COMPRESSED SENSING-BASED FREQUENCY SELECTION FOR CLASSIFICATION OF GROUND PENETRATING RADAR SIGNALS

Wenbin Shao, Abdesselam Bouzerdoum, Son Lam Phung

ICT Research Institute, School of Electrical, Computer and Telecommunications Engineering,
University of Wollongong, Wollongong, NSW, Australia

ABSTRACT

In this paper we present an automatic classification system for ground penetrating radar (GPR) signals. The system extracts the magnitude spectra at resonant frequencies and classifies them using support vector machines. To locate the resonant frequencies, we propose an approach based on compressed sensing and orthogonal matching pursuit. The performance of the system is evaluated by classifying GPR traces from different ballast fouling conditions. The experimental results show that the proposed approach, compared to the approach of using frequencies at local maxima, represents the GPR signal more efficiently using a small number of coefficients, and obtains higher classification accuracy.

Index Terms— compressed sensing, frequency selection, ground penetrating radar, pattern classification

1. INTRODUCTION

Ground penetrating radar (GPR) exploits electromagnetic fields to probe lossy dielectric material. It excels in non-destructive detection of buried objects that are beneath the shallow earth surface or in visually impenetrable structures [1]. GPR has attracted considerable interest in many areas, such as archaeology [2], road construction [1], and mineral exploration and resource evaluation [1].

In railway transportation, the ballast is an essential component for proper railway operation. To ensure safety, regular inspection of rail tracks must be conducted. However, the traditional approach is labor-intensive and time-consuming. Thus, the rail industry is searching for new and more cost-effective approaches. As a non-destructive detection tool, ground penetrating radar has attracted great interest in railway ballast evaluation in recent years [3].

In this paper, we present an automatic classification system to assess railway ballast conditions. The system is based on the extraction of magnitude spectra at resonant frequencies and their classification using support vector machines. Feature selection plays an important role in pattern classification. Over the past three decades, many feature selection techniques have been proposed; for a review, see [4] and references therein. Peng *et al*. presented a two-stage algorithm by combining the minimal-redundancy maximal-relevance criterion and the other approaches such as wrappers [5]. Their experimental results show that, compared to other approaches such as max-dependency and max-relevance, the two-stage algorithm improves feature selection and classification performance. Yang *et al*. proposed a compressed sensing-based algorithm for dimensionality reduction [6]. This approach was evaluated on a pedestrian detection task. The experimental results show that the proposed algorithm achieves competitive performance using a small number of features.

To locate resonant frequencies, we propose a frequency selection approach based on feature selection using compressed sensing. Our method is aimed at reducing feature dimensionality and extracting informative frequencies. The remainder of the paper is organized as follows. In Section 2, a brief introduction to feature selection and compressed sensing is given, and the automatic classification system for GPR signals is introduced. In Section 3, the experimental methods and results are presented, followed by concluding remarks in Section 4.

2. METHODOLOGY

The proposed automatic classification system includes three main stages: pre-processing, feature extraction, and classification. The system block diagram is shown in Fig. 1. When a GPR signal is received, features are extracted from the significant frequencies automatically, and then sent to a trained classifier for assessment of railway ballast condition.

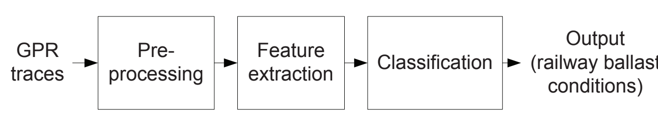

Fig. 1. Block diagram of the proposed automatic classification system.

The pre-processing stage employs basic signal processing techniques to reduce the intrinsic interferences introduced by the GPR and ensure the sampling rate consistency of the time-domain signals. Such preprocessing techniques include pre-processing techniques DC component removal, re-sampling and time shifting.

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 categorize ballast fouling conditions. We have observed empirically that the traces from ballast of different fouling conditions possess different magnitude spectra. Therefore, features are extracted from the magnitude spectra at salient frequencies. The salient frequencies are located where resonances occur. The selection of the salient frequencies, using a compressed sensing approach, is described next.

