Qader [65] developed the following ICA algorithm based on mutual information minimization to detect cancerous tissues in mammographic images.

1. PCA algorithm is used to reduce dimensionality of the data matrix R_{NxM} according to Eq. 77.

$$R^R_{Nxv} = R_{NxM} R_{Mxv} \tag{77}$$

where N represents number of sub-images, M represents size of each square sub-image, v represents number of selected principal components, and R_{Mxv} represents a matrix with the principal components in its columns sorted by descending order according to their variances.

2. The separating matrix W is initialized to the identity matrix. Then, y is calculated using Eq. 78.

$$Y_{vxM} = W_{vxv} (R_{Mxv})^T \tag{78}$$

3. The change in W is calculated using the natural gradient [26].

$$\Delta W = \eta \left[I - G(Y) Y^T \right] W \tag{79}$$

where I is the identity matrix and $G(y)$ is a nonlinear function.

4. The momentum method is used to boost the convergence speed of Eq. 79 using

$$\Delta W_{k+1} = \Delta W_k + \alpha \Delta W_{k-1} \tag{80}$$

where α is in the range [0, 1]. It was found by trial and error that the best value of α is 0.5.

5. The separating matrix is updated and then normalized.

$$W_{k+1} = W_k + \Delta W_k$$
$$W_{k+1} = \frac{W_k}{\| W_k \|} \tag{81}$$

6. Stop the algorithm when W converges or the maximum number of iteration reached.

In case of applying the ICA algorithm to synthetic signals, the impulse response signals and their mixing matrix are known priori. In this case, the separation performance of the used ICA algorithm can be measured using the equation WA which verifies whether the estimated separating matrix W is exactly the inverse of the mixing matrix A. On the other hand, Eq. 82 can be used to compare between two different ICA algorithms in terms of their separation accuracies through measuring the error index (also called rejection ratio) [66]. The algorithm that has the lower error has higher accuracy.

$$E = \frac{2}{N(N-1)}\left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\frac{|p_{ij}|}{\max_j |p_{ij}|}-1\right)+\sum_{j=1}^{N}\left(\sum_{i=1}^{N}\frac{|p_{ij}|}{\max_i |p_{ij}|}-1\right)\right) \quad (82)$$

where $P=\{p_{ij}\}=WA$ is the system (gain) matrix or the cross-global matrix. Similarly, the separation efficiency of any ICA algorithm can be measured using the separation performance [67].

$$SIR_{mean} = \frac{1}{n}\sum_{k=1}^{n} SIR_k \quad (83)$$

where the signal-to-interference ratio for the k[th] separated signal is defined as

$$SIR_k(dB) = 10\log_{10}\frac{\max_i p_{ki}^2}{\sum_{i=1}^{n} p_{ki}^2 - \max_i p_{ki}^2},\ k=1,2,...,n \quad (84)$$

Shi et al. [68] proposed a fixed-point ICA algorithm based on the nonlinear measure of the temporal autocorrelation between the impulse response signals for blind source separation. The proposed method can be summarized as follows.

1. The data is whitened.
2. The weight vector is initialized randomly.
3. The weight vector is updated using

$$w_{k+1} = w_k + E[g(y_k(t))G(y_k((t-\tau)z(t)+G(y_k(t))g(y_k(t-\tau))z(t-\tau)] \quad (85)$$

4. The weight vector is normalized, $w_{k+1} = w_k / \| w_k \|$.
5. Stop when the weight vector converges or maximum number of iterations reached.

In [69], the following second order blind identification (SOBI) algorithm is proposed where the impulse response signals are assumed to have a unit variance.

1. The observed vector x is whitened as shown in Eq. 86.

$$E[Qxx^*Q^T] = QAA^TQ^T = I \tag{86}$$

where Q is the whitening matrix.

2. A numerical algorithm is used to estimate the unitary matrix U that satisfies Eq. 87.

$$A = Q^{-1}U \tag{87}$$

3. The impulse response signals are estimated using

$$y(k) = U^TQz(k) \tag{88}$$

In [70], the following joint approximative diagonalization of eigen matrices (JADE) algorithm is proposed. JADE is based on kurtosis which can be summarized in the following steps. First, the quadri-covariance matrices are constructed using Eq. 89:

$$C_z(M) = E\{(Z^TMZ)ZZ^T\} - R_z tr(MR_z) - R_z MR_z - R_z M^T R_z \tag{89}$$

where R_p and M represent the covariance- and the Eigen-matrices of the whitened data Z and $p = 1,...,m$. Second, the resultant quadri-covariance matrices are decomposed using Eigenvalue decomposition according to $C_p^i(M_i) = U\Lambda_i U^T$ where U is an orthogonal joint diagonalization matrix and $\Lambda_i = diag(k_4(y_1),...,k_4(y_m))$. Finally, the mixing matrix A is estimated using $A = UQ^{-1}$ where Q is the whitening matrix. The JADE algorithm is a computational attractive method with higher separation performance in comparison with other ICA-based methods such as FastICA [70].

In [71], the following ICA algorithm is proposed where the nonlinear function is derived based on the parameterized t-distribution density model. The results indicate that this nonlinear function is robust against outliers and can handle both sub- and super-Gaussian distributions.

1. The linear ICA model is formulated as $y(t) = Wz(t)$ where z is the whitened data.

2. Compute the kurtosis using $k_i = \frac{m_4}{m_2^2} - 3$ where the nth order moment can be calculated using

$$m_j(t) = [1-\eta(t)]m_j(t-1) + \eta(t)y_i^j(t),\ (j = 2,4) \tag{90}$$

3. Form a lookup table using Eq. 91.

$$k_\alpha = \frac{\Gamma(\frac{5}{\alpha})\Gamma(\frac{1}{\alpha})}{\Gamma^2(\frac{3}{\alpha})} - 3 \tag{91}$$

4. Use the lookup table to find the value of α from k_i of step 2.

5. The scaling constant is computed using

$$\lambda_\alpha = [\frac{\Gamma(\frac{3}{\alpha})}{m_2\Gamma(\frac{1}{\alpha})}]^{0.5} \tag{92}$$

6. The nonlinear function is computed using

$$g_i(y_i) = \alpha\lambda_\alpha sign(y_i)|\lambda_\alpha y_i|^{\alpha-1} \tag{93}$$

One of the problems associated with the ICA algorithm is the method of adaptively estimating distribution of the impulse response signals [67]. FastICA algorithms use a fixed nonlinear function which make them non source adaptive. On the other hand, the source densities can be estimated adaptively using parametric methods such as Pearson system model [72], generalized Gaussian distribution model [73], or the extended generalized lambda distribution (EGLD) [74]. Unfortunately, these methods are not completely adaptive and thus fail when source densities do not follow the assumed parametric method. On the other hand, different source adaptive methods are proposed such as using a parametric mixture of logistic distributions [75], using nonparametric source density estimation [76], and kernel ICA [77, 67, 45]. In case the impulse response signals have near-Gaussian densities, kernel ICA methods are more robust than other ICA methods [45].

The literature reported extended methods to ICA such as Topographic ICA [78], Multidimensional ICA [79], Subspace ICA [80], Kernel ICA [77], Tree-dependent component analysis [81], and Subband decomposition ICA (SDICA) [82]. The SDICA assumes that each impulse response signal can be represented as a sum of dependent and independent subcomponents that have different frequency bands.

On the other hand, other methods are proposed in the literature for data decomposition such as non-negative matrix factorization [83], Sparse PCA [84], parallel factor analysis [85], smooth component analysis [86], dual tree wavelets [87], and fast incremental principal non-Gaussian directions analysis algorithm [88].

One of the inherited problems in the ICA estimation is the dependency of the learning rule on the learning rate. However, other methods that do not depend on the learning rate such as relative Newton method [89], JADE algorithm [70], FastICA algorithm [23], and relative trust-region method [90] can be incorporated to overcome such a challenge.

GPR Analysis via Deconvolution

Conducting of GPR in infrastructure testing is still not performed on a routinely basis for many reasons. First, raw GPR data is often complex which requires an experienced and a skilful operator to interpret them. Second, raw GPR scans depend on the unknown dielectric properties of the internal targets of the scanned medium. Third, it is not easy to precisely locate reflected A-scans within a measured B-scan image. Finally, raw GPR data is not an image of the scanned medium (GPR not an imaging method like ultrasound) [17].

Referring to Eq. 2, raw GPR data can be considered (under simplifying assumptions) as a convolutive mixture of the incident pulse (transmitted pulse by the GPR antenna) f and impulse response of the system (i.e., what characterizes the scanned slab, its internal targets, and the measurement system) h as shown in Eq. 94 resulting in one equation and two unknowns. Such a problem can be solved using the deconvolution process.

$$g = f * h = \sum_{j} f(t_{i-j})h(t_j) \tag{94}$$

where * represents the mathematical convolution operation and h represents impulse response of the system. Deconvolving raw GPR data is regarded as a preprocessing step for any automatic target detection algorithm where the original ground response should be recovered from the overlapping reflections. In some cases, reflections from a small defect are masked by reflections from a larger nearby object which makes its detection a difficult task. The minimum vertical distance between two targets should be greater than ¼ of the antenna wavelength to produce non overlapping reflections. For the 1.5GHz antenna, $\lambda = \frac{c}{f} = \frac{3x10^8}{1.5x10^9} = 0.2m = 20cm = 7.8"$. This means that the minimum vertical distance should be $1.95"$ to avoid overlapping. Deconvolution reduces the overlapping between reflections from closely spaced objects by improving the time resolution of the deconvolved signal which aims at simplifying the detection and consequently the depth estimation.

Deconvolution methods can be classified into direct deconvolution and blind deconvolution. Direct deconvolution methods assume a known incident pulse and deconvolve it with raw GPR data in order to recover h. On the other hand, blind deconvolution methods estimate both f and h from raw data with no prior information about f nor h.

Generally speaking, deconvolving raw radar data is considered as non-straight forward problem due to the following reasons. First, most direct deconvolution algorithms assume a stationary radar pulse with minimum phase in order to estimate the ground response which is not the case in most radar pulses. Second, the uneven

towing speed of GPR antenna may cause a horizontal distortion. Third, the unknown velocity of wave propagation through layers of the medium may cause a vertical distortion [91]. Finally, deconvolution artifacts sometimes hide weak deep signals unless a time gaining is applied as a pre processing step for deconvolution. On the other hand, signal characteristics may be changed when applying time gaining as a pre processing step due to its non-linearity.

The convolution process of Eq. 94 can be modeled using matrix operations as shown in equations 95 and 96 where columns of the convolution matrix F are constructed form delayed versions of the incident pulse vector $f = [f(t_1), f(t_2), \cdots, f(t_{nw})]^T$.

$$g = Fh \tag{95}$$

$$F_{n\times n} = \begin{bmatrix} f(t_1) & 0 & 0 & 0 & 0 & 0 \\ \vdots & f(t_1) & 0 & 0 & 0 & 0 \\ f(t_{nw}) & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & f(t_{nw}) & & f(t_1) & 0 & \\ \vdots & 0 & \ddots & \vdots & & 0 \\ 0 & 0 & & f(t_{nw}) & f(t_{nw-1}) \cdots & f(t_1) \end{bmatrix} \tag{96}$$

where n is the number of samples of the GPR trace g, nw is the number of samples of the discrete random vector f, t_i represents time index, and $i = 1, 2, \cdots, nw$.

Since front and tail of the incident pulse have very small values, the matrix F will be a large matrix with near zero main diagonal values and so it is an ill-conditioned matrix that may not have an inverse [92]. If an inverse does exist, F^{-1} may be extremely sensitive to additive noise in g and consequently, h may not be easily estimated. Therefore, conducting simple deconvolution methods such as linear

least squares is not enough to deconvolve raw GPR scans due to noise sensitivity and ground scattering.

Krause and Abdel-Qader [93] estimated the incident pulse of the GSSI antenna of figure 1 by transmitting a pulse into a metal plate (a perfect reflector) and recording the reflected signal as shown in figure 5. Next, direct and blind deconvolution methods will be explored.

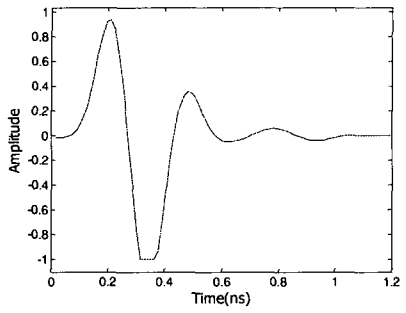

Figure 5. The estimated normalized incident pulse of the GSSI 1.5 GHz antenna.

Direct Deconvolution Methods

In this section, direct deconvolution methods will be explored such as optimization-based methods, conjugate gradient method, singular value decomposition method, subset selection deconvolution algorithm, discrete wavelet methods, correlation-based methods, and Homomorphic method.

Optimization-Based Methods

A set of parameters $\{x_1,...,x_n\}$ is estimated that minimize the rating function $F[\{x_1,...,x_n\}]$. The received GPR signal can be expressed using Eq. 97.

$$g(t) \cong \sum_i a_i p(t-b_i) \tag{97}$$

In this case, the parameter set is the set of a_i's and b_i's where the difference between the current estimate and the received signal indicates the rating function:

$$F[\{a_1,b_1,...a_n,b_n\}] = \sum_k | g(k) - \sum_i a_i p(k-b_i) | \tag{98}$$

There are many methods reported in the literature as optimization algorithms such as Powell's direction set model and annealing and downhill simples [94]. Unfortunately, these algorithms are computationally demanding and are sensitive to initial values (seed) of their parameters. Also, it is not easy to find derivative of the rating function. Finally, these methods may produce incorrect results due to false interpretation of sub-optimal local minima models [95].

Conjugate Gradient Method

It is an iterative algorithm that works with linear systems, mainly used for sparse systems, and works with only symmetrical matrices. Debalina and Tapan [96] developed an algorithm that remotely detects buried objects using impulse radiating GPR. First, a single A-scan is performed. Then, conjugate gradient method is used to

deconvolve the raw radar data. Then, a matrix pencil algorithm is applied that extracts the natural resonance frequencies in order to identify target response.

Singular Value Decomposition - SVD

In general, singular value decomposition (SVD) is used to decompose a real or complex matrix F as the product of an $M \times M$ column-orthogonal matrix U, an M x N diagonal matrixΣ, and an $N \times N$ column-orthogonal matrix V [97]. Using SVD method, the convolution matrix F can be decomposed using Eq. 99.

$$F = U\Sigma V^T \tag{99}$$

where Σ is a diagonal matrix whose entries are the singular values on the diagonal direction. SVD transforms the original matrix into a domain where the covariance matrix is diagonal with singular values given by

$$\sum_{ii} = \sigma_i \text{ and } \sigma_i \geq \sigma_{i+1}$$

Using SVD, the convolution model of Eq. 94 may be restated as

$$g = \sum_i (v_i^T \cdot h)\sigma_i u_i \tag{100}$$

and thus an inverse process can be constructed using matrix operations resulting in an estimate of the impulse response h

$$h = \sum_i \frac{(u_i^T \cdot g)v_i}{\sigma_i} \tag{101}$$

For several reasons, this inverse tends to fail. Some σ_i may equal to zero making the division process impossible. Some σ_i may be smaller than machine

precision resulting in machine errors in division. Additionally, additive noise in g may be greatly amplified if it correlates to unit vectors with a small σ_i. One solution is to replace the division by σ_i with multiplication by a similar yet better-behaved function. The approach can be described as

$$h = \sum_i (u_i^T . g)\, p(\sigma_i)\, v_i \tag{102}$$

where

$$p(\sigma_i) = \begin{cases} \dfrac{1}{\sigma_i} & when \quad \sigma_i \geq threshold \\ 0 & when \quad \sigma_i < threshold \end{cases}$$

The choice of threshold value allows us to avoid division errors and to discard components of g which are determined to be mostly noise. To deconvolve an A-scan of GPR data using SVD method, a convolution matrix F is created based on the transmitted pulse f as modeled in Eq. 96. Matrix F is then decomposed into its U, Σ and V components, and a threshold for values of σ_i is chosen. To process the corresponding GPR trace g, Eq. 102 is used to produce h vector which is composed of signatures of all embedded objects including any possible defects.

On the other hand, PCA algorithm can be implemented using the SVD algorithm. If the matrices F and U are centered, columns of the matrix $U\Sigma$ contain principal components of the matrix F [98]. Another way is to select the non-zero Eigenvalues from the matrix U which results in a new matrix $\tilde{U}$. Then, Eq. 103 can be

used to obtain the principal components of the matrix F [99]. On the other hand, the principal components are equal to ΣV according to [100].

$$Z = \tilde{U}^{T} F \tag{103}$$

Subset Selection Deconvolution Algorithm - SSDA

SSDA assumes that the real data g can be closely approximated by convolving the transmitted pulse f with a small number of delta function pulses [92]. The time delays and amplitudes of these delta functions are chosen to minimize Eq. 104, similar to the linear least squares method,

$$\min_{\tau_i, a_i} \left[\left\| \left(\sum_i a_i \delta (t - \tau_i) \right) * f - g \right\|_2^2 \right] \tag{104}$$

where a_i and τ_i represent amplitude and delay time of the corresponding impulse function. Finding the values of τ_i and a_i which minimize Eq. 104 is a non-trivial challenge especially for large values of i. When a minimum is found, the sought impulse response of the medium is

$$h = \left(\sum_i a_i \delta (t - \tau_i) \right) \tag{105}$$

If the difference in delay times of two separate pulses is under a preselected critical threshold value, SSDA will detect a single pulse whose delay time is roughly the average of the two original pulses and whose amplitude is their sum. Such a false detection obscures the true position of pulses, and generates noise which may obscure low amplitude details. The success of SSDA relies on the initial choices of τ_i and a_i.

To deconvolve a GPR trace g using SSDA, the number of expected pulses i is selected based on prior knowledge of medium and by minimizing the difference given in Eq. 104, the minimum values of τ_i and a_i are estimated allowing for the construction of h per Eq. 105.

Correlation-Based Methods

Krause and Abdel-Qader [93] developed an algorithm to deconvolve raw GPR data. First, a correlation-based iterative decomposition algorithm is used to deconvolve each column of raw data into a list of ordered pairs in order to identify target reflections. Then, the deconvolved image is segmented and then, converted into a set of points and arcs. Arcs are formed by grouping the ordered pairs with high degree of similarity. Finally, an algorithm is used to eliminate, highlight, or add matched objects into the resultant image. The proposed algorithm by [93] successfully deconvolved a number of raw GPR images resulting in an enhanced visual inspection of embedded defects. Unfortunately, it has some drawbacks. First, it optimizes one parameter pair each time instead of optimizing all parameter values. Second, major reflections can be detected only. Third, an error may be produced due to the overlapping between pulses and fake correlation.

Blind Deconvolution Methods

Blind processing methods can be classified into blind source separation (BSS) and blind deconvolution (BD). BSS methods aim to decompose mixed signals into a

new set of signals with no prior information about the source signals or the mixing process while BD methods aim to deconvolve a signal into its original input and impulse response with no prior information on either. In BSS, the data is modeled as a linear mixture while in BD, the data is modeled as a convolutive mixture. Hence, blind deconvolution can be regarded as a particular case of blind source separation [23], since the blind deconvolution can be re-casted into a blind source separation problem if the convolutive mixture is expressed as a linear mixture. All previously mentioned direct deconvolution algorithms of section 3.2.1 estimated the incident pulse of the GPR system and used it to recover the impulse response of the scanned medium.

