Automatic Classification of Ground-Penetrating-Radar Signals for Railway-Ballast Assessment

Wenbin Shao, Abdesselam Bouzerdoum, *Senior Member, IEEE*, Son Lam Phung, *Member, IEEE*, Lijun Su, Buddhima Indraratna, and Cholachat Rujikiatkamjorn

***Abstract*—The ground-penetrating radar (GPR) has been widely used in many applications. However, the processing and interpretation of the acquired signals remain challenging tasks since an experienced user is required to manage the entire operation. In this paper, we present an automatic classification system to assess railway-ballast conditions. It is based on the extraction of magnitude spectra at salient frequencies and their classification using support vector machines. The system is evaluated on real-world railway GPR data. The experimental results show that the proposed method efficiently represents the GPR signal using a small number of coefficients and achieves a high classification rate when distinguishing GPR signals reflected by ballasts of different conditions.**

***Index Terms*—Ground-penetrating radar (GPR) processing, railway-ballast assessment, support vector machine (SVM).**

I. Introduction

THE GROUND-PENETRATING radar (GPR), sometimes called subsurface radar, ground probing radar, georadar, or earth sounding radar, exploits electromagnetic fields to probe lossy dielectric materials [1]–[4]. It can nondestructively detect buried objects beneath the shallow earth surface (less than 50 m) or in a visually impenetrable structure, such as walls and concrete floors. GPR has attracted considerable interest in many areas, such as archaeology [5], road construction [6], glacier and ice sheet investigation [7], and mineral exploration and resource evaluation [8].

As a cost-effective and environment-friendly means of transportation, railways play an important role in daily life. A railway structure typically consists of steel rails, fastening system, sleepers, ballast, subballast, and subgrade [9]. The transverse section of a railway is shown in Fig. 1. The ballast is an essential component for proper railway functioning. To ensure safety, regular inspection of rail tracks must be conducted. Traditionally, track investigation involves drilling to collect ballast samples from the railway sites. The ballast samples are then sent to a laboratory for assessment, which involves fouling index measurement. Finally, maintenance actions are determined based on the evaluation results. The entire procedure is labor intensive and time consuming. Thus, the rail industry is searching for new and more cost-effective approaches. As a nondestructive detection tool, GPR has attracted great interest in railway-ballast evaluation in recent years [10].

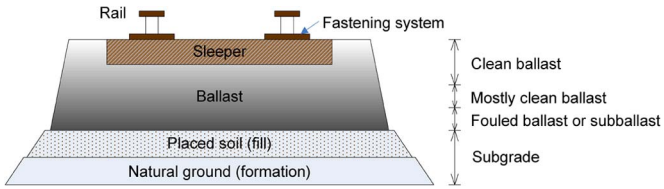


Fig. 1. Railway structure [9], [11].

Despite its commercial success, GPR still faces various fundamental problems. Specifically, processing and interpreting radar profiles are still challenging tasks [12], [13]. In addition to traditional GPR processing techniques, such as dewow and filtering, researchers have employed various signal processing techniques to aid the GPR signal analysis and interpretation [2], [13], [14]. For example, Al-Qadi *et al.* [9] proposed a time-frequency approach to evaluate GPR data for railway-ballast assessment. Their approach utilizes the short-time Fourier transform (STFT). Sinha *et al.* [15] presented a new method for time-frequency map computation for nonstationary signals. Their approach utilizes the continuous wavelet transform (CWT). Experiments on seismic data show that the CWT approach can be used to detect frequency shadows and subtle stratigraphic features. Fujimoto and Nonami [16] suggested a mine detection algorithm based on statistical features, such as Student's t distribution and chi-square distribution. Their algorithm was shown to improve the probability of detection and decrease the probability of false alarm. Zoubir *et al.* [14] compared a number of landmine detection techniques, such as Kalman filtering, background subtraction, matched filter deconvolution, wavelet packet decomposition, and trimmed average power. They evaluated the techniques using receiver operating characteristic curves and computation time. The Kalman filtering approach was found to outperform other methods on detection rate, but it has the highest computational cost. The aforementioned studies

Manuscript received July 21, 2010; revised December 30, 2010; accepted February 20, 2011. Date of publication April 28, 2011; date of current version September 28, 2011. This work was supported in part by a Grant from the Australian Research Council.

W. Shao, A. Bouzerdoum, S. L. Phung, B. Indraratna, and C. Rujikiatkamjorn are with the University of Wollongong, Wollongong, NSW 2522, Australia (e-mail: ws909@uowmail.edu.au; a.bouzerdoum@ieee.org; phung@uow.edu.au; indra@uow.edu.au; cholacha@uow.edu.au).

L. Su is with the School of Civil Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China, and also with the Cooperative Research Centre for Rail Innovation, Faculty of Engineering, University of Wollongong, Wollongong, NSW 2522, Australia (e-mail: lijun@uow.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2011.2128328

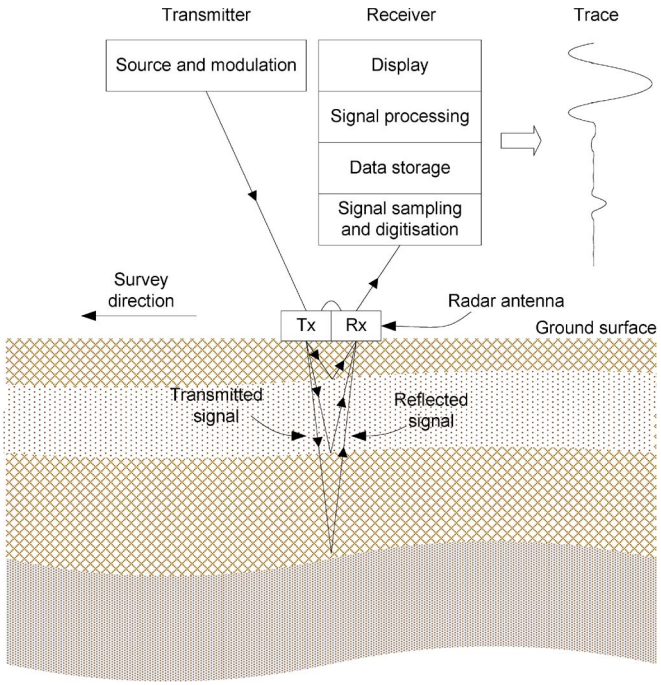


Fig. 2. GPR system components and GPR work process [12].

mainly focus on improving visualization and clarity of GPR signals, and human intervention is still required to interpret the processed signals, which may introduce subjectivity and user dependence into data analysis.

In a GPR survey, because particular resonance frequencies arise in wave propagation, reflected waves from different buried objects or paths present different electromagnetic characteristics. Hence, it is possible to classify the buried objects or underground materials by analyzing the frequency spectra of the received GPR signals. Motivated by this observation, we propose a GPR signal classification system based on magnitude spectrum and support vector machines (SVMs) for ballast fouling assessment. The proposed system is designed so that no human intervention is required. It can automatically extract and select features from GPR railway signals and classify the GPR traces.

The remainder of this paper is organized as follows. In Section II, the proposed classification system is introduced. In Section III, the experimental methods and system implementation are explained. The experimental results are presented in Section IV, followed by some concluding remarks in Section V.

II. Proposed Approach

In this section, we first give an overview of the GPR system and then present the proposed approach for ballast fouling classification.

