Blind Source Separation for Feature Detection and Segmentation in Ground Penetrating Radar (GPR) Imaging of Concrete Bridge Decks for Nondestructive Condition Assessment

Vincent Krause
Western Michigan University, krauser72@gmail.com

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Civil and Environmental Engineering Commons, and the Electrical and Computer Engineering Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/523

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.
BLIND SOURCE SEPARATION FOR FEATURE DETECTION AND SEGMENTATION IN GROUND PENETRATING RADAR (GPR) IMAGING OF CONCRETE BRIDGE DECKS FOR NON-DESTRUCTIVE CONDITION ASSESSMENT

by

Vincent Krause

A dissertation submitted to the Graduate College in partial fulfillment of the requirements for the degree of Doctor of Philosophy
Electrical and Computer Engineering
Western Michigan University
May 2015

Doctoral Committee:
Ikhlas Abdel-Qader, Ph.D., Co-Chair
Osama Abudayyeh, Ph.D., Co-Chair
William Sauck, Ph.D.
Daniel Litynski, Ph.D.
Concrete bridge decks require periodic condition assessment and preventive maintenance to extend their useful lifespan. Nondestructive evaluation methods such as Ground Penetrating Radar (GPR) are slowly beginning to replace or complement the manual (visual) assessment of bridge conditions for detecting defects at their early stages. However, GPR scans of bridge decks are frequently cluttered with high amplitude reflections from known parts of the bridge deck, which make the detection of defects’ low amplitude reflections difficult. One such known part is the embedded steel reinforcement bars known as rebar.

This dissertation presents a novel approach to the automated detection of defects in concrete bridge decks by removing known reflections such as rebar from GPR scans of reinforced concrete bridge decks. The algorithm detects reflections from rebar with a frequency-domain pulse detection method, groups detected pulses into clusters, interpolates synthetic rebar reflections based on each cluster, and subtracts the synthetic rebar reflection from the original GPR scan data. This algorithm will facilitate the automated, non-destructive condition assessment of bridge decks.
Of making many books there is no end, and much study wearies the body.

Ecclesiastes 12:12b
ACKNOWLEDGMENTS

I would like to thank my dissertation advisors, Dr. Ikhlas Abdel-Qader and Dr. Osama Abudayyeh, for their support and guidance throughout my research. They helped me develop a broad and eclectic body of engineering experience.

I would like to thank committee member Dr. Daniel Litynski and the Electrical and Computer Engineering Department at Western Michigan University. The ECE Department has been a nurturing and challenging environment for me, and it will be tough to leave it behind.

I would also like to thank committee member Dr. William Sauck for sharing his Geology experience with a total neophyte like me.

I am grateful for the support of my family and friends. I am especially grateful for my parents’ support. My mother Ellen was always ready to listen as I thought out loud, and at times complained out loud. My father Lester shared his mechanical experience with me, and constructed many contraptions I used in my research.

Vincent Krause
TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ ii

LIST OF FIGURES ........................................................................................................ v

CHAPTER

1. INTRODUCTION .......................................................................................................... 1

1.1. Goals ....................................................................................................................... 2

1.2. Dissertation Organization ...................................................................................... 2

2. BACKGROUND .......................................................................................................... 3

2.1. Concrete Bridge Deck Defects ............................................................................... 3

2.2. Ground Penetrating Radar ................................................................................... 5

2.3. Evaluation of Concrete using GPR ......................................................................... 11

2.4. Mathematical Model of a GPR Scan of Reinforced Concrete ............................. 16

2.5. Methods for Processing GPR Scans of Reinforced Concrete ............................ 18

2.5.1. Migration .......................................................................................................... 18

2.5.2. f-K Filtering .................................................................................................... 21

2.5.3. Deconvolution ................................................................................................. 22

2.5.4. Object Detection Algorithms ......................................................................... 25

2.5.5. Blind Source Separation ................................................................................. 26

3. PROPOSED ALGORITHM ......................................................................................... 28

3.1. Outline .................................................................................................................... 28
Table of Contents - continued

3.2. Preprocessing of GPR Data .......................................................... 30
  3.2.1. Linear Time Invariant Gain Boosting with Exponential Functions .................................................. 30
  3.2.2. Horizontal and Vertical Filtering ......................................... 34
  3.2.3. Rebar Peak Detection and Exemplar Pulse Extraction ........... 35
3.3. Pulse Detection Algorithm .......................................................... 38
  3.3.1. General Model of the Pulse Detection Algorithm ................. 38
  3.3.2. Applying Pulse Detection Algorithm to Detection of Rebar Reflections ........................................... 48
3.4. Hyperbola Detection Through Iterative Application of Search Mask ................................................................. 50
3.5. Generation of Synthetic Rebar Reflections ................................. 54
3.6. Removal of Rebar Reflections Based on Synthetic Model ............ 57
3.7. Postprocessing of GPR Data ....................................................... 57
4. IMPLEMENTATION AND EXPERIMENTAL WORK ....................... 59
  4.1. Test Data Sources .................................................................. 59
    4.1.1. Computer Simulated GPR Scan Data ................................. 59
    4.1.2. Sandbox Test Bed ............................................................. 64
    4.1.3. Concrete Test Slabs .......................................................... 67
5. CONCLUSION .............................................................................. 70
BIBLIOGRAPHY ............................................................................ 73
PUBLISHED WORK ...................................................................... 89
LIST OF FIGURES

1. Model of slab with embedded object and base, showing radar reflection paths ................................................................. 9

2. Model of slab with object and base, showing multipath reflections and refractions .......................................................... 10

3. Concrete deck with rebar perpendicular to travel direction ................. 13

4. Concrete with rebar grid perpendicular and parallel to travel direction .................................................................................. 13

5. Two dimensional slice of model concrete deck .................................. 13

6. Simulated GPR scan of concrete deck shown in Figure 7 ..................... 14

7. Real GPR scan of real 6-inch thick test slab with no embedded defects ..................................................................................... 15

8. Flowchart of proposed algorithm ........................................................... 29

9. GPR scan of six inch concrete slab. Raw data with no gain boosting applied ............................................................................ 31

10. GPR scan of six inch concrete slab with default gain boosting applied by SIRtech base station .................................................. 32

11. GPR scan data of 6-inch reinforced concrete slab ................................ 35

12. GPR scan data of 6-inch reinforced concrete slab with median filtering applied to the top 70 rows .............................................. 35

13. Filtered GPR scan data of 6-inch reinforced concrete slab, with rebar peaks indicated by arrows ......................................................... 37

14. GPR data from column 130 of scan of 6-inch slab ..................................... 37
15. GPR data from column 130 of scan of 6-inch slab. The chosen exemplar pulse is shown in solid line, where the rest of the data is dashed ....... 37

16. Histogram of D(ω) with α=5, τ=0, no interpolation or decimation of pulse p(t), and low noise ................................................................. 40

17. Histogram of D(ω) with α=5, τ=0, no interpolation or decimation of pulse p(t), and medium noise ................................................................. 40

18. Contrasting an exemplar pulse with a copy resampled at a 20:19 ratio ...................................................................................................... 41

19. Complex Plane Histogram results of deconvolving the 20:19 resampled pulse with the original exemplar ................................................. 41

20. Complex Plane scatterplot of deconvolution of time shifted pulse by exemplar pulse .............................................................................. 43

21. Histogram of deconvolution of time shifted pulse by exemplar pulse .................................................................................................. 43

22. Kurtosis of angles of D(ω) for each time shift, when the sample has low noise .................................................................................. 45

23. Kurtosis of angles of D(ω) for each time shift, when the sample has high noise ................................................................................ 45

24. Sample data containing two exemplar pulses ......................................................... 47

25. Scatterplot of deconvolution of the two pulse sample by the exemplar pulse ................................................................................ 47

26. Flowchart of Pulse Detection Algorithm applied to a column of GPR scan data ............................................................................... 49

27. Search Mask for right side of hyperbola ........................................................................ 51

28. Search Mask applied iteratively to scattered hyperbola points .............................................................................. 53

29. Simulated defect with maximum amplitude of 16.9 ........................................ 61
List of Figures - continues

30. Simulated defect as shown in Figure 29, with the addition of simulated rebar with a maximum amplitude of 52.8 .......................................................... 61
31. Output of pulse detection algorithm ................................................................. 62
32. Synthetic rebar impulse response arcs ............................................................ 62
33. Original GPR scan data as shown in Figure 32 minus synthesized rebar reflections .......................................................... 64
34. Concrete simulating sandbox under construction ............................................ 65
35. Rebar grid for concrete simulating sandbox .................................................... 66
36. Concrete simulating sandbox with GPR antenna and base station .................. 66
37. GPR scan of concrete simulation sandbox with one rebar and no embedded defects ........................................................................................................ 67
38. GPR scan of 6-inch slab without embedded defects ........................................ 68
39. Construction of 4-inch slab with embedded defects ....................................... 69
40. GPR scan of 4-inch slab with embedded defects ............................................ 69
Concrete bridge decks require periodic condition assessment and preventive maintenance to extend their useful lifespan. Nondestructive evaluation methods such as Ground Penetrating Radar (GPR) are slowly beginning to replace or complement the manual (visual) assessment of bridge conditions for detecting defects at their early stages.