Blind deconvolution methods can be used to solve the deconvolution problem without knowing the incident pulse. The received raw GPR signal is a convolutive mixture of two different signals resulting in one equation with two unknowns. BD methods recover the original signal (ground response) in an unsupervised mode (with unknown mixing coefficients) where the impulse response signals are assumed to have different statistical properties and are mutually statistically independent.

Many methods have been used in blind deconvolution problems such as independent component analysis [101, 102, 103, 104], super resolution [105], Bussgang filters [106], and optimal sparse representation [107]. Usually these methods require complex computations but they offer robust results. In this section, deconvolution using ICA algorithm will be investigated.

The backscattered reflections from closely spaced layers/targets overlap. Conventional signal processing methods such FFT and matched filtering are not able to resolve their corresponding pulses (spikes of the impulse response) when the time spacing between them is less than $1/B$ where B is bandwidth of the used GPR antenna [108, 109]. Many algorithms have been proposed in the literature to improve time resolution of a GPR trace such as super resolution time-delay estimation methods of [108].

Blind separation of a convolutive mixture can be done in time-domain or in frequency-domain. In the frequency-domain, the convolutive mixture becomes multiplicative mixture of complex signals within different frequency bands [110]. In other words, blind separation of a convolutive mixture in the time-domain is equivalent into blind separation of an instantaneous (memoryless) mixture in each frequency band. In the frequency domain, the independent signals have different frequency representations.

Since frequency-domain ICA methods separate the mixture within each frequency band, the mixing matrix becomes a function of the frequency in the new domain while the basic ICA model assumes it a constant [23]. In other words, all frequency components of each impulse response signal must be grouped. This makes the permutation and scaling of the impulse response signals not consistent across all frequency bands. For this, additional methods to solve the permutation and scaling ambiguity are required to avoid combining contributions from different impulse response signals into a single impulse response when reconstructing the signal in the

time-domain which adds more computational demanding [110]. In the following three subsections, algorithms for blind separation of a convolutive mixture are presented.

Banded-ICA

The impulse response signal of the medium is sparse and should consist of a number of sharp spikes with relatively flat area between them to represent the layered structure of the scanned concrete deck. These spikes can be used to estimate the round-trip travel time of radar waves of embedded targets.

Referring to Eq. 94 and by assuming the impulse response signals have non-Gaussian distribution and statistically independent and identically distributed at different times, the linear ICA model ($x = As$) can be used to model Eq. 94 where the radar trace vector is defined as $x = [g(t_1), g(t_2), \cdots, g(t_n)]^T$, the impulse response vector is defined as $s = [h(t_1), h(t_2), \cdots, h(t_n)]^T$, and the mixing matrix A is defined according to Eq. 96 [23, 109]. Statistically independent and identically distributed at different times mean each signal has the same distribution as the other signals and all signals are mutually independent [23]. Since an impulse response signal of the scanned concrete slab is a sparse signal representing the layered structure of that slab, it has a super-Gaussian distribution and thus it meets the non-Gaussianity requirement condition. Additionally, the impulse response signals extracted from different scans are assumed to be statistically independent, naturally.

As Eq. 96 indicates, the mixing matrix A is a banded matrix with the nonzero elements of its columns represent the unknown incident pulse vector. This prior information about nature of the mixing matrix can be utilized to convert a blind deconvolution problem into a blind source separation problem [111]. However, this represents a single-input single-output instantaneous ICA model which is inadequate since statistics of the independent components cannot be characterized.

In order to form a multiple-input multiple-output ICA model ($X = AS$), time delayed versions of $x(t)$ and $s(t)$ are used to construct multidimensional matrices as shown in Equations 106 and 107 [23].

$$X_{n\times n} = [x(t-n+1)\ x(t-n+2)\ \cdots\ x(t-1)\ \ x(t)]^T \tag{106}$$

$$S_{n\times n} = [s(t-n+1)\ s(t-n+2)\ \cdots\ s(t-1)\ \ s(t)]^T \tag{107}$$

Equations 106 and 107 ensure that every element of the vector $x(t)$ is a convolutive version of $s(t)$ and $f(t)$ according to Eq. 94. In this way, the convolutional model is casted into multidimensional ICA model which converts the problem from blind deconvolution into blind source separation. Unfortunately, these two equations are inadequate for the ICA model as is since the first few rows/columns of S and X have few nonzero elements meaning that statistics of the independent components cannot be constrained due to lack of information. To overcome this challenge, an approximative convolutional model is proposed to solve the lack of information problem by discarding the first few rows of X and S which consequently reduce the number of zero elements in the first few rows [111]. Thus, equations 106 and 107 can be rewritten as

$$X_{m\times n} = [x(t-m+1)\ x(t-m+2)\ \cdots\ x(t-1)\ x(t)]^T \tag{108}$$

$$S_{m\times n} = [s(t-m+1)\ s(t-m+2)\ \cdots\ s(t-1)\ s(t)]^T \tag{109}$$

where $m < n$, n represents length of the corresponding GPR trace while m represents number of independent components to recover. Now we have another problem where Eq. 96 does not represent an exact mapping between X (Eq. 108) and S (Eq. 109) especially for $x_i(t_j)$ in the interval $(i \in \{1\cdots nw\}) \cap (j \in \{(m-1)\cdots n\})$. Fortunately, the mapping is correct for remainder of the mixture x.

In [111], a banded-ICA algorithm was proposed to deconvolve seismic traces. First, raw seismic data x is whitened as shown below.

$$z = \Gamma x \tag{110}$$

Second, the new mixture matrix is formed as

$$\tilde{x} = N_i^T \Gamma^T z,\ i = 1,2,\cdots(m-nw-10) \tag{111}$$

where $\tilde{x}$ contains nw mixtures, $N_i = [0_{(i-1)\times n}; I_{m\times n}; 0_{(n-m-i+1)\times n}]$ represents zero padding matrices, 0 represents a zero matrix, and $i = 1,2,\cdots,n$. The zero padding matrix N_i maps f to a particular column of F. Third, the linear ICA model is formulated as $\tilde{y} = \tilde{W}\tilde{x}$ where $\tilde{W}$ of size $nw \times nw$. ICA recovers a number of independent components equal to rows of $\tilde{W}$. To select the independent component that best represents the desired deconvolved signal, coefficients c_i are estimated such they minimize Eq. 112.

$$\phi(c_i) = \left\| x_k - c_i \tilde{a}_i * \tilde{y}_i \right\|_2^2 \tag{112}$$

where $x_k = [x_k(t_1), x_k(t_2), \cdots, x_k(t_N)]^T$ represents realization of the k[th] mixture, $k = m - nw - 10$, and $\tilde{y}_i = [\tilde{y}_i(t_1), \tilde{y}_i(t_2), \cdots, \tilde{y}_i(t_N)]^T$ represents realization of the i[th] independent component. The best independent component that results in a minimum $\phi(c_i)$ is selected where $i = 1,2,\cdots,nw$. Unfortunately, the estimated independent component is a delayed version of the original synthetic impulse response signal. This is due to the way the mixture matrix *X* is organized. Also, size of the approximative convolutional model (*m*) is determined manually for each case.

In [109], the following blind deconvolution algorithm was developed to estimate thickness of a thin PVC slab.

1. The discrete convolutional model is formed as $g(t_i) = \sum_j f(t_{i-j+1}) h(t_j)$
2. The mixing matrix is formed according to Eq. 96.
3. The *S* and *X* matrices of the linear ICA model are formed as

$$\begin{aligned} S &= [z^{n-1}s, z^{n-2}s, \cdots, zs, s]^T \\ X &= [z^{n-1}x, z^{n-2}x, \cdots, zx, x]^T \end{aligned} \tag{113}$$

   where *z* represents the unit time delay.

4. Since the impulse response signal is a sparse signal, a nonlinear function that is dedicated to deal with super-Gaussian signals is required. Therefore, the hyperbolic tangent nonlinear function is used with the FastICA algorithm. This results in a number of independent components. To select the best

independent component that represents the reflectivity series, each independent component is convolved with its corresponding row of the estimated mixing matrix. Then, it is subtracted from the radar trace and the one that results in the smallest value is selected.

5. The differential time delay between spikes of the estimated impulse response is used to estimate thickness of the pavement.

In the algorithm of [109], the recovered independent component is a delayed version of the impulse response but could not produce accurate depth estimation. In this work, a zero correction step is performed to fix this issue and thus enables depth estimation with respect to the ground line. Also, size of the approximative convolutional model (m) is assumed to be equal to n while in this work best value of m is found in an automated way. Also they used the FastICA method for data decomposition and manually time filtered the raw GPR trace to keep the reflections that correspond to slab and remove the other reflections (air-coupling and substrate). In this work, EFICA method is used since it has higher separation performance than FastICA and is within a completely automated framework.

Al-Qaisi et al. [103] proposed a blind deconvolution method for seismic traces. Their method exploits sparsity of both the mixing matrix (a banded matrix) and the reflectivity sequence (by assuming it has a Bernoulli Gaussian distribution). The proposed algorithm consists of the following steps.

1. Let g represents a seismic trace related to the earth reflection h and to the source wavelet f as $g(t_i) = \sum_j f(t_{i-j})h(t_j)$

2. The mixing matrix is formed according to Eq. 114.

$$A_{2k\times(3k-1)} = \left[fN_0 \;\; fN_1 \;\cdots\; fN_{2k-1}\right]^T \tag{114}$$

where k represents length of the source wavelet f. This ensures that rows of A contain delayed versions of f. The zeros padding matrices are of size k x $(3k-1)$ and defined as $N_i = [0_{k\times i} \;\; I_{k\times k} \; 0_{k\times(2k-1-i)}]$ for $i = 0,1,\cdots,2k-1$.

3. In order to form a multiple input multiple output ICA model ($X = AS$), the following matrices are formed.

$$S_{(3k-1)\times n} = \left[h \;\; zh \;\cdots\; z^{n-3k}h \;\; z^{n-3k-1}h\right]^T \tag{115}$$

and

$$X_{(2k)\times n} = \left[g \;\; zg \;\cdots\; z^{n-3k}g \;\; z^{n-2k}g\right]^T \tag{116}$$

4. The mixture matrix X is whitened using the Eigenvalue decomposition (EVD) of the covariance matrix method. The zero padding matrices are used to exploit sparsity of the mixing matrix. This results in the following whitened matrix $\tilde{X}_{kxn}$.

$$\tilde{X}_{i_{kxn}} = N_i^T T^T Z \tag{117}$$

where $T = D^{-0.5}E^T$ and $Z = TX$. E and $D = diag[d_1, d_2, \cdots, d_n]$ are the eigenvectors and eigenvalues of the covariance matrix $C = \dfrac{XX^T}{n}$, respectively.

5. By assuming the input reflectivity sequence has a Bernoulli Gaussian distribution, the following learning rule was derived.

$$W_i \leftarrow (I - Y_i U_i - U_i U_i^T) W_i \tag{118}$$

where

$$U_i = W \tilde{X}_i + W_0, \ \Delta W_0 = 1 - 2Y_i \tag{119}$$

and the proposed logistic function

$$Y_i = p_i(0.5 + 0.5\tanh(U_i)) + \frac{1-p_i}{2\sqrt{\pi}} erf\left(\frac{U_i}{\sqrt{2\sigma_i^2}}\right) \tag{120}$$

where p_i is the probability of reflections occurrence.

6. The best independent component among *k* recovered independent components can be found using

$$\psi(c_i) = \left\| x_n - c_i \tilde{a}_i * \tilde{d}_i \right\|_2^2 \tag{121}$$

where $\tilde{a}_i$ is a recovered wavelet (transmitted pulse) and $\tilde{d}_i = W \tilde{x}_i$ is the estimated independent component. Eq. 121 has its extreme points when

$$c_i = c_i^{(*)} = \frac{x_n^T (\tilde{a}_i * \tilde{d}_i)}{(\tilde{a}_i * \tilde{d}_i)^T (\tilde{a}_i * \tilde{d}_i)} \tag{122}$$

Performance of the proposed algorithm is compared against performance of FastICA and JADE algorithms. Simulation results indicate that the proposed algorithm is computationally expensive but offer better performance in terms of accuracy (using minimum mean square error), shape, and scaling.

Skewed-Spatiotemporal ICA

The spatial ICA model assumes that each signal is a linear mixture of spatially independent impulse response signals. On the other hand, temporal ICA assumes that each signal is a linear mixture of temporally independent impulse response signals. The spatiotemporal ICA simultaneously minimizes the statistical dependency between the impulse response signals over both space and time. In contrast, conventional ICA methods attain statistical independence over space (Spatial ICA) or time (temporal ICA). On the other hand, skewed ICA assumes skewed density functions for the impulse response signals while conventional ICA methods assume symmetrical density functions [104]. In [104], the following spatiotemporal ICA algorithm was proposed.

1. The SVD algorithm is used to decompose the data vector according to

$$G = U\Sigma V^T \approx \tilde{U}\tilde{V}^T \tag{123}$$

2. Spatiotemporal ICA assumes that $\tilde{U}$ contains a linear mixture of spatially independent components S while $\tilde{V}$ contains a linear mixture of temporally independent components T and

$$\tilde{G} = S\Lambda T^T \tag{124}$$

where Λ is a diagonal scaling matrix, $S = \tilde{U}W_S$ and $T = \tilde{V}W_T$ are the spatial and temporal independent components. W_S and W_T are the separating matrices for the spatial and temporal components. This implies

$$W_S \Lambda W_T^T = I \rightarrow W_T = (W_S^{-1})^T (\Lambda^{-1})^T \quad (125)$$

3. The following function is maximized in order to obtain W_S and W_T.

$$h = \alpha H(Y_S) + (1-\alpha) H(Y_T) \quad (126)$$

where H is the entropy, α is a constant in the range [0, 1] and usually equal to 0.5, $H(Y_S) = \sigma_s(S)$ and $H(Y_T) = \sigma_T(T)$, σ_s and σ_T are the approximations of cumulative density functions of the spatial and temporal independent components, respectively. In [104], $\sigma_s = \sec h^2 y$ and $\sigma_T = \sec h^2 y$.

In case of skewed ICA, the skewed density functions can be estimated using skewness instead of kurtosis as shown in Eq. 127 [104].

$$\sigma_{sk} \propto \exp\left(\frac{a-b}{2} x - \frac{a+b}{2} \sqrt{x^2+1}\right) \quad (127)$$

In case of skewed spatiotemporal ICA, the function of Eq. 126 is maximized while $H(Y_S)$ is replaced with Eq. 127.

In [112], temporal ICA and spatial ICA methods based on natural gradient ICA algorithm were implemented for clutter reduction in order to decompose GPR signals into subspaces of clutter signals and target signals. They followed it with component selection algorithms based on temporal, spatial, and spatiotemporal feature selection. Results indicated that the spatiotemporal selection method produced best performance. Their work can be summarized as follows.

1. The SVD algorithm is used to implement PCA algorithm for dimensionality reduction using $G = U\Sigma V^T$ to obtain the principal components according to $\Sigma_{i,i} V_i$.

2. The temporal ICA algorithm can be summarized as follow.

   - GPR data is projected into a new subspace using $Y = \tilde{U}^T G$ where $\tilde{U}$ represents a vector of size less than number of selected principal components.

   - The temporal ICA is defined as $Y = A_t S_t$.

   - Natural gradient is used to estimate A_t and S_t [26].

   - The original GPR signal is reconstructed as $\tilde{G} = \tilde{U} A_t S_t$.

3. The spatial ICA algorithm is summarized as follow.

   - GPR data is projected into a new subspace using $Y = \tilde{U}^T G$ where $\tilde{U}$ represents a vector of size less than number of selected principal components.

   - The spatial ICA is defined as $\tilde{U}^T = A_s S_s$.

   - Natural gradient is used to estimate A_s and S_s.

   - The original GPR signal is reconstructed as $\tilde{G}^T = Y A_s S_s$.

4. Three component selection algorithms are used to select components that have landmine signatures as follows.

- The temporal selection method reconstructs the original signal using Eq. 128 for temporal and spatial ICA algorithms, respectively.

$$\tilde{G} = W_t \tilde{S}_t, \; \tilde{G}^T = \tilde{W}_s S_s \tag{128}$$

- The spatial selection method reconstructs the original signal using Eq. 129 for temporal and spatial ICA algorithms, respectively.

$$\tilde{G} = \tilde{W}_t S_t, \tilde{G}^T = W_s \tilde{S}_s \tag{129}$$

- The spatiotemporal selection method reconstructs the original signal using Eq. 130 for temporal and spatial ICA algorithms, respectively.

$$\tilde{G} = \tilde{W}_t \tilde{S}_t \text{ and } \tilde{G}^T = \tilde{W}_s \tilde{S}_s \tag{130}$$

Complex ICA

In [113], a spatial complex ICA algorithm is used to extract the components with spatio-temporal dynamics in order to model the change of oxygenated blood flow with neural activity from fMRI recordings of brain activity. The dynamic flow patterns are modeled as a convolutive version of a spatio-temporal source pattern and its time-course of activation. The proposed algorithm is a generalized version of the Infomax ICA [32] where the impulse response signals are assumed as complex random variables with super-Gaussian densities. Their work can be summarized in the following steps.

1. Let x_{ti} represents the mixture where i represents sample index and t represents time index. It is decomposed into different spectral bands using the short-time Fourier transform (STFT).

$$x_{Ti}(f) = \sum_t g_i(T+\tau)R(\tau)e^{-i2\pi f\tau/2k} \tag{131}$$

where $R(\tau)$ is a windowing function (e.g., a Hanning window) centered at time T with finite support in the interval $\tau = -k,...,k-1$ (window length is $2k$) and f is the frequency index $f = 0,...,K$.

In the frequency-domain, Eq. 132 is used to model the mixture for each frequency band f.

$$X_f = A_f S_f \tag{132}$$

2. The following linear projection equation is used to separate the mixture in each frequency band to obtain the complex-valued independent components for each frequency band (bin).

$$Y_f = W_f X_f \tag{133}$$

3. The following learning rule is developed based on natural gradient method.

$$\Delta W_f = \eta(I - E[G(y_f)y_f^H])W_f \tag{134}$$

where G_f is a nonlinear function that can be computed using

$$g_{f_{Ti}} = sign(y_{f_{Ti}})\tanh(|y_{f_{Ti}}|) \tag{135}$$

and

$$sign(y)=\begin{cases}0 \quad if\ y=0,\\ \dfrac{y}{|y|} if\ y\neq 0\end{cases} \tag{136}$$

4. The estimated independent components are reconstructed in the time-domain using inverse STFT by combining the extracted components across several frequency bands.

Results indicate that the proposed algorithm successfully separated the mixture. Unfortunately, artifacts may appear in the reconstructed independent components in the time-domain. Also, the process of solving the permutation and scaling ambiguity when reconstructing the impulse response signals in the time-domain is done manually.