2.1. Frequency selection via compressed sensing

The Shannon sampling theorem states that, in order to reconstruct the original signal perfectly, the sampling frequency must be at least twice the highest frequency of the input signal. Recently, a new theory known as *compressed sensing* (CS) indicates that sparse signals can be reconstructed from a small number of linear and non-adaptive measurements that are under-sampled [7].

Suppose we have a signal $\mathbf{x}$ in $\mathbf{R}^n$. The ℓ_0 norm of $\mathbf{x}$, denoted by $\|\mathbf{x}\|_0$, is defined as the number of non-zero entries in $\mathbf{x}$. If $\|\mathbf{x}\|_0 = k$, the vector $\mathbf{x}$ is called k-sparse. Given m linear measurements of the signal $\mathbf{x}$, $\mathbf{y} = \mathbf{A}\mathbf{x}$, the sparse signal $\mathbf{x}$ can be recovered almost perfectly, even when $m << n$, provided some conditions are satisfied [7]. The vector $\mathbf{y} \in \mathbf{R}^m$ and matrix $\mathbf{A} \in \mathbf{R}^{m \times n}$ are called the *measurement vector* and the *measurement matrix*, respectively. Finding the sparsest $\mathbf{x}$ is equivalent to solving the following ℓ_0 minimization problem,

$$\min \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{1}$$

However, this minimization problem results in high computational cost [7]. Several alternative approaches have been proposed to solve this problem, such as reweighted ℓ_1 minimization [8], gradient projection [9] and orthogonal matching pursuit (OMP) [10].

Our aim is to locate the salient frequencies from the training data. This is equivalent to finding a representative subset of frequencies, which best represents the relationship between the frequency components and the class labels. Therefore, to adopt the compressed sensing paradigm for feature selection, we construct the measurement matrix using the frequency spectra of GPR signals and employ the class label set as the measurement vector.

Consider m training samples that belong to K classes

$$\{(\mathbf{s}_1, y_1), (\mathbf{s}_2, y_2), \ldots, (\mathbf{s}_m, y_m)\},$$

where $\mathbf{s}_i$ is a GPR trace and $y_i \in \{1, 2, \ldots, K\}$ is the class label. In our case, the matrix $\mathbf{A}$ is obtained by applying the n-point discrete-time Fourier transform to each trace; each column of $\mathbf{A}$ represents one frequency component. The $m \times 1$ vector $\mathbf{y}$ contains the class labels of the corresponding traces. Let $\mathbf{x}$ define the sparse vector that represents the salient frequencies. The frequency selection problem can be formulated as a CS problem, Eq. (1). Note that our focus is not the nonzero coefficient values in $\mathbf{x}$ but the coefficient positions, which indicate the selected frequencies.

The greedy algorithm OMP has been adopted to solve the ℓ_0 minimization problem [10]. The main steps of OMP are described as follows:

1. Initialize the iteration index, $t = 1$, a residual signal $\mathbf{r}_0 = \mathbf{s}$, and an empty matrix $\Phi_0 = \emptyset$.
2. For the t-th iteration, locate the atom $\mathbf{a}_t$ (a column of $\mathbf{A}$) that has the strongest correlation with the residual $\mathbf{r}_{t-1}$, and augment the matrix of previously selected atoms as Φ_t, $\Phi_t = [\Phi_{t-1}, \mathbf{a}_t]$.
3. Solve $\mathbf{x}_t = \arg\min_{\mathbf{x}} \|\mathbf{y} - \Phi_t \mathbf{x}\|_2$.
4. Calculate the residual $\mathbf{r}_t$, $\mathbf{r}_t = \mathbf{y} - \mathbf{\Phi}_t \mathbf{x}_t$.

The OMP identifies one significant frequency at a time and approximates the target function iteratively; it ensures the same frequency is not selected twice. Note that, feature selection is dual to sparse approximation. Selected frequencies can be ranked in the same order of OMP iterations. Once the significant frequencies are located by OMP, features can be extracted as the magnitude spectra at these frequency points.