GPR scans of bridge decks are frequently cluttered with reflections from known parts of the bridge deck, which make the detection of defects’ low amplitude reflections difficult. One such known part is the embedded steel reinforcement bars known as rebar. The boundary between steel and concrete produces high amplitude reflections in GPR scans.

GPR scans also produce multiple large computer files. These files are typically processed after the survey crew leaves the survey site. Processing GPR scan data requires human experts to configure computer filters and visually search for defects. This process takes weeks to accomplish.
1.1 Goals

The primary goal of this work is to attenuate the reflections from the top layer of rebar in Ground Penetrating Radar surveys of reinforced concrete without distorting or attenuating reflections from defects. Because reinforced concrete is a man-made structure, it possesses regularity and symmetry that is not typically found in Geological or large scale Civil structures. We will exploit this regularity to detect and remove known structures. Attenuating rebar reflections should make it easier for human inspectors and detection methods to detect and identify defects.

The secondary goal is to automate this process as much as possible. The process should be blind source, where the only a priori information needed is that the data is from a GPR scan of a reinforced concrete bridge deck. Ultimately, the algorithm should require no human interaction.

1.2 Dissertation Organization

Chapter One presents the motivation and goals of this research. Chapter Two presents an overview of ground penetrating radar, the features we are trying to detect in scans of reinforced concrete bridge decks, and tools currently used to process GPR scan data. Chapter Three presents the proposed algorithm. Chapter Four presents results of applying the proposed algorithm to GPR scan data. Chapter Five presents conclusions and future directions for research.
CHAPTER 2

BACKGROUND

2.1 Concrete Bridge Decks Defects

Reinforced concrete bridge decks are a composite structure made of concrete and a reinforcing steel mesh called rebar. Each component of reinforced concrete compensates for the other component’s weakness. The rebar protects concrete from damage due to tension. The concrete insulates rebar from water and air, preventing corrosion. Steel and concrete have nearly the same thermal expansion coefficients, so they expand and contract at the same rate with respect to heat.

Bridge decks deteriorate with time and use. Estimates from the 1990s state that up to 40% of bridges are structurally deficient. [Branco, 2004; Huston, 2002] The early detection and repair of defects is essential to prevent catastrophic failure and to lower maintenance costs over the lifetime of the bridge through preventive maintenance.

There are four major categories of concrete defects. The first concrete defect is spalling. Spalling occurs when water or air penetrate between the concrete and rebar. The penetration forms a layer of rust around the rebar. [Arndt, 2009; He, 2009] This rust expands up to five times more with respect to heat than steel or concrete. This expansion creates pressure within the concrete leading to fine and large scale cracks. These cracks allow more penetration of air and water to the rebar, accelerating the decay.

The second concrete defect is honeycombing. [Branco, 2004] Bad mixing and pouring procedures can lead to the mortar not blending well with the aggregate. This can
cause regions of aggregate surrounded by air. These regions are structurally weak and can lead to further penetration of air and water.

The third concrete defect is delamination. Delamination is the separation of two layers that should be bonded together. This can be caused by poor construction workmanship and by the gradual penetration of air and water between layers.

The fourth concrete defect is voids. Voids are irregularly shaped pockets of air or water within the concrete slab. Voids are often initiated by the previous three defects. Voids are structurally weak and can lead to further penetration.

These defects are part of the normal degradation of concrete over time. The degradation is accelerated in regions with harsh winters. Harsh winters mean seasonal extremes of hot and cold, large amounts of precipitation, and the use of road salts to melt ice in winter. All of these contribute to greater damage due to rust expansion and expansion of ice.

These defects lead to changes in the physical properties of concrete. [He,2009] Some defects gradually alter the chemical, seismic and electromagnetic properties of the concrete. Others introduce new "objects" into the concrete, such as cracks, voids and delaminations.

Evaluation of reinforced concrete includes destructive and non-destructive methods. Non-destructive methods are preferred because they leave the structure intact and they are less time and labor intensive. Non-destructive evaluation methods include visual inspection of the surface, sounding methods including chain dragging and impact echo, thermography methods, and ground penetrating radar. [Branco,2004; Nabulsi,2005] Evaluation typically combines information from multiple methods and a
priori information to determine and confirm quality of pavements and bridge decks. [Kohl, 2005]

2.2 Ground Penetrating Radar

Ground penetrating radar is one of many electromagnetic nondestructive evaluation methods used in geoscience. These methods exploit differences in electrical potential, conductivity, electric permittivity and magnetic permeability of earth materials to determine underground structure. [Telford, 1990] These methods vary greatly in terms of accuracy, resolution, speed, cost, and other practical factors.

GPR is a high frequency method using radio waves transmitted into the earth’s surface. These electromagnetic waves penetrate earth materials, reflect off boundaries between materials with different dielectrics or conductivities, and are detected by a receiver. The transmitter and receiver are moved along a survey line to scan a cross-sectional slice of underground structure. The electromagnetic waves have a complex interaction with objects and boundaries [Chubinski, 2004; Feng, 2009], so the received data is not a simple picture of underground structure. A great deal of processing and interpretation is needed to determine structure. Most tools developed for processing GPR data are based on methods already in use in geoscience, especially methods developed for seismology. [Jol, 2009; Telford, 1990; Hugenschmidt, 2006].

Ground penetrating radar was first used in 1929 to measure the thickness of a glacier in Austria. [Conyers, 2004; Olhoeft, 2002] The method was mostly forgotten until the late 1950s. US Air Force used conventional radar to measure elevation of planes flying over glaciers in Greenland. Instead of reflecting off the top of the glaciers, the
radar penetrated the ice and reflected off the earth beneath. This caused incorrect altitude readings and plane crashes. These accidents inspired research into deliberately using radar to map subsurface structures of ice and soil. GPR systems became commercially available in the 1970s. [Annan, 2002] These systems became more popular with the rise of personal computers. The use of computers for data collection in the field and for data processing makes GPR surveys much more convenient and cost effective.

Ground penetrating radar has been used for many applications [Daniels 2000; Daniels 2004; Jol, 2009]. These applications include:

- Geological mapping
- Hydrogeology, including the mapping of water tables, aquifers, and contaminant plumes
- Archaeology [Conyers, 2004]
- Mining and petroleum exploration
- Military applications, including the detection of unexploded ordinance, clearing mines, and detecting enemy tunnels and bunkers
- Police applications, including forensics and detection of smuggler's tunnels
- Construction and civil engineering applications, including the detection of failure plains, faults and collapse zones; assessing dam and levee integrity; and highway, bridge and airport runway maintenance

A GPR system is comprised of a transmitter antenna which radiates an electromagnetic pulse into the target, a receiver antenna which detects electromagnetic
signals, and a base station which powers the transmitter and records data from the receiver. The transmitter and receiver are typically separated by a fixed distance and moved together along a survey path. Ground penetrating radar is a tomography method. It collects a two dimensional “slice” of data along the straight line path that the antennas are moved along. The X-axis of the scan data represents the position along the scan path. The Y-axis represents the time delay. The numerical value at each point represents the amplitude detected at the receiver.

GPR scans do not generate ‘nice pictures’ of underground structures, as portrayed in television and movies. However, many underground structures generate specific responses in a GPR scan and can be identified based on their signature. The radar response of a given shape can be derived using Snell’s laws and geometry.

When the pulse wavelength is short relative to the distances travelled, the behavior of the radiated pulse can be modelled in terms of ray optics. The radiated pulse is reflected and refracted at the boundaries between different materials. The angle and amplitude of the resulting ray is determined by the difference in conductivity, magnetic permeability and electric permittivity between the two materials. In most GPR applications, we focus on the electric permittivity or dielectric constant of each material.

The refractive index of a medium is

\[ \eta = \sqrt{\epsilon \mu} \]  \hspace{1cm} \text{Eq 2.1}

where \( \epsilon \) is the dielectric constant and \( \mu \) is the magnetic permeability. \( \epsilon \) and \( \mu \) are properties of how a material interacts with electric and magnetic fields. For the
frequencies used by the GPR equipment and the materials being evaluated, we may treat
\( \mu \approx 1 \). [GSSI, 2003] The refractive index may be estimated as

\[ \eta \approx \sqrt{\varepsilon} \]  

Eqn 2.2

The angles of reflection and transmission beams at a boundary are functions of
the refractive indices of the two media. They are governed by Snell's Laws. If the beam
enters with an angle of incidence \( \theta_i \), the angle of reflection \( \theta_R \) is

\[ \theta_R = -\theta_i \]  

Eqn 2.3

and the angle of refraction \( \theta_T \) is

\[ \sqrt{\varepsilon_1 \sin(\theta_i)} = \sqrt{\varepsilon_2 \sin(\theta_r)} \]  

Eqn 2.4

A simple example of a GPR scan is shown in Figure. The white area within the
box represents a slab of uniform material. The dark shaded region represents a layer of
different material beneath the slab. And the lightly shaded region represents an
embedded object. The radar pulse radiates from the transmitter in every direction within
the ‘slice’ of the slab. Reflections occur at the boundaries between different materials.
The reflections that travel from transmitter to receiver must be symmetric around the
normal of the boundary surface. In Figure 1, there is one reflection from the boundary
between top and bottom layers, and another reflection from the boundary between the top
layer and embedded object. The reflections have different travel paths through the top
layer, and will have different travel times. The two reflection pulses will add as they
reach the receiver.

