Savitzky-Golay filtering for Scattered Signal De-noising

M A Selver¹, M Secmen², E Y Zoral¹

¹Dept. of Electrical and Electronics Eng., Dokuz Eylul University, İzmir, TURKEY
²Department of Electrical and Electronics Eng., Yasar University, İzmir, TURKEY

Email: alper.selver@deu.edu.tr, mustafa.secmen@gmail.com, yesim.zoral@deu.edu.tr

Abstract. Extraction of distinguishing features and decision of classifiers are highly influenced by the low signal-to-noise ratio (SNR) rates, when target identification from scattered electromagnetic waves is considered. In order to increase the correct identification rates, smoothing operations, which should increase the SNR without greatly distorting the signal, are occasionally employed. However, this operation is mostly performed to the complete scattered signal via an over-complete basis. On the other hand, Savitzky-Golay filters can de-noise the signal by fitting successive sub-sets of adjacent scattered data points with a low-degree polynomial through the use of linear least squares. Thus in this study, both computational burden and accuracy of Savitzky-Golay filters are compared to three well-established smoothing techniques in time domain, frequency domain and time-frequency analysis. The analyses are performed with both simulated and measured data from various conductor and dielectric targets having different size, geometry and material type.

1. Introduction

Classification of targets from scattered electromagnetic signals become much more difficult as SNR levels decrease. The effect of strong noise cause problems on eliminating the effects of aspect angle and prevents extracting distinguishable features for identification of different targets [1]. Previously, we propose a target classification method, which uses a structural feature set extracted from scattered signals. In those earlier works, various multi-scale approximation methods are performed for de-noising prior to feature extraction [2]. Among them, hierarchical radial basis function network (HRBFN) topology has performed best in suppressing the adverse effects of noise on scattered signals [3]. Unfortunately, still the effectiveness of the structural properties of time domain scattered signals are highly susceptible to the level of noise and distortion. Moreover, the de-noising operations should be performed in real time to satisfy operating conditions. Opposed to well-known wavelet multi-resolution analysis using the discrete wavelet transform [4] or HRBFN strategy, Savitzky-Golay filters can reconstruct the signal in a timely manner (using consecutive frames of a signal) rather than processing the complete signal block by reconstructing the signal via adding and/or removing frequency components [5].

When a neural network (NN) based systems is developed for a non-cooperative system, the system is first trained using a set of references and the classification results is determined by the best match.

There are two key components of such systems, which are 1) the features extracted from scattered signals in order to represent the target characteristics and 2) the classifier which performs the decision
Both of these components are highly sensitive to noise and therefore, the distortions due to low SNR should be handled carefully.

Recently, a target identification system is developed by using a novel feature set [3]. The feature set mainly relies on the structural differences of the scattered signal waveforms and aims to use the structural differences corresponding to the position, amplitude, rise/fall times and the number of the hills/valleys of the scattered signals. In order to obtain these features, a piece-wise linear approximation process is carried out and once the sub-waveforms are represented by triangles; their peaks, widths, slopes are calculated for each of them together with their inter-distance. Then, HRBFN employs a number of Gaussian units to fit to the valleys/hills, which are combined to create an approximation.

However, the above mentioned operation is mostly performed to the complete scattered signal via an over-complete basis. On the other hand, Savitzky-Golay filters can de-noise the signal by fitting successive sub-sets of adjacent scattered data points with a low-degree polynomial through the use of linear least squares. Thus in this study, both computational burden and accuracy of Savitzky-Golay filters are compared to three well-established smoothing techniques in time domain, frequency domain and time-frequency analysis. The analyses are performed with both simulated and measured data from various conductor and dielectric targets having different size, geometry and material type. For evaluation of the proposed technique in target classification, cross validation learning strategy [9] is used. The results show that, up to -10 dB SNR, proposed method can effectively be used for target classification.

2. Scattered Signal Data Set

2.1. Generation of Simulated Data

The simulated signals used in this study are obtained by Matlab based simulation of Hertz and Debye potentials as in [2]. The analytical solutions of these potentials are extracted for a plane wave excitation which is linearly polarized in x-direction and propagates in z-direction (Figure 1). The far field scattered responses are computed over a bandwidth from zero to 12 GHz at 873 frequency sample points with a frequency resolution of 13.75 MHz. The responses are obtained at φ= π/2 plane, with a radial distance of 72 cm from the sphere center for twelve different Bistatic Aspect Angles (BAA), 180° - θ = θb = 10°, 20°, … 180° degrees. The resulting time signals have 1024 sample points with a total time span of 5.115 ns. The noisy scattered time domain signals at all the aspect angles stated above are obtained at the SNR levels of 10, 0 and -10 dB to be used for classifier design and performance testing.

