2007

Subsurface Characterization Using Textural Features Extracted From GPR Data

R. S. Freeland
University of Tennessee, Knoxville

Lameck O. Odhiambo
University of Nebraska-Lincoln, lodhiamo2@unl.edu

Follow this and additional works at: https://digitalcommons.unl.edu/biosysengfacpub

Part of the Bioresource and Agricultural Engineering Commons, Environmental Engineering Commons, and the Other Civil and Environmental Engineering Commons
SUBSURFACE CHARACTERIZATION USING TEXTURAL FEATURES EXTRACTED FROM GPR DATA

R. S. Freeland, L. O. Odhiambo

ABSTRACT. Subsurface conditions can be non-intrusively mapped by observing and grouping patterns of similarity within ground-penetrating radar (GPR) profiles. We have observed that the intricate and often visually indiscernible textural variability found within a complex GPR image possesses important parameters that help delineate regions of similar subsurface characteristics. In this study, we therefore examined the feasibility of using textural features extracted from GPR data to automate subsurface characterization. The textural features were matched to a “fingerprint” database of previous subsurface classifications of GPR textural features and the corresponding physical probing of subsurface conditions. Four textural features (energy, contrast, entropy, and homogeneity) were selected as inputs into a neural-network classifier. This classifier was tested and verified using GPR data obtained from two distinctly different field sites. The first data set contained columnar patterns that indicate the presence or lack of sandstone bedrock in the upper 2 m of a shallow soil profile of fine sandy loam and loam. The second data set contained columnar patterns that correspond to the presence or the lack of vertical preferential flow paths within a deep loessial soil. The classifier automatically grouped each data set into one of the two categories.

Comparing the results of classification using extracted textural features to the results obtained by visual interpretation found 93.6% of the sections that lack sandstone bedrock correctly classified in the first set of data, and 90% of the sections that contain pronounced columnar patterns correctly classified in the second set of data. The classified profile sections were mapped using integrated GPR and GPS data to show ground surface boundaries of different subsurface conditions. These results indicate that textural features extracted from GPR data can be utilized as inputs in a neural network classifier to rapidly characterize and map the subsurface into categories associated with known conditions with acceptable levels of accuracy. This approach of GPR imagery classification is to be considered as an alternative method to traditional human interpretation only in the classification of voluminous data sets, wherein the extensive time requirement would make the traditional human interpretation impractical.

Keywords. Ground-penetrating radar, Neural network, Subsurface characterization, Textural features.
imagery, and found that F-NN can supply accurate soil clustering and classification based on both the arrangement and properties of individual soil horizons.

The need for an automated classification system for GPR data becomes apparent whenever one attempts visual interpretation, as GPR data sets collected during a routine field-scale survey are massive. Visual interpretation of GPR data involves subjective judgment, is laborious, may miss important features, and has no real-time application. These difficulties associated with visual interpretation often limit the use of GPR as a practical, widespread tool for subsurface investigations. A technique that provides automatic characterization of vast quantities of GPR data to classify the subsurface into categories associated with known subsurface conditions would greatly enhance the usefulness of GPR for environmental management, not only by saving time, but also by reducing the probability of misclassification from human error.

OBJECTIVES

In this study, four textural features (energy, contrast, entropy, and homogeneity) based on a co-occurrence matrix, were extracted and used as inputs to a neural network classifier. The classifier was used to partition subsurface profile regions into categories, and the results were matched to a database of previous subsurface classification that relates textural parameters to known subsurface characteristics. We examined the applicability of the procedure to automate subsurface characterization and surface mapping of the ground into categories associated with known subsurface conditions. The method was tested and verified using GPR data sets from two sites.

METHODS AND MATERIALS

DATA COLLECTION

The data used in this study were collected at two sites using a GSSI Subsurface Interface Radar (SIR) System 10-A and a 200 MHz antenna (model 3105, Geophysical Survey Systems, Inc., New Salem, N.H.). This system measures the time that it takes electromagnetic energy to travel from the antenna to an interface and back. The control settings used on the SIR 10-A unit are as shown in table 1. The first site is located at the University of Tennessee Agricultural Experiment Station (Plateau Experiment Station), near Crossville, Tennessee. The soils at this site are fine sandy loam and loam, and are underlain by sandstone bedrock in the upper 2 m of the soil profile. The second site is located at the Ames Plantation near Grand Junction, Tennessee. The soils at this site consist of loess overlying alluvium deposits underlain by tertiary-aged sand deposits. This site was specially prepared for a study of the preferential flow paths by applying water to fill a large ring infiltrometer constructed at the center of the site, and taking GPR surveys in a spiral path around the infiltrometer at 15 min intervals.