Figure 1 - Model of slab with embedded object and base, showing radar reflection paths

The example in Figure 1 has travel paths with single reflections. Figure 2 is an example demonstrating multiple reflections and refractions. In this example, one embedded object could cause three unique travel paths with three corresponding travel times. The radar pulse energy is reduced by the refraction at each boundary.
Assuming no outside interference, the signals detected by ground penetrating radar are those that travel a path starting at the transmitter, reflecting and refracting through the surrounding media, and ending at the receiver. The detection of a radar pulse tells the user that at least one reflection occurred along the travel path, and that path has a delay known as the travel time. It does not reveal the thickness or material type of the media that the pulse traveled through, how many different media the pulse traveled through, or if there were more than one reflection along the travel path.

The ambiguity of the GPR target can resolved by moving the transmitter and receiver antennas to different positions, and taking a radar scan at each position.
scan provides constraints on the possible geometry of the underground structure. The underground object created the first response at antenna position one and the second response at antenna position two. The user may now exclude every possible underground structure that generates one response but not both. More scans provide more constraints and eliminate more possibilities. Taken together, a series of GPR scans can identify the shape, material and depth of an underground structure as well as the properties of the surrounding medium.

There are two ways to vary the position of GPR antennas. The most common is to keep both antennas at a fixed separation and move them both along a straight line path. The alternative is to vary the separation and move the antennas apart along a straight line path. Fixed separation antennas are used for this project.

2.3 Evaluation of Concrete using GPR

Ground Penetrating Radar is a frequently used tool in the nondestructive evaluation of concrete. Civil Engineering applications of GPR often have an advantage over other GPR applications. That is because the target of evaluation was recently constructed by humans to accomplish a specific purpose. There is documentation associated with the structure, if not living witnesses who participated in its construction. This documentation provides \textit{a priori} information that is not available in geological and archeological applications of GPR.

The \textit{a priori} information can simplify the evaluation of concrete bridge decks. The standard structure and components of a concrete bridge deck are known. Based on
this standard model of a deck, a simulated GPR scan can be created offline based on it. This simulated ‘perfect’ GPR scan could be compared with real GPR scans of real slabs.

There are reflectors that occur frequently in GPR evaluation of concrete. The first reflector is horizontal layers. Horizontal layers occur frequently: asphalt on concrete, concrete on soil, concrete over air, concrete on concrete, and whenever there is a change of layer. The normal vector of a horizontal surface points vertically up and down at every point of the surface. The thickness of the layer and the reflection angle are both constants, so the travel time will be a constant. The rebar signature of a horizontal layer is a horizontal band of reflected pulses.

The second reflector is long narrow objects such as rebar, pipes and conduits. Because these objects are narrow, they are treated as point reflectors. The travel time is the distance between the GPR antenna travelling along its straight line scan path and the rebar, times a factor for the speed of light in the medium. The distance between a fixed point and a point travelling along a straight line is a hyperbolic function of the position of the point on the straight line. The rebar signature of a point reflector is a series of reflected pulses forming a hyperbola, with the peak of the hyperbola occurring when the GPR antenna is directly over the point reflector.

Steel rebar must be placed perpendicular to the compression force applied to the concrete. There are two commonly ways to align rebar. Figure 3 shows a rebar placed parallel to each other, perpendicular to both the compression force and the direction of travel on the slab. Figure 4 shows two layers of rebar placed on top of each other, one layer perpendicular and one parallel to the direction of travel.
Figure 3 - Concrete deck with rebar perpendicular to travel direction

Figure 4 - Concrete with rebar grid perpendicular and parallel to travel direction

Figure 5 - Two dimensional slice of model concrete deck
Figure 5 shows a two dimensional cross-section of the model concrete bridge deck shown in Figure 3. This is the tomographic ‘slice’ that the GPR system will interact with. The white region is uniformly mixed concrete. The circles are a cross-section view of steel rebar. And the dark region at the bottom is the layer supporting the concrete slab.

![Figure 5](image)

*Figure 5 - Two dimensional cross-section of the model concrete bridge deck.*

Figure 6 shows a simulated GPR scan of the slab cross-section shown in Figure 3. The band at the top represents ground coupling, a form of above ground interference between transmitter and receiver. The four hyperbolas represent reflections from each rebar. The band at the bottom represents the reflection from the base of the concrete slab. The gaps in the bottom band represent regions where the rebar blocks the reflection path to the slab base, causing a shadowing effect.

![Figure 6](image)

*Figure 6 - Simulated GPR scan of concrete deck shown in Figure 7*
Figure 7 shows a real GPR scan of a real 6-inch thick test slab of reinforced concrete. The real scan has the same general features as the simulation.

Defects in reinforced concrete will also interact with the radar pulse. In practice, however, defects can be hard to see in GPR scan data. The primary reason is that the reflections from rebar have amplitude orders of magnitude greater than reflections from defects. Steel rebar has high conductivity and dielectric compared to the surrounding concrete, so rebar acts as a perfect reflector. Defects often have a lower difference in dielectric compared to the surrounding concrete, so they have lower amplitude reflections. Defects also have irregular shapes and compositions, which can cause multipath reflections which can interfere with each other.

The problem is that GPR scans of reinforced concrete have high amplitude reflections from rebar which are not of interest and low amplitude reflections from
defects which are of interest. The goal of this research is to subtract rebar reflections from GPR scans of reinforced concrete such that reflections from defects are not distorted or obliterated.

2.4 Mathematical Model of a GPR Scan of Reinforced Concrete

The mathematical model of a GPR scan of ideal reinforced concrete is a radiated GPR pulse convolved with a sparse array of delta functions attenuating as a function of depth. The delta functions would match the travel time for reflections from the rebar. Because steel rebar has much greater conductivity and magnetic permeability than the surrounding concrete, we can assume ideal reflections from rebar. [GSSI,2003] The reflection will attenuate based on the dielectric of the surrounding concrete and the travel time.

The proposed model of a GPR scan of a defect in concrete is an irregular reflected pulse. The reflection from a defect should be weakly correlated with the reflection from rebar. This assumption is based on

- A defect will typically have an irregular surface, which will cause self-interference
- A defect will typically have a reflection from its top and bottom surfaces, and these two reflections will interfere
- A defect will typically have irregular composition, with transitions from solid concrete to loose material to void, and vice versa
• The boundary between good concrete and a defect will have less contrast than between concrete and steel, so the reflection will have a lower amplitude.

• Depending on the material within the defect, the polarity of the defect reflection may be opposite to a rebar reflection.

Mathematically, we can model a GPR scan of reinforced concrete as

\[ s(x,t) = p(t) \ast h_x(t) + d_x(t) + n_x(t) \]  

Eqn 2.5

where

• \( s(x,t) \) is the GPR scan data.

• \( p(t) \) is a reflection of the GPR pulse from a high contrast boundary, such as the junction between concrete and steel \( d_x(t) \) is the reflection from a defect at scan position \( x \).

• \( n(t) \) is noise.

• \( h_x(t) \) is the impulse response of the rebar at scan position \( x \).

The impulse response can also be expressed as

\[ h_x(t) = \sum_i e^{-\alpha \tau_i} \delta(t - \tau_i) \]  

Eqn 2.6

This model may be recast in terms of matrix math, with impulse response vector \( h_x \) becoming a Toeplitz matrix \( H_x \), and convolution is performed by multiplying a matrix by a vector.

\[ s_x = H_x p + d_x + n_x \]  

Eqn 2.7
The feature of interest is the reflection from the defect $d$. Unfortunately, the GPR scan data is dominated by the rebar reflection term $H_d p$.

### 2.5 Methods for Processing GPR Scans of Reinforced Concrete

There are two general categories of tools currently used to process ground penetrating scans of concrete. The first category is filtering methods that attempt to rebar reflections. The second category is detection methods that use metrics not based on vision to detect the presence of defects.

#### 2.5.1 Migration

Migration describes methods used in the analysis of propagated wave data to recover the original structure of objects. In GPR, seismology and ultrasound, a wave of energy travels through a volume of material. This could be forward modeled using finite element analysis over the entire volume of the survey. It is often more convenient to use Green's theorem to describe the system in terms of boundary surfaces between media than volume. [Shearer, 1999; Fisher, 1992] If $\Psi_1$ and $\Psi_2$ are two continuous single-valued functions with continuous derivatives, then for a volume $V$ with a closed surface $S$

\[
\int_V \nabla^2 \Psi_1 - \nabla^2 \Psi_2 \, dv = \oint_S \left( \Psi_2 \frac{\partial \Psi_1}{\partial n} - \Psi_1 \frac{\partial \Psi_2}{\partial n} \right) \, dS \quad \text{Eqn 2.8}
\]

In the case of wave propagation, the functions $\Psi_1$ and $\Psi_2$ will satisfy
\[ \nabla^2 \Psi_1 = -k^2 \Psi_1 \]
\[ \nabla^2 \Psi_2 = -k^2 \Psi_2 \]  
Eqn 2.9

and the left hand side of the integral will go to zero.

\[
\oint_S \left( \Psi_2 \frac{\partial \Psi_1}{\partial n} - \Psi_1 \frac{\partial \Psi_2}{\partial n} \right) dS = 0
\]
Eqn 2.10

These equations are initial steps in the derivation of Kirchhoff’s theory of reflection. The point these equations demonstrate is the wave will be directed by the normal of the boundary surface between media. If the boundary between different materials in a GPR scan is horizontal relative to the survey path, then the normal will be a constant vertical vector and the reflection will appear as a horizontal line on the survey. If the boundary is anything other than horizontal, then the reflection will not match the shape of the surface. [Olhoeft, 2003] This makes the interpretation of reflections from subsurface layers and objects in a GPR survey difficult.