A. GPR System Overview

Fig. 2 shows the components of a typical GPR system. It consists of a signal generator (transmitter), transmitting and receiving antennas, and a recording device (receiver) [12], [17]. To detect objects, the transmitter generates a pulse and

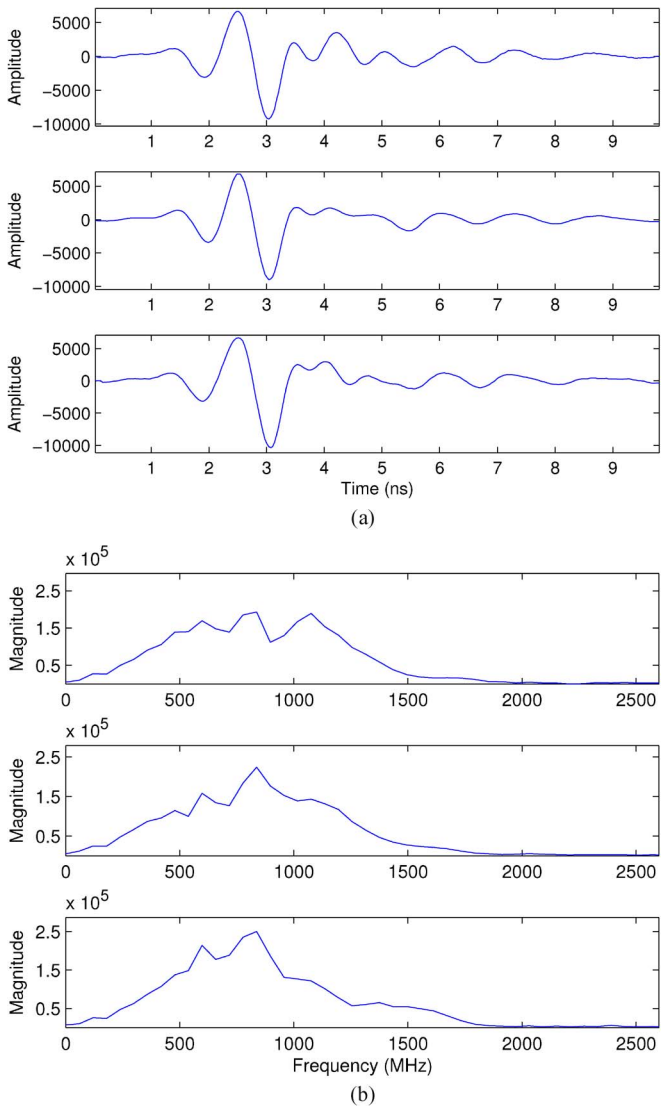


Fig. 3. Three traces from the railway data set. From top to bottom, they are from 50% clay ballast, clean ballast, and 50% coal ballast, respectively. (a) Time-domain waveforms. (b) Frequency magnitude spectra.

delivers it to the transmitting antenna T_x, which radiates an electromagnetic wave. Once the electromagnetic wave hits an object whose electrical properties are different from those of surrounding materials, part of the wave energy is reflected back toward the receiving antenna R_x. The detected energy is then sent to the receiver for storage and display.

Since the GPR device can be mounted on a train, it is possible to conduct a continuous survey without interruption. With GPS devices and signal processing techniques, maintenance decisions can be made on site. A challenging task is how to interpret the GPR signals and assess the ballast condition automatically.

B. GPR Trace Classification System

Because the frequency spectrum of the GPR return reveals the characteristics of the materials on the electromagnetic wave path, we propose to use frequency features to automatically categorize ballast fouling conditions. Three traces from different fouling ballasts are shown in Fig. 3, including

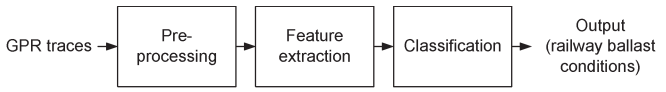


Fig. 4. Block diagram of the proposed automatic classification system.

their time-domain waveforms [Fig. 3(a)] and magnitude spectra [Fig. 3(b)]. It is observed that the traces from ballast of different fouling conditions have different magnitude spectra. For example, the peak in the magnitude spectrum of the 50% clay is lower than the other two. In the frequency range of 800–1200 MHz, the magnitude spectrum of the 50% coal decays more rapidly than that of the clean ballast.

The proposed automatic classification system includes three main stages: preprocessing, feature extraction, and classification. The system block diagram is shown in Fig. 4. When a GPR signal is received, salient features are extracted from it automatically and then sent to a pretrained classifier for assessment of the railway-ballast condition.

C. Preprocessing and Feature Extraction

The preprocessing stage employs basic signal processing techniques, including dc component removal, resampling, and time shifting, to reduce the intrinsic interferences introduced by the GPR and ensure the sampling rate consistency of the time-domain signals; depending on the system, samples located at the end of each trace may be discarded at this stage.

In the proposed system, feature extraction consists of three steps. First, the discrete Fourier transform (DFT) is applied to GPR signals to obtain the magnitude spectra, which are normalized to ensure consistency in magnitude spectrum amplitudes. Second, salient frequencies are determined based on the training data and user-defined parameters. Third, feature vectors are formed by extracting magnitudes of local maxima and arranging them in ascending order of frequencies.

In the first step, the DFT is applied to the time-domain trace. Let $s[n]$ be the discrete-time signal (real or complex) of length L obtained by sampling a continuous-time signal $s(t)$ with a uniform sampling rate f_s. The N-point DFT of $s[n]$ is defined as

$$S[k] = \sum_{n=0}^{N-1} s[n] e^{-j2\pi \frac{k}{N} n}, \qquad k = 0, 1, 2, \ldots, N-1 \quad (1)$$

where $N \geq L$. Note that the analog frequency corresponding to the kth DFT index $f(k)$ is given by

$$f(k) = \frac{k}{N} f_s, \qquad k = 0, 1, 2, \ldots, N-1. \quad (2)$$

In the second step, the salient frequencies are determined. To reduce the dependence on the antenna gain, the magnitude spectrum is normalized as follows:

$$P_k = \frac{|S[k]|}{\sum_{k=0}^{N-1} |S[k]| / N} \quad (3)$$

where $S[k]$ is the DFT coefficient computed in (1). Fig. 5 shows the normalized magnitude spectra of traces obtained

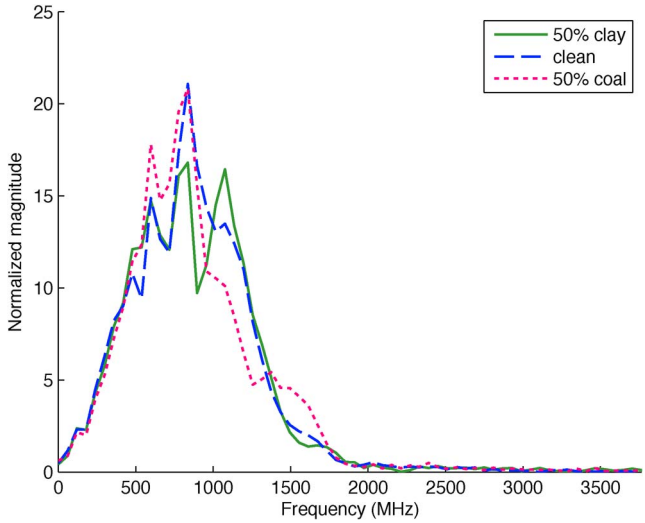


Fig. 5. Normalized magnitude spectra of three different traces obtained with 800-MHz antenna.

with an antenna frequency of 800 MHz. From this figure, it can be observed that the significant frequency components are below 2200 MHz, which is approximately three times the GPR antenna frequency. Similar observations can be made from the magnitude spectra of other GPR signals. The major frequency components of each trace reside mostly in the range $[0, 3f_a]$, where f_a is the antenna frequency. Therefore, the salient features of each trace can be extracted from this frequency range.

There are many frequencies that can be used in the range $[0, 3f_a]$. We choose the local maximum points within the specific frequency range as the salient frequencies. In our algorithm, the local maxima are located via the morphological operation *dilation*. Dilation is used because of its flexibility for local maxima search. Suppose that y is a 1-D discrete time signal and l is a flat structuring element, the dilation of y by l, denoted by $y \oplus l$, is defined as

$$[y \oplus l](x) = \max_{x' \in D_l} \{y(x - x')\} \quad (4)$$

where D_l is the domain of l and the structuring element is centered on x. Consequently, there are two adjustable parameters that determine the number of salient frequencies or the feature vector size: 1) the frequency distance between two adjacent local maxima and 2) the number of instances used to extract salient frequencies.