FEATURE EXTRACTION

Ground-penetrating radar data sets are typically very large and contain a lot of information that is redundant and superfluous for subsurface characterization. The purpose of feature extraction is to reduce the dimensionality of the data and convert it to variables that are more suitable for discrimination between subsurface categories. The GPR data are displayed as a two-dimensional array of numbers, where each value in the array represents the reflective intensity of multivariate soil properties in the soil profile. The vertical direction of such a display is time, which can be converted to depth once the signal velocities are known, and the horizontal direction is linear distance on the ground surface. The reflective intensities are represented in the data by values that range from 0 to 65535, where 0 and 65535 represent the maximum limits of reflection, and the value 32768 represents no reflection, as shown in figure 1.

We have observed that the intricate and often indiscernible textural variability found within a complex GPR image possesses important parameters that help delineate regions of similar subsurface characteristics. Several methods have been used to extract textural features from digital images for use in image classification. Haralick et al. (1973) developed a conceptual framework of measures from which textural features are derived. The framework is based on co-occurrence matrices, which define the spatial relationship of pairs of pixels values. The co-occurrence matrix of a GPR data set, \( P(i, j, d, \theta) \), is the frequency of occurrence in the data set of pairs of reflective intensity levels \((i\) and \(j)\) that are separated by a certain distance \((d)\) and lie along a certain direction \((\angle \theta)\). When the GPR data set is read through a classifier window as it is passed along the linear distance of the GPR display, the unnormalized frequencies for angles quantized to 45° intervals for each window are expressed as follows:

![Figure 1. Typical single waveform showing the maximum limits of radar wave reflections.](image-url)
where \( \# \) denotes the number of elements in the set; \( M \times N \) is the size of the classifier window; \( i,j \) is the number of possible of reflective intensity levels (1 to 256); \( k,m \) is the image width (1 to \( M \)); and \( l,n \) is the image height (1 to \( N \)). The frequencies of occurrence are inherently not invariant under rotations. To alleviate these directional biases, the frequencies were summed as follows:

\[
P_{ij} = P_{i,j,0^0} + P_{i,j,45^0} + P_{i,j,90^0} + P_{i,j,135^0}
\]

Haralick et al. (1973) proposed 14 measures of textural features, which are derived from the co-occurrence matrices, and each represents certain image properties such as coarseness, contrast, homogeneity, and texture complexity. For this study, four commonly used textural features (eqs. 6 through 9) were extracted and used as inputs to the neural network classifier.

**Energy:** Measures texture uniformity or pixel repetition. A high value occurs when the gray level distribution is constant:

\[
f_1 = \sum_{ij} \left( \frac{P_{ij}}{R} \right)^2
\]

**Entropy:** Measures the disorder of an image, and is high when an image is not uniform:

\[
f_2 = \sum_{ij} \left( \frac{P_{ij}}{R} \right) \log \left( \frac{P_{ij}}{R} \right)
\]

**Homogeneity:** Decreases as individual pixel values differ more from their mean:

\[
f_3 = \frac{\sum P_{ij}}{\sum \left| i - j \right|}
\]
The database items were organized into two tables from which data can be accessed and reassembled to determine subsurface categories. Table 2 contains the textural parameters extracted from the two GPR data sets and the assigned classes. Each row contains a unique instance of data for the categories defined by parameters in the columns. The textural parameters \( f_1, f_2, f_3, \) and \( f_4 \) in table 2 are defined in equations 6 through 9. Table 3 contains physical probing information on subsurface condition at the sites and the assigned categories. The database was implemented using a MatLab program developed for extracting textural features from GPR data, characterizing subsurface profile strips using the neural network classifier and a relational database, and mapping using an integrated GPR and GPS data sets to show surface boundaries of different subsurface categories. The method was tested and verified using GPR data collected from the study sites at the Plateau Experiment Station and the Ames Plantation. There is a vast assortment of GPR filters and transforms developed and adapted specifically for GPR image analysis, and are available in GPR processing software packages.

### Table 2. Extracted textural parameters and the assigned classes.

<table>
<thead>
<tr>
<th>Class</th>
<th>Textural Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( f_1 )</td>
</tr>
<tr>
<td>1</td>
<td>0.8675</td>
</tr>
<tr>
<td>2[a]</td>
<td>0.1574</td>
</tr>
<tr>
<td></td>
<td>0.0799</td>
</tr>
<tr>
<td></td>
<td>0.0491</td>
</tr>
<tr>
<td></td>
<td>0.0273</td>
</tr>
<tr>
<td>3</td>
<td>0.0239</td>
</tr>
<tr>
<td>4</td>
<td>0.0949</td>
</tr>
</tbody>
</table>

[a] Class 2 is not a unique category. It represents a conglomeration of several different bedrock depths, thicknesses, and states, i.e., solid and/or fractured.

