Extracting sparse crack features from correlated background in ground penetrating radar concrete imaging using robust principal component analysis technique

Yu Zhang, Tian Xia
yzhang19@uvm.edu, txia@uvm.edu
School of Engineering, University of Vermont, Burlington, VT 05405

ABSTRACT

Crack detection is an important application for Ground penetrating radar (GPR) to examine the concrete road or building structure conditions. The layer of rebars or utility pipes that typically exist inside the concrete structure can generate stronger scattering than small concrete cracks to affect detection effectiveness. In GPR image, the signature patterns of regularly distributed rebars or pipes can be deemed as correlated background signals, while for the small size cracks, their image features are typically irregularly and sparsely distributed. To effectively detect the cracks in concrete structure, the robust principal component analysis algorithm is developed to characterize the rank and sparsity of GPR image. For performance evaluations, simulations are conducted with various configurations.

Keywords: ground penetrating radar, feature extraction, low-rank representation, sparse representation, robust principal component analysis, non-destructive evaluation.

1. INTRODUCTION

Ground penetrating radar (GPR), as a non-destructive evaluation technique, has been effectively utilized to monitor the concrete structure conditions of roads and buildings [1][5]. In such applications, the strong clutter induced by the concrete surface often overwhelms the reflection signal from the cracks inside the concrete. Many methods have been investigated to mitigate the concrete surface clutter or to enhance the target features, including subsection projection approach [6], time gating [7], spatial filtering [8][9], exponential scaling [10][11], pattern matching [12], 2D entropy singular region analysis [13][14], etc. However, for most roadway and building structure, a layer of reinforced steel bars or utility pipes typically exist inside the concrete, which can produce stronger scattering comparing with small cracks inside the concrete. For the regularly distributed rebars or pipes, their signature patterns in GPR image can be deemed as highly correlated background clutter signals, while the small size crack image feature is irregular and sparse. The traditional clutter removal or feature extraction methods are not effective to remove the correlated background clutter produced by these embedded rebars or pipes inside the concrete.

Among many crack detection algorithms, the subsection projection approach [6] is based on the reflection energy difference between the concrete surface and the cracks. Singular value decomposition (SVD) is performed on the data matrix to identify and remove the concrete surface electromagnetic signature. Unfortunately, the correlated background clutter inside the concrete are stronger than crack reflections while weaker than the concrete surface clutter. It is infeasible to determine an appropriate threshold value for the subsection projection approach when no prior-knowledge about the concrete structure is available. A windowing function [7] can be defined to null the signal segments over the time intervals where different signal traces exhibit a high similarity, which facilitates correlated background signal removal. Spatial filtering method [8][9] utilizes the same assumption to filter out the clutter data corresponding to the concrete surface reflection. Considering that reflection signal from cracks with limited spatial extent vary in different A-scan traces, a spatial filter is thus applied along the antenna moving direction to mitigate the spatial zero-frequency and low-frequency components corresponding to concrete surface reflections. Since the correlated while irregularly distributed background clutter signals are also varying at different scan positions, it cannot be distinguished from the cracks with time gating or spatial filtering. Exponential scaling method [10][11] compensates for the signal attenuation loss along the GPR scan range. However, if the cracks have the similar depth with rebars or pipes in the concrete, both the reflection signal of the cracks and the correlated background clutter will be enhanced by amplitude scaling, which inadvertently increases the difficulty of feature isolation. A pattern matching method [12] based...
on the calculation of cross-correlation in the GPR image can deal with the difficult case when cracks and the correlated background clutter overlap in time domain. In the data processing, the signature image pattern of the correlated background clutter, such as rebar, is utilized as the reference for correlated matching pattern search and identification. However, when the crack under test is long and wide, it will also result in a similar image pattern, making it hard to be isolated. 2D entropy analysis \cite{13,14} is developed to detect the singular region in the GPR image by characterizing the statistical distribution of the reflection signal energy distribution assuming rebars or pipes are regularly and uniformly located in the region under inspection. However for tests where rebars are unevenly distributed, the effectiveness of 2D entropy analysis is degraded. Therefore a new effective method to extract the small and weak crack signature from the strong correlated background clutter is important and necessary.

To effectively separate the spares crack signals, the robust principal component analysis (RPCA) technique \cite{15} is applied in this paper which characterizes the rank and sparsity of sensing data matrix. Mathematically, RPCA approach aims to decompose the data matrix into the superposition of two sub-matrix, one low-rank matrix representing the correlated background and the other sparse matrix capturing small irregular features. In GPR imaging, the correlated background clutter produced from the embedded rebar or pipes forms a low-rank matrix, while the small crack leads to a sparse matrix on top of the correlated background clutter. The crack features extraction problem can be transformed into a RPCA problem. Upon the decomposition, the sparse matrix representing the small and weak crack feature can be extracted.