In [110], a frequency-domain version of the FastICA algorithm is used according to Eq. 137.

$$\begin{aligned} W_{f+1} &= W_f + E[Z(W_f Z)^* G(|W_f Z|^2)] - E[G(|W_f Z|^2) + |W_f Z|^2 G'(|W_f Z|^2)]W_f \\ W_{f+1} &= \frac{W_f}{\|W_f\|} \end{aligned} \tag{137}$$

where Z is the whitened data. A suitable choice for the nonlinear function can be as follows

$$G'(y) = \tanh(\Re\{y\}) + j\tanh(\Im\{y\}) \tag{138}$$

Since a high correlation exists between spectrums of the mixture in adjacent frequency bands, resultant separation matrices should have no great change in their coefficients. This means by initializing the separation matrix in the current frequency band W_f to the final value obtained in the previous frequency band, the separation

matrices in adjacent frequency bands will have same permutation order. Thus the permutation ambiguity is solved. On the other hand, Eq. 139 can be used to solve the scaling ambiguity.

$$Y_f = R_f W_f X_f \tag{139}$$

where X is the mixture and R is defined as $R_f = diag\{C_f\}$ and

$$C_f = W_f^{-1} = \begin{bmatrix} c_{11} & \cdots & c_{1M} \\ \vdots & \ddots & \vdots \\ c_{M1} & \cdots & c_{MM} \end{bmatrix} \tag{140}$$

where M is number of impulse response signals. The proposed algorithm has a fast execution time with fast convergence speed. On the other hand, the proposed algorithm requires a high resolution in the frequency-domain to solve the permutation and scaling ambiguity.

In [114], a similar approach was proposed to solve the permutation ambiguity by applying frequency coupling between adjacent frequency bands according to Eq. 141. This equation is similar to the momentum method of Eq. 16.

$$\Delta W_{(f+1)} = \Delta W_f + \alpha \Delta W_{f-1} \tag{141}$$

where α is in the range [0, 1].

On the other hand, using a constant step size for all frequency bands may affect the separation performance at certain frequency bands. In [114], the following fixed-point ICA algorithm is proposed based on Newton optimization.

$$\begin{aligned} \Delta W_f &= D[diag(-\alpha_i) + E\{G_f(y) Y_f^H\}] W_f \\ W_{f+1} &= W_f [W_f^H W_f]^{-0.5} \end{aligned} \tag{142}$$

Comparing Eq. 142 to the natural gradient equation, we notice that the identity matrix is replaced with an adaptive term $diag(-\alpha_i)$ to increase the convergence speed of the algorithm [23]. Also, the fixed learning rate is replaced with a matrix D that adapts to the data according to Eq. 143 [114].

$$D = diag(\frac{1}{\alpha_i - E\{G^{'}(y_i)\}}) \tag{143}$$

where $\alpha_i = E\{y_i G(y_i)\}$. There are other approaches to solve the permutation ambiguity such as information maximization [115], prior smoothness information [116], high-dimensional optimizations [117], measuring distance between components across frequency bands and matching component pairs [118], direction of arrival estimation [119], and inter-frequency dependency relation [120].

Target Detection in GPR Scans

Karlsen et al. [99] implemented an algorithm based on selective ICA for mine-like objects detection where the SVD algorithm was used as a whitening step. Four different ICA linear mixture models were tested. The time-time model assumes time independence of GPR time signals. The time-spatial model assumes a spatial independence of time signals. In the frequency domain, frequency independence is assumed in the frequency-frequency model and the frequency-spatial model assumes a spatial independence. Two ICA algorithms based on the Infomax ICA of Eq. 42 and delayed-decorrelations ICA [121] were investigated. The independent components with high non-Gaussianity structure were selected as they have mine reflections

where kurtosis is used to measure non-Gaussianity. The buried mines have weak reflections. The decorrelation ICA is based on decorrelation of delayed-time signals. Since the time-spatial model assumes a spatial independence of time signals, both ICA methods seek to enhance the spatial independence of the mixture. Decorrelation ICA failed to separate the time-spatial mixture because decorrelation enhances spatial signature of both objects (detected mines) and clutter simultaneously when it is supposed to enhance signatures of mines and suppress clutter signatures. The frequency-frequency model assumes the GPR trace as a linear mixture of independent frequency spectrum signals. Both Infomax and decorrelation ICA methods try to enhance frequency independence of the mixture. Results indicate that the decorrelation ICA has better separation performance of the frequency-frequency mixture. However, the natural gradient method can be used to enhance the Infomax ICA and thus improving its results.

Lotsch et al. [122] applied the FastICA to deconvolve remotely sensed image sequences. Results indicate that ICA successfully extracts spatial and temporal components of the data, separate them, and recognize data artifacts due to GPR instrumentation and data processing.

In [123], four ICA algorithms were applied to the GPR detection of non-metallic land mines to decide on their suitability for GPR data. These are the FastICA algorithm, the Infomax ICA based on maximum likelihood estimation, the SOBI algorithm, and the JADE algorithm. The clutter reduction efficiency is estimated using the Receiver Operating Characteristics (ROC) curves and used to compare

between performances of the implemented algorithms. The experimental results indicate that both JADE and SOBI algorithms have a better performance than FastICA and Infomax ICA.

In [124], a comparison between ICA and Blind Instantaneous Signal Separation (BISS) is established for ground bounce removal while preserving signature of deep objects in the landmine detection problem. The received GPR signal is assumed as a linear mixture of strong ground bounce signal, landmine signals, and noise. In case of ICA, the independent components are selected based on the non-homogenous detector (NHD) method. The cumulant-based BISS algorithm selects few impulse response signals from a large number of observed signals based on the NHD method to determine number of impulse response signals to recover. Results, based on one GPR image, show that the BISS used less number of recovered components than ICA and thus it is less computational demanding.

Defect Detection in GPR Scans

In [125], a simple algorithm is proposed to detect subsurface defects such as knots, decays, and embedded metals in wooden logs in real time. The metal detectors employed by the saw mills cannot detect all kinds of defects and thus their saw blades can be damaged by the undetected defects. A GSSI 900MHz antenna is used to scan the wood logs. Two different methods were compared for this problem.

1. In the first method, raw GPR data is processed using the RADAN software. First, a time gaining is applied to enhance reflections from deep objects. Then,

position of the whole B-scan is adjusted vertically to enable depth estimation with respect to the ground line. Next, background and noise are removed. The previous steps result in a B-scan with clear defect signatures. Finally, whenever a change in signal bands is encountered, a defect is declared. Unfortunately, analyzing the output image in RADAN requires an expert operator and it is time consuming.

2. In the second method, the Surf plot of raw GPR data is obtained with the round-trip travel time shown in the horizontal axis and amplitude shown in the vertical axis. Next, surface reflections are clipped until the second positive peak. Then, a top view is taken with 90° clockwise rotation. Next, echo reflections from bottom of the wooden log are reduced. Then, a thresholding is applied to detect subsurface defects. Finally, output of the algorithm is used as an input for the CNC sawing machine.

In [126], an enhanced version of the previous algorithm is proposed to detect subsurface defects in wooden logs in real time. The proposed algorithm can be summarized in the following steps.

1. A linear time gaining is applied to enhance reflections from deep objects. Unfortunately, this also enhances the ringing bands.
2. Zero correction: in this step, part of the signal is deleted until the first positive peak (ground band reflection). First, position of maximum points is found for all traces of the scan. Then, an average position is found and finally data points till that average position are deleted.

3. Extended zero correction: in this step, the signal is further clipped till either the first minimum (to detect both internal- and surface-defects) or to the second maximum (to detect only internal defects).
4. Clipping the signal: first, average of the scan is subtracted from the entire scan. Next, data points in every column (trace) are deleted till the first positive pixel to the right of the corresponding column minimum. Finally, data points are deleted from the entire scan till the average maximum position of all columns.
5. Bottom reflection removal: since the lowest amplitude value after the 250^{th} point represents bottom reflection, average of all minimum values for all columns after the 250^{th} point is found, then, data points are deleted from this position till end of the scan.
6. Filtering the signal: A low pass filter (300 MHz cutoff) is used followed by a high pass filter (1800MHz cutoff) to eliminate the remaining noise.
7. Thresholding the signal: since average of the entire scan is zero, it cannot be used as a threshold. Alternatively, average value of maximum values of all columns is found and used as a threshold.
8. To prevent declaring a defect in a normal log, the previous threshold value is adjusted by adding or subtracting a constant value.
9. Depth of a detected defect can be found in two steps. First, total depth is estimated using Eq. 144.

$$d_t = \frac{ct_{wood}}{2\sqrt{\varepsilon'_{wood}}} \tag{144}$$

where d_t represents depth of the detected defect in meters, c represents speed of light, t_{wood} represents the round trip travel time in seconds, and ε'_{wood} represents the dielectric constant of the wood (10-26) depending on moisture content of the wood.

Second, depth of the defect is found using

$$d = d_t \times \frac{Defect\ peak\ row number}{total\ number of\ rows in the scan} \tag{145}$$

A second way of estimating depth of a defect can be done as follows, 1) the difference between first positive peak and the last negative peak in each column is estimated, 2) the average of these differences is computed which represents the number of rows in the scan, and 3) depth is estimated by dividing the defect row number over the number of rows in the scan and multiplying the resulting ratio by the user defined average diameter of the wooden log.

Fractal-Based Algorithms

There are different features extraction algorithms reported in the literature. Some GPR objects have regular shapes such as rebar. However, other objects have irregular shapes such as defects. Generally speaking, these irregular objects can be detected using methods that require training and testing phases. Fractal-based

techniques can be used to differentiate between defects and other embedded objects, thus enabling the automation of defect detection process.

A complex trace $z(t)$ consists of a real seismic trace and an imaginary trace (Hilbert transform of the real trace) as shown in Eq. 146 [56].

$$z(t) = f(t) + j\bar{f}(t) \tag{146}$$

where $f(t)$ is the real trace and $\bar{f}(t)$ is the imaginary trace. The Hilbert transform is defined according to Eq. 147 which results in a $90°$ phase shift to the input signal.

$$G(w) = -j\,\mathrm{sgn}(w) = \begin{cases} -j, & w > 0 \\ j, & w < 0 \\ 0, & w = 0 \end{cases} \tag{147}$$

In the frequency domain,

$$Z(w) = F(w) + j\bar{F}(w) = \begin{cases} 2F(w), & w > 0 \\ F(w), & w = 0 \\ 0, & w < 0 \end{cases}$$

where $F(w)$ is the Fourier magnitude spectrum of $f(t)$. Then, $z(t)$ is reconstructed using the inverse Fourier transform of $Z(w)$ where three seismic attributes can be extracted from it. Amplitude is found using Eq. 148 while phase is found using Eq. 149 and the frequency using Eq. 150 which represents rate of change of the time dependent phase.

$$A(t) = \sqrt{f^2(t) + \bar{f}^2(t)} \tag{148}$$

$$\theta(t) = \tan^{-1}[f^2(t) + \bar{f}^2(t)] \tag{149}$$

$$\bar{w}(t) = \frac{\partial \theta(t)}{\partial t} \tag{150}$$

Nath and Dewangan proposed the following algorithm to detect strength and location of reflections in seismic traces [56]. First, a sliding window is moved progressively along the corresponding attribute (amplitude or phase) and its corresponding fractal dimension within the window is estimated using the Divider and the Hurst Methods and then plotted. Whenever there is an object reflection, a sharp change should occur to both phase and amplitude attributes which cause a change in the graphed fractal dimension. One of the challenging tasks is the determination of the optimum window length. Results indicate that the amplitude attribute can be used to estimate strength of the reflection while the phase attribute can be used to estimate location of the reflection.

Oleschko et al. [49] developed a mathematical model to extract soil structure from a recorded radar trace using the Hurst method and a wavelet-based method. A strong relation between the fractal dimension of the detected radar traces and heterogeneity of the soil physical properties is shown. The proposed algorithm shows that the Hausdorff dimension of the radar signal is the same as the mass fractal dimension of the soil structure. The proposed fractal dimension mapping method is able to detect small changes in the mechanical and physical properties of the soil.

A seismic B-scan can be considered as a two-dimensional image and its fractal Brown movement (fBm) model can be developed [127]. Then, fractal reconstruction phase is used to enhance significant information and suppress non-

useful information in the raw seismic data since the image reconstruction phase can inherit texture characteristics of the raw image.

Zhao et al. [48] proposed an algorithm based on fBm to extract reflections of pipelines and soil interfaces from a raw GPR B-scan of a tunnel across a river. First, the GPR time series (A-scan) is proved to satisfy certain requirements of fBm characteristics since it can be regarded as a self-affine fractal set. Second, an fBm model is constructed for a radar B-scan by regarding it as a 2D image to extract its fractal and statistical features: the Hurst exponent H (connected with fractal dimension) and the standard deviation of time series σ (connected with GPR data distribution). Then, the successive random additions method is used to reconstruct the GPR B-scan which should contain the extracted significant information. In other words, the low-frequency components of the raw scan are suppressed while the useful high-frequency components are enhanced in the reconstructed scan.

Dogaru and Carin [128] analyzed the time-domain electromagnetic waves scattered from a target using three rough-surface statistical models in order to investigate effect of surface roughness on the detection performance. These statistical models are exponential, Gaussian, and fractal surfaces. The multi-resolution time-domain method is used to model physics of the underlying GPR wave. The fractal surface realized the greatest target-signature randomization which achieved best target detection performance where the target is assumed to reside under a randomly rough air-ground interface.

Don and Revathy used the fractal dimension and the fractal-signature techniques as feature extraction methods in order to describe the degree of self-similarity between pixel values within the clusters in mammographic images where K-means algorithm is used as a classification algorithm [129].

Depth Estimation

Velocity of radar waves through layers of the scanned medium can be approximated using Eq. 151.

$$V = \frac{c}{\sqrt{\varepsilon}} \tag{151}$$

where V is the speed in m/s, c is speed of light (2.998 x 10^8 m/s), and ε is the dielectric constant. Depth of a target can be estimated using ground truth (velocity analysis), dielectric table, or hyperbolic shape analysis [16]. Depth of targets deeper than 1.5" under the surface can be accurately measured by the 1.5 GHz antenna [16]. This is because reflections of targets within 1.5" from the surface are masked by the direct coupling.

Dielectric Table Method

Since the dielectric constant of concrete is in the range of [4.5,9], the dielectric table method assumes a dielectric constant value of 6.25 (dry concrete) or 9 (moister concrete) and estimates velocity of radar waves through the medium using Eq. 151. The depth of a defect can be determined using:

$$d_d = \frac{t\,c}{2\eta} = \frac{tc}{2\sqrt{\mu\varepsilon}} \approx \frac{tc}{2\sqrt{\varepsilon}} \approx \frac{Vt_d}{2} \qquad (152)$$

where the round-trip travel time to a target is estimated from the deconvolved signal (as demonstrated by Eq. 213), η represents refractive index of the target medium, ε represents the dielectric constant of the scanned medium, and μ is its relative permeability. In case of bridge decks, μ is approximately equal to one and thus $\eta \approx \sqrt{\varepsilon}$. It is worthy to state here that the dielectric table method is the simplest but the least accurate method [16].

Velocity Analysis Method

It approximates velocity of radar waves based on a known depth of a target. In bridge decks, depth of rebar is known and can be used for this purpose. Other applications include drilling a target at known depth or using the concrete slab thickness in case of a visible slab bottom. In case of having several layers inside the scanned medium and a known target depth, the computed velocity is the average velocity within these layers. However, for accurate results, the thickness ratio of these layers should not change along the scan. The velocity analysis method will be explained in proposed framework.

Hyperbolic Shape Analysis (Migration) Method

In the hyperbolic shape analysis method, the velocity of radar waves is estimated based on shape of the hyperbolic reflections using the Migration function in RADAN. If the materials in the scanned medium have low dielectric constant, radar waves propagate with high velocity resulting in wide hyperbolae. This method does not require prior knowledge of a target's depth (such as rebar) to estimate the velocity, but requires a skillful operator to work with RADAN to manually analyze the GPR scans.

Summary and Conclusions

A variety of ICA-based algorithms have been utilized to solve the blind source separation problem. One of the inherited problems in the ICA estimation is the dependency of the learning rule on the learning rate (step size). In case of GPR, the learning rate should be estimated for each scan where most likely the estimated value is not the optimal one resulting in a less reliable ICA algorithm due to higher execution time and degraded separation performance. Therefore, several solutions to this problem were proposed that does not depend on the learning rate such as relative Newton method [89], JADE algorithm [70], FastICA algorithm [23], and relative trust-region method [90].

The FastICA algorithm is the most reported method in the literature since it has the following properties [35, 38, 39]: easy implementation since it does not depend on the learning rate, has high separation performance, and has fast

convergence speed. In this work, EFICA is selected for the ICA estimation since it was reported with slightly higher computational demands than FastICA algorithm, but it has superior separation performance in comparison with JADE and FastICA algorithms [39].

Methods to estimate the fractal dimension such as the Hurst method [56], the Divider method [52, 56], the differential box counting method (DBC) [54], fractal Brownian motion method (fBm) [48], to name a few, have been reported. Applying these methods to the same data of a specific problem, such as a GPR scan, does not necessarily result in the same estimated value of the fractal dimension. These differences are due to the estimation algorithm used by a particular method. Therefore, choice of the suitable method is an application dependent. Since no previous work was done to compare between these methods suitability for GPR data, a comparison between most used methods is presented in this work, namely: fBm, DBC, and Hurst methods.

Since reflections from closely spaced targets overlap in any GPR application, conventional signal processing methods, such FFT and matched filtering, could not recover their corresponding pulses (spikes of the impulse response) when the time spacing between them is less than $1/B$ (0.625 ns for the 1.5GHz antenna) where B is bandwidth of the GPR antenna.

Since front and tail of the incident pulse have very small values, the convolution matrix F will be a large matrix with near zero main diagonal values and so it is an ill-conditioned matrix that may not have an inverse [92]. If an inverse does

exist, F^{-1} may be extremely sensitive to additive noise in g and consequently, h may not be easily estimated. Therefore, conducting simple deconvolution methods such as linear least squares is not enough to deconvolve raw GPR scans due to noise sensitivity and ground scattering [92].

Deconvolution methods can be divided into direct and blind deconvolution. Direct deconvolution methods assume a known incident pulse f and attempts to recover the impulse response of the scanned medium h assuming it is following a linear system model, $g = f * h$, where g is the reflected raw GPR data.

On the other hand, blind deconvolution methods estimate both f and h from raw data with no prior information about f or h. Direct deconvolution methods are simpler, less computationally demanding, but less accurate in comparison with blind deconvolution methods [130].

Deconvolution can be performed using the data in the frequency-domain or in time-domain. In frequency-domain, the convolutive mixture becomes multiplicative mixture of complex signals within different frequency bands [110]. In the frequency domain, the independent signals have different frequency representations. Since frequency-domain ICA methods separate the mixture within each frequency band, the mixing matrix becomes a function of the frequency in the new domain while the time-domain ICA model assumes it to be a constant [23]. In other words, all frequency components of each impulse response signal must be grouped. This makes the permutation and scaling of the impulse response signals lose consistency across all frequency bands. For this, additional methods to solve the permutation and scaling

ambiguity are required to avoid combining contributions from different impulse response signals into a single impulse response when reconstructing the signal in the time-domain which adds more computational demanding [110]. The literature reported different methods for blind deconvolution using ICA such as banded-ICA [23, 111, 103, 109], skewed-spatiotemporal ICA [104], and complex-ICA [113]. Both banded-ICA and skewed-spatiotemporal ICA work in time-domain while complex-ICA works in frequency domain.

To avoid the previously mentioned difficulties associated with frequency-domain, deconvolution in the time-domain will be considered in this work. Banded-ICA deconvolution algorithm resulted in robust results [111, 103, 109] and outperformed both singular value decomposition (SVD) and subset selection deconvolution (SSDA) algorithms in deconvolving GPR scans of simulated concrete bridge decks [130]. The estimated independent component was modeled as a delayed version of the original impulse response signal which prevented accurate depth estimation since their estimates were not with respect to the ground-line. Also, they determined the value of the approximative convolutional model m manually for each case or they assumed it equal to the length of the corresponding GPR trace, n [111, 103, 109, 130].