In the third step, the spectrum amplitudes at the selected frequencies are retrieved and arranged in ascending order of frequencies to form a feature vector. In preliminary experiments, another frequency range $[0, 2f_a]$ was considered for feature extraction; however, using the same parameters, the classification rate was reduced for the frequency range $[0, 2f_a]$ compared with the frequency range $[0, 3f_a]$. Thus, $3f_a$ was chosen as the frequency boundary. On average, about half of the extracted features are found in the range $[2f_a, 3f_a]$.

D. Classification Using SVMs

There are many methods available for pattern classification, such as discriminant analysis [18], decision trees [19],

k-nearest neighbors (k-NN) [18], Bayesian classifier [20], neural networks [21], and SVMs [22]. Here, we choose SVMs as the classification tool because they have been found to perform well in various practical applications [23]–[25]. SVMs are originally formulated for two-class classification problems. In SVMs, the decision boundary is obtained from the training data by finding a separating hyperplane that maximizes the margins between the two classes. This learning strategy is shown to increase the generalization capability of the classifier. We can apply SVMs to complex nonlinear problems by projecting the data onto a high-dimensional space using kernel methods.

Consider M training samples

$$\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_M, y_M)\}$$

where $\mathbf{x}_i \in R^n$ is a feature vector and $y_i \in \{1, -1\}$ is the class label. If the classes are linearly separable in the input space, the decision function can be written as

$$\begin{cases} \langle \mathbf{w}, \mathbf{x}_i \rangle + b \geq 1, & \text{for } y_i = 1 \\ \langle \mathbf{w}, \mathbf{x}_i \rangle + b \leq 1, & \text{for } y_i = -1 \end{cases} \tag{5}$$

or

$$y_i \left(\langle \mathbf{w}, \mathbf{x}_i \rangle + b \right) \geq 1 \tag{6}$$

where $\mathbf{w}$ is the vector normal to the hyperplane, b is a bias term, and $\langle \mathbf{w}, \mathbf{x} \rangle$ is the dot product of the vectors $\mathbf{w}$ and $\mathbf{x}$.

There are many hyperplanes that can separate the data [Fig. 6(a)]. However, only one hyperplane, called *optimal separating hyperplane*, can achieve *maximum margin* [represented with the solid line in Fig. 6(b)]. The margin perpendicular to the hyperplane can be expressed as $2/\|\mathbf{w}\|$. Consequently, the problem is to find $\mathbf{w}$ and b that maximize the margin. This is equivalent to minimizing

$$J(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 \tag{7}$$

subject to

$$y_i \left(\langle \mathbf{w}, \mathbf{x} \rangle + b \right) \geq 1, \qquad i = 1, \ldots, M. \tag{8}$$

If the classes are not separable, it is necessary to introduce nonnegative slack variables ξ_i into constraint (8)

$$y_i \left(\langle \mathbf{w}, \mathbf{x} \rangle + b \right) \geq 1 - \xi_i. \tag{9}$$

A classifier that generalizes well can be found by minimizing

$$\tau(\mathbf{w}, \boldsymbol{\xi}) = \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i=1}^{M} \xi_i \tag{10}$$

subject to

$$y_i \left(\langle \mathbf{w}, \mathbf{x} \rangle + b \right) \geq 1 - \xi_i, \qquad i = 1, \ldots, M \tag{11}$$

where C is a constant representing the tradeoff between margin maximization and training error minimization. This is a constrained optimization problem. By introducing nonnegative Lagrange multipliers α_i and β_i, the problem can be expressed as

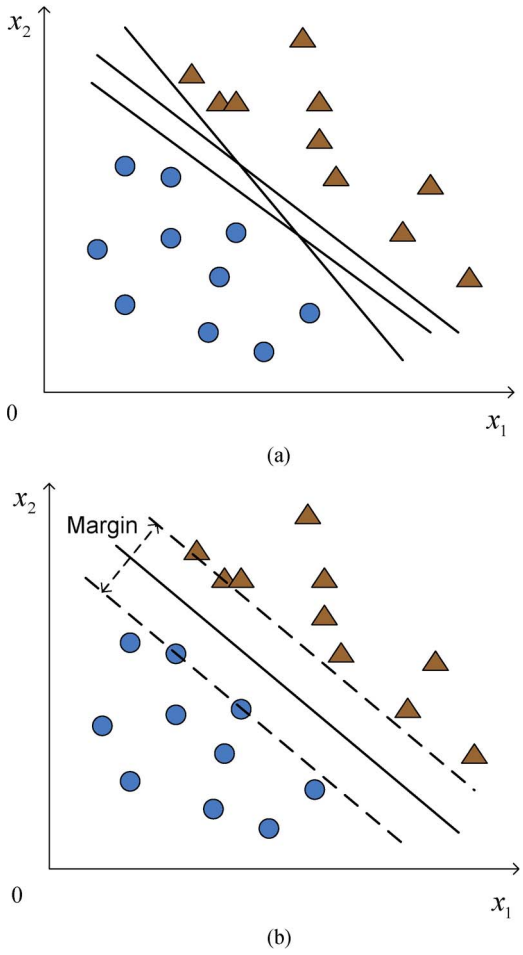


Fig. 6. SVM optimal hyperplane for a two-class problem. (a) Data can be separated by many hyperplanes. (b) Only one hyperplane achieves the maximum separation.

$$\begin{aligned} \min_{\mathbf{w},b} \max_{\boldsymbol{\alpha},\boldsymbol{\beta}} L(\mathbf{w}, b, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = & \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i=1}^{M} \xi_i \\ & - \sum_{i=1}^{M} \alpha_i \left[y_i \left(\langle \mathbf{w}, \mathbf{x}_i \rangle + b \right) - 1 + \xi_i \right] \\ & - \sum_{i=1}^{M} \beta_i \xi_i. \end{aligned} \tag{12}$$

The optimal solution should satisfy the following Karush–Kuhn–Tucker conditions [23], [26]

$$\frac{\partial}{\partial \mathbf{w}} L(\mathbf{w}, b, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = 0 \tag{13}$$

$$\frac{\partial}{\partial b} L(\mathbf{w}, b, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = 0 \tag{14}$$

$$\frac{\partial}{\partial \boldsymbol{\xi}} L(\mathbf{w}, b, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = 0 \tag{15}$$

$$\alpha_i \left[y_i \left(\langle \mathbf{w} \ldots \mathbf{x}_i \rangle + b \right) - 1 + \xi_i \right] = 0, \qquad i = 1, \ldots, M \tag{16}$$

$$\beta_i \xi_i = 0, \qquad i = 1, \ldots, M \tag{17}$$

$$\alpha_i \geq 0, \ \beta_i \geq 0, \xi_i \geq 0, \qquad i = 1, \ldots, M. \tag{18}$$

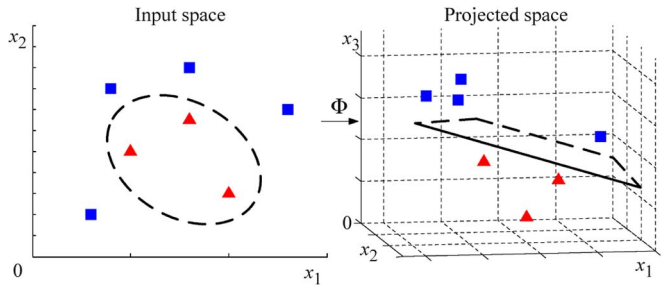


Fig. 7. By mapping data from the input space to a higher dimensional space via Φ, it is possible to find a nonlinear decision boundary in the original input space.

Equations (13)–(18) lead to

$$\mathbf{w} = \sum_{i=1}^{M} \alpha_i y_i \mathbf{x}_i \tag{19}$$

$$\sum_{i=1}^{M} \alpha_i y_i = 0 \tag{20}$$

$$\alpha_i + \beta_i = C. \tag{21}$$

Substituting (19)–(21) into (12), the primal variables $\mathbf{w}$ and b can be eliminated and a dual optimization problem is obtained

$$\text{maximizing } Q(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} \alpha_i \alpha_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \tag{22}$$

subject to

$$0 \le \alpha_i \le C, \qquad i = 1, \ldots, M \tag{23}$$

$$\sum_{i=1}^{M} \alpha_i y_i = 0, \qquad i = 1, \ldots, M. \tag{24}$$

In real-world applications, classes are usually not linearly separable in the input space and the classifiers obtained in the original input space may not have high generalization ability for unknown data. Therefore, the data samples from the input space are usually projected onto a higher dimensional dot product space via a mapping function Φ. The linear decision boundary constructed in the projected space yields a nonlinear decision boundary in the input space (see Fig. 7).