The rest sections of the paper are organized as followings. Sec. 2 describes the principle of RPCA technique. Sec. 3 models the crack pattern extraction in GPR concrete inspection as a RPCA problem. In Sec. 4, the proposed concrete crack extraction method is evaluated using finite-difference time-domain (FDTD) simulations. Sec. 5 summarizes the concluding remarks.

2. LOW-RANK AND SPARSE REPRESENTATION IN RPCA

The RPCA \cite{15} interprets the observed data matrix $D \in \mathbb{R}^{m \times n}$ as a superposition of a low-rank matrix $L \in \mathbb{R}^{m \times n}$ and a sparse matrix $S \in \mathbb{R}^{m \times n}$, where $L$ represents the correlated background, while $S$ models the target features on top of the correlated background. The mathematical expression is $D = L + S$. In the GPR imaging application, $m$ represents the number of data points in each A-Scan trace, while $n$ represents the total number of A-Scan traces in the B-Scan image.

Decomposing the data matrix $D$ into $L$ and $S$ is an optimization problem. Through Lagrangian reformulation, it can be expressed as:

$$\min_{L,S} \text{rank}(L) + \lambda \|S\|_0 \quad s.t. \quad D = L + S$$  \hspace{1cm} (1)

In the general rectangular case, where $m \geq n$, if

$$\text{rank}(L) \leq \rho_n \frac{n}{(\log m)^{2}}$$  \hspace{1cm} (2)

and

$$\|S\|_0 \leq 0.1 \times mn$$  \hspace{1cm} (3)

matrix $L$ and $S$ can be uniquely reconstructed by solving Eq. (1). In Eq. (2), $\rho_n$ is a positive constant coefficient.

Unfortunately, Eq. (1) is a highly nonconvex optimization problem subsuming both the low rank matrix completion problem and the $l^0$-minimization problem, which are both NP-hard. By replacing the $l^0$-norm with the $l^1$-norm, and the rank of $L$ with the nuclear norm $\|L\|_* = \sum_i \sigma_i(L)$, a tractable optimization problem can be obtained \cite{15}:

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1 \quad s.t. \quad D = L + S$$  \hspace{1cm} (4)

where $\|L\|_*$ is the nuclear norm or sum of singular values of matrix $L$, $\|S\|_1$ is the $l^1$-norm or sum of absolute values of the entries of $S$, and $\lambda$ is a tuning parameter that balances the contribution of the $l^1$-norm term relative to the nuclear norm term. Choice of $\lambda = 1/\sqrt{\max(m,n)}$ is universal for solving the optimization problem in Eq. (4).

3. CRACK FEATURES EXTRACTION IN GPR IMAGE
The GPR concrete inspection model is depicted in Figure 1. The transceiver antennas move along the concrete surface. The rebars or utility pipes inside the concrete produce similar image patterns, while their distributions along the antenna scanning axis can be irregular.

\[ x_n(t) = b_n(t) + l_n(t) + s_n(t) \]  

where \( b_n(t) \) is the reflection signal from the concrete surface, \( l_n(t) \) is the correlated background clutter produced from the embedded rebar or pipes inside the concrete, and \( s_n(t) \) is the reflection signal from the cracks inside the concrete. Various methods have been developed to mitigate or remove the concrete surface reflection signal \( b_n(t) \). The resulting signal model upon such processing can be expressed as

\[ d_n(t) = l_n(t) + s_n(t) \]  

In this paper, RPCA technique is investigated to separate the correlated background clutter \( l_n(t) \) and the image signature of cracks \( s_n(t) \).