In this work, a modified version of the banded-ICA algorithm is developed to overcome all limitations that prevented automation of the detection process. A zero-correction step is applied to the estimated independent components to allow depth

estimation with respect to the ground line. Also, the estimated incident pulse is utilized to select the best independent component among m estimated ones.

Different methods for depth estimation such as dielectric table method, velocity analysis method, and hyperbolic shape analysis method are reported in the literature [3, 6, 16]. The dielectric table method is the simplest but the least accurate method while the hyperbolic shape analysis method does not require prior knowledge of a target's depth (such as rebar) to estimate the velocity. However, it requires a skillful operator to work with RADAN to manually analyze the GPR scans. Velocity analysis method was reported with significantly high depth estimation accuracy [3, 6] and since it does not require a skillful operator or additional post processing, I used it as the tool for defects depth estimation in this dissertation.

CHAPTER 4

A NEW FRAMEWORK FOR DETECTING EMBEDDED DEFECTS

The framework consists of four stages to identify, localize, and characterize defects in concrete as shown in Figure 6: a fractal-based feature extraction stage to detect defective regions and localize them horizontally, a banded-ICA stage for the deconvolution of the defective region traces, a velocity analysis stage to estimate the depth of defects, and a classification stage to characterize detected defects. As a preprocessing step, a zero-correction is performed on raw B-scans that involves the deletion of first part of the B-scan all the way to the first positive peak of the ground band reflection. This step is intended to enable depth estimation with respect to the ground line. Next, the proposed defect detection algorithm will be summarized.

Fractal-Based Feature Extraction

In this work, a fractal-based feature extraction (FBFE) algorithm is proposed and applied to each A-scan extracted from the B-scan. FBFE can be summarized in the following steps after normalizing all traces to zero mean,

1. Feature vectors can be constructed according to the scanning method (as will be shown by Figure 7A). In the case of having the survey line perpendicular to the rebar, the dominant reflections are from rebar and targets above it as will be shown in Figure 15. When the survey line is parallel to the rebar, most of

the reflections are from deeper objects with rebar reflections being minimized as will be shown in Figure 11E. These differences caused us to seek different feature vectors to characterize traces.

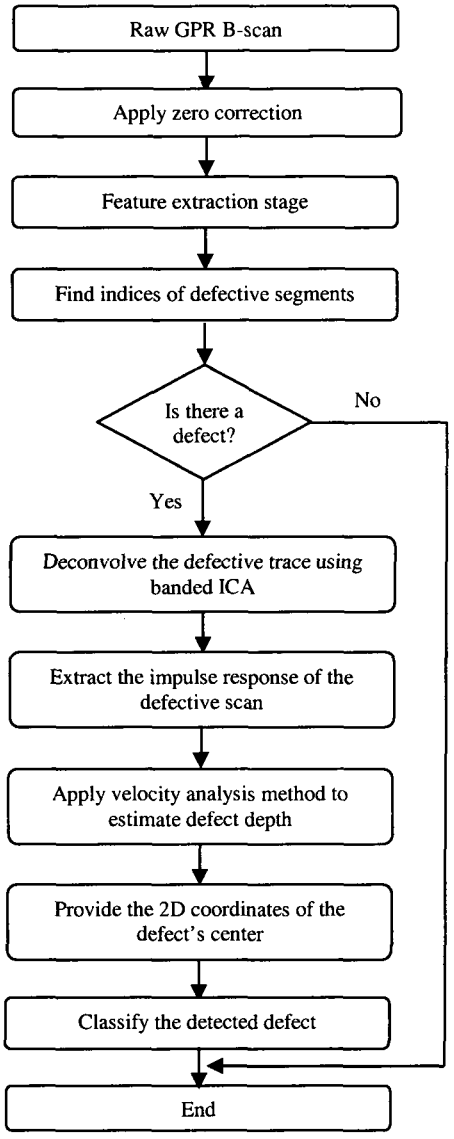

Figure 6. Block diagram of the proposed framework for defect detection.

- A: When the scan is conducted perpendicular to the rebar, the first feature vector, f_G consists of time-domain statistical features such as: fractal dimension (FD), root mean square (RMS), energy, and number of local maximum points (peaks). The RMS and energy are determined from the GPR trace and from its DFT real part values. The second feature vector, f_L contains the value of the summation of peak-to-peak value at each local maximum point.
- B: When the scan is parallel to the rebar, the first feature vector, f_G is constructed in the same way as in A above, but without the energy feature since it proved to be not efficient for defective and healthy traces while the second vector feature, f_L, containing the square mean root (SMR) of each GPR trace which is viewed as a reasonable compromise between the geometric and the arithmetic means.

2. Three methods are used to estimate the fractal dimension of the GPR trace $x(t)$,

- In fractal Brownian motion, the logarithmic plot of $E[x(t+\Delta t)-x(t)]^2$ versus step size, Δt is obtained for various values of the step size Δt and the fractal dimension, *FD* is estimated as:

  $FD = 2 - H$

  where H is the Hurst exponent obtained from the slope of the graphed line.

- In the differential box counting, the GPR trace is covered with three boxes. The first box completely covers the trace; the second box covers the first half of the trace while the third box covers the second half of the trace, leading to the following FD:

$$FD = \frac{\log(N_1 + N_2) - \log N_3}{\log 2}$$

where N_i is the maximum difference of data within its corresponding i^{th} box.

- In the Hurst method, windows of different sizes are used where the maximum difference of data within each window is computed and plotted against its corresponding window size in the logarithmic space. Fractal dimension is estimated as $FD = 2 - H$

3. Each feature vector is a column in the feature matrices F_G and F_L.

4. The mean and standard deviation vectors, ($f_{G\mu}, f_{L\mu}, f_{G\sigma}, f_{L\sigma}$) are computed for all rows of the feature matrices F_G and F_L.

5. The index of segments of lengths greater than a threshold *T*, corresponding feature vectors in F_G that are greater than $f_{G\mu} + \alpha f_{G\sigma}$, and corresponding feature vectors in F_L matrix that are smaller than $f_{L\mu} + \alpha f_{L\sigma}$ where $\alpha \in [0,1]$ are identified as defective regions.

6. The center trace within each defective segment is labeled as by the defect's column number. Dividing it over the total number of columns in the B-scan

and multiplying it by the length of the B-scan, the horizontal location of defect (distance) is determined as:

$$X_d = \frac{D_c}{T_c} \times L_s \tag{153}$$

where D_C represents defect's column number, T_C represents total number of columns in the B-scan (350-550 depending on the scan), and L_S represents length of the B-scan in *inches*.

A number of statistical measures are used in this work to evaluate performance of the proposed FBFE algorithm. Precision is used as a measure of fidelity and recall as a measure of completeness. False negative *FN* also used as the case of failure to detect a defect and false positive *FP* as the case of declaring a non-existing defect. Precision is defined according to Eq. 154 while recall is defined according to Eq. 155 as follows:

$$Pc = \frac{TP}{TP + FP} \tag{154}$$

$$Rc = \frac{TP}{TP + FN} \tag{155}$$

where *TP* represents number of correctly classified defects.

Deconvolution Using Banded-ICA

The proposed banded-ICA algorithm is a modified version of a previously developed algorithm in [130]. In the current version, we intended to enhance sparseness of the estimated independent components. Also, the independent

component with highest sparseness was selected as the best candidate while in the current version a more effective criterion is used to select the best candidate. It can be summarized in the following steps.

1. Let x represents a defective trace marked by the FBFE algorithm.
2. The mixture matrix X is constructed according to Eq. 108:

   $$X_{m\times n} = [x(t-m+1)\ x(t-m+2)\ \cdots\ x(t-1)\ \ x(t)]^T$$

3. Since the impulse response signal to be estimated is sparse and consists a number of sharp spikes with relatively flat area between them (representing the layered structure of the scanned concrete deck), the separating matrix $W_{m\times m}$ is initialized to the identity matrix.
4. Mahalanobis transformation is used to whiten the mixture matrix X as follows:

   $$Z = C^{-0.5}(X - \overline{X})$$

   where $\overline{X}$ represents the mean of the mixture matrix X and C represents its covariance matrix. The whitening step is a decorrelation process that results in an identity covariance matrix.
5. In this step, three ICA algorithms are used to decompose the whitened mixture which consequently recovers m independent components. They are FastICA, EFICA, and Pearson ICA.

   5.1. The FastICA algorithm has the following learning rule (Eq. 37). The FastICA algorithm does not depend on the learning rate and thus it offers easy implementation, has high accuracy, and fast convergence speed as was

reported in the literature [23]. Since the impulse response signal to be estimated is a sparse signal, the hyperbolic tangent is selected for the nonlinear function g since it is more suitable for super-Gaussian signals [23].

$$w_{k+1} = w_k + E[zg\ (w_k^T z)] - E[g'(w_k^T z)]w_k$$

$$w_{k+1} = \frac{w_k}{\| w_k \|}$$

5.2. Efficient FastICA (EFICA) is a statistical efficient version of the FastICA algorithm [39] with an added computationally complexity. However, it has superior separation performance as was reported in the literature. First, the FastICA is executed using the hyperbolic tangent function until convergence of the weight matrix or the maximum number of iterations reached. Then, the FastICA algorithm is used with the following nonlinear function for the kth estimated independent component in order to refine the estimated independent components.

$$g_k(y) = \begin{cases} ye^{-3.348|y|}, & \mu_{4k} > 3 \\ sign(y).|y|^{\min\{\alpha_k - 1.14\}}, & 1.8 < \mu_{4k} \le 3 \\ sign(y).|y|^{14}, & \mu_{4k} \le 1.8 \end{cases}$$

where μ_{4k} represents the fourth-order moment of the kth independent component and the parameter α_k is defined as $\alpha_k = [0.29\sqrt{\mu_{4k} - 1.8} - 0.185(\mu_{4k} - 1.8)]^{-1}$.

5.3. The Pearson ICA algorithm uses few parameters to model wide class of distributions. Pearson-ICA can be summarized in the following steps:

I. Moments of the independent components are computed.

II. For each independent component, parameters of its nonlinear function are estimated based on the computed moments.

III. FastICA algorithm based on likelihood maximization is used with the estimated nonlinear functions.

IV. The previous steps are repeated until convergence.

6. The selected independent component is to satisfy the following conditions,

- It should have a minimum number of spikes n_s (since the defect may mask reflections from targets beneath it) as defined by:

$$n_s = n_i - 1 \tag{156}$$

where n_i is the number of interfaces in the scanned concrete slab.

- It should lead to the minimum mean square error (MMSE) according to:

$$MSE_i = E(g - f * y_i)^2 \tag{157}$$

where g is the GPR trace and $*$ represents the convolution process between the estimated incident pulse of the GPR antenna f and the current estimated independent component or its horizontally flipped version, y_i, for all $i = 1,\ldots,m$.

Depth Estimation Using Velocity Analysis

The velocity-analysis based depth estimation method can be presented as:

1. The initial total range of a GPR trace is 10ns. After applying zero-correction step, the new total range of the GPR trace is found using:

$$T_n = \frac{L_z}{L_o} \times T_r \tag{158}$$

where L_z represents the length of the trace after zero correction, L_o represents the original length of the trace, and T_r represents the total range.

2. Velocity of radar waves through the scanned medium can be approximated using Eq. 159 and noting that the depth of the rebar, d_r in the concrete bridge decks is known,

$$V = \frac{2d_r}{t_r} \tag{159}$$

where V is the speed in inch/s, d_r is rebar depth in inches, and t_r is the round-trip travel time to the rebar.

3. Using Eq. 152, the actual depth of the defect is estimated. The spikes in the deconvolved trace are on a one-to-one correspondence with the ground line, defect, rebar, and bottom of the simulated deck. The round-trip travel time of the object (rebar or defect) t_{rd} is determined as:

$$t_{rd} = \frac{S_{rd}}{L_{tr}} \times T_n$$

where S_{rd} represents the corresponding spike location of the object, L_{tr} represents the deconvolved trace length, and the new total range T_n was defined in Eq. 158.

Classification Using PCA and Euclidean Distance

The proposed defect classification algorithm can be summarized in the following steps.

1. Eight defective sub-images are cropped from five raw scans and scaled into 45x45 pixels based on center of each defect to form the training set. Four sub-images are extracted from delamination defects while the other four are extracted from air-void defects.
2. Principal component analysis algorithm is used to identify main features from the training set and the labeled defective region.
3. Euclidean distance is used as a similarity measure to match the selected features from the defective region with those of a sub-image from the training set to classify the defective region into air-void or delamination defect.

CHAPTER 5

EXPERIMENTAL RESULTS

Experimental Setup

Six concrete slabs were constructed to simulate bridge decks: three with several embedded defects of known dimensions and locations as shown in Figure 7A and Tables 1 to 3, and three with no defects. PVC pipes and Styrofoam blocks were used to model air-void and delamination defects, respectively [3]. The slabs are 45 inches long and 45 inches wide with one bottom layer of rebar (number 5 steel) in both directions spaced at 6 inches on center. Two slabs are of 3.54 inches thickness with the rebar placed 1.3 inches and 1.93 inches deep with rebar cover of 1 inch. The other two slabs are of 5.51 inches thickness with the rebar placed 2.53 inches and 3.15 inches deep with rebar cover of 1.74 inches. The final set of slabs is of 7.32 inches thickness with the rebar placed 4.33 inches and 4.96 inches deep with rebar cover of 1.74 inches. Figure 7 shows a 4-inch slab during the construction phase, Figure 8 shows a schematic diagram of a 6-inch slab, Figure 9 shows a plan view of a 4-inch slab, and Figure 10 shows a plan view of 6- and 8-inch slabs [3].

A 1.5 GHz (GSSI model 5100) bistatic antenna is used to scan the concrete slabs [18]. Thirty one scans (labeled as scans 1 through 31) of length 37-38 inches were collected of which eleven for the healthy slabs and twenty for the defective ones as shown in Table 4. Twenty six scans (scans 1 through 26) are used to test

performance of the FBFE and deconvolution algorithms while five scans (scans 27 through 31) are used as training set for the classification stage.

The dielectric constant of concrete is in the range [4.5, 9]. An initial scan is conducted and the migration function in RADAN is used to estimate the velocity of radar waves and to find the dielectric constant (as will be shown in the parameter analysis section). This analysis resulted in a dielectric constant of 6.25 for the simulated concrete bridge decks. Therefore, during the data collection process, the dielectric constant is set to 6.25. As Table 4 shows, the scans are collected for the 4-, 6-, and 8-inch concrete slabs with different acquisition parameters to test performance of the proposed defect detection algorithm with different acquisition parameters.

After collecting the scans, RADAN is used to remove the gain applied to data during the data collection process. Therefore, the proposed defect detection algorithm processes raw GPR scans.

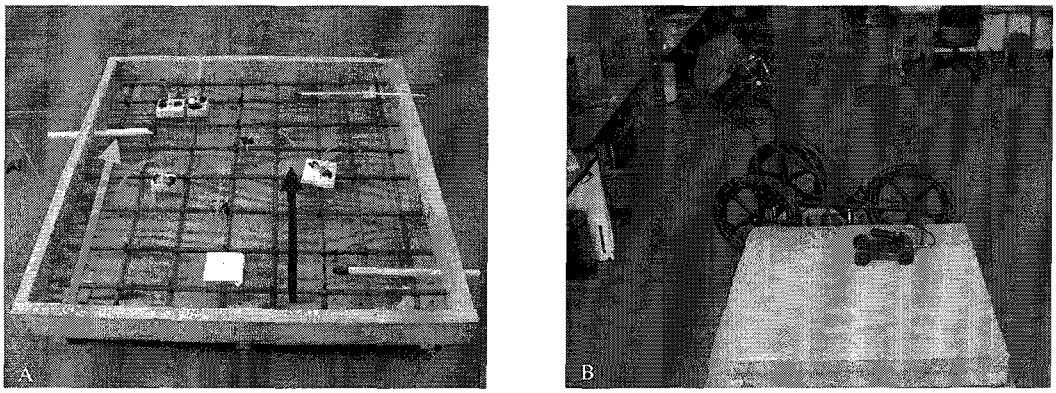

Figure 7. Simulated 4-inch concrete bridge deck with embedded defects. (A) during the construction phase. The green arrow indicates a parallel survey line to the first rebar from left while the red arrow indicates a perpendicular survey line to the rebar. (B) the finished slab.

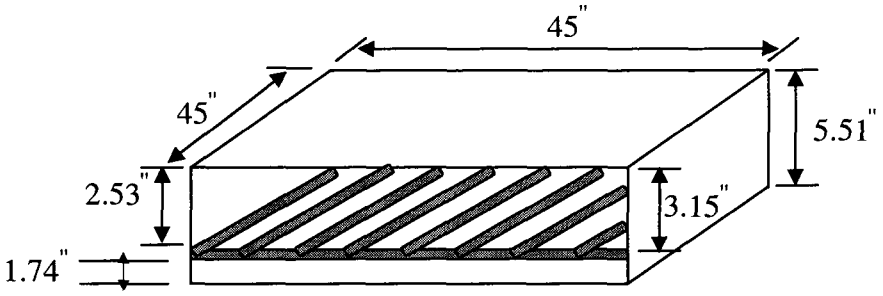

Figure 8. Schematic diagram of a 6-inch simulated concrete bridge deck.

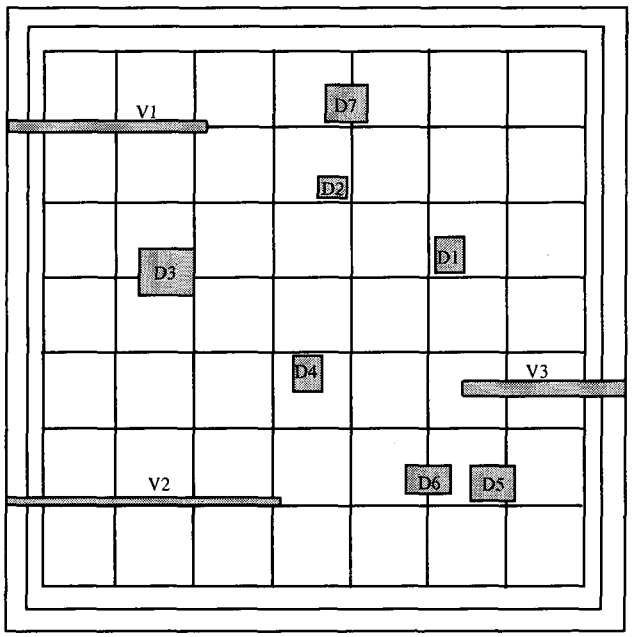

Figure 9. Plan view for the 4-inch slab.

Table 5 shows results of applying the fBm, DBC, and the Hurst methods to the fifteen defective scans. The fBm method detects fourteen of the fifteen defective scans while the Hurst method detects partially (not all the defective regions within the scan) thirteen of the fifteen defective scans and the DBC method detects partially ten

defective scans only. In the case of the fBm, the average difference between the estimated and actual horizontal location is 0.76 inches which indicates the applicability of the proposed algorithm in this field. In case of the DBC method, the average difference is1.42 inches while the Hurst method has an average difference of 0.93 inches. Finally, the proposed FBFE algorithm is able to detect and mark the defective regions using only the underlying B-scan with no need for the number of B-scans for algorithm training.