However, the projection is usually computation intensive. To simplify the projection, a positive semidefinite kernel H is employed

$$H(\mathbf{x}, \mathbf{x}') = \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}') \rangle . \tag{25}$$

Using the kernel, the dual problem in (22) is expressed as

$$\text{maximizing } Q(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} \alpha_i \alpha_j y_i y_j H(\mathbf{x}_i, \mathbf{x}_j) \tag{26}$$

subject to the constraints in (23) and (24).

Compared with several other kernels (linear and polynomial), the radial basis function (RBF) kernel has been chosen because it performs nonlinear mapping and has less hyperparameters than the polynomial kernel; it is given by

$$H(\mathbf{x}, \mathbf{x}') = e^{-\gamma \|\mathbf{x} - \mathbf{x}'\|^2} \tag{27}$$

(a)

(b)

Fig. 8. Experimental setup. (a) Existing railway track used for GPR data collection. (b) GPR data collection system.

where γ is a positive scalar. In this paper, we focus on SVMs with RBF kernels.

III. Experimental Methods

The GPR operation along the railway can be affected by many factors, such as cross winds due to high speed rail, high electromagnetic interference, and radio frequency interference from railway communications and automation, geomagnetic storms, and thunderstorms [27]. To collect real-world data for system evaluation, we conducted GPR surveys along an existing railway track at Wollongong station in New South Wales, Australia. We have collected 25 920 GPR traces, of which 5 896 are with known ground truth. In this section, first we introduce the railway track and experimental setup; then, we explain the implementation of the proposed system.

A. Railway-Ballast Data Collection Using GPR

Fig. 8 shows the data collection equipment and the railway track where the experiments were conducted. This track is parallel to several tracks that are in service. Considering the time and cost, three sections with known ground truth of the

TABLE I
RADAR PARAMETER CONFIGURATIONS FOR 800-MHz ANTENNA USED IN THE SURVEYS

Antenna height (mm)	Sampling frequency (MHz)
200	16477
	20401
	25201
	30601
300	16477
	20401
	25201
	30601
400	20401

railway track were chosen for the GPR data. Each section has a length of 2.0 m and a depth of 0.55 m; the width is equivalent to the existing ballast width. We excavated the long-standing ballast from these sections and then filled them with different types of ballast that were premixed. Each section contained only one type of ballast. The sleepers were not reinstalled, and the rails remained untouched.

The three ballast types are chosen based on the most common ballast fouling conditions: 1) 50% clay fouling; 2) clean; and 3) 50% coal fouling. Here, the percentage of fouling is a *relative ballast fouling ratio*; it represents the proportion of fouling particles to ballast particles [28]. Compared with the traditional *fouling index* and *percentage void contamination*, the relative ballast fouling ratio can reveal the effect introduced by specific gravity and gradation of fouling materials.

Before GPR surveys are conducted, a proper GPR system must be chosen based on the survey environment, budget, and GPR system availability. To obtain quality GPR profiles, several factors were considered in choosing the appropriate antenna system, such as the depth to the bottom of ballast, resolution for ballast gravel grain size and fouling, operation environment, and the antenna height to avoid trash, sensors, and switches [27]. Preliminary experiments were conducted to select the proper GPR antenna [29]. Different GPR systems from different companies were evaluated. Based on the results, the GPR system from MALÅ Geoscience was selected for data acquisition. The preliminary results also showed that the time–distance records from an 800-MHz antenna were clearer than those from a 1.2-GHz antenna. Therefore, our surveys mainly used the MALÅ 800-MHz antenna. The railway data collection system is shown in Fig. 8(b), and the parameters of the GPR used in the experiments are listed in Table I. Note that the bandwidth is approximately equal to the center frequency (antenna frequency).

The Wollongong railway data set consists of two parts: one collected under dry ground condition and the other gathered under wet condition.

1) The dry ground data samples were acquired during sunny weather conditions; the materials filled in the three sections were also dry. Two antennas of center frequencies 800 MHz and 1.2 GHz from MALÅ Geoscience were deployed, each at two different heights: 200 and 300 mm. The antenna elevations can prevent collision of the GPR with a variety of devices along the railway. Different GPR configuration parameters, including antenna height, time window, and sampling frequency, were utilized. Twenty-four GPR profiles were collected with the antenna frequency of 800 MHz and 12 profiles with 1.2 GHz. Each profile contains the GPR signals for an entire section (50% clay, clean, or 50% coal).
2) The wet ground data set was obtained under cloudy weather conditions; heavy rains from the previous night saturated the materials. Only the antenna of center frequency 800 MHz was used. All radar profiles shared the same GPR configuration parameters. The antenna height was lifted to 400 mm to avoid obstacles along the railway track.

TABLE II
NUMBERS OF AVAILABLE TRACES IN COMBINED 800-MHz DATA SET

Condition	Dry		Wet
Antenna height	200 mm	300 mm	400 mm
Section clay	469	470	745
Section clean	477	478	642
Section coal	436	438	705
Total	1382	1386	2092

A summary of the Wollongong railway data set using the 800-MHz antenna is presented in Table II. This data set, namely, the combined 800 MHz data set, can be divided into three subsets based on the antenna heights:

1) 200-mm data subset;
2) 300-mm data subset;
3) 400-mm data subset.

Each data subset consists of GPR traces from three different types of ballast. To reduce the border effects, the first and last 15% traces of each GPR profile were discarded.

B. System Implementation

In the preprocessing phase, an automatic dc offset is applied to each trace to obtain a zero-mean signal. Next, every GPR trace is resampled to ensure data consistency. Then, each trace is shifted according to the position of the global maximum point. The shifting reduces the effects of antenna height variations; a few samples may be discarded from the end of each trace, based on the minimum trace length after resampling.

For feature extraction, the fast Fourier transform algorithm is applied to obtain the amplitude spectra. After normalization, several traces are selected to find the feature points, i.e., the salient frequencies in the range $[0, 3f_a]$. The magnitude spectrum features are extracted at these points to form the feature vector, which is fed to the classifier. Consider the three example traces in Fig. 5, representing three different ballast types. Each trace has a length of 308 in the discrete time domain. The magnitude spectra of the tree traces and the salient frequencies are shown in Fig. 9. In the figure, each vertical dotted line indicates a frequency where a magnitude feature is extracted. There are 17 feature points in this example; hence, each trace is represented by a feature vector of size 17.

To train and test the SVM classifiers, the *LIBSVM* tool, developed by Chang *et al.* [30], was used. When building SVM classifiers, an exhaustive search for optimal SVM training parameters is computation intensive. Thus, a hierarchical approach is applied in our system to reduce the computation cost.

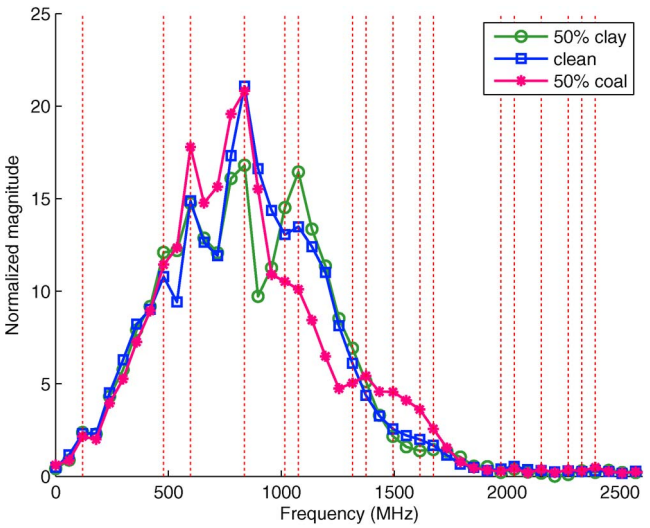


Fig. 9. Feature points of the three traces shown in Fig. 3(a). Each vertical dotted line represents a feature point.

