NONSTATIONARY SIGNAL CLASSIFICATION USING PSEUDO POWER SIGNATURES

Vidya Venkatachalam and Jorge L. Aravena

Department of Electrical and Computer Engineering,
Louisiana State University, Baton Rouge, LA 70810, USA.
Phone: (504) 388-5537, Fax: (504) 388-5200.
email: aravena@gate.ee.lsu.edu

ABSTRACT

This paper deals with the problem of classification of nonstationary signals using signatures which are essentially independent of the signal length. We develop the notion of a separable approximation to the Continuous Wavelet Transform (CWT) and use it to define a power signature. We present a simple technique which uses the Singular Value Decomposition (SVD) to compute such an approximation, and demonstrate through an example how it is used to perform the classification process. This example serves to show both the effectiveness and the limitations of the approach. Our main result is an alternate approach which develops the idea of using orthogonal projections to refine the approximation process, thus allowing for the definition of better signatures.

1. INTRODUCTION

This research is motivated by a classification problem, common in non-intrusive subsurface exploration, and introduced here as an event detection situation: *There is a known class of events, $\{C_k; k = 1, \ldots, n\}$, which may appear in a given scene for a variable time interval. Using a probe one collects data about the scene. The objective is to analyze the probe signal to determine what events are present and the duration of the occurrence of each of these events.*

In this paper we consider a simple case. Suppose one has collected the data as a signal $x(t); t_l \le t \le t_h$, and it is known that only one event may be present at any given time. Then there is an unknown partition $P_x = \{t_l \le t_1 \le t_2 \ldots \le t_r \le t_{r+1} \ldots \le t_h\}$, of *transition times* marking the start and end times of an event. Our goal is to determine the transition times and the events occurring in each time interval. This process is called *classification of the signal* $x(t)$.

The Short Time Fourier Transform ($STFT$) might be considered as a possibility to determine the type of event taking place, but it has resolution limitations in determining the transition times. For this reason, we consider the CWT as the tool for the classification process. However, the fact that each event may have an unknown time support presents difficulties in the conventional classification using time-frequency distributions, motivating this research.

2. PSEUDO POWER SIGNATURES

In this section, we introduce a methodology for signal classification that is essentially **independent** of the actual duration of each event. We achieve this objective by using the concept of spectral energy distribution and developing a representation that allows us to define an "instantaneous energy distribution" which we call *pseudo power signature.*

Consider any $x \in L^2(\Re)$ with CWT, $c_\psi^x(a,b)$, where ψ is an admissible wavelet; i.e., $C_\psi = 2\pi \int_\omega \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty$, where $\Psi(\omega)$ is the Fourier Transform of the wavelet. It is well known ([1]) that the associated scalogram can be interpreted as a time-scale energy density function since one can write

$$\int |x(t)|^2 dt = C_\psi^{-1} \int_b \int_a |c_\psi^x(a,b)|^2 \frac{da db}{a^2}$$

Hence, the function

$$P_\psi^x(a,b) = \frac{|c_\psi^x(a,b)|^2}{a^2}$$

can be viewed as the corresponding time-scale power density function and the function $P_\psi^x(\cdot, b)$ as the "scale power distribution at time b". Notice that one can estimate this function using the mean value theorem as follows

$$(b_2 - b_1) P_\psi^x(a, b_1, b_2) = \int_{b_1}^{b_2} \frac{|c_\psi^x(a,b)|^2}{a^2} db$$

If the wavelet has compact support, then as the scale decreases, the value of $P_\psi^x(a, b_1, b_2)$ is essentially independent of the values of $x(t)$ outside the interval $b_1 \le t \le b_2$. Thus, the scale power distribution can be estimated by the scalograms of small segments of the signal, and for low scales, it would be essentially independent of the length of the record. The lower the scale that one can use, the smaller the segments that are required.

The ideal situation would arise if one could define a wavelet such that for a given class of signals the corresponding wavelet transforms are separable, i.e.

$$c_\psi^x(a,b) = s_\psi^x(a) r_\psi^x(b)$$

Then the power distribution would always be proportional to $|\frac{s(a)}{a}|^2$ and one could use a suitably normalized version to characterize the signal in a manner that is independent of duration.