Results of the Defect Detection Algorithm

Table 6 shows false positive, false negative, accuracy, precision, and recall for the fBm, DBC, and Hurst methods applied the twenty six scans. The fBm method has the highest accuracy, precision, and recall, and the lowest false negative and false positive rates. The DBC method results in the highest false negative rate while the Hurst method results in the highest false positive rate.

The DBC method estimates the fractal dimension of a signal through covering it with three boxes and taking the maximum difference between data within each box. This is an approach that is easy to implement but it does not take into account all neighbor-to-neighbor pixel variations, causing DBC results to be the least accurate. The Hurst method follows a similar approach but with increased number of windows that cover the signal. This covers more neighbor-to-neighbor pixel variations than with DBC method and thus it should have a clear improvement in the results. The fBm method takes the difference between a shifted version of the signal and the

original signal which covers more neighbor-to-neighbor pixel variations than the other two methods and consequently improves the results.

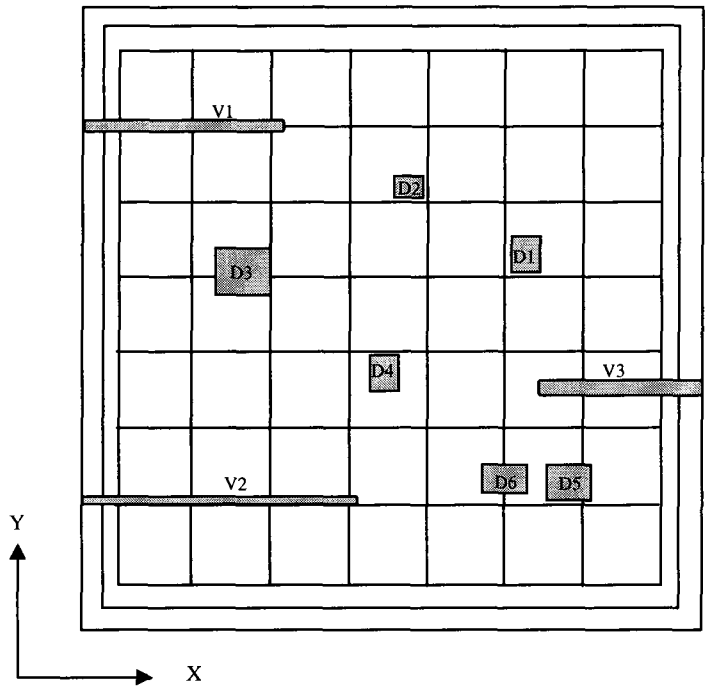

Figure 10. Plan view for the 6- and 8-inch slabs.

Table 1

Defects map of the 4-slab

Defect	Length	Width	Thickness	Coordinates (X,Y)		Depth
D1	2	3	2	33.5	26	1
D2	2	1.5	1	26	30.5	0.75
D3	4	4	0.5	13	24.5	0.75
D4	1.5	1.5	1.5	23	18	1.25
D5	4	3	2	34.5	10	1.25
D6	3	3	1.5	31.2	10.2	1.5
D7	3	3	0.5	24.5	39	1.43
V1	12	NA	0.5	NA	39.25	1.58
V2	16	NA	0.25	NA	7.5	1.02
V3	12	NA	1.0	NA	16	1.65

Table 2

Defects map of the 6-slab

Defect	Length	Width	Thickness	Coordinates (X,Y)		Depth
D1	2	3	2	30.0	26	1
D2	2	1.5	1	23	32.5	3
D3	4	4	0.5	9	25.5	2.03
D4	1.5	1.5	1.5	22	19	2
D5	4	3	2	34.5	8.5	1.15
D6	3	3	1.5	30.25	9	1.65
V1	12	NA	0.5	NA	37	1.42
V2	16	NA	0.25	NA	7	4.49
V3	12	NA	1.0	NA	16	0.39

Table 3

Defects map of the 8-slab

Defect	Length	Width	Thickness	Coordinates (X,Y)		Depth
D1	2	3	2	34.0	25.5	2.75
D2	2	1.5	1	23.5	30.5	3.96
D3	4	4	0.5	10.5	25	3.83
D4	1.5	1.5	1.5	23.5	20.5	3.46
D5	4	3	2	36	9	2.33
D6	3	3	1.5	31	9.25	2.83
V1	12	NA	0.5	NA	37.5	2.91
V2	16	NA	0.25	NA	7	3.62
V3	12	NA	1.0	NA	16.5	2.6

Since the best results are obtained using the fBm method, the marked defective traces from this method were used for the deconvolution process to reduce overlapping between reflections from adjacent targets and estimate the round-trip travel time to and from the embedded defects and rebar. The FastICA, EFICA, and Pearson ICA methods are used for the ICA decomposition within the deconvolution algorithm to test their performance for the depth estimation accuracy and select the best one. The velocity analysis method is used to estimate depth (in inches) of

detected defects as shown in Table 7. As the table indicates, the EFICA method has better results than FastICA and Pearson ICA methods. For EFICA, the estimated depth in all scans is within 0.85 inches from the actual depth with average difference between the actual and estimated depth of 0.37 inches. Also, the proposed EFICA-based deconvolution algorithm is able to detect shallow defects of depth 0.39 inches. The antenna wavelength is $\lambda = c/f = 3x10^8/1.5x10^9 = 0.2m = 20cm = 7.8$ inches. This indicates that the deconvolution algorithm is able to resolve overlapping reflections of targets spaced at $0.43/7.8 = 6\%$ of the antenna wavelength which indicates robustness of the proposed deconvolution algorithm.

In some cases, the FBFE algorithm detects portion of the defective region. Therefore, the defective trace to be deconvolved did not represent peak of the defect. This indicates that the estimated depth of that defect is not its actual depth. Also, the defect maps in tables 1 through 3 are not accurate since some of the embedded defects moved a little bit from their actual location indicated by the tables during the concrete casting process. Also, there is a depth measurement error. All these factors affect the depth estimation accuracy.

Tables 8 and 9 summarize results of Table 5 and Table 7 for each concrete slab. The percent error is obtained by dividing the average difference between the estimated and actual horizontal location over scan length (Table 8) and the average difference of the depth over the corresponding slab thickness (Table 9). In case of fBm-based feature extraction, best results of the horizontal location estimation accuracy are obtained for the 8-inch slab while the 6-inch slab has the best results for

the depth estimation accuracy. The estimated horizontal location is within ±2.4% error in all slabs while the estimated depth is within ±12.45% error in all slabs.

Table 4

Collected scans from 6, 4, and 8-inch slabs with different acquisition parameters

Scan	Slab Thickness	Defects		Samples/scan	Scans/inch
1	6-inch	None		512	8.33
2	6-inch	Air-void (V1)		512	8.33
3	6-inch	None		512	8.33
4	6-inch	None		512	8.33
5	6-inch	Delamination (D1)	Delamination (D3)	512	8.33
6	6-inch	Air-void (V3)		512	8.33
7	6-inch	None		512	8.33
8	6-inch	Air-void (V3)		512	8.33
9	6-inch	Delamination (D6)	Delamination (D1)	512	8.33
10	6-inch	None		512	8.33
11	6-inch	Delamination (D4)		512	8.33
12	6-inch	Delamination(D3)	Air-void (V1)	512	8.33
13	4-inch	Delamination (D7)	Air-void (V1)	512	8.33
14	4-inch	None		512	8.33
15	4-inch	None		512	8.33
16	4-inch	None		512	8.33
17	4-inch	Delamination (D1)		512	8.33
18	4-inch	Delamination (D1)	Delamination (D3)	256	10
19	4-inch	Air-void (V3)		256	10
20	4-inch	Delamination (D5)	Delamination (D6)	256	10
21	4-inch	None		256	10
22	8-inch	Delamination (D5)	Air-void (V3)	256	10
23	8-inch	None		256	10
24	8-inch	None		256	10
25	8-inch	Delamination (D3)		256	10
26	8-inch	Delamination (D5)	Delamination (D6)	256	10
27	6-inch	Air-void (V1)		512	8.33
28	6-inch	Delamination (D6)	Air-void (V1)	512	8.33
29	4-inch	Delamination (D3)		512	8.33
30	4-inch	Delamination (D7)	Air-void (V1)	256	10
31	8-inch	Delamination (D5)	Air-void (V3)	256	10

Table 5

Actual and estimated horizontal location of defects*

Scan Number	Actual Location	Estimated Location (fBm)	Diff	Estimated Location (DBC)	Diff	Estimated Location (Hurst)	Diff
Scan2	33.75	34.28	0.53	32.02	1.73	ND	ND
Scan5	8.75	8.71	0.04	10.98	2.23	8.83	0.08
	28.75	29.19	0.44	28.4	0.35	29.76	1.01
Scan6	3.5	3.83	0.33	ND	ND	3.94	0.44
Scan8	10.75	9.63	1.12	ND	ND	9.63	1.12
Scan9	3.75	3.45	0.3	4.16	0.41	2.26	1.49
	19	17.61	1.39	ND	ND	18.8	0.2
Scan11	14.5	ND	ND	ND	ND	ND	ND
Scan12	19.5	21.61	2.11	21.61	2.11	21.61	2.11
	32	32.76	0.76	ND	ND	32.76	0.76
Scan13	13.75	12.57	1.18	ND	ND	13.68	0.07
	33.75	35.26	1.51	ND	ND	34.56	0.81
Scan17	4.75	4.2	0.55	7.58	2.83	3.07	1.68
Scan18	5.25	4.38	0.87	8.16	2.91	6.95	1.7
	24.75	24.82	0.07	ND	ND	24.63	0.12
Scan19	3	4.62	1.62	ND	ND	4.62	1.62
Scan20	3.25	2.99	0.26	1.91	1.34	2.75	0.5
	7.05	8.03	0.98	ND	ND	6.55	0.5
Scan22	2.75	2.28	0.47	2.22	0.53	5.26	2.51
	9.2	8.24	0.96	ND	ND	8.24	0.96
Scan25	29	29.64	0.64	29.64	0.64	29.64	0.64
Scan26	3.25	3.03	0.22	1.95	1.3	3.03	0.22
	7.05	7.36	0.31	7.65	0.6	ND	ND
Average			**0.76**		**1.42**		**0.93**

* All measurements are in inch and *ND* means the defective segment is not detected

Table 6

False positive, false negative, accuracy, precision, and recall for the fBm, DBC, and Hurst algorithms

Algorithm	FP Rate	FN Rate	Accuracy	Precision	Recall
fBm	7.69%	2.56%	89.74%	88%	95.65%
DBC	12.82%	28.21%	58.97%	70.59%	52.17%
Hurst	20.51%	7.69%	71.8%	71.43%	86.96%

Table 7

Actual and estimated depth of the detected defects*

Scan Number	Actual Depth	Estimated Depth (EFICA)	Diff	Estimated Depth (FastICA)	Diff	Estimated Depth (Pearson)	Diff
Scan2	1.42	1.68	0.26	1.42	0	1.21	0.21
Scan5	1	1.11	0.11	2.81	1.81	0.14	0.86
	2.03	2.53	0.5	2.05	0.02	1.09	0.94
Scan6	0.39	0.48	0.09	0.53	0.14	0.21	0.18
Scan8	0.39	0.43	0.04	0.39	0	0.20	0.19
Scan9	1.65	1.7	0.05	1.48	0.17	0.77	0.88
	1	1.36	0.36	0.16	0.84	0.10	0.90
Scan11	2	NA	NA	NA	NA	NA	NA
Scan12	2.03	1.59	0.44	1.17	0.86	1.78	0.25
	1.42	1.18	0.24	1.13	0.29	1.95	0.53
Scan13	1.43	1.76	0.33	1.23	0.2	0.74	0.69
	1.58	1.35	0.23	2.17	0.59	0.78	0.80
Scan17	1	1.62	0.62	1.9	0.9	0.71	0.29
Scan18	1	1.16	0.16	0.16	0.84	0.10	0.90
	0.75	1.01	0.26	1.39	0.64	0.14	0.61
Scan19	1.65	0.92	0.73	1.08	0.57	0.22	1.43
Scan20	1.25	0.74	0.51	0.38	0.87	0.21	1.04
	1.5	0.8	0.7	0.25	1.25	0.29	1.21
Scan22	2.33	2.28	0.05	2.08	0.25	2.49	0.16
	2.6	2.29	0.31	2.29	0.31	3.50	0.90
Scan25	3.83	3.22	0.61	3.22	0.61	3.72	0.11
Scan26	2.33	1.58	0.75	3.68	1.35	4.29	1.96
	2.83	3.68	0.85	3.68	0.85	3.90	1.07
Average			**0.37**		**0.61**		**0.73**

* All measurements are in inch, *NA* means the defective segment is not detected

Figures 11A, 11B, 11C, and 11D show healthy scans from a 6-inch concrete slab (scan1, scan3, scan4, and scan7) with the survey line perpendicular to the rebar while Figure 11E (scan10) shows a healthy B-scan with the survey line parallel to the rebar. The three interfaces as shown in each B-scan are: air-concrete, concrete-rebar, and concrete-substrate. No defective segments are declared by the proposed fractal-based feature extraction algorithms after applying them to the aforementioned scans.

Table 8

Average difference between actual and estimated location of the corresponding slab defects

Slab Thickness	Average Diff (fBm)	Percent Error	Average Diff (DBC)	Percent Error	Average Diff (Hurst)	Percent Error
4-inch	0.88	2.38%	2.36	6.39%	0.88	2.38%
6-inch	0.78	2.05%	1.37	3.61%	0.9	2.37%
8-inch	0.52	1.37%	0.77	2.03%	1.08	2.84%

Table 9

Average difference between actual and estimated depth of the corresponding slab defects

Slab Thickness	Average Diff (EFICA)	Percent Error	Average Diff (FastICA)	Percent Error	Average Diff (Pearson ICA)	Percent Error
4-inch	0.44	12.43%	0.79	22.32%	0.87	24.58%
6-inch	0.23	4.17%	0.38	6.9%	0.55	9.98%
8-inch	0.51	6.97%	0.85	11.61%	0.84	11.48%

Figures 12A and 12B show healthy scans from a 4-inch concrete slab (scan14 and scan15) with the survey line parallel to the rebar while Figures 12C and 12D (scan16 and scan21) show healthy scans with the survey line perpendicular to the rebar. 512 samples are collected per scan for Figures 12A, 12B, and 12C while 256 samples are collected per scan for figure 12D. This means that every trace of scan21 contains half the number of the samples that traces of the other scans contain. Therefore, these scans show more visible reflections from deep objects with better resolution in comparison with scan21. On the other hand, scan21 takes roughly half the execution time required by the other scans. No defective segments are declared by

the proposed fractal-based feature extraction algorithms after applying them to the aforementioned scans.

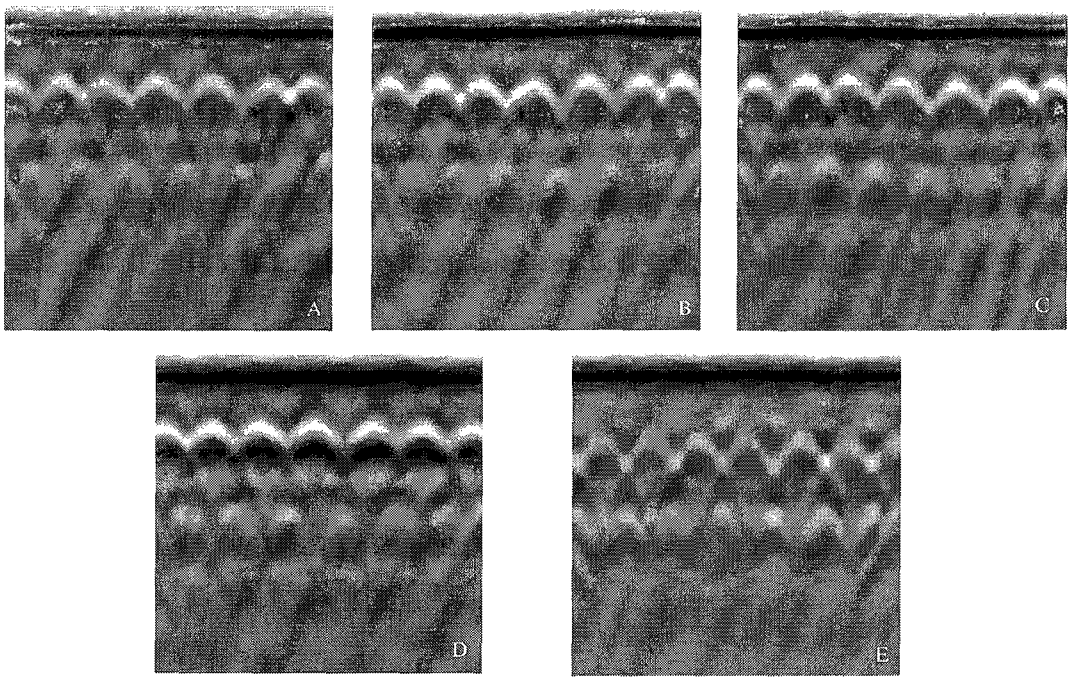

Figure 11. A, B, C, D, and E are healthy scans from a 6-inch concrete slab.

Figures 13A and 13B show healthy scans from an 8-inch concrete slab (scan23 and scan24). Since the width to thickness ratio of the 8-inch concrete slab is not large enough, there is a diagonal arc under the rebar represents interference caused by the reflections from the right side boundary of the slab. This diagonal arc may mask reflections from deeper objects. One possible solution to reduce the interference is by using the migration method. No defective segments are declared by the proposed fractal-based feature extraction algorithms after applying them to the aforementioned scans.

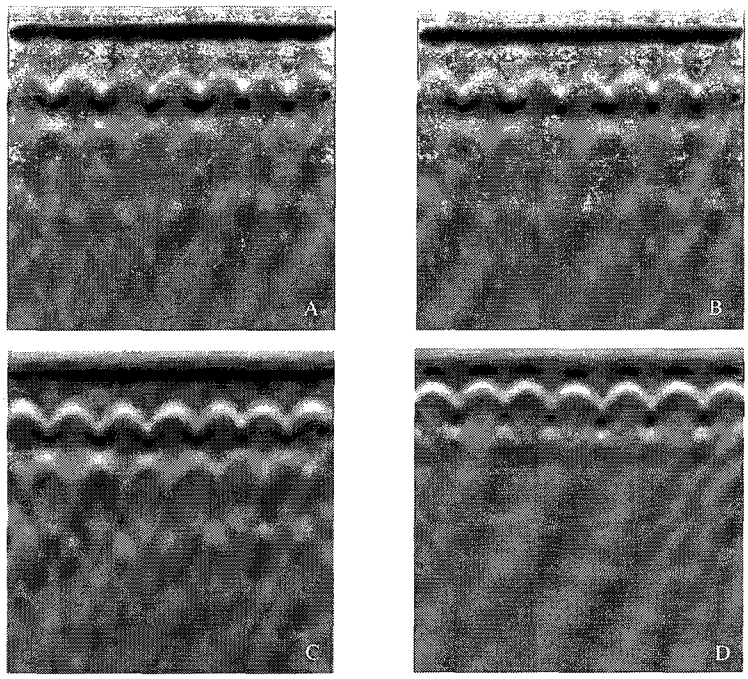

Figure 12. A, B, C, and D are healthy scans from a 4-inch concrete slab.