First, the parameter search is performed on a coarse grid. Once a possible region containing the optimal parameters is identified, a finer search is applied within the identified region. Compared with the exhaustive search, the two-level hierarchy reduces the training time by half. During training, the system sometimes finds more than one set of optimal parameters. To solve this, we simply construct a number of classifiers using the chosen parameters and form them as an SVM pool [31]. Whenever a test sample is input into the system, it will be evaluated by every SVM classifier in the pool; a majority voting strategy is then applied to obtain the overall classification result.

SVMs utilize explicit decision functions and are formulated for two-class problems. It is necessary to extend the SVM formulation to handle multiclass problems. There are several ways to extend SVMs; one-versus-all and pairwise are two common approaches. In this paper, we focus on the one-versus-all approach and give results of the pairwise SVM approach only for comparison purposes.

1) In the one-versus-all approach, a k-class problem is decomposed into k two-class problems [23], [32]. Each SVM is trained with all the training samples. For the ith SVM, where $i \leq k$, samples in the ith class are labeled as positive, and samples in all other classes are labeled as negative. Note that the classifier parameters that yield high generalization are automatically selected using fivefold cross-validation on the training set.
2) The pairwise approach requires $k(k-1)/2$ two-class SVM classifiers to solve a k-class problem. Each SVM classifier is trained with samples from two classes. Let c_{ij} be the SVM classifier that is trained on data from the ith and jth classes. In the test phase, the SVM classifier $c_{ij}(i < j)$ divides all the data into class i and class j. The final classification results of pairwise SVMs are obtained by combining all two-class classifiers with a majority voting scheme. For an input instance $\mathbf{x}$, if a pairwise SVM classifier categorizes $\mathbf{x}$ in the kth class, then the vote for class k is increased by one. Once all classifiers have voted, the pattern $\mathbf{x}$ is assigned to the class that has the highest voting score.

To evaluate the generalization ability of the classifiers, *cross-validation* is used. There are several methods of cross-validation; in the proposed system, we employ fivefold cross-validation. The entire data set is randomly divided into five partitions of approximately equal sizes. Four partitions are used to train, and the remaining partition is used to validate the classifier. The step is repeated five times until all partitions have been evaluated. Finally, the average classification rate across five folds is computed and used to measure the system performance.

IV. Results and Analysis

The proposed system is used to classify ballast fouling conditions. In Section IV-A, we present the experimental results using different numbers of salient frequencies with one-versus-all SVMs trained and tested on the entire 800-MHz data set. In Section IV-B, we present the experimental results using the three data subsets. In Section IV-C, we show the system performance on the 1.2-GHz data. In Section IV-E, we compare the one-versus-all SVMs with pairwise SVMs and discuss the advantages and disadvantages of these two multiclass SVM approaches. Then, the system is compared with the k-NN algorithm and the Mahalanobis distance classifier using the proposed magnitude feature. A comparison is also made between the proposed feature extraction method and the STFT spectrogram.

A. Classification Performance on the Combined Data Set

In the first experiment, the proposed classification system is trained and tested on the combined data subsets collected with the 800-MHz antenna at different heights: 200, 300, and 400 mm. The proposed feature extraction approach searches for local maximum points in the magnitude spectra; these points determine the corresponding salient frequencies. Our experiments show that it is not necessary to use all local maxima for classification. Thus, in the following, we analyze how the number of salient frequency points affects the system performance. Note that the number of frequency points is equivalent to the feature vector size.

There are two parameters that control the number of prominent frequencies: the distance between peaks and the number of traces used. In system evaluation, these two factors are both varied from 3 to 18. If there exist more than one pair of parameters that bear the same number of salient frequencies, the median classification rate is reported. The classification rate is the percentage of test samples that are correctly classified.

The classification performance on the combined 800-MHz data set using fivefold cross-validation is shown in Table III. The proposed system can achieve a classification rate of 99.5% with 7 salient frequencies and 99.7% with 14 frequencies.

TABLE III
CLASSIFICATION RATES FOR DIFFERENT NUMBERS OF SALIENT FREQUENCIES ON THE COMBINED 800-MHz DATA SET

Number of salient frequencies	7	8	10	11	14	20	24
Overall classification rate (%)	99.5	99.6	99.6	99.8	99.7	100.0	100.0
Number of salient frequencies	29	30	31	32	33	34	-
Overall classification rate (%)	100.0	100.0	100.0	100.0	100.0	100.0	-

TABLE IV
CLASSIFICATION RATES FOR DIFFERENT NUMBERS OF SALIENT FREQUENCIES. DATA SET: $f_a = 800$ MHz; $h = 200$ mm

Number of salient frequencies	2	3	4	5	6	10	14	16
Overall classification rate (%)	70.1	77.7	78.4	90.4	93.3	97.1	99.1	99.5
Number of salient frequencies	17	18	19	20	22	23	24	25
Overall classification rate (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

TABLE V
CLASSIFICATION RATES FOR DIFFERENT NUMBERS OF SALIENT FREQUENCIES. DATA SET: $f_a = 800$ MHz; $h = 300$ mm

Number of salient frequencies	3	4	5	6	7	9	12
Overall classification rate (%)	84.4	84.6	90.5	96.1	96.6	97.5	99.8
Number of salient frequencies	13	21	23	24	26	27	28
Overall classification rate (%)	99.8	99.8	99.8	99.8	99.9	100.0	100.0

TABLE VI
CLASSIFICATION RATES FOR DIFFERENT NUMBERS OF SALIENT FREQUENCIES. DATA SET: $f_a = 800$ MHz; $h = 400$ mm

Number of salient frequencies	8	10	13	15	17	18	20	24	27
Overall classification rate (%)	99.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Number of salient frequencies	34	36	41	44	45	46	47	49	50
Overall classification rate (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

B. Classification Performance Versus Antenna Height

Further experiments have been conducted to explore the system performance on the three data subsets of different antenna heights. The three experiments using the 800-MHz data set are as follows:

1) training and testing on the 200-mm data subset;
2) training and testing on the 300-mm data subset;
3) training and testing on the 400-mm data subset.

Since the salient frequency points are determined from the training data, the feature vectors are different for each experiment.

1) The system classification performance on the 200-mm data subset as a function of the number of salient frequencies is given in Table IV. The system performance improves when more frequency points are used. When fewer than five frequency points are used, the classification rate is below 80.0%. When the number of frequency points reaches five, the classification rate increases to 90.4%. Once the feature size reaches 14, the system performance remains stable with a classification rate above 99.0%. Perfect classification is achieved with 17 frequencies or higher.
2) Table V shows the classification rates when the system is trained on the 300-mm data subset. The classification rate improves steadily with increasing number of salient frequencies. When the number of salient frequencies reaches 12, the system is able to classify the test set with a classification rate of 99.8%.
3) For 400-mm antenna height data, the system achieves an overall classification rate of 99.7% with only eight salient frequencies (see Table VI); the classification rate reaches 100.0% with ten features.

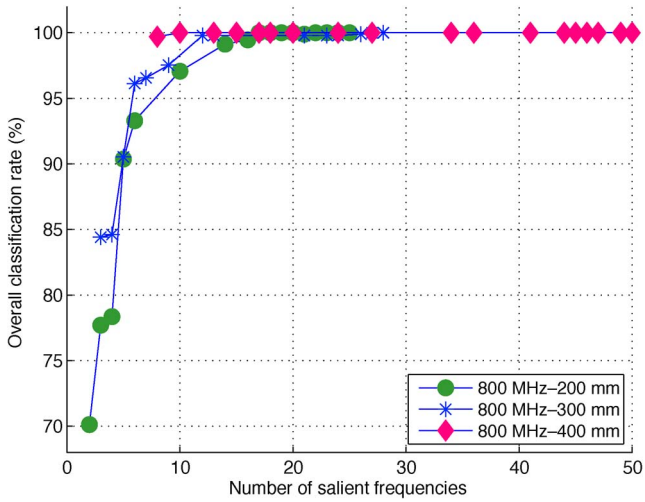


Fig. 10. Classification rates for different feature vector sizes and antenna heights.

