# Sparse signal decomposition for ground penetrating radar

Wenbin Shao, Abdesselam Bouzerdoum, and Son Lam Phung
School of Electrical, Computer and Telecommunications Engineering,
University of Wollongong, Australia
Email: ws909@uowmail.edu.au, a.bouzerdoum@ieee.org, phung@uow.edu.au

**Abstract**

In this paper, we present an adaptive approach for sparse signal decomposition, in which each GPR trace is decomposed into elementary waves automatically. A sparse feature vector is extracted from the decomposition and used for classification of railway ballast. The experimental results show that the proposed approach can represent the GPR signals efficiently, and effective features can be extracted for pattern classification.

## I. Introduction

Ground penetrating radar exploits electromagnetic fields to image the subsurface areas. Buried objects such as water tables and pipes, which are beneath the shallow earth surface or in a visually impenetrable structure, can be detected by GPR non-destructively [1]–[3]. Through visual inspection of the radar return, pseudo-imaging and signal processing, it is possible to identify the detailed characteristics of the objects. These advantages have facilitated considerable GPR applications in many areas, such as monitoring and investigating soil, underground water [4], coastal environments, glacier and ice sheet [4], [5], and railways and airfields [6].

Processing and interpreting GPR profiles is a difficult and challenging task [4], [5]. To date, most processing techniques for ground penetrating radar signals, such as frequency filtering and spectrum analysis [2], require visual inspection by an experienced human operator. However, human intervention may produce subjectivity and user-dependency into data elucidation. In this paper, we propose an adaptive approach of sparse signal decomposition to aid GPR analysis and interpretation. Compared to other time-frequency approaches, such as the discrete wavelet transform and the short-time Fourier transform, our approach has a clear advantage in reconstruction error and sparsity. Based on the signal decomposition, we also suggest a sparse feature extraction technique, which is used to classify the traces in the radar profile.

The paper is organized as follows. In Section II, the GPR processing techniques are reviewed. In Section III, first the proposed sparse signal decomposition approach is explained and compared with the discrete wavelet transform. Then, the feature extraction method is described. In Section IV, the experimental results are presented. Finally in Section V, concluding remarks are given.

## II. GPR PROCESSING TECHNIQUES

Ground penetrating radar relies on electromagnetic signals for subsurface investigation. The embryo of applying electromagnetic signals to detect subsurface objects can be traced backed to a German patent in the $1900$s [2]. However, GPR was not commercialized until 1970s; it became popular in the $1980$s because of digital data acquisition [4], [7].

The main components of a GPR system are the signal generator (transmitter), transmitting and receiving antennas, and recording device (receiver) [4], [8]. The general process of object detection using GPR is as follows.

1) The transmitter generates a pulse of electromagnetic energy and delivers it to the transmitting antenna $T_x$.
2) The electromagnetic wave radiates from the transmitting antenna into the subsurface.
3) If, on the path of the wave propagation, there exists an object whose electrical properties are different from those of surrounding materials, part of the wave energy is reflected.
4) The reflected energy is detected by the receiving antenna $R_x$.
5) The receiving antenna sends the received signal to the receiver for storage and display, where real-time signal processing may be applied.

The general objective of GPR data processing is to enhance the 2-D time-distance record so that it can be interpreted by the human operator [2]. The basic processing techniques usually consist of the following:

- **Dewow** is aimed at removing the low-frequency components from the raw GPR data and reducing the mean of each trace to zero [2], [5].
- **Time varying gain** is applied to compensate for the attenuation effects during wave propagation [2].
- **Filtering** techniques are applied to improve the signal to clutter ratio and improve the visual quality of the radar data [2], [7].
- **De-convolution** tends to remove the source wavelet effect.
- **Migration** is applied when the reflected electromagnetic waves are geometrically and spatially distorted because of antenna characteristics and electrical properties of the ground.

Researchers also apply other signal processing techniques to aid GPR signal analysis. Al-Qadi and colleagues [9] proposed a time-frequency approach to evaluate GPR data for railway ballast assessment. Their approach utilizes short-time Fourier transform (STFT). The energy attenuation of STFT images is visually inspected to assess ballast conditions.

Sinha and colleagues presented a new method for time-frequency map computation for non-stationary signals [10]. The traditional short-time Fourier transform has a limit on time-frequency resolution because the window length is pre-defined. This problem can be overcome by employing the continuous wavelet transform (CWT). Their experiments on seismic data show that the CWT approach can be used to detect frequency shadows and subtle stratigraphic features.