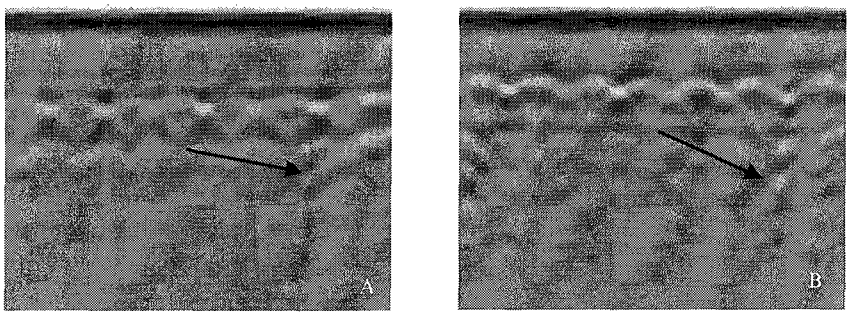

Figure 13. A and B are healthy scans from an 8-inch concrete slab.

Applying the fBm-based feature extraction algorithm on a given raw B-scan followed by the EFICA-based deconvolution algorithm, results in marking the defective regions. Figure 14A shows raw B-scan (scan2) with an embedded air-void

defect that lies above the sixth and seventh hyperbolae. Figure 14B shows the processed scan where the defective region is partially marked. Figure 15A shows raw B-scan (scan5) with two embedded delamination defects. The first defect is shallower than the other one and is partially overlapped with the black section of the ground coupling band while the second defect lies right above the sixth hyperbola. Figure 15B shows the processed scan where the two defective regions are successfully marked. Figure 16A shows a raw B-scan (scan6) with a shallow embedded air-void defect that is overlapped with the black section of the ground coupling band. Figure 16B shows the processed scan where the defective region is successfully marked.

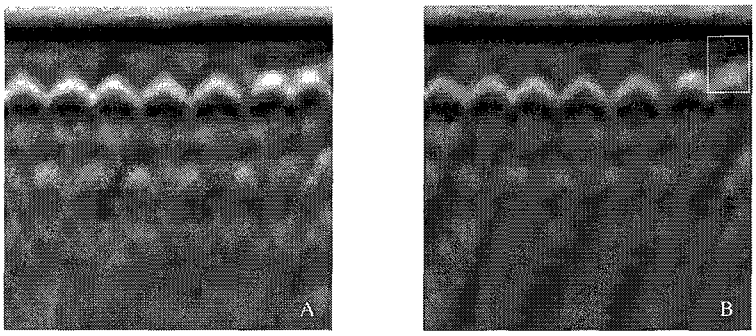

Figure 14. A and B are raw and processed scans from a 6-inch slab with embedded air-void defect. The white rectangle area marks partially the air-void defect.

Figure 17A shows raw B-scan (scan8) with shallow embedded air-void defect that is partially overlapped with the black section of the ground coupling band. Figure 17B shows the processed scan where the defective region is successfully marked.

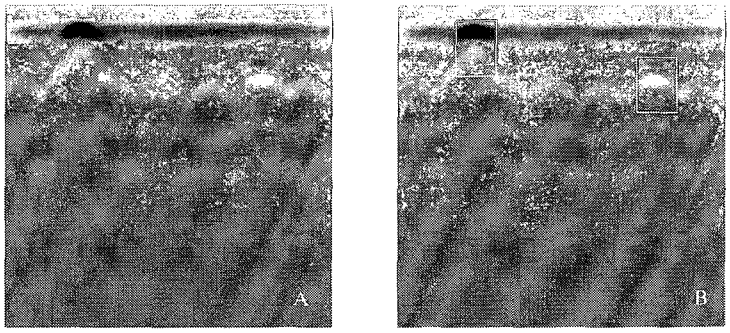

Figure 15. A and B are raw and processed scans from a 6-inch slab with two embedded delamination defects. The white rectangle areas mark the two defects.

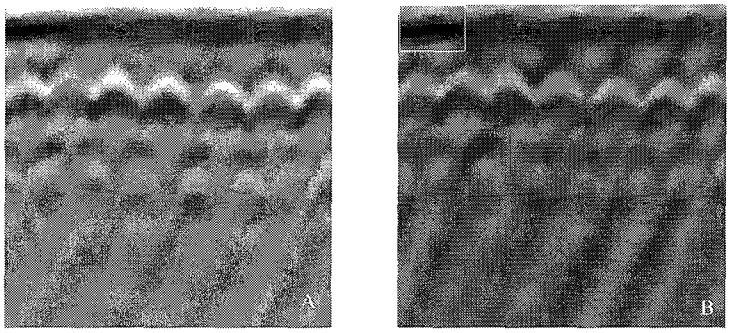

Figure 16. A and B are the raw and processed scans from a 6-inch slab with embedded air-void defect. The white rectangle area marks the air-void defect.

Figure 18A shows raw B-scan (scan9) with two embedded delamination defects. Figure 18B shows the processed scan where the first and second defective regions are successfully and partially marked, respectively. Figure 19A shows raw B-scan (scan11) with embedded delamination defect that masked reflections from the third rebar. Signature of the delamination defect in is barely visible with low contrast from its surrounding background. Therefore, the proposed method is unable to detect

this weak signature defect as shown in Figure 19B. Figure 20A shows raw B-scan (scan12) with embedded delamination and air-void defects. Figure 20B shows the processed scan where both defects are partially marked.

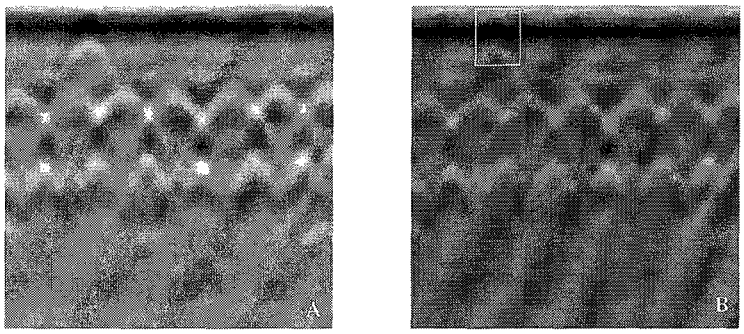

Figure 17. A and B are the raw and processed scans from a 6-inch slab with embedded air-void defect. The white rectangle area marks the air-void defect.

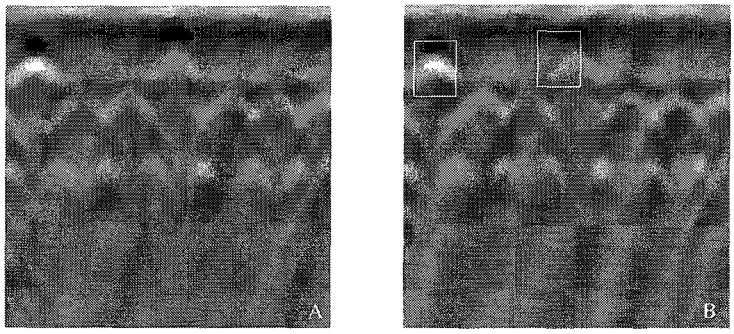

Figure 18. A and B are the raw and processed scans from a 6-inch slab with two embedded delamination defects. The white rectangle areas mark completely and partially the two defects.

Figure 21A shows raw B-scan (scan13) with embedded delamination and air-void defects. Figure 21B shows the processed scan where the first and second defective regions are partially marked. Figure 22A shows raw B-scan (scan17) with

embedded delamination defect. Figure 22B shows the processed scan where the defective region is successfully marked.

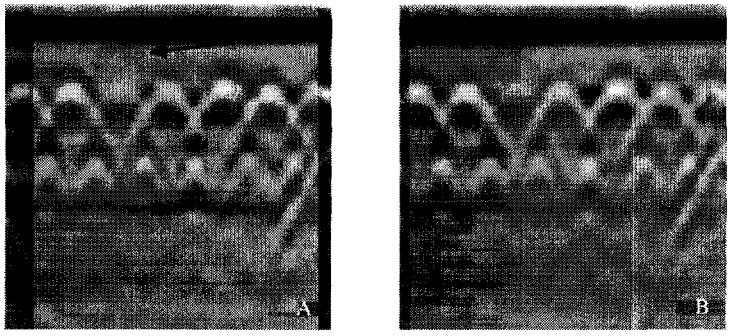

Figure 19. A and B are the raw and processed scans from a 6-inch slab with embedded delamination defect.

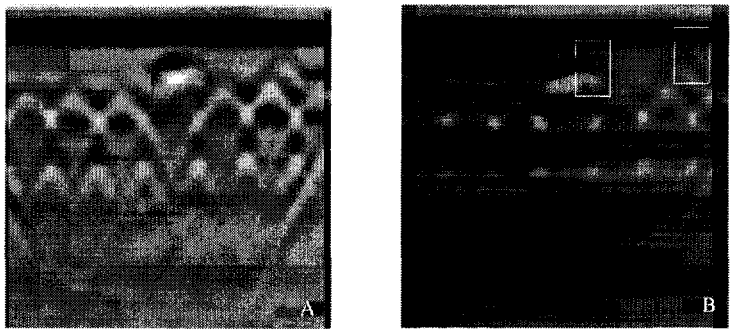

Figure 20. A and B are the raw and processed scans from a 6-inch slab with embedded delamination and air-void defects. The white rectangle areas mark partially the two defects.

Figure 23A shows raw B-scan (scan18) with two embedded delamination defects. Figure 23B shows the processed scan where the two defective regions are successfully marked. Figure 24A shows raw B-scan (scan19) with embedded air-void defect. Figure 24B shows the processed scan where the defective region is partially marked. Figure 25A shows raw B-scan (scan20) with two embedded delamination

defects. Figure 25B shows the processed scan where the two defective regions are successfully marked.

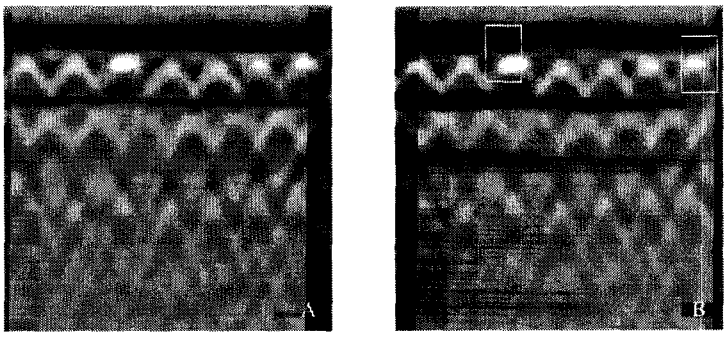

Figure 21. A and B are the raw and processed scans from a 4-inch slab with embedded delamination and air-void defects. The white rectangle areas mark partially the two defects.

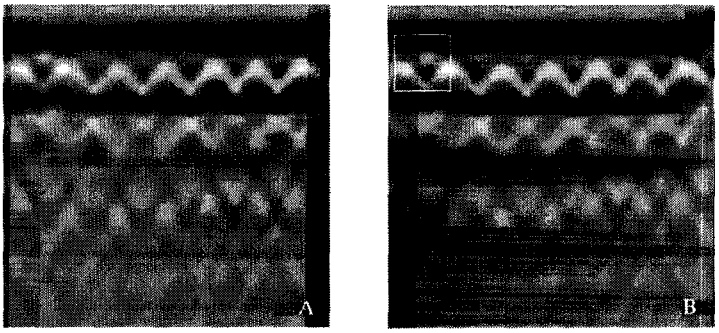

Figure 22. A and B are the raw and processed scans from a 4-inch slab with embedded delamination defect. The white rectangle area marks the defect.

Figure 26A shows raw B-scan (scan22) with embedded delamination and air-void defects. Figure 26B shows the processed scan where the defective regions are successfully and partially marked. Figure 27A shows raw B-scan (scan25) with embedded delamination defect. Figure 27B shows the processed scan where the

defective region is partially marked. Figure 28A shows raw B-scan (scan26) with two embedded delamination defects. Figure 28B shows the processed scan where the two defective regions are partially and successfully marked.

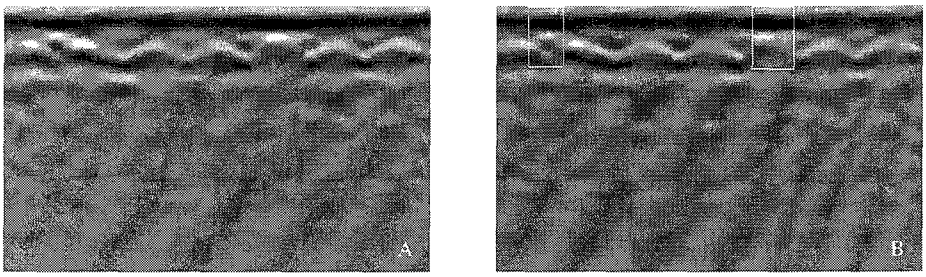

Figure 23. A and B are the raw and processed scans from a 4-inch slab with two embedded delamination defects. The white rectangle areas mark the two defects.

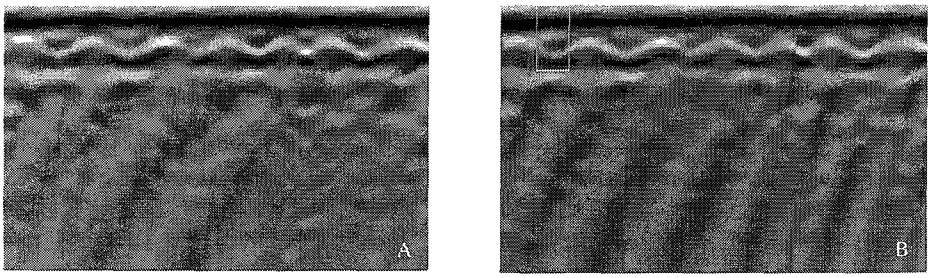

Figure 24. A and B are the raw and processed scans from a 4-inch slab with embedded air-void defect. The white rectangle area marks partially the defect.

In summary, the integrated fBm-based feature extraction and EFICA-based deconvolution framework fully identified and labeled twelve of the twenty three defective regions. It was also able to partially identify and label ten additional defective regions. The labeled defective region is extracted from the raw GPR scan as

a sub-image and used as an input to the classification stage. Table 10 shows a classification accuracy of 90.91%. As the table shows, all the delamination defects are correctly classified while two air-void defects are not correctly classified. This is because an air-void defect has two different signatures depends on being parallel or perpendicular to the survey line while a delamination defect has nearly same signature regardless of the survey line direction. When the air-void defect is parallel to the survey line, it has a flat reflection similar to that of the delamination defect. For example, a parallel air-void defect to the survey line (Figures 16) has a different signature when it is perpendicular to the survey line (Figure 17).

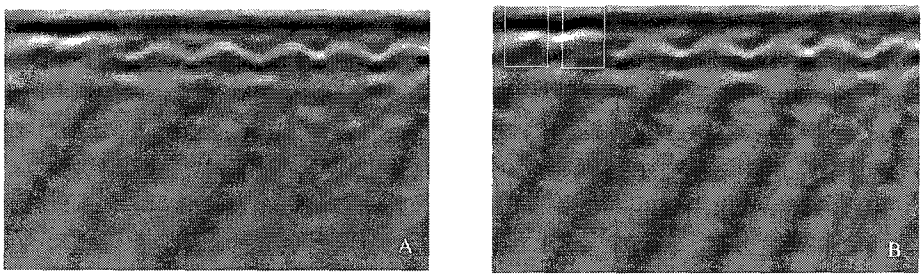

Figure 25. A and B are the raw and processed scans from a 4-inch slab with two embedded delamination defects. The white rectangle areas mark the two defects.

Parameter Analysis

Several parameters that impact performance accuracy of the proposed defect detection algorithm are addressed in the following subsections. These parameters include, towing speed of the GPR antenna, dielectric constant, number of samples per

trace, minimum defective segment length, and characteristics of the used GPR antenna such as frequency and physical size.

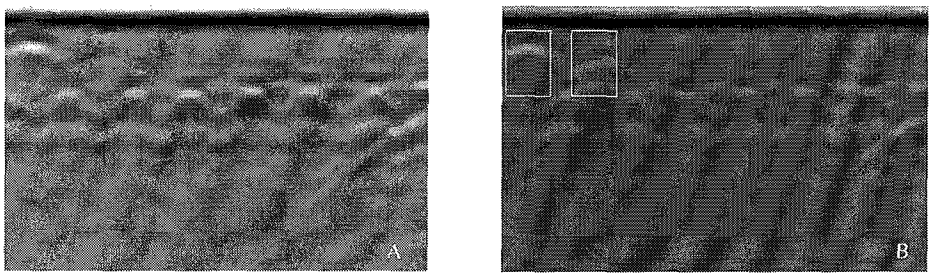

Figure 26. A and B are the raw and processed scans from an 8-inch slab with embedded delamination and air-void defects. The white rectangle areas mark partially the two defects.

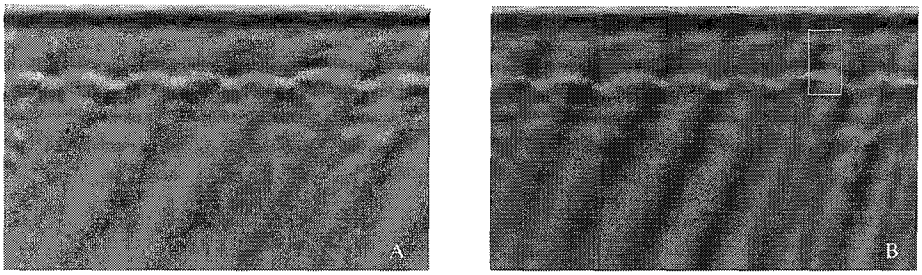

Figure 27. A and B are the raw and processed scans from an 8-inch slab with embedded delamination defect. The white rectangle area marks partially the defect.

Towing Speed of the GPR Antenna

During the data collection process, the GPR antenna is towed by hand. This non-uniform towing speed determines number of scans per inch which results in non-equal horizontal distance between rebar peaks. This affects shape of the detected objects and defects. A calibration procedure should be performed prior to conducting

the scans in order to reduce this effect and thus improves the horizontal distance estimation and classification accuracies.

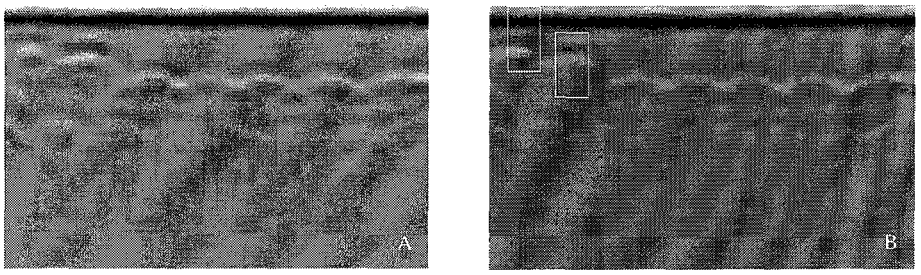

Figure 28. A and B are the raw and processed scans from an 8-inch slab with two embedded delamination defects. The white rectangle areas mark partially and completely the two defects.

If a whole bridge to be scanned, the GPR antenna should be towed behind a vehicle moving at a constant speed to achieve accurate results. Another way to enhance distance estimation and classification accuracies is by using the distance normalization function in RADAN which establishes a constant horizontal scale (equal number of scans per inch) between marks where the marks are entered during the data collection process every fixed distance (3 feet for example).

Dielectric Constant

The dielectric constant reflects velocity of radar waves through the scanned medium. Higher value of the dielectric constant indicates a slower travel time and thus shallower penetration. The depth estimation accuracy depends on the proper choice of the dielectric constant as Eq. 152 indicates. If the used value of the

dielectric constant ε' during the data collection process is different from the proper value ε in the amount of Δ, then the estimated depth will be different from the actual depth as shown in Eq. 160.