Migration algorithms are the inverse of wave propagation models. Given reflection data from a survey, a migration algorithm will backsolve to recover a surface that could have created that data. Every migration method requires assumptions regarding the uniformity and dielectric of the medium above the boundary surface. If these assumptions do not match the actual physical properties, then migration may add noise to the image instead of reducing it. [Porsani, 2007]

Many migration algorithms have been developed. The most used is Kirchhoff migration, which was initially developed for seismology. [Shearer, 1999] Other methods,
such as those developed by [Leuschen,2001] and [Streich,2007], are developed specifically for electromagnetic waves from GPR equipment. The algorithm proposed by [Leuschen,2001] is based on Born scattering theory, a first order approximation of Green's theorem. The algorithm proposed by [Streich,2007] uses exact field calculation of wave propagation and exact antenna radiation properties to perform migration more accurately.

Most migration methods assume a single uniform upper layer of soil or pavement. This is not appropriate for complex soils and pavements. A method proposed by [Sena,2006] uses split-step Fourier techniques to compensate for multiple layers and non-uniformities. [Di,2004] proposed a migration algorithm based on finite element analysis instead of using Green's theorem. This method may be more computationally expensive, but it may be well suited for migrating surveys over soil with continuously varying dielectrics and conductivities. [Zhuge,2010] proposed a MIMO array algorithm for migration that can be used for both ground coupled and air coupled GPR antennas.

[Nemeth,1999] developed a modification to Kirchhoff migration that makes it possible to migrate data even if survey data is incomplete. Lateral gaps in seismic data due to coarse sampling or equipment failure are often padded with zeros. These zeros will create artifacts when migrated with normal Kirchhoff migration. The proposed least-squares migration method successfully migrates incomplete data. This algorithm may be useful for GPR surveys performed using global positioning systems (GPS).

[Porsani,2007] presents a case study of detecting two storage tanks buried in soil. Because the tanks are curved and of modest size relative the wavelength of the GPR antenna, their reflections appear as hyperbolas in the survey data. Each tank produces its
own clear hyperbola. A third hyperbola is also visible; it is an artifact created by a reflection path traveling from the transceiver, reflecting off the first tank, reflecting again off the second tank, then returning to the transceiver. This artifact hyperbola appears to have a different dielectric than the two real hyperbolas. Choosing migration parameters than collapse the two real hyperbolas may make the artifact hyperbola more distorted. It may even flip the artifact hyperbola upside-down, creating a prominent 'smile' in the image.

[Sava,2005] developed an algorithm to find dielectric and true shape of boundaries using migration. The algorithm varies the velocity parameter until the resulting migrated image has the highest focus of energy. It is assumed that the migrated curve with the highest focus of energy matches the original boundary, and its velocity parameter provides the dielectric of the upper layer.

[Zhou,2008] developed a method for migrating GPR data from non-homogeneous soils and aquifers. The algorithm uses finite element analysis and genetic algorithms to find the most likely underground structure that could have generated a given set of GPR data.

**2.5.2 f-K Filtering**

f-K filtering is a filter applied to the two dimensional Fourier transform of the GPR scan data. The axes of the GPR data are position along the X axis and time/depth along the Y axis. When the two dimensional Fourier transform is applied to the GPR data, the X axis becomes the wavenumber $k$ (the number of wavelengths per unit length), and the Y axis becomes frequency $f$. 
If an image has high amplitude diagonal lines, then the two dimensional Fourier transform of that image will have prominent diagonal components along the same angle. The diagonal lines in the image can be removed by filtering values along that angle in the Fourier domain and applying the inverse transform.

Reflections from rebar in a GPR scan of reinforced concrete have a hyperbolic shape. They are curved at their peak, but asymptotically approach diagonal lines as you move away from the peak. F-k filtering can be used to remove the diagonal components of the rebar reflections within the asymptotic region with minimal impact to horizontal components.

A drawback of F-k filtering is that reflections from defects may have the same diagonal components as the rebar reflections. Filtering out the rebar may result in filtering out the details of interest.

2.5.3 Deconvolution

When a pulse of energy is transmitted into a material, multiple copies of that pulse are reflected back to the receiver. These pulses have a finite time width. If reflections are separated by a time width less than the pulse width, they interfere with each other and cannot be distinguished as individual reflections. Deconvolution describes methods that attempt to improve the resolution of scan data by replacing wide transmitted pulses with sharper spikes.

Many models of deconvolution assume perfectly identical echoes from every boundary. This is not necessarily an accurate model. Many materials can cause frequency and phase filtering to the pulse. [Chubinski,2004] These effects occur most
dramatically in materials with gradually changing dielectrics. In this case, reflections from different objects will produce markedly different pulses. Boundaries between materials with high contrast in dielectric or conductivity typically do not produce this effect.

Deconvolution serves two purposes in the processing of GPR, seismic and ultrasound data. The first is improving the resolution of data by converting broad received pulses into sharp spikes. The second is removing multiple echoes caused by pulse energy oscillating within a single layer. These multiple echoes are known as ringing. The same algorithms are often used to accomplish both purposes.

Few time-domain deconvolution algorithms require a copy of the initial pulse to deconvolve scan data. [R. Liu, 2008; Krause, 2007]. Most other algorithms are blind or semi-blind source separation methods. [Chahine, 2009] notes that the performance of deconvolution methods improves dramatically with the inclusion of a priori information when compared with pure blind source separation techniques.

[Peacock, 1969] explains the use of least-squares inverse filters and Wiener filters for deconvolution. [Alam, 1981] proposes the use of Gram-Schmidt orthogonalization to estimate a filter purely from the data, as opposed to the use of auto-correlation within the data as used in other methods such as [Rickard, 2007]. [Ulrych, 1991] notes that the phase of the transmitted pulse may limit the effectiveness of deconvolution. In particular, minimum phase wavelets can be shortened to a narrow spike, while non-minimum phase wavelets tend to have long tails with multiple reverberations.

Many predictive deconvolution methods are based on independent component analysis (ICA). ICA requires components with different statistical properties. [J. Liu,
[2007] proposed a method that exploits the non-Gaussian nature of the error, the difference between the original signal and the reconstructed signal based on the deconvolution. [Chahine,2009] proposed a method that uses the sparse nature of the reflectivity term as a non-Gaussian component.

Echo removal is an important area of research. Ringing echoes can occur in any high contrast material where the radiated pulse reflects off the top and bottom boundaries multiple times. A prominent example of this in both seismic and GPR surveys is caused by water. [Morely,1983; Schoenberger,1998]. It also occurs in thin layered materials such as organ walls with ultrasound [Kling,1993; Chang,2008] and thin delaminations in pavements with GPR [Chahine,2009].

As with spike deconvolution, there are many algorithms for echo removal. Many use autocorrelation and cross-correlation [Chang,2008] or least squares decomposition and Wiener filters [Morely,1983] to build prediction filters. [Hornbostel,1999] proposes the creation of a 'noise optimized objective' operator, a filter that maximizes the echo noise in scan data. The inverse of the NOO operator would be expected to remove echo noise from the scan data. The deconvolution it typically performed iteratively, but [Porsani] proposes a single step method requiring very large matrices. Some researchers use transforms to other domains to isolate echo noise. Fourier domain frequency based filtering has long been used for this purpose. [Massier,1997; Conyers,2004] The algorithm proposed by [Duquet,1999] uses the discrete Radon transform to separate echo noise from desired signals.

One interesting approach was explored by [Kling,1993] for medical ultrasound. It has been noted that ringing multiples from flat boundary targets are frequency
independent, while ringing multiples from small scatterers are frequency dependent. The paper proposes performing two ultrasound scans at two distinct frequencies over the same area. Any ringing multiples in common between the two scans are subtracted, leaving only the reflections from small scatterers. This approach may be applicable to GPR surveys of earth materials that produce ringing multiples.

2.5.4 Object Detection Algorithms

The interpretation of GPR scans requires a great deal of human processing and human interpretation. [Conyers, 2004; Daniels, 2004; Olhoeft, 2002] The goal of object detection algorithms is to automate the detection and classification of buried targets within a GPR scan.

Object detection algorithms apply a metric to regional subsets of GPR scan data. If the subset’s metric is above a threshold, that region is declared “bad”; otherwise, it is declared “good”. Multiple metrics can be applied simultaneously to isolate and classify regions of interest.

Reflections from shallowly buried objects may be completely obscured by the ground coupling signal in a GPR scan. [J. Liu, 2006 and 2007] This is especially difficult when the surface of the ground is not completely flat and smooth. Simple averaging filters may not be enough to remove ground coupling. [Wu, 2001] proposes a nonlinear frequency based filter to remove time shifted ground coupling from scan data. [Gao, 2006] proposes the use of principle component analysis (PCA) and kernel independent component analysis (KICA) to resolve faint features obscured by ground
bounce. KICA is better than conventional ICA at dealing with non-linear mixing and when multiple components are nearly Gaussian.