### Table 3. Physical probing information on subsurface conditions and their assigned categories.

<table>
<thead>
<tr>
<th>Subsurface Conditions</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>Sandstone bedrock absent</td>
</tr>
<tr>
<td></td>
<td>Sandstone bedrock present</td>
</tr>
<tr>
<td>Site 2</td>
<td>Preferential flow paths</td>
</tr>
<tr>
<td></td>
<td>No preferential flow path</td>
</tr>
</tbody>
</table>

### Implementation and Results

The crossover parameter \( \psi \) was manually adjusted through iteration for each site based on texture type and the desired number of output classes. The database of the two study sites consisted of GPR images having different types of texture, and therefore, a common \( \psi \) value was rather difficult to find. At the Plateau Experiment Station, the data show
underlying bedrock, which contains features that are associated with three known subsurface conditions: solid bedrock, fractured bedrock, and no bedrock. The data for this site were divided into 604 profile strips (510 pixel depth × 100 pixel width), which were classified to identify areas with no bedrock in the 2 m depth from the rest of the area using a $q$ value of 1.2. The results are shown in figure 4. Figure 4a shows how the program separated the subsurface profile into two categories (areas without bedrock, and areas with bedrock). Out of the 604 profile strips, 47 were classified as having no bedrock in the 2 m depth. Careful visual interpretation of the data found 44 profile strips indicating the absence of bedrock in the 2 m depth. Comparing the results of classification using extracted textural features to visual interpretation found 93.2% of the profile strips lacking sandstone bedrock correctly classified and 6.8% misclassified. Figure 4b shows how markers on the GPR data were assigned to the different categories, and figure 4c is a surface map showing areas without and with sandstone bedrock in the 2 m depth, which matched physical probing to the top of the bedrock using a hand tile probe and GPR imagery from a previous study (Freeland et al., 2002b).

At the Ames Plantation site, sections of the GPR data show pronounced columnar patterns occurring in and around the alluvium/tertiary sand interface. These columnar patterns have been associated with vertical preferential flow paths (Freeland et al., 2002a). The data for this site were divided into 305 subsurface profile strips (510 pixel depth × 100 pixel width), which were classified to identify areas with pronounced columnar patterns from the rest of the area using a $q$ value of 0.5. The results are shown in figure 5. Figure 5a shows how the program separated the subsurface profile into two categories: areas that exhibited pronounced columnar patterns, and areas that exhibited few or no columnar patterns. Out of the 305 profile strips, 126 were classified as exhibiting pronounced columnar patterns, while careful visual interpretation identified 140 as exhibiting pronounced columnar patterns. Comparing the results of classification using extracted textural features to visual interpretation found 90.0% of the profile strips having pronounced columnar patterns correctly classified and 10.0% misclassified. Figure 5b shows how markers on the GPR data were assigned to the different categories, and figure 5c is a surface map showing areas with pronounced columnar patterns, and areas with few or no columnar patterns.
CONCLUSIONS

The results of this study indicate that textural features extracted from GPR data can be used to automate subsurface characterization. The method was demonstrated using fairly simple GPR data obtained from two distinctly different field sites. The subsurface conditions were determined by matching the extracted textural features to “fingerprints” in a relational database of previous subsurface classifications of GPR textural features and the corresponding physical probing of subsurface conditions. Only four textural parameters were used in this study, but the effects of additional textural parameters on the accuracy of prediction could be investigated in future studies. The relational database can be expanded when new data on subsurface categories become available. A neural network classifier was used to assign data to the known subsurface categories. The $\varphi$ values are optimized based on the texture, so no single $\varphi$ value is applicable to all types of GPR data. This implies that for each subsurface category, a $\varphi$ value must be defined. The results of subsurface characterization using extracted textural features was found to be in close agreement with results obtained by careful visual interpretation of the data (93.6% correct classified for site 1, and 90% correct classified for site 2). The classified subsurface profile sections were mapped using integrated GPR and GIS data to show surface boundaries of different subsurface categories.

This approach of GPR imagery classification is to be considered as an alternative method to traditional human interpretation only in the classification of voluminous data sets, wherein the extensive time requirement would make the traditional human interpretation impractical.

ACKNOWLEDGEMENTS

We thank the Tennessee Agricultural Experimental Station for financial and technical support. We gratefully acknowledge support from the Trustees of the Hobart Ames Foundation, Ames Plantation, Grand Junction, Tennessee.

REFERENCES