2.2. Generation of Measured Data (Dielectric Rods and Model Aircrafts)

The measured data includes four dielectric rods and three model aircrafts. The dielectric rods have 15 cm length, 1.25 cm radii and different permittivity values of εr = 3, 4, 5 and 5.5 The target dimensions, between 0.5λ and 6λ, are again in the resonance region. The VV-polarized rods with the same length
and the same radius but different permittivity values. Frequency responses of targets are measured at the aspect angles of $\theta=0^\circ$ to $170^\circ$ with $10^\circ$ steps, where at $0^\circ$ the rod lies along $z$ axis.

The second measured target set of model aircrafts contains three conducting small-scale aircrafts targets including a Boeing 747, a DC-10 and a Boeing 767. The models are scaled by $1/500$ for Boeing 747 and DC-10; but, by $1/600$ for Boeing 767. The body, wing and tail lengths of each target are 14.5 cm, 12.7 cm and 4.8 cm, for Boeing 747; 12.48 cm, 12.54 cm and 5 cm, for Boeing 767; and 12.7 cm, 11.4 cm and 5.25 cm, for DC10, respectively. The measurement setup is given in Figures 2 and 3 while examples of simulated and measured scattered signals are given in Figures 4 and 5.

Figure 2. The photographic views of the measurement setup: the antennas and target configuration

Figure 3. The photographic views of the measurement setup: The antennas and vector network analyzer configuration.

Figure 4. Scattered signals at $90^\circ$ and 1.8cm distance with (a) without noise, (b) SNR=20 dB, (c) SNR=10 dB (d) SNR= -10 dB (e) SNR= -20 dB (Electric Field (V/m) vs time)
Figure 5. (top) The scattered time signals of the small-scale aircraft targets Boeing 747, DC-10, Boeing 767 at the aspect angle of $\theta=90$ degree for the frequency band of (1-12 GHz). (bottom) Scattered time signals of the dielectric rods with $\varepsilon_r = 3$ and 4 at the aspect angle of $\theta=90$ degree for the frequency band 1-12 GHz.
3. De-Noising Scattered Signals

When the scattered signals are corrupted by noise, their waveform structures change significantly. This distortion affects the identification algorithms that rely on structural time-domain properties. In order to prevent the deteriorations in performance, the scattered signal should be recovered. Previously, three well-established multi-resolution techniques are compared for the task of de-noising scattered signals at low SNR conditions [8]. The differences between those three methods, namely wavelet decomposition (WD), HRBFN approximation and matching pursuit (MP), have two folds. The first one is the differences at the type of base functions. Wavelet and HRBFN use a single basis function and re-use it by changing its scale, amplitude and other parameters while MP uses multiple bases. The second difference is the de-noising strategy such that WD performs a fine-to-coarse approximation by eliminating high frequency components step-by-step while HRBFN employs a coarse-to-fine strategy by utilizing bases with increasing frequency.

On the other hand, these techniques have a property in common. They all handle the complete time-domain signal at once requiring all the signal to be acquired before processing. However, there might be cases during which the time span of the scattered signal is long and the de-noising procedure require real time analysis. In such cases, it would be inefficient to wait for the complete signal acquisition. Thus, in this study, Savitzky-Golay filters, which allows processing and de-noising sub-frames (i.e. time intervals).

3.1. Wavelet Decomposition (WD) with Gabor Basis

Discrete WD is proven to be very effective for analyzing non-stationary signals in various applications [15]. In this application, the Gabor wavelet is chosen as the motherlet since it is tunable specific frequencies and allows detection of high frequency information together with noise filtering in a single step [14]. The Gabor wavelet is composed of a complex exponential multiplied by a Gaussian function is defined as

$$g(x) = A \exp\left( -\frac{(x-\mu_x)^2}{2\sigma_x^2} \right)$$  \hspace{1cm} (1)

where, $\mu_x$ is the center and $\sigma_x$ is the width and $A$ is the amplitude. A Gabor wavelet function can be defined as

$$\psi(x) = A \exp\left( -\frac{(x-\mu_x)^2}{2\sigma_x^2} \right) \exp\left( jk(x-\mu_x) \right)$$  \hspace{1cm} (2)

where k determines the frequency of the complex exponential. Briefly, WD based de-noising corresponds to application of a group of Gabor filters followed by down-sampling as the levels of the decomposition proceed.