The classification rates for the three data subsets are compared in Fig. 10. The experimental results show that the system performance varies with different numbers of salient frequencies; the classification rate tends to increase when more frequency points are used. When fewer salient frequency points are used, the system trained with 400-mm-antenna-height data performs better than the ones trained with 200- and 300-mm antenna height data. A possible explanation is that the 400-mm data were collected under a water-saturated condition. The higher dielectric permittivity of the water results in a stronger reflection than the dry ballast. Although the distance between peaks and the number of traces are both varied from 3 to 18 for each experiment, the system is able to detect more points in the 400-mm data subset. For example, using the same range of

TABLE VII
CLASSIFICATION RATES FOR DIFFERENT NUMBERS OF SALIENT FREQUENCIES. DATA SET: $f_a = 1.2$ GHz; $h = 200$ mm

Number of salient frequencies	4	7	8	9	10	11	12	15	16	19	21
Overall classification rate (%)	48.3	77.5	77.2	75.9	83.2	86.2	84.8	88.1	88.1	93.0	95.7
Number of salient frequencies	24	25	26	27	28	29	30	31	32	34	-
Overall classification rate (%)	95.4	96.7	94.1	97.1	94.4	97.6	99.4	99.4	98.9	99.4	-

parameters, more frequency points are extracted from the 400-mm data subset than from the other two subsets: 25 frequency points are extracted from the 200-mm data subset, 28 from the 300-mm subset, and 50 from the 400-mm subset. This can also be explained by the stronger reflection of the 400-mm data.

We also analyzed the system performance when it was trained on data collected with one antenna height and tested on data collected with another antenna height. The results show that the classification rate decreases. However, the system performed well when it was trained and tested on mixed data of different antenna heights (see Table III). This shows that the proposed system can operate at different antenna heights, provided that the training data set is representative.

C. Analysis of Operating Antenna Frequency

As mentioned in Section III-A, the MALÅ 1.2-GHz antenna was employed during the first survey. For comparison purposes, the classification performance for this antenna is shown in Table VII. When fewer than 16 frequency points are used, the classification rate is below 90.0%. When the feature vector size reaches 21, the classification rate reaches 95.7%. A classification rate of 99.0% requires a feature vector size of 30 or more. The results show that when a small number of frequency points are used, the classification rate for the 1.2-GHz data is lower than the classification rate for the 800-MHz data. For example, the 1.2-GHz system requires 19 salient frequencies to achieve a classification rate of 93.0%, whereas the 800-MHz 200-mm system needs only 6 salient frequencies to obtain a similar classification rate (see Table IV). In GPR, low antenna frequencies penetrate deeper than high frequencies, while high frequencies provide finer resolution than the low frequencies [33]. The choice of antenna frequency is a tradeoff between the required depth and resolution. In this case, the results indicate that the 1.2-GHz antenna is not as good as the 800-MHz antenna.

D. Analysis of SVM Design

This section compares the performances of one-versus-all and pairwise SVMs. With the one-versus-all SVM approach, if a sample is classified as positive by more than one classifier or negative by all classifiers, it will be labeled as unclassified. The unclassifiable regions of the one-versus-all approach are shown in Fig. 11. Pairwise SVMs, on the other hand, have a smaller unclassifiable area compared with one-versus-all SVMs [23]. When a new ballast class is added to the system, the one-versus-all approach requires retraining all the classifiers, while the pairwise approach involves training new classifiers between the added class and existing classes only.

Consider samples that do not carry sufficient resonances. The one-versus-all system will not classify these samples into the predefined classes (50% clay, clean, and 50% coal). However, the pairwise system will assign incorrect class labels to these samples. The overall classification rates of the two SVM systems on the combined 800-MHz data set are shown in Fig. 12. The performances of one-versus-all and pairwise SVMs are nearly the same.

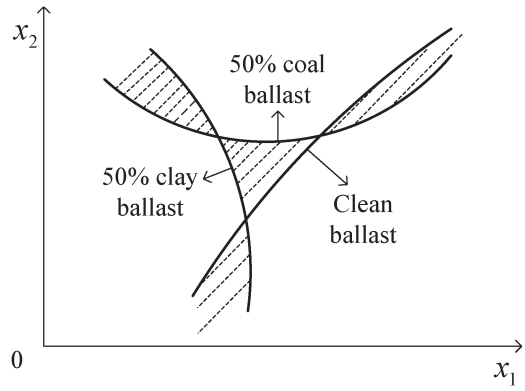


Fig. 11. Example of unclassifiable regions using one-versus-all SVMs. The solid lines are the class boundaries, and the shaded regions represent the unclassifiable areas.

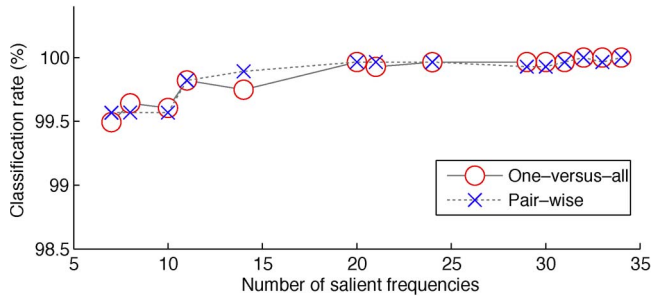


Fig. 12. Comparison of classification rates between one-versus-all and pairwise SVMs.

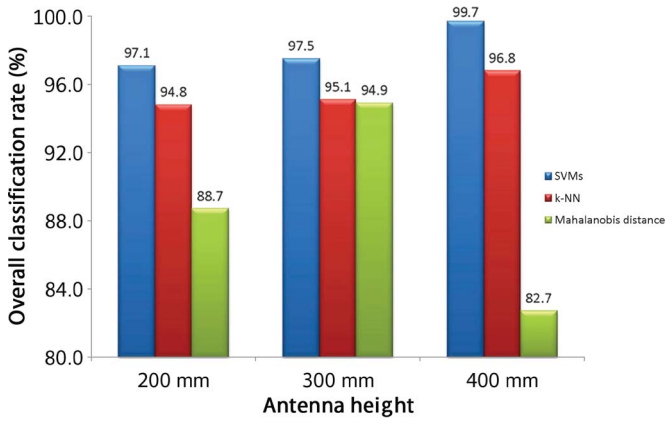


Fig. 13. Comparison of SVM, k-NN ($k = 15$), and Mahalanobis distance classifiers. Data set: $f = 800$ MHz.

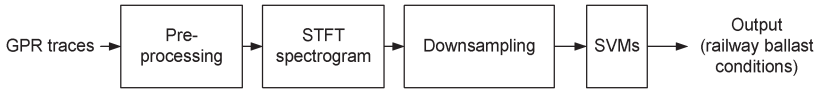


Fig. 14. Block diagram of the STFT spectrogram implementation.

E. Comparison With Other Approaches

In this section, we compare the performance of the proposed classification system with those of the k-NN and the Mahalanobis classifiers, using the same data set. We also compare the proposed magnitude spectrum features with features extracted from the STFT spectrogram.

1) Comparison With k-NN and Mahalanobis Distance Classifier: The k-NN classifier is a supervised learning algorithm based on sample distances [18]. It classifies a new sample by searching for the closest training samples. The label of the new sample is decided via a majority voting scheme based on the labels of the k nearest neighbors. In our implementation, k was varied from 1 to 17 in steps of 2.