## III. Sparse signal decomposition and classifiers

We propose a system based on sparse signal decomposition to aid the analysis of GPR signals. The system is composed of four major stages: pre-processing, sparse signal decomposition, feature extraction, and classification. The preprocessing stage consists of DC component removal, re-sampling and time shifting. In this section, we first describe the proposed sparse signal decomposition in Section III-A, then present a comparison between the sparse signal decomposition approach and the discrete wavelet transform in Section III-B. Next, we explain a new feature representation approach using the sparse signal decomposition in Section III-C, followed by a description of the classification tool in Section III-D.

### *A. Sparse signal decomposition*

In the proposed approach, every GPR trace is decomposed into its fundamental constituent waves. Consider that a radar trace $s(t)$ is a linear combination of elementary signals, or waves, expressed as

$$s(t) = \sum_{i=1}^{N} \alpha_i \varphi_i(t), \tag{1}$$

where $\varphi_i(t)$ is an elementary wave and $\alpha_i$ is a scalar constant, or weight. We should note that the signals $\varphi_i(t)$ are chosen from an over complete non-orthogonal dictionary. Hence in general Equation (1) does not represent an orthogonal signal decomposition. Here we consider the elementary waves $\varphi_i(t)$ as shifted, or time delayed Ricker and Gabor wavelets:

$$\varphi_i(t) = g_i(t - \tau_i), \tag{2}$$

where $g_i(t)$ is a basic wavelet and $\tau_i$ is a time delay to be determined. Therefore, the original trace $s(t)$ may be expressed as

$$s(t) = \sum_{i=1}^{N} \alpha_i g_i(t - \tau_i). \tag{3}$$

The approach we follow is that of finding a sparse signal approximation in (1). The time delay of each constituent wave is assumed unknown, and hence must be determined adaptively. Suppose that $s$ is a GPR trace of $N$ samples in the discrete-time domain. The procedure for sparse signal decomposition can be described as follows:

1) Form a dictionary of fundamental waves or atoms, $G = [g_1,\ g_2,\ \cdots,\ g_M]$ with all the functions $g_i$ having unit norm, based on the parameters of ground penetrating radar used, such as antenna frequency and sampling frequency.
2) Initialize the iteration index, $k = 1$, a residual signal $s_0^* = s$, and an empty matrix $\Phi_0$.
3) For the $k$-th iteration, compute the cross-correlation $r_{ki}(\tau)$ of the function $g_i \in G$ and the residual signal $s_{k-1}^*$.
4) Find the atom $g_{i^*}$ that gives the highest correlation in absolute value, where

$$i^* = \arg\max_i \left[ \max_\tau |r_{ki}(\tau)| \right], \tag{4}$$

and determine the corresponding time delay $\tau_k$.

5) Let $\varphi_k = g_{i^*}[n - \tau_k]$, and update the matrix $\Phi_k$:

$$\Phi_k = [\Phi_{k-1}, \varphi_k],$$

where $\varphi_k$ is the $k$-th column of $\Phi_k$.

6) Compute the weight vector $\boldsymbol{\alpha}_k = [\alpha_1, \ldots, \alpha_k]^T$:

$$\boldsymbol{\alpha}_k = ({\Phi_k}^T \Phi_k)^{-1} {\Phi_k}^T s. \tag{5}$$

7) Update the residual signal $s_k^* = s - \Phi_k \boldsymbol{\alpha_k}$.

8) Repeat Steps 3 to 7 until $k$ reaches a pre-defined value or the residual satisfies

$$\frac{||s_k^*||_2}{||s||_2} < \epsilon, \tag{6}$$

where $\epsilon$ is a predefined tolerance.

### *B. Efficiency of the sparse signal decomposition*

To illustrate the efficiency of the proposed sparse signal decomposition for GPR traces, we compare the proposed approach with the discrete wavelet transform using the normalized root mean squared error (NRMSE) as the criterion. NRMSE is a measure that indicates the difference between the approximated signal and the original signal, defined as