Even with the ground coupling removed, reflections from buried objects are typically weak and can be obscured by reflections from clutter and background noise. [Angelova, 2008] and [Lou, 2005] propose methods using Kalman filters to separate possible targets from background and noise. The classification of targets is performed by a Bayesian filter trained on models of mines and other common objects. [Benedetto, 2005] proposes the use of the Neyman-Pearson radar test to detect local changes which may be defects. [van der Merwe, 2000] proposes a frequency based iterative filtering algorithm to components most likely associated with noise and clutter, similar to the frequency based method proposed by [Kling, 1993]. [Zhou, 2005] proposes the use of wavelet packet transform to clarify the GPR data and neural networks to classify land mines and other targets. [Feng, 2009] proposes the use of migration to convert irregular reflection data back to the surface that likely produces the reflection.

2.5.5 Blind Source Separation

Radar signals are an example of an echoic mixture. A pulse radiates out from a transmitter, reflects off of targets, and returns to the receiver. The majority of received pulses are exact copies of the radiated pulse. These echo pulses differ in amplitude and time shift, based on the two-way travel time to each target. [Jol, 2009]

Blind source separation of echoic mixtures has been studied acoustic and speech processing. One promising approach is the Degenerate Unmixing Estimation Technique algorithm developed by [Rickard, 2005]. The DUET algorithm was originally developed
to separate echoic mixtures in multiple-source-multiple-receiver problems in speech processing.

Suppose we have two signals, a pulse and a time shifted copy of the pulse. In the Fourier domain, that time shift is expressed by multiplying the function by $e^{-i\omega t}$.

$$x_1(t) = p(t)$$
$$x_2(t) = a_2 p(t - \tau)$$  \hspace{1cm} \text{Eq 2.11}

$$X_1(\omega) = P(\omega)$$
$$X_2(\omega) = a_2 P(\omega) e^{-i\omega \tau}$$  \hspace{1cm} \text{Eq 2.12}

We can solve for $a$ and $\tau$ for each frequency $\omega$ in the Fourier domain.

$$a(\omega) = \frac{|X_1(\omega)|}{|X_2(\omega)|}$$  \hspace{1cm} \text{Eqn. 2.13}

$$\tau(\omega) = -\frac{1}{\omega} \angle \frac{X_1(\omega)}{X_2(\omega)}$$  \hspace{1cm} \text{Eqn. 2.14}

If $x_1$ and $x_2$ are time shifted copies of each other, then the values of $a(\omega)$ and $\tau(\omega)$ should form a tight cluster in a histogram in $a$-$\tau$ space. If the two signals are unrelated, then $a(\omega)$ and $\tau(\omega)$ should be widely scattered. The clusters in the histogram should reveal the time-shift and amplitude of every ‘echo’ of the original pulse.

A limitation of this model is that the angle operator is a wrapped function. $\omega \tau$ increases linearly with $\omega$, but $e^{-i\omega t}$ cycles between $\pm \pi$ as $\omega$ increases. This means that phase information is permanently lost. Phase unwrapping is an open field of research with application specific solutions. [Gdeisat, 2005] A phase unwrapping method must be developed to implement this algorithm.
PROPOSED MODELING AND METHODS

3.1 Outline

The proposed algorithm is intended to remove rebar reflections from GPR scans of reinforced concrete. This will be accomplished by exploiting our a priori knowledge of the structure of reinforced concrete and GPR’s response to that structure. Figure 8 is the general flowchart of this algorithm.

- **Stage I: Preprocessing of the GPR scan data.** Preprocessing covers vertical and horizontal filtering, gain boosting, rebar peak detection, and extraction of an exemplar pulse.

- **Stage II: Pulse detection algorithm.** For each column of GPR scan data, the pulse detection algorithm compares a windowed sample of the GPR scan against the exemplar pulse and adds the results to a histogram. An impulse response for the column is constructed based on high threshold peaks in the histogram.

- **Stage III: The hyperbola detection algorithm.** For each rebar peak, the hyperbola detection algorithm collects nearby points that are likely to be part of a hyperbola. This algorithm must correctly bridge gaps where points are missing and exclude points from noise and neighboring objects.
Figure 8 - Flowchart of proposed algorithm

- Stage IV: The generation of synthetic rebar arcs. These synthetic arcs are based on the hyperbolas detected in the previous stage. Missing points in the impulse response are interpolated based on the parameters of the detected points. The impulse response is then convolved with the exemplar pulse.
• Stage V: The subtraction of the synthetic rebar reflections from the original GPR scan data. The rebar reflections should be greatly reduced in the resulting image, with no change in other features. Defects should appear more clearly in the resulting image.

• Stage VI: Postprocessing. This includes any gain boosting and filtering needed before the GPR scan data is used by a Civil Engineer or an object detection algorithm. Postprocessing includes functions available in commercially available GPR processing software such as GSSI’s RADAN.

3.2 Preprocessing of GPR Data

Preprocessing the GPR data involves three major steps. The first is removing previously applied gain boosting from the GPR data, and applying exponential gain boosting if necessary. The second is the application of horizontal and vertical filters to remove obvious forms of noise and unwanted features. The third is the detection of rebar peaks and choosing of an exemplar pulse for use in the pulse detection stage.

3.2.1 Linear Time Invariant Gain Boosting with Exponential Functions

The amplitude of ground penetrating radar signals decays exponentially with travel time. The amplitude is also reduced by every reflection and refraction in its travel path. Because of this, the amplitude of reflections from deep features is often very weak. To counter this, GPR systems apply gain boosting (or gain ranging) to amplify the weak signals. Gain boosting is multiplying each column of GPR scan data by a time based function.
Useful gain functions are typically monotonically increasing splines made from polynomials or exponentials. Gain functions can be automatically generated by the GPR data collection system or be set by the user with GPR processing software.

\[ s_{GB}(x,t) = s(x,t)g(t)\]  

Eqn. 3.1

Figure 9 - GPR scan of six inch concrete slab. Raw data with no gain boosting applied

Figure 9 shows raw data from a GPR scan of a defect free six inch concrete slab. The highest amplitude parts are the horizontal ground coupling line and the reflections when directly over a rebar. Reflections from beneath the top level of rebar are faint in comparison. Figure 10 shows an example of gain boosting automatically applied by GPR scanning equipment. It is same scan of the six-inch concrete slab as Figure 1, but with gain boosting applied by the SIR-3000 Data Acquisition system. Note that the amplitude of the ground coupling is greatly reduced and reflections from deeper objects are much more prominent.
Gain boosting seems like a panacea for all amplitude problems. If signals from a certain depth are faint, just multiply that depth by a larger number. However, gain boosting can break the time-invariance of reflections from a single object. Making the data time-variant will negatively affect our ability to perform convolution, deconvolution and correlation on GPR scan data.

\[ s(t) = \sum_i a_i p(t - \tau_i) \]  
Eqn. 3.2
The gain boosted version of the GPR scan data is

\[ s_{\text{gain}}(t) = \sum_{i} g(t) a_i p(t - \tau_i) \]  

Eqn. 3.3

\[ s_{\text{gain}}(t) = \sum_{i} a_i \frac{g(t)}{g(t - \tau_i)} g(t - \tau_i) p(t - \tau_i) \]  

Eqn. 3.4

The new amplitude and pulse terms are

\[ \tilde{a}_i = a_i \frac{g(t)}{g(t - \tau_i)} \]  

Eqn. 3.5

\[ \tilde{p}(t) = g(t) p(t) \]  

Eqn. 3.6

The new function for the gain boosted GPR data becomes:

\[ s_{\text{gain}}(t) = \sum_{i} \tilde{a}_i \tilde{p}(t - \tau_i) \]

This new function is time invariant only if \( \frac{g(t)}{g(t - \tau_i)} \) is a constant for all \( t \). The only function that satisfies this relationship is \( g(t) = e^{bt} \)

\[ \tilde{a}_i = a_i e^{b\tau_i} \]  

Eqn. 3.7

\[ \tilde{p}(t) = e^{bt} p(t) \]  

Eqn. 3.8

Preprocessing of GPR scan data for this algorithm must include removing any previously applied gain boosting, and replacing it with exponential gain boosting only. This will maintain the linear time invariance of the data, and preserve the convolutive nature of the rebar reflections.
3.2.2 Horizontal and Vertical Filtering

Many filters are used in the processing of ground penetrating radar data. These include finite impulse response, frequency domain, and statistical filters, and can be applied along one or two dimensions. Some of these filters may be applied to GPR scan data before the pulse detection algorithm.

Horizontal filtering is used to remove reflections from large scale horizontal objects. The primary use is to remove ground coupling at the top of the GPR scan data. This horizontal filtering may be applied to specific rows, and not across the entire GPR dataset. This allows the removal of ground coupling near the top of the GPR scan data without affecting lower horizontal features near the bottom.

Care must be taken to not distort the rebar reflections while applying a horizontal filter. Radar pulses approach zero mean when summed over time, but the rebar reflection may be non-zero when summed over varying positions with constant time. An averaging filter may distort the rebar reflection and create new background noise. This is especially true at the rebar reflection peaks. Median filters may detect the higher occurrence rate of the background, and thus subtract the correct value.