For GPR imaging, \( d_n(t) \), \( l_n(t) \) and \( s_n(t) \) are recorded as \( M \times 1 \) vectors \( d_n, l_n \), and \( s_n \) respectively. \( M \) is the number of samples collected at each scan position. Assembling the data of all \( N \) scan positions leads to the following \( M \times N \) data matrices:

\[ D = [d_1, d_2, d_3, ..., d_N] \]  

\[ L = [l_1, l_2, l_3, ..., l_N] \]  

\[ S = [s_1, s_2, s_3, ..., s_N] \]  

According to Eq. (6), it has

\[ D = L + S \]  

Since \( L \) contains the correlated background data, the rank of \( L \) matrix is low. While for the object data matrix \( S \), the crack features in concrete are spatially sparse, therefore \( S \) is a sparse matrix. Based on the analysis elaborated in Section II, if the crack features are sparser than the distribution of the correlated background clutter, the crack features \( S \) can be extracted from the correlated background signal \( L \) by solving Eq. (4). The procedures are summarized below:

1. Pre-processing: Remove the concrete surface reflection signal using a 2-D high pass filter \[9\].
2. Decompose the processed GPR data matrix \( D \in \mathbb{R}^{M \times N} \) into a low-rank matrix \( L \in \mathbb{R}^{M \times N} \) and a sparse matrix \( S \in \mathbb{R}^{M \times N} \) by solving Eq. (4) with tuning parameter \( \lambda = 1/\sqrt{\max(M,N)} \).
3. The sparse matrix \( S \) represents the crack features in the concrete.

4. Experimental results

In order to evaluate the proposed correlated background removal method, experiments on GPR concrete inspection data are conducted. The test utilizes the data set synthesized with the simulation program GprMax \[16\], and the optimization problem in Eq. (4) is solved utilizing the mathematical toolbox TFOCS \[17\]. Two types of rebar distributions are examined in the experiments: one is a layer of evenly distributed rebars in the concrete, the other is a layer of unevenly distributed rebars in the concrete. For each rebar distribution configuration, three types of crack settings are tested respectively: (1) One rebar
crack filled with air; (2) One crack filled with water; (3) One crack filled with air and one crack filled with water. In each setup, the crack is modeled as a tiny cylinder parallel to the rebar.

4.1 Evenly distributed rebars test

The geometry for evenly distributed rebars test data is: GPR transceiver antennas are located 2.5 mm above the concrete surface. The concrete is modeled as a homogeneous layer of 25 cm thickness with 6.0 dielectric constant. Eleven steel reinforce bars of 1.25 cm radius are evenly located inside the concrete. For three types of crack settings, the crack in the concrete is modeled as a tiny cylinder of 1.25 mm radius respectively. For the first crack setting, the material of the crack is set as air. For the second crack setting, the material of the crack is set as water. For the third crack setting, two cracks are placed in the concrete, where the material of the first crack is set as air, and the material of the second crack is set as water. The test geometry detail for each test case is depicted in Figure 2(a), Figure 3(a) and Figure 4(a) respectively.

In the FDTD simulation, the GPR waveform is generated as a Ricker waveform (i.e. negative normalized second derivative of a Gaussian pulse) with its center frequency being 2.3 GHz. GPR A-Scan traces are collected from left to right along the horizontal direction uniformly at every 2 cm distance. The size of the data matrix is 678 × 156. The raw GPR B-scan image for each test case is plotted in Figure 2(b), Figure 3(b) and Figure 4(b) respectively. In each raw B-Scan image, the concrete surface reflection is shown as a horizontal line, and the rebars inside the concrete produce hyperbolic patterns. The reflection signals from the cracks are barely visible in the raw B-Scan image for being masked by the strong clutter.

After the pre-processing, the concrete surface reflection signal is removed, and the pre-processed GPR B-Scan image for each test case is shown as Figure 2(c), Figure 3(c) and Figure 4(c) respectively, where the crack features are still barely visible. Using the tuning parameter λ = 1/√max(678, 156) ≈ 0.038, the sparse crack features can be extracted from the pre-processed B-Scan data matrix. The processed B-Scan image upon the RPCA for each test case is shown in Figure 2(d), Figure 3(d) and Figure 4(d) respectively, where the crack features are much pronounced.
Figure 2. Evenly distributed rebars in concrete with single crack: (a) Geometry data; (b) Raw B-scan image; (c) Pre-processed B-Scan image; (d) Extracted crack features.

Figure 3. Evenly distributed rebars in concrete with single crack filled with water: (a) Geometry data; (b) Raw B-scan image; (c) Pre-processed B-Scan image; (d) Extracted crack pattern.
Figure 4. Evenly distributed rebars in concrete with two cracks – one consists air and one consists water: (a) Geometry data; (b) Raw B-scan image; (c) Pre-processed B-Scan image; (d) Extracted crack features.