The Mahalanobis distance is a statistical distance measure that takes into account correlation between variables. First, the mean $\mathbf{m}_i$ and the covariance matrix $\mathbf{C}_i$ of each class are computed from the training population. For an observation $\mathbf{x}$ to be classified, the Mahalanobis distance between $\mathbf{x}$ and each class is computed as follows:

$$D_i(\mathbf{x}, \mathbf{m}_i) = \sqrt{(\mathbf{x} - \mathbf{m}_i)\mathbf{C}_i^{-1}(\mathbf{x} - \mathbf{m}_i)^{\mathrm{T}}} \tag{28}$$

where i denotes the class index. The sample $\mathbf{x}$ is assigned to the class with the smallest Mahalanobis distance, i.e., the index of the winning class i^* is given by

$$i^* = \arg\min_i (D_i). \tag{29}$$

For comparison, fivefold cross-validation was applied. The numbers of frequencies for the three 800-MHz data subsets were 10, 9, and 8, respectively. Parameter k for the k-NN classifier was chosen based on the training data set.

The results are shown in Fig. 13. For the 200- and 400-mm antenna heights, the overall classification rates of the k-NN classifier are superior to those of the Mahalanobis distance classifier. With the 300-mm antenna height data, the k-NN classifier and the Mahalanobis distance classifier have close performance. For all the data subsets, the one-versus-all SVMs outperform both the k-NN and the Mahalanobis distance classifiers in terms of overall classification rate. For example, on the 300-mm data subset, the SVM classifier achieves a classification rate of 97.5%, while the k-NN and the Mahalanobis distance classifiers reach 95.1% and 94.9%, respectively.

2) Comparison With STFT Spectrogram: In [9], Al-Qadi *et al.* proposed a time-frequency approach using STFT. The energy attenuation of the STFT spectrogram is utilized to assess ballast conditions. However, their approach requires visual inspection. Here, we are interested in the classification performance of the STFT spectrogram features when used in the proposed system.

Our STFT spectrogram implementation is shown in Fig. 14. The GPR traces are preprocessed, and the discrete-time STFT is then applied to obtain the spectrogram. The discrete-time STFT is defined as

$$X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n} \tag{30}$$

where $X(m, \omega)$ is the STFT of windowed data, $x[n]$ is a GPR trace, and $w[n]$ is a window function. The spectrogram is represented by a 2-D matrix, whereas the SVMs accept a 1-D feature vector only. Therefore, the spectrogram is converted into a row vector. Furthermore, considering the computational complexity, we downsample the row vector to a feature vector of 16, 32, 64, 128, or 256 elements. Next, the extracted feature vectors are used as inputs to one-versus-all SVMs.

TABLE VIII
CLASSIFICATION RATES FOR STFT SPECTROGRAM FEATURE. DATA SET: COMBINED 800-MHz DATA SET

Feature vector size	16	32	64	128	256
Overall classification rate (%)	68.3	70.8	76.0	92.9	88.1

The results on the combined 800-MHz data set are shown in Table VIII. The STFT spectrogram requires 128 frequency points to achieve an overall classification rate of 92.9%, while the proposed magnitude spectra yield a classification rate of 99.5% using only seven frequency points.

V. CONCLUSION

Compared with the traditional approach, GPR provides a nondestructive and mobile means for fouling assessment of railway ballast. In this paper, we have presented an automatic classification system for GPR traces. The proposed system is based on magnitude spectrum analysis and SVMs; it automates the entire GPR signal processing and interpretation. Real-world railway data of three common ballast fouling conditions (clean ballast, 50% clay ballast, and 50% coal ballast) were collected to evaluate the proposed system. We have made the comparison between the proposed salient magnitude spectra and the STFT spectrogram and between SVMs and other two common classifiers. The experimental results indicate the following: 1) The proposed salient spectrum amplitudes are an efficient representation of GPR signals; 2) the system performs well in ballast fouling classification; for example, on the combined 800-MHz data set, the system can achieve a classification rate of 99.5% using 7 salient frequencies; and 3) the system can operate with different antenna heights, such as 200, 300, and 400 mm, provided that the training data set is representative of antenna height variations.

ACKNOWLEDGMENT

The railway GPR data were collected as part of the Rail CRC-AT5 project, sponsored by Cooperative Research Centre Rail for Innovation. We thank the anonymous reviewers for their constructive comments.

REFERENCES

[1] A. P. Annan, "GPR—History, trends, and future developments," *Subsurf. Sens. Technol. Appl.*, vol. 3, no. 4, pp. 253–270, Oct. 2002.

[2] D. J. Daniels, Ed., *Ground Penetrating Radar*, 2nd ed. London, U.K.: IEE, 2004.

[3] D. J. Daniels, D. J. Gunton, and H. F. Scott, "Introduction to subsurface radar," *Proc. Inst. Elect. Eng.—F, Radar Signal Process.*, vol. 135, no. 4, pp. 278–320, Aug. 1988.

[4] A. Neal, "Ground-penetrating radar and its use in sedimentology: Principles, problems and progress," *Earth-Sci. Rev.*, vol. 66, no. 3/4, pp. 261–330, Aug. 2004.

[5] M. Skolnik, Ed., *Radar Handbook*, 3rd ed. New York: McGraw-Hill, 2008.

[6] A. Denis, F. Huneau, S. Hoerl, and A. Salomon, "GPR data processing for fractures and flakes detection in sandstone," *J. Appl. Geophys.*, vol. 68, no. 2, pp. 282–288, Jun. 2009.

[7] J. J. Degenhardt, Jr., "Development of tongue-shaped and multilobate rock glaciers in alpine environments—Interpretations from ground penetrating radar surveys," *Geomorphology*, vol. 109, no. 3/4, pp. 94–107, Aug. 2009.

[8] J. Francke, "Applications of GPR in mineral resource evaluations," in *Proc. 13th Int. Conf. GPR*, 2010, pp. 1–5.

[9] I. L. Al-Qadi, W. Xie, and R. Roberts, "Time-frequency approach for ground penetrating radar data analysis to assess railroad ballast condition," *Res. Nondestruct. Eval.*, vol. 19, no. 4, pp. 219–237, Oct. 2008.

[10] J. P. Hyslip, S. S. Smith, G. R. Olhoeft, and E. T. Selig, "Assessment of railway track substructure condition using ground penetrating radar," in *Proc. Annu. Conf. AREMA*, Chicago, NY, 2003.

[11] A. Loizos and C. Plati, "Ground penetrating radar: A smart sensor for the evaluation of the railway trackbed," in *Proc. IEEE Instrum. Meas. Technol. Conf.*, 2007, pp. 1–6.

[12] J. M. Reynolds, *An Introduction to Applied and Environmental Geophysics*. New York: John Wiley, 1996.

[13] H. M. Jol, Ed., *Ground Penetrating Radar Theory and Applications*, 1st ed. Amsterdam, The Netherlands: Elsevier, 2009.

[14] A. M. Zoubir, I. J. Chant, C. L. Brown, B. Barkat, and C. Abeynayake, "Signal processing techniques for landmine detection using impulse ground penetrating radar," *IEEE Sens. J.*, vol. 2, no. 1, pp. 41–51, Feb. 2002.

[15] S. Sinha, P. S. Routh, P. D. Anno, and J. P. Castagna, "Spectral decomposition of seismic data with continuous-wavelet transform," *Geophysics*, vol. 70, no. 6, pp. 19–25, Nov./Dec. 2005.

[16] M. Fujimoto, K. Nonami, A. Tatsuo, Y. Shigeru, and M. Kazuhisa, "Mine detection algorithm using pattern classification method by sensor fusion-experimental results by means of GPR," in *Systems and Human Science*. Amsterdam, The Netherlands: Elsevier, 2005, pp. 259–274.

[17] B. J. Allred, J. J. Daniels, and M. R. Ehsani, Eds., *Handbook of Agricultural Geophysics*. Hoboken, NJ: CRC Press, 2008.

[18] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006.

[19] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.

[20] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color pixel classification: Analysis and comparison," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 1, pp. 148–154, Jan. 2005.

[21] S. L. Phung and A. Bouzerdoum, "A new image feature for fast detection of people in image," *Int. J. Inf. Syst. Sci.*, vol. 3, no. 3, pp. 383–391, 2007.