Figure 11 shows GPR scan data from a 6-inch reinforced concrete slab with no defects. The high amplitude bands at the top of the data are due to ground coupling. Figure 12 shows the same GPR scan data after a median filter is applied to the first 70 horizontal rows. The ground coupling is mostly removed, leaving the rebar arcs as the most prominent feature in the data.
Vertical filtering is typically based on frequency. The assumption is that the pulses of interest fill a specific frequency range, and anything outside that range may be treated as noise. FIR and frequency domain filters may be constructed and applied to pass the frequencies of interest and block frequencies outside that band. Any filter that does not break the linear time invariance is allowed.

Two dimensional filtering in either the position-time domain or in frequency-wavelength domain is generally not compatible with the proposed algorithm. It can cause horizontal averaging filtering, which as stated earlier can create new background noise and distort rebar reflection signals.

### 3.2.3 Rebar Peak Detection and Exemplar Pulse Extraction

The peaks of rebar reflections create high amplitude local extrema in the GPR scan data. These are straightforward to detect, especially after horizontal interference has
been removed. Figure 13 shows the filtered GPR scan data, with arrows indicating the local maxima.

One of these rebar peaks is chosen as the exemplar pulse. This is the pulse used for all comparisons in the pulse detection algorithm as shown in Figure 14. This selection can be automated or done manually. Rebar peaks form high threshold local maxima in GPR scan data. The algorithm could choose one of these maxima and take a vertical windowed sample around it as the exemplar pulse. If the arc under the rebar peak seems misshapen or otherwise suspect, the human operator should choose a different rebar peak for sampling. For this example, the second peak at x=130 is chosen. Figure 14 shows the amplitudes in the chosen column. The exemplar pulse is a subset of this data which must include the maximum. In Figure 15, the lobe containing the maximum and its two side lobes are selected out of the column data.

The horizontal position of rebar peaks is information that can be reused in the hyperbola reconstruction algorithm later on. It is not necessary, but it is a useful seed value.
Figure 13 - Filtered GPR scan data of 6-inch reinforced concrete slab, with rebar peaks indicated by arrows.

Figure 14 – GPR data from column 130 of scan of 6-inch slab

Figure 15 - GPR data from column 130 of scan of 6-inch slab. The chosen exemplar pulse is shown in solid line, where the rest of the data is dashed.
3.3 Pulse Detection Algorithm

The pulse detection algorithm generates a histogram of the frequency of amplitude and time shift pairs based on the results of deconvolution in the Fourier domain. This histogram is used to verify the existence of pulses, and find the amplitude and time shift of each pulse. Section 3.3.1 covers the derivation and properties of the pulse detection algorithm. Section 3.3.2 covers the application of this algorithm to the task of detecting rebar reflections in GPR scan data.

3.3.1 General Model of the Pulse Detection Algorithm

The pulse detection algorithm must find instances of the exemplar pulse within each column of GPR scan data. For each instance detected, the algorithm must return the time shift and amplitude of the pulse. This algorithm will populate a histogram based on the results of Fourier deconvolution of the windowed sample by the exemplar pulse. The impulse response for that column will be constructed based on high threshold time-amplitude points within each column’s histogram.

Suppose a simple example of an exemplar pulse $p(t)$ and a sample $q(t)$, where $q(t)$ may equal

$$q(t) = ap(t) + n(t)$$

Eqn. 3.9

where $a$ is a constant multiplier and $n(t)$ is noise uncorrelated to $p(t)$. The Fourier transform can be applied to $q(t)$.

$$Q(\omega) = aP(\omega) + N(\omega)$$

Eqn. 3.10
In the Fourier domain, deconvolution is simply division of the two sequences for each omega.

\[
D(\omega) = \frac{Q(\omega)}{P(\omega)} = a + \frac{N(\omega)}{P(\omega)}
\]

Eqn. 3.11

The constant \(a\) in the complex plane has an absolute amplitude of \(a\) and a phase angle of zero radians. When \(D(\omega)\) is plotted in an amplitude-angle histogram, a large cluster of points is expected around \((a,0)\) with perturbations due to the noise term \(N(\omega)\).

Figure 16 demonstrates the histogram of \(D(\omega)\) when \(a=5\) and uniformly distributed noise that is -25 decibels of the original signal is added. Figure 17 demonstrates the histogram of \(D(\omega)\) when \(a=5\) and uniformly distributed noise that is -10 decibels of the original signal is added. Note that the cluster is more concentrated and has greater height with low noise, and becomes more spread with lower height as noise increases.

This Fourier domain histogram method has several useful features. One is that amplitude in the time domain has little effect on the clustering in the Fourier domain. A low amplitude pulse will return a low correlation coefficient in the time domain. In the Fourier domain histogram, a low amplitude pulse produces the same sized peak at a different amplitude-angle point.
Another feature is that it rejects pulses that are visually similar but have different Fourier domain properties. If a pulse is stretched and distorted in time, a correlation method may return a high correlation coefficient between the distorted pulse and the exemplar. In the Fourier domain histogram, the distorted pulse will return scattered points with no clustering. Figure 18 shows an exemplar pulse with a distorted pulse that has been upsampled by a factor of 20 and decimated by a factor of 19. They are quite similar to the naked eye.
Figure 18 – Contrasting an exemplar pulse with a copy resampled at a 20:19 ratio

Figure 19 – Complex Plane Histogram results of deconvolving the 20:19 resampled pulse with the original exemplar

Figure 19 shows the Fourier domain histogram of the distorted pulse deconvolved by the exemplar pulse. The deconvolution results are scattered across angle-amplitude space, forming no tall peak anywhere. If the threshold level were set at 10 occurrences, then this would be considered a “No pulse detected” situation.
Now consider a more complex example, where the pulse within \( q(t) \) may be time shifted by \( \tau \).

\[
q(t) = ap(t - \tau) + n(t)
\]

Eqn. 3.12

In the Fourier domain, this becomes

\[
Q(\omega) = ae^{-i\omega\tau} P(\omega) + N(\omega)
\]

Eqn. 3.13

and the deconvolution product becomes

\[
D(\omega) = \frac{Q(\omega)}{P(\omega)} = ae^{-i\omega\tau} + \frac{N(\omega)}{P(\omega)}
\]

Eqn. 3.14

The \( e^{i\omega\tau} \) term corkscrews through phase space as \( \omega \) increases. When plotted in the real-imaginary complex plain, this forms a circle. This is shown in Figure 20. When plotted in angle-amplitude complex plane, it appears as a low height line where amplitude equals \( a \). This is shown in Figure 21.

Phase angle is a wrapped function. The \( \omega\tau \) term can range from minus infinity to plus infinity. However, angle of \( e^{i\omega\tau} \) ranges between \(-\pi\) and \(+\pi\). Time shift \( \tau \) cannot be recovered from phase angle through algebra alone.

Looking at Fourier domain histograms for different time shifts reveals the distribution of \( D(w) \) values are roughly uniform across angle when \( \tau \) is not zero. When \( \tau \) is zero, then the angles of \( D(w) \) cluster tightly around zero radians.

This could be resolved by creating a library of \( p_j(t) \), where

\[
p_j(t) = p(t - j)
\]

Eqn. 3.15
and \( j \) is every physically possible time shift. Deconvolution could then be performed using every \( P_j(w) \)

\[
D_j(\omega) = \frac{Q(\omega)}{P_j(\omega)} = ae^{-i\omega(\tau-j)} + \frac{N(\omega)}{P(\omega)e^{-i\omega j}} \tag{Eqn. 3.16}
\]

The \( D_j(w) \) with the highest concentration of points with an angle of zero radians could be considered the case where \( j = \tau \). A measurement is needed to rate the distribution of points across angles of \( D_j(w) \), with one rating for a uniform distribution and a different rating for a highly peaked distribution.

---

**Figure 20** – Complex Plane scatter plot of deconvolution of time shifted pulse by exemplar pulse

**Figure 21** – Histogram of deconvolution of time shifted pulse by exemplar pulse
The measurement used in the proposed algorithm is kurtosis. Kurtosis is the fourth standardized moment about an average, divided by the second moment squared.

\[
K = \frac{\sum_{n} (X_n - \mu)^4}{\left(\sum_{n} (X_n - \mu)^2\right)^2}
\]

Eqn. 3.17

Kurtosis is a dimensionless measurement used in statistics to measure the ‘peakedness’ of a distribution. A low kurtosis value indicates that few values occur near the distribution’s average. A high value indicates that the values at or near average value occur very frequently. For example, the kurtosis of a uniform distribution is 1.8. The kurtosis of a Gaussian distribution is 3, and the kurtosis of a Laplacian double exponential distribution is 6.

For the proposed algorithm, the kurtosis of the angles of \(D_j(w)\) is measured across all \(\omega\). The time shift \(j\) with the highest kurtosis is then treated as the best guess for time shift \(\tau\). This \(D_\tau(w)\) is evaluated in a histogram as before. If the pulse matches the exemplar pulse, then a high threshold peak is expected in the corresponding histogram for that time shift. If high kurtosis is generated coincidentally by noise given tau, the histogram based on that tau is expected to lack a high threshold peak.