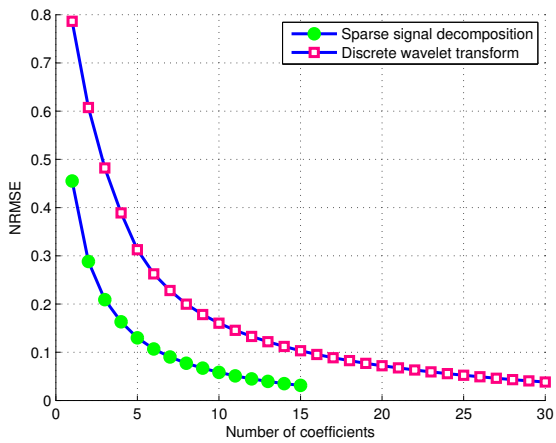
$$\text{NRMSE} = \frac{\sqrt{\sum_{i=1}^{N} (s_i - p_i)^2 / N}}{\sigma_s}, \tag{7}$$

where $s_i$ is the $i$-th sample of the signal $s$, $p$ is the signal approximation, and $\sigma_s$ is the standard deviation of $s$.

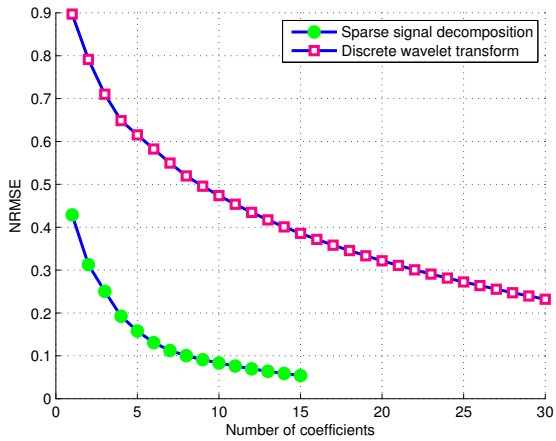
In the sparse signal decomposition, we consider the elementary waves $\varphi_i(t)$ as shifted Gabor functions. In the wavelet processing, we first apply the discrete wavelet transform with Daubechies wavelets of order $6$ to the trace. Then we threshold the wavelet coefficients: only a few wavelet coefficients remain, the other coefficients are assigned to $0$.

Two real-world data sets have been used for the evaluation: GPR data collected from from Windmill Islands (Windmill Islands data set) [11] and data gathered from railway at Wollongong station, Australia (Wollongong railway data set) [12]. From each GPR data set, there are $100$ traces arbitrarily selected for comparison.

The overall NRMSE results of both data sets are shown in Fig. 1. For the Windmill Islands data, using the sparse signal decomposition approach, the average NRMSE is $0.21$ with $3$ coefficients, and $0.03$ with $15$ coefficients. To achieve similar performances, the discrete wavelet transform requires $8$ and $34$ coefficients, respectively. For the Wollongong railway data set, our approach has an NRMSE of $0.25$ with $3$ coefficients and $0.05$ with $15$ coefficients. In comparison, for the wavelet transform, $28$ and $93$ coefficients are required to obtain similar overall errors, respectively. A summary of the comparison is given in Table I. The results show that the sparse decomposition can represent the GPR signal more efficiently.




(a)




(b)

Fig. 1. Overall NRMSE of sparse signal decomposition and discrete wavelet processing on real-world data. (a) GPR data from Windmill Islands. (b) Wollongong railway data set.

TABLE I

OVERALL NRMSE OF SPARSE SIGNAL DECOMPOSITION (SSD) AND DISCRETE WAVELET PROCESSING (DWT) ON REAL-WORLD DATA.

| Number of coefficients | | **1** | **3** | **7** | **11** | **15** |
|---|---|---|---|---|---|---|
| Windmill Islands data | **SSD** | 0.46 | 0.21 | 0.09 | 0.05 | 0.03 |
| | **DWT** | 0.79 | 0.48 | 0.23 | 0.15 | 0.10 |
| Wollongong railway data | **SSD** | 0.43 | 0.25 | 0.11 | 0.08 | 0.05 |
| | **DWT** | 0.90 | 0.71 | 0.55 | 0.45 | 0.39 |

### C. Feature extraction

During feature extraction, we utilize Gabor functions to construct the dictionary. The Gabor function is defined as the product of a Gaussian kernel and a sinusoidal function:

$$\begin{cases} e^{-\frac{t^2}{2\sigma^2}}\sin(2\pi f t) & \text{odd Gabor,} \\ e^{-\frac{t^2}{2\sigma^2}}\cos(2\pi f t) & \text{even Gabor.} \end{cases} \tag{8}$$

By varying the values of $\sigma$ and $f$ in (8), a Gabor dictionary can be constructed.