3.2. B. Hierarchical Radial Basis Function Network (HRBFN)

Despite fine to coarse approach of WD, HRBFN uses several Gaussian units having decreasing variances and amplitudes as given in (3) to perform a coarse to fine approximation.

$$g(x) = \exp\left( -\frac{(x-c)^2}{\sigma^2} \right) \left( 1/\sigma \right)^2 \left( x-c \right)$$  \hspace{1cm} (3)

where, \{c $\in \mathbb{R}$\}, \{\sigma $\in \mathbb{R}$\} are the center and the width of the $j^{th}$ Gaussian unit at the $i^{th}$ layer. If $M$ denotes the total number of Gaussians and \{w $\in \mathbb{R}$\} denotes the associated amplitude, the approximated function can be expressed as

$$f(x) = \sum_{j=1}^{M} w_j g(x \cdot c; \sigma_j)$$  \hspace{1cm} (4)
3.3. Matching Pursuit (MP)

The MP uses multiple bases from an over-complete dictionary [12]. Since the scattered signals may contain abrupt changes that can be represented with multiple bases especially under low SNR conditions. Since the quality of approximation depends on how well the signal characteristics are matched by the bases, MP offers advantages over single basis approximations. This property of the MP might be particularly useful when the noise (i.e. unwanted information) has a narrow frequency spectrum. In such cases, MP might be considered as a superior alternative to the wavelets, which are more effective on determination of isolated discontinuities. The use of the MP is carried out by using trigonometric, exponential, and polynomial bases, which are selected through an extensive experimentation process.

3.4. Savitzky-Golay (SG)

Savitzky–Golay is a digital smoothing filter that can be used to increase the SNR of a corrupted signal without distorting it very much. This is achieved by fitting successive time frames (or windows) with a pre-defined degree polynomial via linear least squares. As in the case of scattered signals, when the data points are equally spaced, an analytical solution can be derived in the form of a single set of convolution coefficients. A demonstrative application result is given in Figure 6.

4. Application & Results

The four de-noising techniques in Section III are applied to the targets described in Section II via a time domain classification method, which extracts a feature set that collects the structural differences of the scattered signal waveforms such as the position, amplitude, rise/fall times and the number of the hills/valleys after a triangularization process [3].

During the tests, cross-validation techniques, specifically 9-fold splitting [9], is employed in order to compare the results with earlier studies. Only the signals in the testing fold are corrupted with noise such that the classifier is trained against the features of the high SNR scattered signals. An MLP (i.e. Multi-Layer Perceptron) network having six hidden and four output neurons is trained by back-propagation with adaptive learning rate [10]. The training goal is chosen to be 0.001.

Considering the results presented in Table I, the de-noising has no significant contribution to the result at 0dB SNR since the triangularization process can be performed more accurately when the distortions on the scattered signal do not alter its waveform. However, at lower -10 dB SNR, the de-noising methods seem to increase the performance of the classification results. In all cases, the highest improvement is achieved when the HRBFN method is used while the SG filter performs the second best.

When the dielectric rods, which represent the target set with different material properties, and small scale aircraft models, which represent the targets having different geometries are considered, the contribution of the SG filter becomes more apparent. Although, the results show that the HRBFN is superior to other, the SG filter results are the closest to HRBFN on improving the classification performance (Table II).

Overall, the results for all target types (i.e. targets having different sizes, geometries and dielectric) show that, under low SNR conditions, the employment of de-noising techniques prior to feature extraction and classification process significantly improves the correct classification rates. Although the coarse-to-fine approach of HRBFN seems to outperform the other methods, SG filter performs significantly close to HRBFN and provide superior performance compared to wavelets and MP. Together with the advantages of convergent decomposition and real time processing, SG can be a useful alternative to HRBFN especially when the scattered signal acquisition is long.
Table 1. Target recognition results of spherical targets (Accuracy %)