Table 10

Classification results of the detected defects

Scan number	Defect type	Classification result
Scan2	Air-void	Air-void
Scan5	Delamination	Delamination
	Delamination	Delamination
Scan6	Air-void	Air-void
Scan8	Air-void	Air-void
Scan9	Delamination	Delamination
	Delamination	Delamination
Scan11	Delamination	Not detected
Scan12	Delamination	Delamination
	Air-void	**Delamination**
Scan13	Delamination	Delamination
	Air-void	**Delamination**
Scan17	Delamination	Delamination
Scan18	Delamination	Delamination
	Delamination	Delamination
Scan19	Air-void	Air-void
Scan20	Delamination	Delamination
	Delamination	Delamination
Scan22	Delamination	Delamination
	Air-void	Air-void
Scan25	Delamination	Delamination
Scan26	Delamination	Delamination
	Delamination	Delamination
Accuracy		**90.91%**

$$d_t' = \frac{ct}{2\sqrt{\varepsilon'}} = \frac{ct}{2\sqrt{\varepsilon \pm \Delta}} \tag{160}$$

One possible way of finding the proper value of ε is by measuring velocity of radar waves through the scanned medium using the Migration function in RADAN.

From Figures 29-31, the estimated velocity of radar waves for the 4-, 6-, and 8-inch slabs is 4.7244 inches/ns. Therefore, the dielectric constant is

$$\varepsilon = \left(\frac{c}{v}\right)^2 = \left(\frac{3\times10^8\times100}{2.54\times4.7244\times10^9}\right)^2 = 6.25$$

This is the same value used during the data collection process which indicates that the estimated depth in Table 7 is accurate. Since the concrete slab has homogenous materials, the dielectric constant should not change significantly with depth (as opposed to the soil). Therefore, assuming a constant value of the dielectric constant for the concrete slabs should not affect the depth estimation accuracy [18].

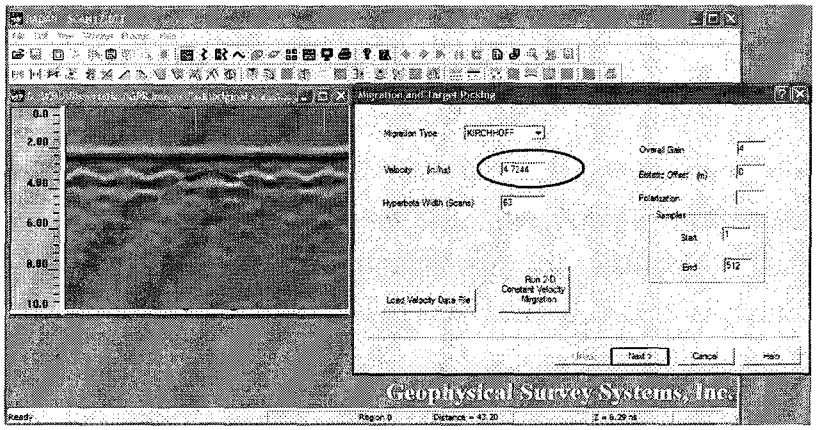

Figure 29. Measuring velocity of radar waves using the migration function in RADAN for a 4-inch slab.

Number of Samples per Trace

Each trace consists of a number of individual data points called samples. Generally speaking, the vertical resolution of the B-scan will be improved by increasing number of samples per scan. Since the concrete slabs are shallow, there is

no need for a large number of samples per scan. To investigate effect of number of samples per scan on depth estimation accuracy, scans from Table 7 with the lowest accuracies are re-sampled to 1024 samples per scan using RADAN. Then, the fBm-based feature extraction, EFICA-based deconvolution algorithm, and velocity analysis method are applied to these scans. The new scans take considerably longer execution time than the execution time needed for the B-scans with 512 and 216 samples per scan. This is because each GPR trace has 1024 samples resulting in large mixture matrix for the deconvolution algorithm to decompose which is difficult to handle using a regular personal computer with a limited memory. Also, using a large number of samples to represent shallow medium will degrade the depth estimation accuracy as shown in Table 11.

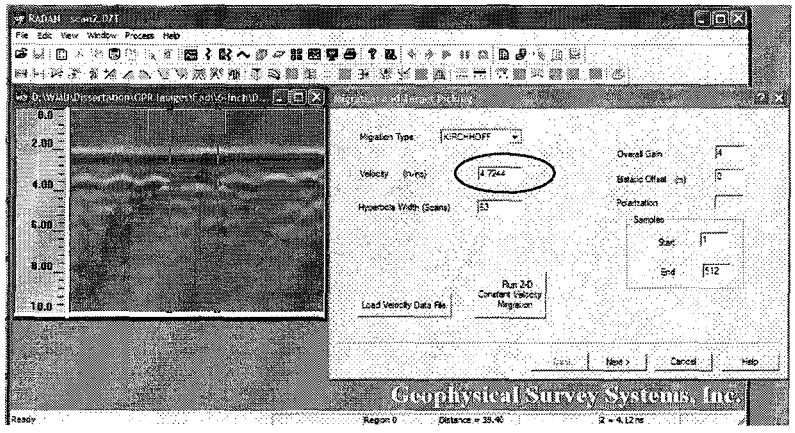

Figure 30. Measuring velocity of radar waves using the migration function in RADAN for a 6-inch slab.

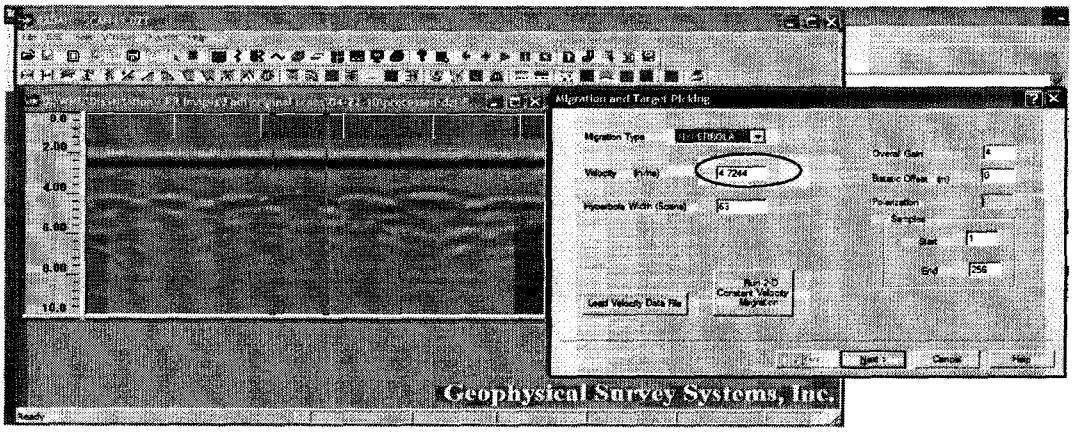

Figure 31. Measuring velocity of radar waves using the migration function in RADAN for an 8-inch slab.

Table 11

Actual and estimated depth of defects using different samples per scan

Scan	Actual Depth	S	Estimated Depth	Diff	S	Estimated Depth	Diff
Scan22	1.25	256	0.74	0.51	1024	0.32	0.93
	1.5		0.8	0.7		2.63	1.13
Scan29	2.33	256	1.58	0.75	1024	3.1	0.77
	2.83		3.68	0.85		1.51	1.32
Average				**0.7**			**1.04**

Minimum Defective Segment Length

The densest recommended number of scans per inch for concrete structures using the 1.5 GHz antenna is 10 scans per inch [16]. Therefore, the GPR antenna is set to 8.3 scans per inch (100 scans per foot) for scans 1 through 17 and 10 scans per inch (120 scans per foot) for scans 18 through 26 during the data collection process. The smallest defect embedded in the slabs is 1.5 inches wide. Ideally, T should be

8.3x1.5=12.5 for scans 1 through 17 and 10x1.5=15 for scans 18 through 26. For both cases, T should be 12 to detect the smallest embedded defect. Since the proposed FBFE algorithm may detect portion of the defective region (see Figure 20 for example) due to that fact that a defective region may have stronger reflections in some parts than other parts, the threshold value, T is set to 7 to be able to detect defective regions as small as 0.79 inches (scans 1 through 17) and 0.7 inches (scans 18 through 26). This user-defined parameter can be adjusted depending on the application needs.

Characteristics of the Used GPR Antenna

The used GPR antenna has three factors that affect the data collection process and consequently accuracy of the results: physical size of the antenna, the offset between the transmitter and receiver, and its frequency.

Bridge deck condition assessment requires using a high frequency antenna since it is more suitable for shallow surfaces. High frequency antennae have small physical size and consequently small distance between the transmitter and the receiver.

The 1.5GHz antenna has a small physical size (10 inches long) [18]. Using a small size antenna allows the detection of small size defects. Since the transmitted radar signal attenuated drastically away from center of the antenna, the antenna has to go directly over the small defect to detect it. This slows the data collection process and becomes non practical in case of scanning large bridges.

The offset (distance) between the transmitter and receiver of the 1.5 GHz antenna is 2.3 inches. Therefore, targets of less than 2.3 inches deep may appear slightly off from their actual depth in raw scans. This is because radar signals travel at an angle from transmitter to the receiver rather than a straight line down and back from center of the GPR antenna [16]. Since all the embedded defects and rebar in the 4-inch slab are within 1.65 inches from the surface, this slab has the lowest horizontal location and depth estimation accuracies.

In summary, it is recommended to use a high frequency antenna for bridge deck condition assessment [16, 18, 5]. The price comes at a little degradation in distance and depth estimation accuracies of shallow defects but with accurate results for deep defects. Also, the data collection process will be slow. However, the high frequency antenna has the capability of detecting small defects and providing high resolution scans that facilitate the automation of defect detection.

CHAPTER 6

CLOSURE

Summary and Conclusions

In this dissertation, a framework has been developed to automate the detection, localization, and characterization of subsurface defects inside bridge decks. The framework consists of four algorithms: 1) a fractal-based feature extraction algorithm to detect and horizontally label defective regions; 2) a banded-ICA deconvolution algorithm to reduce overlapping reflections from closely spaced targets and to recover travel time to and from detected defects and rebar; 3) a velocity analysis method to estimate depth of detected defects; and 4) a classification algorithm using principal component analysis to identify main features in defective regions. This framework was implemented and tested using real GPR scans of simulated concrete bridge decks of varying thicknesses and with several embedded defects of different types, dimensions, and locations.

Attempting to investigate Fractals for detecting and horizontally labeling defective regions, I presented a comparison between three different fractal methods to determine the most suitable one. Results indicate that fractal Brownian motion based feature extraction algorithm has 89.74% accuracy of detecting defects and localizing them horizontally where the average difference between the actual and estimated horizontal locations is 0.76 inches resulting in the highest accuracy, recall, and

precision among the three methods. Results also demonstrate that Brownian motion algorithm has the lowest false negative and false positive rates. EFICA-based deconvolution algorithm estimates defects depths with significant accuracy where the average difference between the actual and estimated depths is 0.37 inches. Also, the deconvolution algorithm is able to detect defects of a depth of 0.39 inches which indicate that the deconvolution algorithm is able to resolve overlapping reflections of targets spaced at 6% of the antenna wavelength which indicates robustness of the proposed deconvolution algorithm.

The integrated fBm-based feature extraction, EFICA-based deconvolution, and velocity analysis framework fully identified and labeled twelve of the twenty three defective regions. It was also able to partially identify and label ten additional defective regions. Results of the classification phase indicate that the algorithm has 90.91% classification accuracy for delamination and air-void defects. All the delamination defects are correctly classified while two air-void defects are not correctly classified. This is because an air-void defect has two different reflections (flat or arch) depending on being parallel or perpendicular to the surveying line while a delamination defect shape is independent of the surveying line direction (flat reflection). When the air-void defect is parallel to the survey line, it has a flat reflection similar to that of the delamination defect and thus it may lead to erroneous classification results.

The depth estimation accuracy depends on the proper choice of the dielectric constant, the proper choice of the number of samples per scans, the whole detection

of the defective region by the fractal-based feature extraction algorithm, the deconvolution algorithm, and the accurate measurement of rebar depth.

Contribution

The work presented in this dissertation has several contributions to the current body of knowledge in ground penetrating radar detection analysis techniques and in concrete bridge deck condition assessment. The contributions include the following:

- Developed a complete framework that detects, localizes, and classifies subsurface defects inside concrete bridge decks.
- Presented a comparison between the most common fractal methods to determine the most suitable one for bridge deck condition assessment.
- Introduced a fractal-based feature extraction algorithm that is capable of detecting and horizontally labeling defective regions with reasonable computational demands and using only the underlying GPR B-scan without the need for a training dataset as required by other algorithms.
- Developed an EFICA-based deconvolution algorithm that is able to detect embedded defects in bridge decks. This demonstrates that the deconvolution algorithm resolved the overlapping reflections of adjacent objects problem.
- Introduced an automated identification methodology of defective regions which can be integrated into a CAD system that allows for better visual assessment by the maintenance engineer of size and cost of repairs needed.
- Presented a completely automated framework that should eliminate human

interpretation errors and reduce condition assessment time and cost.

- Presented an investigation and a successful attempt to classify the common defects in bridge decks and produced highly accurate classification results.
- Presented a framework that can, with little modification, be used in other applications as well, including new materials or structures.

Future Work

An important extension of this work is testing and validating the work using data of real bridges. The EFICA-based deconvolution algorithm as presented is robust but it is computationally demanding. Therefore, future work may include work on improving its execution time by reducing its computational complexity. This can be accomplished by exploiting other methods to construct the mixture matrix and/or considering parallel processing. Future work may also include efforts to improve the proposed FBFE algorithm to be able to detect the whole defective region and thus improving the depth estimation accuracy. Furthermore, detecting boundaries of the defective regions to enable the maintenance engineer to estimate the size of repair needed, can be an excellent addition to this work. Investigating the use of a more robust classifier that can handle a variety of defect types can be another valuable extension. Finally, contour maps can also be integrated into this framework to quantify percentage of the detected defects.

BIBLIOGRAPHY

[1] B. Cross and J. Finke, "2006 structural engineering and public safety," Proceedings of the 2006 Structures Congress, pp. 18-21, May 2006.

[2] The American Society of Civil Engineers' (ASCE) 2009 Report Card for America's Infrastructure, pp. 1-11, January 2009.

[3] S. Nabulsi, O. Abudayyeh, S. Yehia, and I. Abdel-Qader, "The detection of common concrete bridge deck defects using thermography, impact echo, and ground penetrating radar," Western Michigan University, MI, USA, Technical Report CEM-05-01, January 2005.

[4] G. Washer, R. Fenwick, N. Bolleni, and J. Harper, "Effects of environmental variables on infrared imaging of subsurface features of concrete bridges," Transportation Research Record: Journal of the Transportation Research Board, vol. 2108, no. 2009, pp. 107-114, 2009.

[5] C.L. Hing, "Nondestructive Evaluation of Fiber Reinforced Polymer Bridge Decks Using Ground Penetrating Radar and Infrared Thermography," Dissertation, West Virginia University, Morgantown, West Virginia, 2006.

[6] O. Abudayyeh, S. Yehia, I. Abdel-Qader, and A. Zalt, "GPR imaging for bridge deck condition assessment," Bridge Structures: Assessment, Design and Construction. Taylor and Francis Publisher, vol. 4, no. 2, 2008.

[7] J. Geotech. and Geoenvir, "Inspection and condition assessment using ground penetrating radar," Journal of Geotechnical and Geoenvironmental Engineering, vol. 136, no. 1, pp. 207-214, 2010.

[8] W. Kim, "Ground penetrating radar application for non-destructive testing: bridge deck inspection and dowel bar detection," Dissertation, University of Missouri-Rolla, Missouri, 2003.

[9] Z. Zhao, "Application of ground penetrating radar to locate subsurface voids in urban areas," Dissertation, University of Louisville, Louisville, Kentucky, December 2005.

[10] H. Liu, "An efficient concrete bridge disease identification system based on sample database," International Symposium on Computer Science and Computational Technology, pp. 184-188, 2008.

[11] S.N. Shoukry, D. Martinelli, S.T. Varadarajan, and U.B. Halabe, "Radar signal interpretation using neural network for defect detection in concrete," Materials Evaluation, vol. 54, no. 3, pp. 393-397, 1996.

[12] H.J. Shin, "Probabilistic bridge deck condition analysis using ground penetrating radar (GPR)," Dissertation, Rensselaer Polytechnic Institute, Troy, New York, June 2004.

[13] S. Lahouar, "Development of Data Analysis Algorithms for Interpretation of Ground Penetrating Radar Data," Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, October 2003.

[14] K.M. Belli, "Ground penetrating radar bridge deck investigations using computational modeling," Dissertation, The Department of Mechanical and Industrial Engineering, Northeastern University, Boston, Massachusetts, April 2008.

[15] U. Halabe, H. Chen, V. Bhandarkar, and Z. Sami, "Detection of sub-surface anomalies in concrete bridge decks using ground penetrating radar," ACI Materials Journal, vol. 94, no. 5, pp. 396-408, 1997.

[16] GSSI Handbook for RADAR Inspection of Concrete. Geophysical Survey Systems, Inc, 2006.

[17] I. Al-Qadi and S. Lahouar, "Ground penetrating radar: state of the practice for pavement assessment," Materials evaluation, vol. 62, pp. 759-763, 2004.

[18] TerraSIRch SIR System-3000 User's Manual, Geophysical Survey Systems Inc., Salem, NH, 2004.

[19] J. Paik, C. Lee, and M. Abidi, "Image processing-based mine detection techniques: a review," Subsurface Sensing Technologies and Applications, vol. 3, no. 3, pp. 153-202, 2002.

[20] A.P. Annan, "Ground penetrating radar," Workshop Notes, Sensors & Software Inc., September 2001.

[21] D. Potin, E. Duflos, and P. Vanheeghe, "Landmines ground penetrating radar signal enhancement by digital filtering," IEEE Transactions on Geosciences and Remote Sensing, vol. 44, no. 9, 2006.

[22] S. Greenhalgh, and L. Marescot, "Modeling and migration of 2-D georadar data: a stationary phase approach," IEEE Transactions on Geosciences and Remote Sensing, vol. 44, no. 9, 2006.

[23] A. Hyvarinen, J. Karhunen, and E. Oja, "Independent component analysis," John Wiley & Sons Inc; 2001.

[24] F. Vrins, J.A. Lee, M. Verleysen, V. Vigneron, and C. Jutten, "Improving independent component analysis performances by variable selection," NNSP'2003 proceedings neural networks for signal processing, pp. 359-368, 2003.

[25] P. Tichavsky, Z. Koldovsky, and E. Oja, "Asymptotic performance of the FastICA algorithms for independent component analysis and its improvements," IEEE/SP 13th workshop on statistical signal processing, pp. 1084-1089, 2005.

[26] S. Amari, "Natural gradient works efficiently in learning," Neural Computation, vol. 10, no. 2, pp. 251-276, 1998.

[27] D. Rumelhart and J. McClelland, "Parallel distributed processing," vol. 1, 1986.

[28] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," IEEE Transactions on Neural Networks, vol. 10, no. 3, pp. 626–634, 1999.

[29] A. Cichocki and R. Unbehauen, "Robust neural networks with on-line learning for blind identification and blind separation of sources," IEEE Transactions on Circuits and Systems, vol. 43, no. 11, pp. 894-906, 1996.