2.2. Classification using SVMs

There are many methods available for pattern classification, such as discriminant analysis, decision trees, k-nearest neighbors, Bayesian classifier, neural networks and support vector machines (SVMs). Here, we choose SVMs as the classification tool because they have good generalization ability [11]. SVMs utilize explicit decision functions and are formulated for two-class problems. It is necessary to extend the SVM formulation to handle multi-class problems. There are several ways to extend SVMs; one-versus-all and pair-wise SVMs are two common approaches. In this paper, we adopt one-versus-all SVMs because this approach can handle effectively samples that do not carry sufficient resonances.

3. RESULTS AND ANALYSIS

In this section, we first explain the experimental data, then present the experimental results, including the selected significant frequencies and classification performance. Finally, we compare the OMP-selected features with local maxima features. The classifier generalization ability is evaluated using five fold *cross-validation*.

The data set used in the experiments was collected along an existing railway track in Wollongong station, Australia [12]. It consists of data from three different types of ballast based on the most common ballast fouling conditions: (i) 50% clay fouled ballast, (ii) clean ballast, and (iii) 50% coal fouled ballast. This data set can be divided into three subsets based on the antenna heights: 200 mm, 300 mm, and 400 mm data subsets. Each data subset consists of GPR traces from three different types of ballast. In total, there are 1382 traces in the 200 mm subset, 1386 traces in the 300 mm subset, and 2092 traces in the 400 mm subset; refer to [12] for more details.

3.1. Analysis of frequency selection and classification rate

In this section, we present the overall classification rates using different numbers of frequency points. The classification rate is the percentage of test samples that are correctly classified. Figure 2 shows the most significant 25 frequencies found using the OMP algorithm. It is clear that all the 25 frequencies are in the range of $[0,\ 3f_a]$, where f_a represents the antenna centre frequency (800 MHz). Moreover, the frequencies with highest weights are close to the antenna centre frequency in all data subsets.

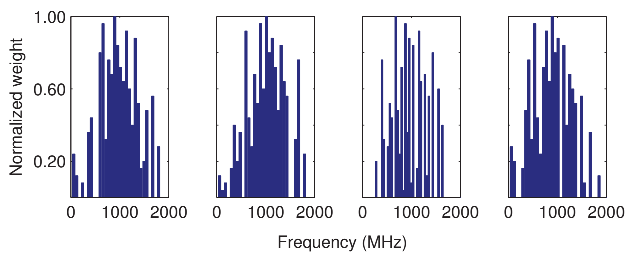

Fig. 2. Salient frequencies selected by OMP and their weights, for different antenna heights. The bar height indicated the weight of the corresponding frequency. From left to right, the data used are 200 mm subset, 300 mm subset, 400 mm subset, and combined data set.

Figure 3 compares the classification rates using different numbers of significant features from the 200 mm, 300 mm, and 400 mm data subsets, respectively. The system performance improves steadily with increasing number of significant frequencies.

- On the combined data set, using only 3 frequency points, the classification rate reaches 93.0%. When the number of frequency points reaches 6, the classification rate increases to 99.6%. Once the feature size reaches 10, the system performance remains stable with a classification rate of 100.0%.

- On the 200 mm data subset, when there are fewer than 4 frequency points, the classification rates are below 90.0%. When the number of frequency points reaches 4, the classification rate increases to 95.1%.

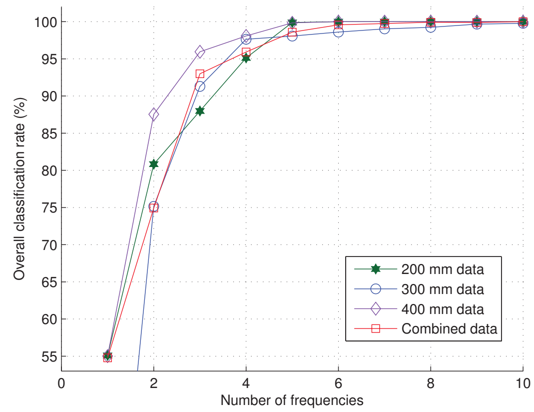

Fig. 3. Overall classification rates using different numbers of significant frequencies.

- On the 300 mm data subset, the system is able to classify the test set at a classification rate of 97.6% with 4 significant frequencies.