Assuming $K$ iterations are used for the sparse signal representation, then the acquired time delays can form a vector $\boldsymbol{\tau} = \{\tau_1, \tau_2, \cdots, \tau_K\}$. We should note that $\tau_k$ is an integer. Correspondingly, there exist two vectors: the frequency vector $\mathbf{f} = \{f_1, f_2, \cdots, f_K\}$ and $\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, \cdots, \sigma_K\}$, where $f_k$ and $\sigma_k$ are the parameters extracted in the $k$-th iteration. Let $s$ denote a trace and $N$ represent the number of samples. Taking the feature set $\boldsymbol{\sigma}$ as an example, the sparse feature vector for each trace is constructed as follows.

1) Create a vector $V_s$ with $N$ zero entries.
2) Assign $\sigma_k$ to the $\tau_k$-th entry of $V_s$, i.e. $V_s(\tau_k) = \sigma_k$.
3) The resulting vector $V_s$ forms a sparse feature vector.

Likewise, for the frequency feature, assign $f_k$ instead of $\sigma_k$ in Step 2.

### D. Classification

In this paper, we choose support vector machines (SVMs) as the classification tool because of their excellent performance in various practical applications [13]. For comparison purposes, we also implement the $k$-nearest neighbor classifier.

SVMs are a supervised learning algorithm formulated for two-class problem. Suppose we have $M$ training samples $\mathbf{x}_i$ ($i = 1, 2, \ldots, M$) and each sample is associated with a class label $y_i$ ($y_i \in \{1, -1\}$). If the data are linearly separable in the input space, the decision function is given by

$$y_i(\langle \mathbf{w}, \mathbf{x_i} \rangle + b) \geq 1, \tag{9}$$

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

Among the many hyperplanes that can separate the data, there is only one *optimal separating hyperplane* that achieves maximum margin. The maximum margin perpendicular to the hyperplane is expressed as $2/\|\mathbf{w}\|$. Thus, the classification problem is equivalent to finding $\mathbf{w}$ that maximizes the margin.

To cope with data that are usually not linearly separable, non-negative slack variables are introduced into (9):

$$y_i(\langle \mathbf{w}, \mathbf{x} \rangle + b) \geq 1 - \xi_i \text{ for } i = 1, \ldots, M. \tag{10}$$

To construct the optimal separating hyperplane, we minimize

$$\tau(\mathbf{w}, \boldsymbol{\xi}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{M}\xi_i, \tag{11}$$

subject to (10), where $C$ is a constant representing the trade-off between margin maximization and training error minimization. This is a constrained optimization problem. By introducing non-negative Lagrange multipliers and applying the Karush-Kuhn-Tucker conditions, a dual optimization problem is obtained, that is,

$$\text{maximizing } Q(\boldsymbol{\alpha}) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} \alpha_i \alpha_j y_i y_j \left\langle \mathbf{x_i}, \mathbf{x_j} \right\rangle, \tag{12}$$

subject to constraints

$$0 \leq \alpha_i \leq C \text{ for } i = 1, \ldots, M, \tag{13}$$

$$\sum_{i=1}^{M} \alpha_i y_i = 0 \text{ for } i = 1, \ldots, M. \tag{14}$$

In practice, the data samples from the input space are projected to a higher-dimensional dot product space via a mapping function $\Phi$; a positive semidefinite kernel $H$ is usually employed to simplify the projection, mathematically,

$$H(\mathbf{x_i}, \mathbf{x_j}) = \langle \Phi(\mathbf{x_i}), \Phi(\mathbf{x_j}) \rangle. \tag{15}$$

Therefore, the optimization problem in Eq. (12) becomes

$$\text{maximizing } Q(\boldsymbol{\alpha}) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} \alpha_i \alpha_j y_i y_j H(\mathbf{x_i}, \mathbf{x_j}), \tag{16}$$

subject to the constraints in (13) and (14).

## IV. Experiments and analysis

The proposed sparse signal representation is applied to the classification of railway ballast fouling conditions on the Wollongong railway data set. The Wollongong railway data set was collected using $800$ MHz center frequency. It consists of GPR signals from three common ballast types: (i) $50\%$ clay fouling, (ii) clean, and (iii) $50\%$ coal fouling. The fouling material is measured using relative ballast fouling ratio. The whole data set can be divided into three subsets based on the antenna heights: $20$ cm data subset, $30$ cm data subset, and $40$ cm data subset. The $20$ cm and $30$ cm data subsets were collected under dry ground condition: sunny weather and dry materials. The $40$ cm data subset was acquired under wet condition: cloudy weather and water-saturated materials. More details of the data set can be found in [14].