Unfortunately, such wavelets do not exist. [1] It is known that in order to be the CWT of an L^2 signal, the function, $c_\psi^x(a,b)$, must belong to a closed subspace, M, of the Hilbert space $H = L^2(\Re^2, C_\psi^{-1}\frac{da\,db}{a^2})$ [2]. We show in [3] that functions of the form $s(a)r(b)$ do exist in the space H but it is impossible to have an element of this form in the closed subspace M. This leaves us with the problem of finding suitable separable *approximations* to the CWT of the form $s_\psi^x(a)r_\psi^x(b)$. We denote the function $s_\psi^x(a)$ as the *pseudo power signature* of x. In this paper, we present two approaches to solve this problem. Note that this representation is dependent on the wavelet used, and so a related problem is to determine the wavelet that provides the most discriminating signatures for a given class of signals.

3. SVD APPROACH

Our first approach to the generation of pseudo power signatures uses the decomposition of the CWT as a sum of separable terms. This decomposition is the natural extension of the SVD analysis, and effectively determines the closest separable approximation in H to the CWT given by $c_\psi^x(a,b)$. This is based on the following :
The CWT can always be expressed as

$$c_\psi^x(a,b) = \sum_i \sigma_i s_i(a) r_i(b)$$

where $s_i(a) \in S = L^2(\Re, C_\psi^{-1}\frac{da}{a^2})$, *and* $r_i(b) \in R = L^2(\Re, db)$ *for each* i.

This result follows from the observation that H is isomorphic to the tensor product of the above two Hilbert spaces, i.e. $S \otimes R$. Suppose ϕ_k and ψ_l are orthonormal bases, ONB, for each of the two Hilbert spaces S and R respectively. Then, the collection $\phi_k \psi_l$ is an ONB for the Hilbert space H ([4]), and we can write

$$c_\psi^x(a,b) = \sum_{k,l} \alpha_{kl} \phi_k(a) \psi_l(b)$$

Using the notation $A = [\alpha_{kl}]$, $\Phi = [\phi_k]$, $\Psi = [\psi_l]$, we get

$$c_\psi^x = \Phi A \Psi^T$$

Applying the SVD to the matrix, A, the result can be easily derived.
For all practical purposes, c_ψ^x has compact support in the time-frequency plane. For a suitably chosen ONB, (where the basis elements themselves have compact support), this implies that $\exists \;\; m,n$ such that $\alpha_{kl} = 0, \; \forall \;\; k > m$ and $\forall \;\; l > n$. Hence the SVD problem can be made finite dimensional. Observe that the functions s_i and r_i thus defined belong to the Hilbert spaces S and R, respectively. The term corresponding to the principal component, namely $s_1(a)$, is used to define the pseudo power signature of x.

3.1. Simulation Results

Here we present computer results showing the application, and limitations, of the SVD approach to some artificially generated signals. The signals are the simple modulated sinc functions $\{x1, x2, x3\}$ shown in Fig. 1. Their frequency spectra $\{f1, f2, f3\}$ (the axis is expressed as a fraction of π) and their pseudo power signatures $\{S1, S2, S3\}$ (the axis is expressed as a logarithmic function of the scale on a dyadic grid) are also shown in the same figure. These signatures were generated using the $Db4$ wavelet. [2] We used Shensa'a algorithm ([5]) to compute the discretized CWT coefficients with the scale varying on a dyadic grid. The pseudo power signatures were then readily obtained by taking the principal component of the SVD of the coefficient matrix.

Now consider a signal created by concatenating segments of each signal class: $x1$ over the interval [-125:-50], $x2$ over the interval [-50:50] and $x3$ over the interval [50:115]. The composite signal, its $STFT$, and its discretized CWT are shown in Fig. 2. Observe that merely examining the signal, its $STFT$, or the CWT is not sufficient to identify either the component signals or the transition points. Furthermore, direct comparison of the $CWTs$ of each signal class with the CWT of the composite signal is also not feasible because the CWT support is dependent on the signal duration which is, in general, unknown. For classification purposes, we need a representation which is more intrinsic to each signal class, and is independent of the signal support. These conditions are satisfied by the power signatures shown in Fig. 1.