- For data of 400 mm antenna height, when the feature vector size is as small as 5, the system achieves an overall classification rate of 99.9%. When the feature vector size increases, the classification rate reaches 100.0%.

3.2. Comparison with local maxima features

We compare the OMP-selected features with the features extracted at local maxima. The local maxima are located via the morphological operation *dilation* in the frequency range $[0,\ 3f_a]$ [12].

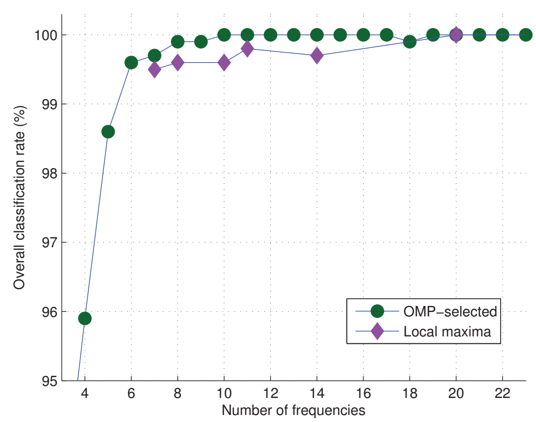

Fig. 4. Compare OMP-selected features with features at local maxima on the combined 800 MHz data set.

Figure 4 shows that the OMP-selected features achieve higher classification rates when fewer than 20 feature points are used. When the feature size reaches 20, both approaches

Table 1. Compare OMP-selected features with features at local maxima on the 200 mm, 300 mm, and 400 mm data sets, respectively.

Data set	*Features*	*Number of frequencies required for the CR of*		
		95.0%	99.0%	100.0%
200 mm	OMP-selected	4	5	6
	Local maxima	10	14	17
300 mm	OMP-selected	4	7	19
	Local maxima	6	12	27
400 mm	OMP-selected	3	5	6
	Local maxima	na	8	10

reach a classification rate of 100.0%. The comparison on the data subsets is given in Table 1. The results show that the OMP-selected features can achieve a high classification rate using fewer features than the local maxima approach.

To compare the computational cost, we evaluate the execution time of both approaches on a computer with Intel Q9400 (2.66 GHz) CPU and 3.23 GB RAM. The experiments are conducted on 10 data sets, each containing 20 traces. The average time of 10 sets is given in Table 2. The results show that the OMP search has an advantage over the local maxima approach in terms of computational cost.

Table 2. Computation time of OMP and local maxima search algorithms (in milliseconds).

Number of features	8	11	14	17
Local maxima	7.67	7.68	7.70	7.72
OMP search	0.17	0.26	0.37	0.56

4. CONCLUSIONS

In this paper, we have presented a frequency selection method based on feature selection using compressed sensing. This approach is integrated into an automatic GPR classification system, and evaluated using real-world railway GPR data. The experimental results indicate that the feature selection scheme is able to find a compact representation of ground penetrating radar signals. Furthermore, the selected feature set performs well in ballast fouling classification, and the CS-selected features outperform the local maxima features.

Acknowledgment

This work is supported in part by a grant from the *Australian Research Council*. The Wollongong railway GPR data were collected as part of the Rail CRC-AT5 project, sponsored by *CRC Rail for Innovation*.

References

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

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

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

[4] A. K. Jain, R. P. W. Duin, and M. Jianchang, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.

[5] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

[6] J. Yang, A. Bouzerdoum, F. H. C. Tivive, and S. L. Phung, "Dimensionality reduction using compressed sensing and its application to a large-scale visual recognition task," in *International Joint Conference on Neural Networks*, 2010, pp. 1–8.

[7] O. Scherzer, Ed., *Handbook of Mathematical Methods in Imaging*, 1st ed. New York: Springer Science+BusinessMedia, 2011.

[8] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[9] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[10] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

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

[12] W. Shao, A. Bouzerdoum, S. L. Phung, L. Su, B. Indraratna, and C. Rujikiatkamjorn, "Automatic classification of ground-penetrating-radar signals for railway-ballast assessment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3961–3972, 2011.