To quantitatively evaluate the performance of crack extraction method, signal-to-clutter ratio (SCR) is used as a metric for characterizing the power ratio between the backscattering signal from the cracks under test and the clutter. Let \( c_{t,n} \) be the correlated background clutter data at the time index \( t \) and scan axis \( n \), and \( s_{t,n} \) be the crack features data at the time index \( t \) and scan axis \( n \). The correlated background clutters are within the region \( R_c = \{(t,n) | t \in (t_1,t_2), n \in (n_1,n_2)\} \), and the crack features are within the region \( R_s = \{(t,n) | t \in (t_3,t_4), n \in (n_1,n_2)\} \). The SCR is calculated as

\[
SCR = 10 \log_{10} \left( \frac{\sum_{(t,n) \in R_s} \| s_{t,n} \|^2}{\sum_{(t,n) \in R_c} \| c_{t,n} \|^2} \right)
\]

The SCR of both the Pre-processed B-Scan images and the correlated background removed B-Scan images are calculated and summarized in Table 1. The SCR improvement for each test case upon RPCA is also calculated and summarized in the last column of Table 1. The quantitative analysis results indicate that the proposed correlated background removal can dramatically enhance the crack features in GPR concrete B-Scan image.

Table 1. SCR of each B-Scan Image

<table>
<thead>
<tr>
<th>Rebar Distribution</th>
<th>Crack Configuration</th>
<th>SCR of Pre-Processed B-Scan (dB)</th>
<th>SCR of RPCA B-Scan (dB)</th>
<th>SCR Improvement (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evenly</td>
<td>Air</td>
<td>-2.47</td>
<td>16.46</td>
<td>18.93</td>
</tr>
<tr>
<td></td>
<td>Water</td>
<td>-1.84</td>
<td>28.29</td>
<td>30.13</td>
</tr>
<tr>
<td></td>
<td>One Air &amp; One Water</td>
<td>-7.74</td>
<td>29.29</td>
<td>37.03</td>
</tr>
<tr>
<td>Unevenly</td>
<td>Air</td>
<td>-3.67</td>
<td>22.58</td>
<td>26.25</td>
</tr>
<tr>
<td></td>
<td>Water</td>
<td>-1.36</td>
<td>29.41</td>
<td>30.77</td>
</tr>
<tr>
<td></td>
<td>One Air &amp; One Water</td>
<td>-8.89</td>
<td>29.30</td>
<td>38.19</td>
</tr>
</tbody>
</table>

4.2 Unevenly distributed rebars test

The geometry structure set up for producing the unevenly distributed rebars test data is similar to the setup in Sec. 4.1, where the only difference is that eleven steel reinforce bars of 1.25 cm radius are unevenly located inside the concrete. The test configuration for each test case is shown in Figure 5(a), Figure 6(a) and Figure 7(a) respectively.

The raw B-scan image for each test case is plotted as Figure 5(b), Figure 6(b) and Figure 7(b) respectively. In these three setups, the hyperbolas in B-Scan image are repeatedly but not periodically distributed. The pre-processed B-Scan image for each test case is shown in Figure 5(c), Figure 6(c) and Figure 7(c) respectively, where the crack features are barely visible.
For this test data processing, the tuning parameter in Eq. (4) is calculated as \( \lambda = 1 / \sqrt{\max(678, 156)} \approx 0.038 \), and the sparse crack features are extracted. The processed B-Scan image upon RPCA for each test case is depicted as Figure 5(d), Figure 6(d) and Figure 7(d) respectively, where the crack features clearly stand out.

Figure 5. Unevenly distributed rebars in concrete with single crack: (a) Geometry data; (b) Raw B-scan image; (c) Pre-processed B-Scan image; (d) Extracted crack features.
Figure 6. Unevenly distributed rebars in concrete with single crack filled with water: (a) Geometry data; (b) Raw B-scan image; (c) Pre-processed B-Scan image; (d) Extracted crack features.

Figure 7. Unevenly distributed rebars in concrete with two cracks – one consists of air and one consists of water: (a) Geometry data; (b) Raw B-scan image; (c) Pre-processed B-Scan image; (d) Extracted crack features.

The SCR of the B-Scan images before and after RPCA are all calculated and summarized in Table 1 for quantitative validation purpose. The analysis results demonstrate that the proposed crack features extraction method based on RPCA
technique can improve the SCR of B-Scan image significantly and is robust to the uneven distribution of the embedded rebars or pipes in the concrete.

5. Conclusions

In this paper, the crack features extraction from the correlated background in GPR concrete inspection using RPCA technique has been investigated. Experiments with the synthetic data indicate that the proposed method can automate the extraction of the crack features in concrete, without the need of prior knowledge about the test scenario. Also, the proposed approach is insensitive to the distribution of the correlated background matrix, which makes it effective for practical use.

References