[30] J. Eriksson, "Contributions to theory and algorithms of independent component analysis and signal separation," Helsinki University of Technology, Signal Processing Laboratory, Espoo 2004.

[31] S. Douglas and S. Amari, "Natural-gradient adaption," Unsupervised Adaptive Filtering, vol. 1, pp. 13-61, 2000.

[32] A. Bell, T. Sejnowski, "An information maximization approach to blind separation and blind deconvolution," Neural Computation, vol. 7, pp. 1129-1159, 1995.

[33] J.F. Cardoso, "The invariant approach to source separation," In Proceedings of the International Symposium on Nonlinear Theory and Applications NOLTA, vol. 1, pp. 55–60, 1995.

[34] E.G. Learned-Miller and J. Fisher, "ICA using spacings estimates of entropy," Journal of Machine Learning, vol. 4, pp. 1271–1295, 2003.

[35] E. Hyvarinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," International Journal of Neural Systems, vol. 10, pp. 1-8, 2000.

[36] S. Amari, "Blind signal separation: mathematical foundations of ICA, Sparse component analysis and other techniques," Proceedings of SPIE, pp. 1-10, 2005.

[37] K. Zhang, "An adaptive method for subband decomposition ICA," Neural Computation, vol. 18, pp. 191-223, 2006.

[38] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," Neural Networks, vol. 13, pp. 411-430, 2000.

[39] Z. Koldovsky, P. Tichavsky, and E. Oja, "Efficient variant of algorithm FastICA for independent component analysis attaining the Cramer-Rao lower bound," IEEE Transactions on Neural Networks, vol. 17, no. 5, pp. 1265-1277, 2006.

[40] M. Plumbley, "Conditions for nonnegative independent component analysis," IEEE Signal Processing Letters, vol. 9, no. 6, pp. 177-180, 2002.

[41] Z. Yuan, "Advances in independent component analysis and nonnegative matrix factorization," Helsinki University of Technology. Dissertations in information and computer science, Espoo 2009.

[42] A. Chicocki and S. Amari, "Adaptive blind signal and image processing, John Wiley & Sons, 2002.

[43] A. Rudra, "Prasanta chandra mahalanobis: a biography," Oxford University Press, 1996.

[44] P. C. Mahalanobis, "Decorrelation in statistics: the mahalanobis transformation. Added material to data compression: the complete reference,"pp. 1-6, 2000.

[45] Q. Gao, T. Li, and R. Wu, "A novel KICA method for ground bounce removal with GPR," International Conference on Radar, pp. 1-4, 2006.

[46] Q. Du, I. Kopriva, "Dependent component analysis for blind restoration of images degraded by turbulent atmosphere," Neurocomputing, vol. 72, pp. 2682-2692, 2009.

[47] K. Falconer, "Fractal geometry mathematical foundations and applications," John Wiley & Sons Ltd, second edition, 2003.

[48] Y. Zhao, J. Wu, and M. Wan, "Fractal technique for abstracting high-resolution information from GPR section," Tenth international conference on ground penetrating radar, pp. 21-24, June 2004.

[49] K. Oleschko, G. Korvin, A. Balankin, R.V. Khachaturov, L. Flores, B. Figueroa, J. Urrutia, and F. Brambila, "Fractal scattering of microwaves from soils," Physical Review Letters, vol. 89, no. 18, 2002.

[50] H. Schepers, J. Beek, and J. Bassingthwaighte, "Four methods to estimate the fractal dimension from self-affine signals," IEEE Engineering in Medicine and Biology, vol. 92, pp. 57-64, 1992.

[51] J.Z. Zhang, "Fractals," Beijing: QingHua University Publisher, pp. 56-111, 1995.

[52] B. Klinkenberg, "A review of methods used to determine the fractal dimension of linear features," Mathematical Geology, vol. 26, pp. 23-46, 1994.

[53] D. Russel, J. Hanson, and E. Ott, "Dimension of strange attractors," Physical Review Letters, vol. 45, no. 14, pp. 1175-1178, 1980.

[54] J. Ehlers, "Fractal adaptive moving average," www.mesasoftware.com, 2005.

[55] B. Chaudhuri and N. Sarkar, "Texture segmentation using fractal dimension," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 1, pp. 72-77, 1995.

[56] S.K. Nath and P. Dewangan, "Detection of seismic reflections from seismic attributes through fractal analysis," Geophysical Prospecting, vol. 50, pp. 341-360, 2002.

[57] J.X. Xu, "Modern theory of nonlinear dynamics-bifurcation, chaos, fractal," Xi'an Jiaotong University Press, Xi'an, China, 1996.

[58] H.S. Greenside, A. Wolf, J.B. Swift, and T. Pignataro, "Impracticality of a box counting algorithm for calculating the dimensionality of strange attractors," Physical Review, vol. 25, pp. 2342-2345, 1982.

[59] H. Liu and W. Yang, "Radar target recognition using a modified FastICA algorithm plus GAs," IEEE Region 10 Conference on TENCON, pp. 1-4, 2006.

[60] J. De-La-Rosa, C. Puntonet, R. Piotrkowski, A. Moreno, and J. Gorriz, "Comparison of two ICA algorithms in BSS applied to non-destructive vibratory tests," Simulated Evolution and Learning: 6th International Conference, vol. 4247, pp. 750-755, 2006.

[61] H. Yang, S. Amari, and A. Cichocki, "Information theoretic approach to blind separation of sources in non-linear mixture," Signal Processing, vol. 64, no. 3, pp. 291-300, 1998.

[62] Y. Singh and C. Rai, "A simplified approach to independent component analysis," Neural Computation and Application, vol. 12, pp. 173-177, 2003.

[63] Y. Baram and Z. Roth, "Multidimensional density shaping by sigmoids," IEEE Transactions on Neural Networks, vol. 7, no. 5, pp. 1291-1298, 1996.

[64] H. Park, J. Lee, S. Oh, and S. Lee, "Blind deconvolution with sparse priors on the deconvolution filters," Independent Component Analysis and Blind Signal Separation: 6th international conference, ICA, vol. 3889, pp. 658-665, 2006.

[65] F. Abu-Amara and I. Abdel-Qader, "Hybrid mammogram classification using rough set and fuzzy classifier," International Journal of Biomedical Imaging, Hindawi Publishing Company, vol. 2009, article ID 680508, 2009.

[66] S. Amari, A. Cichocki, and H. Yang, "A new learning algorithm for blind signal separation," Advances in Neural Information Processing Systems, vol. 8, pp. 757-763, 1996.

[67] Y. Xue, Y. Wang, and J. Yang, "Independent component analysis based on gradient equation and kernel density estimation," Neurocomputing, vol. 72, pp. 1597-1604, 2009.

[68] Z. Shi, Z. Jiang, and F. Zhou, "A fixed-point algorithm for blind source separation with nonlinear autocorrelation," Journal of Computational and Applied Mathematics, vol. 223, pp. 908-915, 2009.

[69] A. Beloucharni, M. Amin, and K. Abed-Meriam, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices," IEEE Signal Processing Letters, vol. 4, September 1997.

[70] A. Cichocki and S. Amari, "Adaptive blind signal and image processing: learning algorithms and applications," John Wiley & Sons, West Sussex, UK, 2003.

[71] J. Cao, N. Murata, S. Amari, A. Cichocki, and T. Takeda, "Independent component analysis for unaveraged single-trial MEG data decomposition and signle-dipole source localization," Neurocomputing, vol. 49, pp. 255-277, 2002.

[72] J. Karvanen, J. Eriksson, and V. Koivunen, "Pearson system based method for blind separation," Proceedings of the Second International Workshop on Independent Component Analysis and Blind Signal Separation, pp. 585–590, 2000.

[73] S. Choi, A. Cichocki, and S. Amari, "Flexible independent component analysis," Journal of VLSI Signal Process, vol. 26, pp. 25–38, 2000.

[74] J. Eriksson, J. Karvanen, and V. Koivunen, "Source distribution adaptive maximum likelihood estimation of ICA model," Proceedings of the Second International Workshop on Independent Component Analysis and Blind Signal Separation, pp. 227–232, 2000.

[75] L. Xu, C. Cheung, and S. Amari, "Learned parametric mixture based ICA algorithm," Neurocomputing, vol. 22, pp. 69–80, 1998.

[76] R. Boscolo, H. Pan, and V. Roychowdhury, "Independent component analysis based on nonparametric density estimation," IEEE Transactions on Neural Networks, vol. 15, no. 1, pp. 55–65, 2004.

[77] F. Bach and M. Jordan, "Kernel independent component analysis," Journal of Machine Learning, vol. 3, no. 1, pp. 1–48, 2002.

[78] Y. Matsuda and K. Yamaguchi, "A fixed-point algorithm of topographic ICA," 16th International Conference on Artificial Neural Networks, ICANN 2006, vol. 4132, pp. 587-594, 2006.

[79] J.F. Cardoso, "Multidimensional independent component analysis," IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, pp. 1941-1944, 1998.

[80] A. Sharma and K. Paliwal, "Subspace independent component analysis using vector kurtosis," Pattern Recognition, vol. 39, no. 11, pp. 2227-2232, 2006.

[81] F. Bach and M. Jordan, "Beyond independent components: trees and clusters," Journal of Machine Learning Research, vol. 4, pp. 1205-1233, 2003.

[82] C. Yi, J. Wei, W. Ling, Z. Lei, and Y. Hong, "A Gabor subband decomposition ICA and MRF hybrid algorithm for infrared image reconstruction from subpixel shifted sequences," Optics Communications, vol. 282, no. 5, pp. 786-797, 2009.

[83] D. Lee and H. Seung, "Learning of the parts of objects by non-negative matrix factorization," Nature, vol. 401, pp. 788-791, 1999.

[84] C. Chennubhotla and A. Jepson, "Sparse PCA extracting multi-scale structure from data," Proceedings of the IEEE International Conference on Computer Vision, vol. 1, pp. 641-647, 2001.

[85] F. Miwakeichi, E. Martnez-Montes, P. Valds-Sosa, N. Nishiyama, H. Mizuhara, and Y. Yamaguchi, "Decomposing EEG data into space-time-frequency components using parallel factor analysis," NeuroImage, vol. 22, no. 3, pp. 1035-1045, 2004.

[86] R. Szupiluk, P. Wojewnik, and T. Zabkowski, " Smooth component analysis as ensemble method for prediction improvement," 7th International Conference on independent Component Analysis and Signal Separation, vol. 4666, pp. 277-284, 2007.

[87] I. Kopriva and D. Seršic, "Robust blind separation of statistically dependent sources using dual tree wavelets," IEEE International Conference on Image Processing, vol. 1, pp. 1433-1436, 2007.

[88] I. Dagher and R. Nachar, "Face recognition using IPCA-ICA algorithm," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 6, pp. 996-1000, 2006.

[89] M. Zibulevsky, "Blind source separation with relative newton method," Proceedings of the ICA, Nara, Japan, pp. 897–902, 2003.

[90] S. Choi and H. Choi, "A relative trust-region algorithm for independent component analysis," Neurocomputing, vol. 70, pp. 1502–1510, 2007.

[91] G. Olhoeft, "Maximizing the information return from ground penetrating radar," Journal of Applied Geophysics, vol. 43, pp. 175–187, 2000.

[92] V. Lijn, F. Roth, and M. Verhaegen, "Estimating the impulse response of buried objects from ground penetrating radar signals," Detection and Remediation Technologies for Mines and Minelike Targets VIII, Proceedings of SPIE, vol. 5089, 2003.

[93] V. Krause and I. Abdel-Qader, "Deconvolution and segmentation of ground penetrating radar images," Western Michigan University, MI, USA, Technical Report 2007-103, 2007.

[94] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, "Numerical recipes in C: the art of scientific computing," second edition, Cambridge University Press, Cambridge, England, 1992.

[95] S. Lahouar, "Development of data analysis algorithms for interpretation of ground penetrating radar data," Doctor of Philosophy thesis, Virginia Polytechnic Institute and State University, 2003.

[96] G. Debalina and S. Tapan, "Extraction of the signature of a buried object using GPR," CIE International Conference of Radar Proceedings, IEEE Radar Conference, pp. 296-301, 2006.

[97] D. Rothwell and W. Sun, "Time domain deconvolution of transient radar data," IEEE Transactions on Antennas and Propagation, vol. 38, pp. 470-475, 1990.

[98] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning: data mining, inference and prediction," New York, Springer-Verlag, 2001.

[99] B. Karlsen, H. Sørensen, J. Larsen, and K. Jakobsen, "GPR detection of buried symmetrically shaped mine-like objects using selective independent component analysis," Proceedings of the 2003 Detection and Remediation Technologies for Mines and Mine-Like Targets, AeroSense, vol. 5089, pp. 375-386, 2003.

[100] B. Karlsen, J. Larsen, H. Sorensen, and K. Jakobsen, "Comparison of PCA and ICA based clutter reduction in GPR systems for anti-personal landmine detection," Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing, pp. 146-149, 2001.

[101] C. Liu and H. Wechsler, "Independent component analysis of gabor features for face recognition," IEEE Transactions on Neural Networks, vol. 14, no. 4, pp. 919-928, 2003.

[102] H. Yin and I. Hussain, "Independent component analysis and non-gaussianity for blind image deconvolution and deblurring," Integrated Computer-Aided Engineering, vol. 15, pp. 219-228, 2008.

[103] A. Al-Qaisi, W. Woo, and S. Dlay, "Novel statistical approach to blind recovery of earth signals and source wavelet using independent component analysis," WSEAS Transactions on Signal Processing, vol. 4, no. 4, pp. 231-240, 2008.

[104] J.V. Stone, J. Porrill, N.R. Porter, and I. Wilkinson, "Spatiotemporal independent component analysis of event-related fmri data using skewed probability density functions," NeuroImage, vol. 15, no. 2, pp. 407-421, 2002.

[105] F. Šroubek, J. Flusser, and G. Cristóbal, "Super-resolution and blind deconvolution for rational factors with an application to color images," Computer Journal, vol. 52, no. 1, pp. 142-152, 2009.

[106] M. Elsabrouty, "Fast converging blind signal separation algorithm using the bussgang cost function and the natural gradient," EEE International Conference on Signal Processing and Communications, pp. 229-232, 2007.

[107] M. Bronstein, A. Bronstein, M. Zibulevsky, and Y. Zeevi, "Blind deconvolution of images using optimal sparse representation," IEEE Transactions on Image Processing, vol. 14, no. 6, pp. 726-736, 2005.

[108] C.L. Bastard, V. Baltazart, Y. Wang, and J. Saillard, "Thin pavement thickness estimation using gpr with high and super resolution methods," IEEE Transactions on Geosciences and Remote Sensing, vol. 45, no. 8, pp. 2511-2519, 2007.

[109] K. Chahine, V. Baltazart, Y. Wang, and X. Dérobert, "Blind deconvolution via independent component analysis for thin pavement thickness estimation using GPR, Non-Destructive Testing in Civil Engineering, Nantes, France, June 30th – July 3rd, 2009.

[110] P. Xie and S. Grant, "A fast and efficient frequency-domain method for convolutive blind source separation," IEEE Region 4 Conference, pp. 1-4, 2008.

[111] S.T. Kaplan and T.J. Ulrych, "Blind deconvolution and ICA with a banded mixing matrix," Online Proceeding of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation, Nara, Japan, pp. 591-596, April 2003.

[112] B. Karlsen, H. Sorensen, J. Larsen, and K. Jakobsen, " Independent component analysis for clutter reduction in ground penetrating radar data," Proceeding of SPIE, vol. 4742, 2002.

[113] J. Anemuller, J. Duann, T. Seijnowski, and S. Makeig, "Spatio-temporal dynamics in fMRI recordings revealed with complex independent component analysis," Neurocomputing, vol. 69, pp. 1502-1512, 2006.

[114] N. Mitianoudis and M. Davies, "Audio source separation of convolutive mixtures," IEEE Transactions on Speech and Audio Processing, vol. 11, no. 5, pp. 489-497, 2003.

[115] S. Douglas, A. Chichocki, and S. Amari, "Self-whitening algorithms for adaptive equalization and deconvolution," IEEE Transactions on Signal processing, vol. 47, pp. 1161–1165, 1999.

[116] L. Parra and C. Spence, "Convolutive blind source separation of non-stationary sources," IEEE Transactions on Speech and Audio Processing, vol. 8, pp. 320-327, 2000.

[117] H. Attias and C. Schreiner, "Blind source separation and deconvolution: the dynamic component analysis algorithm," Neural Computation, vol. 10, no. 6, pp.1373-1424, 1998.

[118] J. Anemuller, T. Sejnowski, and S. Makeig, "Complex independent component analysis of frequency-domain electroencephalographic data," Neural networks, vol. 16, pp. 1311-1323, 2003.

[119] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," In Proceedings of International Conference on Independent Component Analysis and Blind Source Separation, pp. 505-510, 2003.

[120] I. Lee, T. Kim, and T. Lee, "Fast fixed-point independent vector analysis algorithms for convolutive blind source separation," Signal Processing, vol. 87, pp. 1859-1871, 2007.

[121] L. Molgedey and H. Schuster, "Separation of independent signals using time-delayed correlations," Physical Review Letters, vol. 72, no. 23, pp. 3634-3637, 1994.

[122] A. Lotsch, M. Friedl, and J. Pinzón, "Spatio–temporal deconvolution of NDVI image sequences using independent component analysis," IEEE Transactions on Geosciences and Remote Sensing, vol. 41, no. 12, pp. 2938-2942, 2003.

[123] F. Abujarad and A. Omar, "Comparison of independent-component-analysis (ICA) algorithms for GPR detection of non-metallic land mines," Image and signal processing for remote sensing, XII. Proc. Of SPIE, vol. 6365, pp. 636516-1, 2006.

[124] J. Liu, B. Zhang, R. Wu, "GPR ground bounce removal methods based on blind source separation," Progress in Electromagnetic Research Symposium, pp. 256-259, 2006.

[125] D. Devaru, U. Halabe, B. Gopalakrshnan, S. Agrawal, and S. Grushecky, "Algorithm for detecting defects in wooden logs using ground penetrating radar," Proceeding of SPIE, vol. 5999, pp. 59990B-1, 2005.

[126] D. Devaru, U. Halabe, B. Gopalakrshnan, and S. Agrawal, "Ground penetrating radar (GPR) based system for nondestructive detection of interior defects in

wooden logs," International Journal of Manufacturing Research, vol. 3, no. 4, pp.425-451, 2008.

[127] S.G. Yang, "The random fractal interpolation of seismic section," Computing Techniques for Geophysics and Geochemical Exploration, vol. 20, no. 3, pp. 218-222, 1998.

[128] T. Dogaru, L. Carin, "Time-domain sensing of targets buried under a gaussian, exponential, or fractal rough interface," IEEE Transactions on Geoscience and Remote Sensing, vol. 39, no. 8, pp. 1807-1819, 2001.

[129] S. Don and K. Revathy, "Classifying mammogram images using fractal features," International Conference on Computational Intelligence and Multimedia Applications, pp. 539-543, 2007.

[130] I. Abdel-Qader, F. Abu-Amara, O. Abudayyeh, and V. Krause "Comparative Study of Deconvolution Algorithms for GPR Bridge Deck Imaging," Technical Report ITIA-2010-01, Information Technology and Image Analysis (ITIA) Center, College of Engineering and Applied Sciences, Western Michigan University, January 2010.