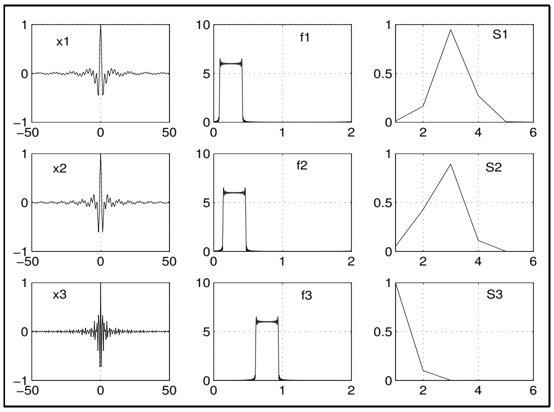

Figure 1: The 3 signal classes and their corresponding signatures

[1] If one moves away from L^2 signals, one can find functions whose formal CWT is separable. Consider the power signal $x(t) = Ae^{-j\theta t}$. If $\psi(t)$ is an admissible wavelet with Fourier transform, $\Psi(\omega)$, the function $c_\psi^x(a,b) = \int x(t)\frac{1}{\sqrt{a}}\overline{\psi(\frac{t-b}{a})}dt$ is defined for all values of $a \neq 0, b \in R$. Observe then that $c_\psi^x(a,b) = A\sqrt{a}\overline{\Psi(a\theta)}e^{j\theta b} = s(a)r(b)$.

[2] This is one of Daubechies' compact support wavelets, and is defined through a two scale equation with 8 coefficients.

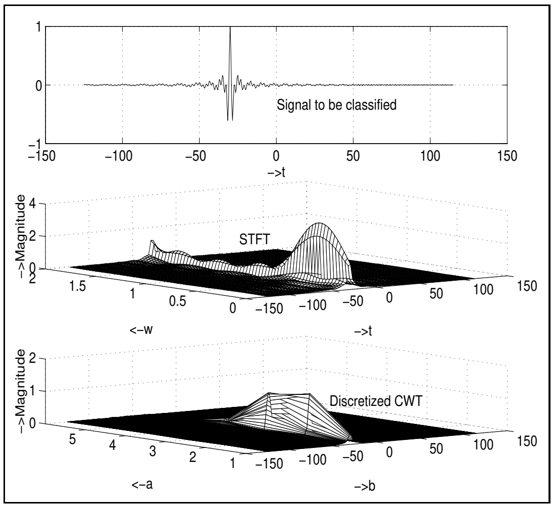

Figure 2: The signal, its $STFT$, and its CWT

In our first approach, we divided the CWT time axis into 10 segments $\{p1, p2, ..., p10\}$, obtained the principal component Pj of each segment, and determined its correlation with each Si. The result is shown in Fig. 3, where $C1, C2, C3$ represent the correlation graphs of $S1, S2, S3$ with the principal components of each segment. From these graphs, the only unambiguous conclusion we may draw is that $x3$ is present in segments $8, 9, 10$. From the nature of the problem, where it is known that each of the 3 signal classes is present, and that no two classes are present at the same time, we may also conclude that $x1$ is present in segments $1, 2, 3$, and $x2$ in segments $4, 5, 6, 7$. Though we could identify and establish approximately the duration of each signal present using this approach, there is some ambiguity about the validity of the classification, particularly with respect to the identification of the highly similar signal classes $x1$ and $x2$.

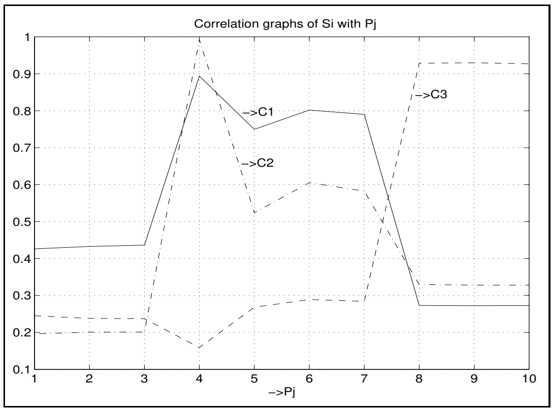

Figure 3: Correlation graphs of each signature with the principal components of the different segments