Figure 22 and Figure 23 show kurtosis measurements for each time shift of a pulse shifted by \(\tau =13\). Figure 22 is an example with low noise and a high kurtosis peak.
Figure 23 is an example with medium noise. The added noise creates more outliers, which in turn reduces the kurtosis.

A data set that does not contain the exemplar pulse will still have a maximal kurtosis at some time shift. A minimum threshold may be applied to reject tests where the maximum kurtosis is too close to the uniform distribution. Or the maximum kurtosis may be blindly trusted and a histogram generated based on that time shift $\tau$. The resulting histogram should be scattered, and have no peaks high enough above the threshold to trigger a detection.

**Figure 22 – Kurtosis of angles of $D(\omega)$ for each time shift, when the sample has low noise**

**Figure 23 - Kurtosis of angles of $D(\omega)$ for each time shift, when the sample has high noise**
The pulse detection algorithm fails when two pulses occur within one sample window. Figure 24 shows a sample containing two pulses and no noise. Figure 25 shows a scatterplot of the deconvolution product in phase space. Where the presence of one pulse generates a circle in the complex plane, the presence of two pulses generates a more intricate spiraling pattern. The pulse detection algorithm varies a single tau parameter to maximize kurtosis to detect a single pulse. The two pulse case would require varying two tau parameters and the ratio of amplitudes between the two pulses simultaneously to find the pulses’ properties.

A two pulse detection algorithm is not necessary in this application. By using sliding and overlapping sample window, the algorithm may detect the first pulse by itself in one sample window and the second pulse by itself in a different sample window. If this does not succeed, the two pulses will be interpolated later by the hyperbola reconstruction algorithm.
Figure 24 – Sample data containing two exemplar pulses

Figure 25 – Scatterplot of deconvolution of the two pulse sample by the exemplar pulse
3.3.2 Applying Pulse Detection Algorithm to the Detection of Rebar Reflections

The model for GPR scans of reinforced concrete as covered in Section 2.4 indicates that reflections from rebar are expected to be the same pulse repeated with different time shifts and amplitudes. This means that the pulse detection algorithm may be useful in processing GPR scan data.

One concern is that the GPR scan data will have instances where rebar reflections overlap. In these instances, the pulse detection algorithm will return a false negative for both of those rebar reflections. The stages after the pulse detection algorithm must compensate for this tendency. The hyperbola detection algorithm is built to “bridge gaps” where pulses are not detected, and the hyperbola synthesis stage will interpolate missing pulses.

The flowchart for detecting pulses in GPR scan data is shown in Figure 26. Windowed samples from each column of GPR scan data are processed using the pulse detection algorithm. The results of the algorithm are time shifted to match the time of the windowed sample and added to a histogram for that column. After the entire column has been sampled, the histogram is examined to find high threshold local maxima. Each local maximum represents the time shift and amplitude of the impulse response of a rebar reflection. These maxima are used to create an impulse response for the column. When all the columns are processed, an impulse response image for rebar reflections is assembled.
Figure 26 – Flowchart of Pulse Detection Algorithm applied to a column of GPR scan data
The windowed samples are chosen to have overlap. Overlap has several helpful features. First, it creates the possibility of capturing one pulse within one sample while excluding nearby pulses. This can reduce the occurrence of false negatives. Second, windowed samples can capture the same pulse with different slices of background noise and features. If background noise randomly obscures a pulse in one sample, it may not obscure the pulse in the next sample. This improves performance in regions where the rebar reflection has low amplitude relative to the background.

### 3.4 Hyperbola Detection Using Search Mask

The impulse response matrix generated by the pulse detection algorithm is a sparsely populated matrix. If the matrix was displayed as an image, a human observer would see a collection of scattered points. The observer would see arcs in the data, and could “connect the dots” to find hyperbolas. This task is tedious and must be automated. That is the purpose of the hyperbola detection algorithm.

Rebar reflection peaks were found in the original GPR scan data during the preprocessing stage. Rebar peaks in the impulse response should have the same horizontal $x$ position as in the original GPR scan data. The vertical $y$ position should be shifted upward in the impulse response due to deconvolution. The impulse response positioned directly above the position of a rebar peak in the GPR scan data can be treated as the rebar peak in the impulse response.
With the rebar peaks known, the points forming the hyperbola beneath each peak must be found. This is accomplished by the iterative application of a search mask. The search mask is a matrix of ones and zeros. The placement of ones and zeros is determined by the curve of the expected hyperbola. If the search mask is places adjacent to a known hyperbola point, then positions where other hyperbola points are expected to be found are set to one and regions where points are excluded are set to zero. The shape has to accommodate the plateau region at the peak of the hyperbola and the asymptotic region as we move away from the peak. Figure 27 shows an example search mask. Because rebar hyperbola are two sided, a mask may be constructed and for the right side and mirrored copy of the mask used for the left side.

Figure 27 - Search Mask for right side of hyperbola
Figure 28 shows an example of the search mask applied iteratively to a set of points to detect the hyperbola. The rebar peak is declared to be the “current point”. A square region adjacent to the current point is copied and multiplied by the search mask. The region is searched by column from left to right for a non-zero entry. When such a point is found, the position and amplitude of the point is added to the list of points under that rebar peak. The newly discovered point is declared the current point, and the process is repeated until no more points are found. When the right side of the hyperbola is finished, the process is repeated and mirrored to the left.

The shape of the search mask has to accommodate the asymptotic region of each hyperbola. This is necessary because this is the region where overlapping pulses and false negatives will occur. It must bridge the gap where points are missing and exclude points from other arcs and from noise. Currently, the lower region of the search mask must be set by the operator to the correct angle to capture these points.

The hyperbola detection algorithm produces a database of rebar peaks, and of the points associated with each peak.
Figure 28 - Search Mask applied iteratively to scattered hyperbola points
3.5 Generation of Synthetic Rebar Reflections

The parameters of each rebar arc may be derived using the database of rebar peaks and points associated with each peak. This is accomplished using our a priori knowledge of rebar reflections in a GPR scan of reinforced concrete. The positions of points associated with a rebar peak should match a hyperbola; therefore there must be a set of best fitting parameters for that hyperbola. The amplitude of points should decay exponentially with depth; again, there must be best fitting parameters for that exponential function.

The function for a hyperbola is

\[ y_i = A\sqrt{(x_i - x_0)^2 + B} + y_0 \]  

Eqn. 3.18

where \( x_0, y_0, A \) and \( B \) are the parameters of the hyperbola. A cost function \( V_{\text{Hyperbola}} \) is created to sum the squared error between the observed data points \([x_i, y_i]\) and the estimate hyperbola. The hyperbola parameters are varied to minimize the cost \( V_{\text{Hyperbola}} \).

\[ V_{\text{Hyperbola}} = \sum_i \left[ A^2 (x_i - x_0)^2 + A^2 B - (y_i - y_0)^2 \right]^2 \]  

Eqn. 3.19

Using MATLAB function \textit{lsqcurvefit}, I was able to find the parameters that minimize a user defined function over a set of data points. The function \textit{lsqcurvefit} is used to find the best fit hyperbola parabola for each detected rebar arc. The function \textit{lsqcurvefit} requires initial parameters as a starting point. These initial parameters must be chosen with care because the cost function is nonlinear. Some starting points will diverge.
away from the optimum parameters. In practice, setting \( x_0 \) to the \( x \) position of the rebar peak, \( A=1, B=0 \) and \( y_0=0 \) consistently converges.

Setting \( x_0 \) to \( x_{\text{Peak}} \) makes the observed and initial estimate hyperbolas share the same conic axis. This allows the algorithm optimize three parameters rather than four.

Taking partial derivatives of \( V_{\text{Hyperbola}} \) with respect to \( B \) and \( x_0 \) and solving for \( B \) and \( x_0 \) involves division by \( A \). If the initial value of \( A \) is critically close to zero, the algorithm will push \( B \) and \( x_0 \) to huge values outside of the convergence region for the function. The algorithm will either never converge from this bad start or converge extremely slowly. Setting the initial value of \( A \) to 1 is sufficiently far from zero to avoid this problem.

Parameter fitting provides a way to judge the quality of a prospective rebar arc. If a collection of points does not form a hyperbola, then the parameters will have impossible values such as complex numbers. Complex valued parameters will occur when the right side of a prospective rebar arc is not the mirror of the left side. Arcs with impossible valued parameters may be removed from the database of rebar arcs.

This best fit process is repeated for the exponential amplitude decay as a function of time depth \( y \).

\[
z_i = d_0 e^{-d_1 y_i}
\]

Eqn. 3.20
where $z_i$ is the amplitude of a detected impulse response, $d_0$ is the constant multiple, $y_i$ is the delay time of the detected impulse response and $d_1$ is the decay rate. The parameters that best fit the data will minimize the squared error cost function.

$$V_{Decay} = \sum_i \left[ z_i - d_0 e^{-d_1 y_i} \right]^2$$

Eqn. 3.21

Again, Matlab’s lsqcurvefit function is used to find the best fit parameters for the exponential decay rate.

The best fit parameters of the rebar reflection hyperbola are now used to generate the synthetic arc. The parameters for a rebar reflection are applied to the hyperbola function and decay rate functions. For each $x$ within the range of the GPR scan, an impulse response is placed at $(x,y(x))$ with amplitude $z(y(x))$. This is repeated for each separate rebar reflection.