| Raw Signal | SNR Level | 0 dB | -10 dB |
|------------|-----------|------|--------|
|            | S-1  | S-2  | S-3  | S-4  | S-1  | S-2  | S-3  | S-4  |
| D.C.       | Waveform Classification (WFC) | 97.4 | 95.1 | 92.9 | 95.2 | 74.2 | 66.4 | 62.6 | 72.3 |
| DWT        | 93.4 | 84.6 | 81.8 | 88.6 | 72.2 | 63.6 | 60.4 | 66.1 |
| DWT+WFC    | 94.5 | 93.8 | 91.1 | 94.7 | 81.9 | 77.8 | 69.4 | 72.9 |
| HRBFN+WFC  | 96.9 | 94.7 | 91.8 | 95.0 | 88.4 | 82.8 | 83.0 | 87.4 |
| MP+WFC     | 95.4 | 93.6 | 92.9 | 95.4 | 86.7 | 80.1 | 80.4 | 82.3 |
| SG+WFC     | 95.1 | 93.9 | 93.2 | 94.9 | 85.1 | 81.3 | 82.0 | 80.8 |
| DWT+SG     | 93.3 | 89.8 | 79.3 | 86.9 | 75.7 | 64.2 | 57.0 | 68.3 |
| DWT+MP     | 94.0 | 90.5 | 80.7 | 85.3 | 83.6 | 72.4 | 69.1 | 77.2 |
| DWT+SG+WFC | 94.4 | 91.3 | 81.2 | 85.8 | 81.3 | 71.5 | 66.9 | 74.8 |
| DWT+HRBFN  | 93.9 | 90.4 | 81.6 | 86.0 | 82.9 | 70.5 | 67.4 | 75.7 |
| DWT+HRBFN  | 93.9 | 90.4 | 81.6 | 86.0 | 82.9 | 70.5 | 67.4 | 75.7 |

Table 2. Target recognition results of dielectric rods and model aircrafts (Accuracy %)

| SNR | WFC after MSA Based Recovery | Dielectric Rods | Small Scale Aircrafts |
|-----|-------------------------------|-----------------|-----------------------|
|     |                               | ε_r= 3  | ε_r= 4  | ε_r= 5  | ε_r= 5.5| B-747 | DC-10  | B-767 |
| 0 dB| DWT                           | 85.8    | 80.5    | 82.4    | 82.2    | 84.4  | 90.0   | 87.4  |
|     | HRBFN                         | 88.4    | 83.1    | 84.8    | 82.8    | 85.9  | 91.2   | 87.5  |
|     | MP                            | 87.8    | 78.9    | 81.3    | 84.4    | 83.2  | 92.0   | 86.1  |
|     | Savitzky-Golay                | 88.1    | 81.8    | 82.7    | 82.3    | 85.7  | 90.9   | 87.2  |
| -10 dB| DWT                          | 82.8    | 70.6    | 71.7    | 70.7    | 76.7  | 82.8   | 78.5  |
|     | HRBFN                         | 86.2    | 81.5    | 81.3    | 80.4    | 84.7  | 90.0   | 86.8  |
|     | MP                            | 80.4    | 72.5    | 74.3    | 71.1    | 78.5  | 87.4   | 80.0  |
|     | Savitzky-Golay                | 84.8    | 82.1    | 81.8    | 78.6    | 82.3  | 88.1   | 83.5  |

Figure 6. (a) Noisy scattered signal (b) de-noised signal using SG filter (Electric Field (V/m) vs time)
References

[1] Secmen M and Turhan-Sayan G 2009 *IET Radar Sonar and Navigation* 3 583-595
[2] Turhan-Sayan G 2005 *IEEE Trans. On Antennas and Propagation* 53 766-776
[3] M. A. Selver, M. M. Taygur, M. Secmen and E.Y. Zoral, “Hierarchical Reconstruction and Structural Waveform Analysis for Target Classification,” *IEEE Transactions on Antennas and Propagation*, 64(7), pp. 3120-3129, July 2016.
[4] Mallat S G 1999 *A Wavelet Tour of Signal Processing* San Diego CA Academic Press 2nd ed.
[5] Savitzky, A.; Golay, M.J.E. (1964). "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". Analytical Chemistry. 36 (8): 1627–39. doi:10.1021/ac60214a047.
[6] G. Turhan-Sayan, K. Leblebicioglu, S. Inan “Input signal shaping for target identification,” *Microwave and Optical Technology Letters*, vol. 17, no. 2, pp. 128-132, Feb. 1998.
[7] G. Turhan-Sayan, K. Leblebicioglu, T. Ince, “Electromagnetic target classification using time-frequency analysis and neural networks,” *Microwave Opt. Tech. Lett.*, vol. 21, no. 1, pp. 63-69, April 1999.
[8] Selver, M. A., Secmen, M., & Zoral, E. Y. (2017, March). Comparison of scattered signal waveform recovery techniques under low SNR for target identification. In *Antennas and Propagation (EUCAP), 2017 11th European Conference on* (pp. 1091-1095). IEEE.
[9] Raudis S J and Jain A K 1991 *IEEE Trans. On PAMI* 13 252-264
[10] Haykin S 1998 *Neural Networks: A Comprehensive Foundation* Upper Saddle River NJ: Prentice Hall 2nd ed.