In order to get a more accurate picture of the signal composition, we determined the correlation of each Si with the discretized CWT of the composite signal for each b. The results are presented in Fig. 4. Observe that the results show quite clearly that there are 2 transition points in the signal, (the first around -50, and the second around 50), a situation which is not very evident upon examination of the signal. Here, we can make the legitimate assumption that the correlation values must remain fairly constant over a range for the signal to be classified as having support in that range. Again, based on our underlying assumptions, we can conclude from the graphs that the support of $x1$ is $[-125 : -50]$, that of $x2$ is $[-50 : 50]$, and that of $x3$ is $[50 : 115]$. Note that we disregarded the high correlation values of $S1$ in the range $[-50 : 50]$ because we assume in this classification that only one signal can be present at any given time, and $S2$ has a higher correlation in that range than $S1$, and is more likely to be present in the range $[-50 : 50]$ than anywhere else.

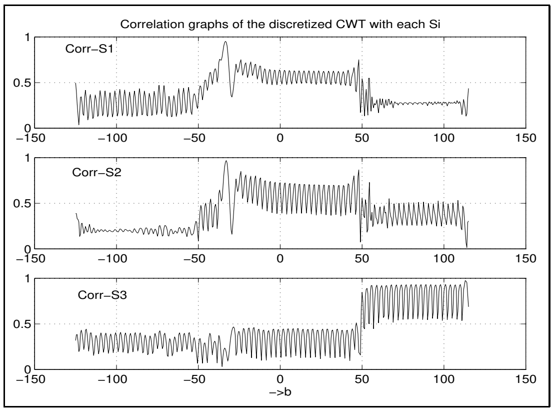

Figure 4: Correlation graphs of the discretized CWT

It is clear from the results presented that the simplistic process of taking the principal component of the SVD of c_ψ^x as the power signature of a signal class can, at times, lead to ambiguous interpretations. This example shows that while power signatures are indeed more discriminating than the Fourier spectra, and more robust than the CWT, we need a more sophisticated technique to find pseudo power signatures which can capture even fine distinctions between different signal classes.

4. PROJECTOR APPROACH

We conjecture that the principal component does not create the best signature because the separable element in H is not a wavelet transform. Hence, we propose to create signatures by finding *the separable approximation in H whose projection onto M is the closest to a given CWT*. This development is based on the following.

We have shown ([3]) that there exists an orthogonal projector, $\mathcal{K}$, in H whose range is M. This projector is defined as follows. Given any $c \in H$,

$$\mathcal{K}\left[c\right](a,b) = C_\psi^{-1} \int_\alpha \int_\beta \overline{c_\psi^{\psi_{ab}}(\alpha,\beta)} c(\alpha,\beta) \frac{d\beta d\alpha}{\alpha^2}$$

Moreover, if the wavelet transform is characterized as a map $\Gamma : L_2 \longrightarrow H$, then $\mathcal{K} = \Gamma\Gamma^*$. *Consider now a function* $c(a,b) \in H$ *given by* $c(a,b) = s(a)r(b)$, *where* $r(b) \in L^2(\Re, db)$ *and* $s(a) \in L^2(\Re, C_\psi^{-1}\frac{da}{a^2})$. *There exists a function* $\hat{x}(t) \in L^2$ *whose* CWT *is given by the projection of* $c(a,b)$ *on the closed subspace* M, *i.e.* $c_\psi^{\hat{x}}(a,b) = \mathcal{K}\,[c]\,(a,b)$.

This result associates to every separable element of H a unique CWT in M. Conversely, for a given signal $x \in L^2$ and its corresponding CWT, c_ψ^x, we can associate with it an element, $\hat{c}_\psi^x(a,b) = s_\psi^x(a)r_\psi^x(b) \in H$ such that $\mathcal{K}\left[\hat{c}_\psi^x\right](a,b)$ is *as close* as possible to the wavelet transform of x. The function $s_\psi^x(a)$ is denoted as the *pseudo power signature* of x. If there is more than one separable element with the same projection, we guarantee uniqueness in the signature by using the element with the minimal norm in H.