[22] W. Shao, G. Naghdy, and S. L. Phung, "Automatic image annotation for semantic image retrieval," in *VISUAL2007*, vol. 4781, *Lecture Notes in Computer Science*, G. Qiu, Ed. Berlin, Germany: Springer-Verlag, 2007, pp. 372–381.

[23] S. Abe, *Support Vector Machines for Pattern Classification*. New York: Springer-Verlag, 2005.

[24] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[25] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 1998.

[26] B. Schlkopf and A. J. Smola, *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, 1st ed. Cambridge, MA: MIT Press, 2001.

[27] G. R. Olhoeft, "Working in a difficult environment: GPR sensing on the railroads," in *Proc. IEEE Antennas Propag. Soc. Int. Symp.*, 2005, vol. 3B, pp. 108–111.

[28] B. Indraratna, L. Su, and C. Rujikiatkamjorn, "A new parameter for classification and evaluation of railway ballast fouling," *Can. Geotech. J.*, vol. 48, no. 2, pp. 322–326, Feb. 2011.

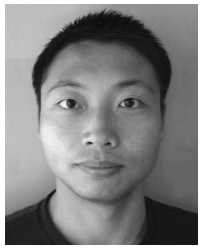
[29] W. Shao, A. Bouzerdoum, S. L. Phung, L. Su, B. Indraratna, and C. Rujikiatkamjorn, "Automatic classification of GPR signals," in *Proc. XIII Int. Conf. Ground Penetrating Radar*, Lecce, Italy, 2010, pp. 1–6.

[30] C.-C. Chang and C.-J. Lin, LIBSVM: A Library for Support Vector Machines, 2007. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[31] W. Shao, "Automatic Annotation of Digital Photos," M.S. thesis, Univ. Wollongong, Wollongong, Australia, 2007.

[32] V. N. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. New York: Springer-Verlag, 2000.

[33] C. Hauck and C. Kneisel, Eds., *Applied Geophysics in Periglacial Environments*. Leiden: Cambridge University Press, 2008.



Wenbin Shao received the B.Eng. degree in communication engineering from Northwestern Polytechnical University, Xi'an, China, in 2003 and the M.Eng. degree from the University of Wollongong, Wollongong, Australia, in 2007, where he is currently working toward his Ph.D. degree.

Abdesselam Bouzerdoum (M'89–SM'03) received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle.

He is a Professor of Computer Engineering and the Associate Dean (Research) of the Faculty of Informatics, University of Wollongong, Wollongong, Australia. He joined Adelaide University, Adelaide, Australia, in July 1991, and he was promoted to Senior Lecturer (Senior Assistant Professor) in January 1995. In 1998, he was appointed Associate Professor with Edith Cowan University, Perth, Australia. In 2004, he moved to the University of Wollongong to take up the position of Professor of Computer Engineering and Head of the School of Electrical, Computer and Telecommunications Engineering. In 2007, he was appointed Associate Dean (Research) in the Faculty of Informatics. Since 2009, he has been a Member of the Australian Research Council College of Experts and the Deputy Chair of the Engineering, Mathematics and Informatics Discipline since 2010. He held several appointments as Visiting Professor at Institut Galilée, University of Paris-13 (2004, 2005, 2007, 2008, and 2010), Hong Kong University of Science and Technology (2007), and Villanova University (2010). He has published over 260 technical articles and graduated 16 Ph.D. and 7 research Masters students. Currently, he is an Associate Editor of three international journals.

Dr. Bouzerdoum is the recipient of numerous awards and distinctions as acknowledgment of his research contributions. Among these, the most notable are a Distinguished Researcher (Chercheur de Haut Niveau) Fellowship from the French Ministry of Scientific Research in 2001 to conduct research at the Laboratoire d'Analyse et d'Architectures des Systèmes/Centre National de la Recherche Scientifique in Toulouse, two Vice Chancellor's Distinguished Researcher Awards, two awards for Excellence in Research Leadership, three awards for Excellence in Research Supervision, and the 2004 Chester Sall Award for Best Paper in IEEE TRANSACTIONS ON CONSUMER ELECTRONICS (3rd prize). He served as the Chair of the IEEE WA Section Signal Processing Chapter in 2004 and was the Chair of the Chair of the South Australian Section Neural Network Region Interest Group from 1995 to 1997. From 1999 to 2006, he served as Associate Editor of IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS. He is currently a member of the governing board of the Asia Pacific Neural Network Assembly. He is a member of the Optical Society of America and INNS.

Son Lam Phung (M'02) received the B.Eng. (first-class honors) and Ph.D. degrees in computer engineering from Edith Cowan University, Perth, Australia, in 1999 and 2003, respectively.

He joined the University of Wollongong, Wollongong, Australia, as a Research Fellow in 2005 and is currently a Senior Lecturer with the School of Electrical, Computer and Telecommunications Engineering. His general research interests are in the areas of image and video processing, neural networks, pattern recognition, and machine learning. He was the recipient of the University and Faculty Medals in 2000.

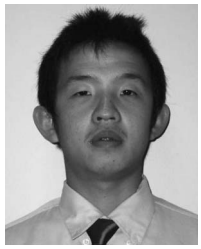
Lijun Su received the B.Eng. and M.Eng. degrees from Xi'an Jiaotong University, Xi'an, China, in 2000 and 2002, respectively, and the Ph.D. degree in geotechnical engineering from Hong Kong Polytechnic University, Kowloon, Hong Kong, in 2006.

He is a Research Fellow with the Department of Civil, Mining and Environmental Engineering, University of Wollongong, Wollongong, Australia. He is also with the School of Civil Engineering, Xi'an University of Architecture and Technology, China. His areas of expertise include constitutive modeling of geotechnical materials and nondestructive inspection of underground conditions. He has published over 30 articles in international journals and conferences.

Buddhima Indraratna received the B.S. degree in civil engineering and the M.S. degree in soil mechanics from Imperial College, University of London, London, U.K., and the Ph.D. degree in geotechnical engineering from the University of Alberta, Edmonton, Canada.

He is an internationally acclaimed Geotechnical Researcher and Consultant. He is currently a Professor and the Head of the School of Civil, Mining and Environmental Engineering, University of Wollongong, Wollongong, Australia. His outstanding professional contributions encompass innovations in railway geotechnology, soft clay engineering, ground improvement, environmental geotechnology, and geohydraulics, with applications to transport infrastructure and dam engineering. Under his leadership, the Centre for Geomechanics and Railway Engineering at the University of Wollongong has evolved to be a world class institution in ground improvement and transport geomechanics, undertaking national and international research and consulting jobs. He is the author of 4 other books and over 350 publications in international journals and conferences, including more than 30 invited keynote lectures worldwide.

Dr. Indraratna is a Fellow of the Institution of Engineers, Australia, the American Society of Civil Engineers, and the Geological Society of U.K. Recognition of his efforts is reflected by numerous prestigious awards, such as the 2009 EH Davis Memorial Lecture by the Australian Geomechanics Society for his outstanding contributions to the theory and practice of geomechanics and the Australian Commonwealth Government-hosted 2009 Business Higher Education Round Table Award for Rail Track Innovations among others. In the past, several of his publications have received outstanding contribution awards from the International Association for Computer Methods and Advances in Geomechanics, Canadian Geotechnical Society, and Swedish Geotechnical Society. He is also a Chartered Engineer in U.K. and a Professional Engineer in Australia.



Cholachat Rujikiatkamjorn received the B.Eng. degree in civil engineering from Khonkaen University, Khonkaen, Thailand, the M.Eng. degree from the Asian Institute of Technology, Pathumthani, Thailand, and the Ph.D. degree in geotechnical engineering from the University of Wollongong, Wollongong, Australia.

He is a Senior Lecturer in civil engineering with the University of Wollongong. His key areas of expertise include ground improvement for transport infrastructure and soft soil engineering. He has published over 50 articles in international journals and conferences.

Dr. Rujikiatkamjorn was the recipient of an award from the International Association for Computer Methods and Advances in Geomechanics for an outstanding paper by an early career researcher in 2009. He is also the recipient of the 2006 Wollongong Trailblazer Award for innovations in soft soil stabilization for transport infrastructure.