We first extract feature vectors from the decomposition approach, then feed the features to a pair-wise support vector machine and a $k$-nearest neighbor classifier ($k$-NN) for classification. In the pair-wise SVM approach, $k(k-1)/2$ two-class SVM classifiers are constructed to solve a $k$-class problem. Each SVM classifier is trained with samples from two classes.

### A. Analysis of sparse feature vectors

To analyze the classification performance of the sparse feature vector, two types of experiments have been conducted: (i) using single feature, such as the frequency $\mathbf{f}$, or the standard deviation $\boldsymbol{\sigma}$; (ii) using a combined feature vector with $\mathbf{f}$ and $\boldsymbol{\sigma}$. The features are normalized as follows:

1) The frequency feature is divided by the antenna frequency.
2) The standard deviation feature is divided by the mean of the standard deviation values across the Gabor dictionary.

The classification performance is shown in Table II.

The results show that, using only one feature ($\mathbf{f}$ or $\boldsymbol{\sigma}$), SVMs have similar classification rates on the $20$ cm and $30$ cm data subsets; likewise, $k$-NN classifiers have similar results. Both SVM and $k$-NN classifiers achieve their best performance on the $40$ cm data. For example, when the feature $\boldsymbol{\sigma}$ is used, the SVM achieves $96.0\%$, $95.3\%$, and $97.7\%$ on the $20$ cm, $30$ cm, and $40$ cm data subsets, respectively; the $k$-NN reaches a classification rate of $89.7\%$, $87.4\%$, and $95.4\%$, correspondingly. This phenomenon may be attributed to the fact that the $40$ cm data subset was collected under water-saturated condition. The higher dielectric permittivity of the water results in a stronger reflection than the dry ballast, and therefore improves the classification rate.

TABLE II
CLASSIFICATION RATES (%): SPARSE FEATURES.

| **Classifier** | **SVMs** | | | **$k$-NN** | | |
|---|---|---|---|---|---|---|
| Features | *20 cm* | *30 cm* | *40 cm* | *20 cm* | *30 cm* | *40 cm* |
| $\mathbf{f}$ | 90.8 | 91.0 | 97.3 | 75.7 | 76.3 | 91.4 |
| $\boldsymbol{\sigma}$ | 96.0 | 95.3 | 97.7 | 89.7 | 87.4 | 95.4 |
| $\mathbf{f}$ and $\boldsymbol{\sigma}$ | 96.4 | 94.6 | 98.4 | 90.4 | 86.9 | 96.8 |

For all the data subsets, the SVMs outperform the $k$-NN classifier. For example, on the $30$ cm data subset, SVM classifier achieves $91.0\%$ and $95.3\%$ classification rates using frequency and standard deviation features, while the $k$-NN classifier reaches $76.3\%$ and $87.4\%$, respectively.

### *B. Refined sparse feature vector*

Besides normalization, the sparse feature vector can be refined further. Utilizing the time delay found through the first iteration of the decomposition approach, some information can be discarded.

The signal decomposition approach always discovers the time delay $\tau_1$ with maximum correlation coefficient in the first iteration. From the second iteration onwards, it is possible to find a time delay $\tau_k$ smaller than $\tau_1$. We consider any $\tau_k$, where $\tau_k < \tau_1$, as non-significant information. For the entire data set, the global minimum $\tau_{min}$ is used as a threshold to trim the sparse feature vector. For example, if there are three traces and a $4$-iteration decomposition is performed, we will have the time delay matrix

$$\begin{bmatrix} \tau_{11} & \tau_{21} & \tau_{31} \\ \tau_{12} & \tau_{22} & \tau_{32} \\ \tau_{13} & \tau_{23} & \tau_{33} \\ \tau_{14} & \tau_{24} & \tau_{34} \end{bmatrix}.$$

The minimum time delay is given by

$$\tau_{min} = \min(\{\tau_{11},\ \tau_{21},\ \tau_{31}\}). \tag{17}$$

Any feature with time delay $\tau_{ij}$, where $\tau_{ij} < \tau_{min}$, is rejected during feature extraction.

Using the thresholding approach to refine the sparse feature vector, for the $20$ cm data subset, on average $20.2\%$ of the $15$ iterations are discarded. For the $30$ cm and $40$ cm data subsets, on average $18.6\%$ and $22.0\%$ are discarded, respectively. The overall classification rates in Table III show that the system can achieve similar classification rates with fewer significant features.