The solution to determining this unique signature lies in solving the following optimization problem:
For a given $c_\psi^x \in M$, *find the decomposition* $s_\psi^x r_\psi^x \in H$ *that minimizes the index*

$$J(s_\psi^x, r_\psi^x) = \min \parallel c_\psi^x - \mathcal{K}[s_\psi^x r_\psi^x] \parallel_M^2 + \parallel s_\psi^x r_\psi^x \parallel_H^2$$

where $\mathcal{K}$ *is the orthogonal projection operator defined earlier.*

We claim that such a pair $s_\psi^x(a)r_\psi^x(b)$ is effectively the best separable approximation to the CWT of x. In effect, the $\hat{x} \in L^2$ associated with it minimizes the norm $\parallel x - \hat{x} \parallel_2^2$. Hence, we can expect that it will better represent the intrinsic properties of the signal x.

For computational feasibility we need to reduce this infinite dimensional nonlinear minimization problem to a finite dimensional one. The most difficult step is the computation of the projected value, $\mathcal{K}[sr]$, of a given separable element. We solve this problem by applying the inverse Shensa algorithm to a suitable discretization of the vector, $s(a)r(b)$, followed by application of the Shensa algorithm to the resulting sequence. The discretization can be interpreted as restricting the separable elements $s(a)r(b)$ to be *piecewise constant* on the plane. We have shown in [3] that the approach computes exactly the samples of a continuous wavelet transform. Hence the only approximation lies in assuming that elements in H can be well approximated by piecewise constant functions. Formally, this approach assumes that any element $c \in H$ can be approximated as

$$c(a,b) \approx \sum_{m=1}^{M}\sum_{n=1}^{N} c(2^m, n) q_{m,n}(a,b)$$

where

$$q_{m,n}(a,b) \quad = \quad \begin{cases} 1 & 2^m \leq a < 2^{m+1}, n \leq b < n+1 \\ 0 & elsewhere \end{cases}$$

Based on this assumption, we have also developed a computational algorithm to solve this nonlinear minimization using iterative techniques. The algorithm guarantees that at each step the cost function is reduced. Our initial implementations of the algorithm are not yet efficient. We are currently working on a more efficient technique to solve this nonlinear optimization problem. We propose to compare the performance of the SVD approach with this approach to the creation of the pseudo power signatures in terms of the discriminating capability of each, computational complexity, and the ease of implementation.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we introduced the idea of signal signatures which are essentially independent of the signal length. The determination of such signatures is based on using separable approximations to the CWT of the signal. First, we presented a simple approach using the Singular Value Decomposition to generate these signatures. We tested this approach on several examples with good results. However, these signatures were limited by a lack of fine discriminating capability as was demonstrated through the example shown in this paper.
Next, we proposed a more sophisticated method to create signatures. This is the projector approach which essentially involves solving an inverse projection problem. This method gives a much better approximation according to our preliminary results.
An interesting empirical observation made in both the approaches was that better performance is obtained when the wavelet used to generate the signatures matches the signals as closely as possible.
In conclusion, we have formulated a concept which is very useful for signal classification problems and permits the separation of highly correlated signals. Moreover, the signatures are vectors of small dimension. It appears feasible to extend the concept and define signatures for *classes of signals*. This has potential applications in areas like oil exploration, target detection, objection recognition, and system identification.

6. REFERENCES

[1] O. Rioul and M. Vetterli," Wavelets and Signal Processing", Signal Processing Magazine, vol.8, pp. 14-38, Apr. 1991.

[2] I. Daubechies, " Ten Lectures on Wavelets", Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[3] Vidya Venkatachalam and Jorge. L. Aravena, "Nonstationary Signal Analysis using Pseudo Power Signatures", In preparation (available from authors).

[4] M. Reed and B. Simon, " Methods of Modern Mathematical Physics 1 : Functional Analysis", Academic Press, Inc., San Diego, 1980.

[5] Mark J. Shensa, " The Discrete Wavelet Transform: Wedding the À Trous and Mallat Algorithms", IEEE Transactions On Signal Processing, vol. 40, pp. 2464-2482, Oct. 1992.