The hyperbola function $y(x)$ outputs decimal values even when $x$ is constrained to integer values. The impulse response for this $(x,y(x))$ pair must be placed at an integer location on the impulse response matrix. Rounding $y(x)$ to the nearest integer value can create significant artifact error, especially when the remainder of the rounding is near $\pm0.5$. This stage will round $y(x)$ to the nearest integer. The error associated with that will be dealt with in the following stages.

These synthetic arcs will interpolate missing impulse responses that were skipped as false negatives by the pulse detection algorithm. We now have the impulse response of rebar reflections for the GPR scan data.
3.6 Removal of Rebar Reflections

The main algorithm starts with a selected exemplar pulse as shown in Figure 15. That exemplar pulse was sufficient to find similar pulses, but it may be narrower than the actual rebar reflection pulse. It also may include background noise. The synthetic impulse response is now used to create an improved exemplar pulse. The improved exemplar pulse will be used in convolution and image subtraction.

An improved pulse may be constructed by deconvolving columns of the original GPR scan data with corresponding columns of the impulse response.

The synthetic rebar reflection is generated by convolving the improved exemplar pulse with each column of the impulse response image. The filtered GPR scan is generated by subtracting the synthetic rebar reflection from the original GPR scan data.

3.7 Postprocessing of GPR Data

The filtered GPR scan data will have greatly lowered amplitudes for rebar reflections. However, it may need additional processing before it is ready for use by a Civil Engineer or an object detection algorithm.

The filtered GPR scan data may have high amplitude artifact noise. This occurs mainly at intersections where two rebar reflections overlap. These artifacts may have the highest amplitude of any feature in the GPR scan data. One way to deal with this is with amplitude clipping. Select a maximum and minimum allowable amplitude for the data.
Any point whose amplitude exceeds that range is “clipped” to the boundary value that it exceeded. Clipping can attenuate high amplitude artifact noise without affecting lower amplitude features.

Another approach to removing artifact noise is applying a low-pass horizontal filter to remove high frequency components. When viewed horizontally, the artifact noise forms spikes within the rows of data. A low-pass horizontal filter may reduce these spikes while leaving more slowly changing features intact.

After the effects of artifact noise are dealt with, the filtered GPR scan data can be processed like any other GPR scan data. The standard techniques of filtering and gain boosting may be applied.
4.1 Test Data Sources

We implemented the proposed method using GPR scan data from three sources. The first source is computer simulations of GPR scans. This includes fully computer generated scans and real GPR scans of reinforced concrete with computer generated defects added. The second source is from GPR scans of a sandbox constructed as a concrete slab substitute while the third source is GPR scans of a slab of reinforced concrete that has known defects embedded within it.

The data sources are treated as having dimensionless amplitude. This is justified in part because GPR equipment conceals absolute voltage levels from casual users. In addition, Civil Engineers and Geologists interpret GPR scan data by viewing dimensionless relative differences in amplitude, rather than voltage levels. This dissertation work will continue the practice of dimensionless amplitudes in GPR scans.

4.1.1 Computer Simulated GPR Scan Data

Computer simulations of GPR scan data were generated as test inputs for the rebar removal algorithm. Figure 29 through Figure 33 demonstrate one computer
simulated GPR scan and the results of processing the data through the proposed algorithm.

Figure 29 shows a simulated defect with a pulse that is exponentially decaying sine wave with a maximum amplitude of 16.9. The horizontal shape would suggest a delamination or a horizontal air void if found in real GPR scan data. An exponentially decaying sine wave is used because it is different from the pulse that will be used for the rebar reflections.

Figure 30 shows the same defect after the addition of high amplitude rebar reflections. The absolute maximum value of the rebar reflections is 52.8. The grayscale image is scaled so that 52.8 is mapped to 256, zero is mapped to 128, and -52.8 is mapped to zero. In this image, the embedded defect is barely visible, depending on the quality of monitor or printer used to view it.

Figure 31 shows the output of the pulse detection algorithm. The detected pulses match the curves and amplitudes of the rebar arcs in Figure 30. Pulses are frequently not detected in regions where rebar arcs overlap.

Figure 32 shows the synthetic rebar impulse response, based on the results of the hyperbola detection algorithm and best fitting hyperbola parameters.
Figure 29 - Simulated defect with maximum amplitude of 16.9

Figure 30 - Simulated defect as shown in Figure 29, with the addition of simulated rebar with a maximum amplitude of 52.8
Figure 31 - Output of pulse detection algorithm

Figure 32 - Synthetic rebar impulse response arcs
Figure 33 shows the result of subtracting the synthesized rebar reflection from the original GPR scan data. The amplitude of the rebar reflections has been reduced to an absolute value of 1.4. That is a reduction of -31dB. The added defect is now visible again. There is significant artifact noise in regions where the rebar arcs cross. The amplitude of this artifact noise has a maximum of 10.7, which is a reduction of -10dB from the original amplitude.
4.1.2 Sandbox Test Bed

The drawback of experimenting with concrete is that concrete is resistant to change. If you wish to change the configuration of a defect embedded in the slab, you must create an entirely new slab. Thus, concrete does not work well with the rapid prototyping that Electrical Engineers are accustomed to.

One key feature of Ground Penetrating Radar is that it does not actually react to concrete. It reacts to the dielectric constant of concrete. If the concrete was replaced with another material that has the same dielectric constant, it would appear the same as concrete on a GPR scan.

Figure 33 - Original GPR scan data as shown in Figure 32 minus synthesized rebar reflections
Concrete typically has a dielectric constant of 6. Dry sand typically has a
dielectric constant of 5. [Conyers] Dry sand is a passable substitute for concrete in this
case.

A sandbox was constructed. It has a 4 foot by 4 foot base, to match the 4 foot by
4 foot concrete slabs. The sides are 1X12 planks. The box was filled with dry sand to a
depth of 10 inches.

Figure 34 - Concrete simulating sandbox under construction
Figure 35 - Rebar grid for concrete simulating sandbox

Figure 36 - Concrete simulating sandbox with GPR antenna and base station
Figure 37 - GPR scan of concrete simulation sandbox with one rebar and no embedded defects

4.1.3 Concrete Test Slabs

Reinforced concrete test slabs have been constructed by WMU’s Civil and Construction Engineering Department and have been made available for GPR research. [Nabulsi] The slabs have been constructed in 4, 6 and 8 inch depths.
An additional 4 inch slab has been constructed with embedded objects and simulated defects. PVC pipes and artificial sponges were added before the concrete was poured. After the concrete cured, the sponges were dissolved using acetone and drained. The empty space left behind is an air void, similar to what is created by real-world wear.
Figure 39 - Construction of 4-inch slab with embedded defects

Figure 40 - GPR scan of 4-inch slab with embedded defects
CHAPTER 5

CONCLUSION

This dissertation proposed an algorithm for the segmentation of features of GPR scans of reinforced concrete bridge decks. GPR scans of bridge decks are cluttered with high amplitude reflections from rebar and other known structures. These high amplitude signals can obscure the lower amplitude reflections from small scale defects. Traditional filtering methods can remove the rebar reflections, but these methods attenuate and distort the signals we are trying to detect.

The proposed algorithm exploited a priori knowledge of the structure of reinforced concrete bridge decks to detect and isolate reflections from rebar in GPR scan data. When the rebar-only data was subtracted from the original GPR scan data, the remaining image showed defects and background features clearly with no attenuation or distortion. The algorithm reduced the amplitude of rebar reflections by -30dB in simulated GPR scans.

The proposed algorithm automates the processing of GPR scan data. The algorithm requires a small set of parameters to configure, and uses features contained within the GPR scan data to detect defects. Further optimization will make this algorithm a utility that can be applied in the field while GPR scans of bridge decks are performed.
This will greatly reduce the time and cost associated with performing GPR scans. The algorithm may be used to aid in human processing and interpretation of GPR scan data. It may also be a first stage of automated defect detection algorithms.

The algorithm detects pulses in the frequency domain. This approach works best when the frequency distribution of the pulses is different than the frequency distribution of noise, or when the noise is a low amplitude perturbation to the pulses. The algorithm’s performance suffers when either of those conditions is not met. The algorithm may be improved by extrapolating rebar arcs based on reflections from low noise amplitude regions.

The proposed algorithm may be expanded to process GPR scans using multi-antenna arrays. The geometry of an antenna array can provide even more a priori information for the segmentation algorithm. This can expand the object detection capabilities beyond geometrically simple objects such as rebar. This can allow us to reconstruct irregularly shaped embedded objects and calculate their dielectric properties.

A core motivation for this algorithm is that single globally applied filters are often unsuccessful in processing an additive mixture of signals. A globally applied filter intended for removing rebar reflections will distort non-rebar reflections to a measurable extent. Segmenting features from the data into separate images will allow us to apply specific filters to each image appropriate for its features. That in turn will allow us to recover the properties and composition of embedded objects. This concept of data
segmentation, separate processing of segments, and recombination can be extended to other non-destructive evaluation methods.


F. Branco, J. Brito, Handbook of Concrete Bridge Management, ASCE Press, 2004


PUBLISHED WORK