TABLE III

CLASSIFICATION RATES (%): REFINED FEATURES.

| **Classifier** | **SVMs** | | | **$k$-NN** | | |
|---|---|---|---|---|---|---|
| Features | *20 cm* | *30 cm* | *40 cm* | *20 cm* | *30 cm* | *40 cm* |
| $\mathbf{f}$ | 89.5 | 87.7 | 95.9 | 70.2 | 72.9 | 89.0 |
| $\boldsymbol{\sigma}$ | 95.5 | 91.1 | 97.9 | 88.2 | 83.9 | 92.1 |
| $\mathbf{f}$ and $\boldsymbol{\sigma}$ | 95.6 | 92.6 | 97.9 | 87.1 | 83.9 | 94.0 |

Figure 2 shows the comparison between the sparse features and the refined features on the $20$ cm data set when SVMs are applied. It can be seen that though $20.2\%$ of the sparse features are rejected, the system performances have not dropped considerably. The classification rates vary only from $0.4\%$ to $1.3\%$.

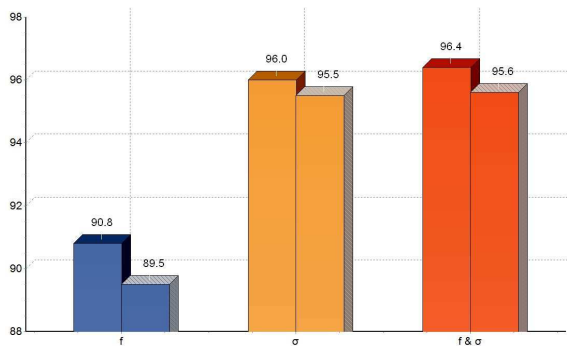



Fig. 2. Compare the sparse features (left bar) with the refined features (right bar) on the 20 cm data subset using SVMs.

The comparison on the same data set using $k$-NN is shown in Fig. 3. The results represent that the performance varies from $1.5\%$ to $5.5\%$. Compared to the classification rates of SVMs, the classification rates of $k$-NN have larger variation.

## V. CONCLUSION

Based on the characteristics of the ground penetrating radar signals, this paper proposes an adaptive sparse signal decomposition approach for GPR signal analysis. We also suggest a sparse feature vector representation for

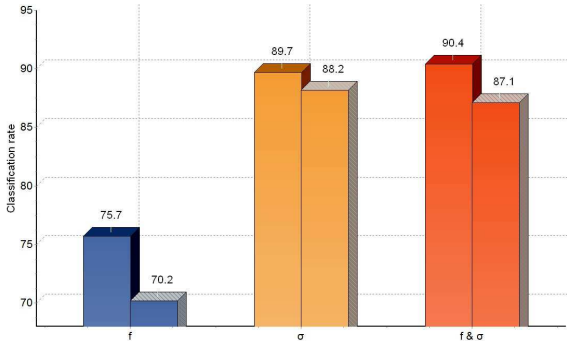



Fig. 3. Compare the sparse features (left bar) with the refined features (right bar) on the 20 cm data subset using $k$-NN.

classification. The experimental results demonstrate that a sparse vector can efficiently and effectively represent a GPR trace. In future research, we aim to further explore the signal decomposition approach, and investigate other time-frequency techniques for GPR analysis.

## 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] A. P. Annan, "GPR - history, trends, and future developments," *Subsurface Sensing Technologies and Applications*, vol. 3, no. 4, pp. 253–270, 2002.

[2] D. J. Daniels, Ed., *Ground Penetrating Radar*, 2nd ed. London: The Institution of Electrical Engineers, 2004.

[3] A. Neal, "Ground-penetrating radar and its use in sedimentology: principles, problems and progress," *Earth-Science Reviews*, vol. 66, no. 3-4, pp. 261–330, 2004.

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

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

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

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

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

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

[10] 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, 2005.

[11] J. Pettersson, *Ground Penetrating Radar Data from Windmill Islands - Miscellaneous Data*, 2006. [Online]. Available: http://data.aad.gov.au/aadc/metadata/

[12] W. Shao, A. Bouzerdoum, S. L. Phung, L. Su, B. Indraratna, and C. Rujikiatkamjorn, "Automatic classification of GPR signals," in *XIII International Conference on Ground Penetrating Radar*, Lecce, Italy, 2010.

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

[14] 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," *to appear in IEEE Transactions on Geoscience and Remote Sensing*, 2011.