International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:7, 2008

# Classification of Non Stationary Signals Using Ben Wavelet and Artificial Neural Networks

Mohammed Benbrahim, Khalid Benjelloun, Aomar Ibenbrahim, and Adil Daoudi

***Abstract*—**The automatic classification of non stationary signals is an important practical goal in several domains. An essential classification task is to allocate the incoming signal to a group associated with the kind of physical phenomena producing it. In this paper, we present a modular system composed by three blocs: 1) Representation, 2) Dimensionality reduction and 3) Classification. The originality of our work consists in the use of a new wavelet called "Ben wavelet" in the representation stage. For the dimensionality reduction, we propose a new algorithm based on the random projection and the principal component analysis.

***Keywords*—**Seismic signals, Ben Wavelet, Dimensionality reduction, Artificial neural networks, Classification.

## I. INTRODUCTION

THE classification of non stationary signals is a difficult and much studied problem. On one hand, the non-stationarity precludes classification in the time or frequency domain; on the other hand, nonparametric representations such as time-frequency or time-scale representations, while suited to non-stationary signals, have high dimension. In order to overcome these problems, several works have been developed, we note: [1], [2], [3].

In this paper, we present a modular system for the classification of seismic signals. Three blocs compose this system (see Fig. 1): **Representation, Dimensionality reduction, Classification**. The advantage of this system is the ability to profit from the existent evolutions of each module in any time.

This work was supported by PROTARS III (CNRST, Rabat, Morocco) under Grant D48.

M. Benbrahim is with the Ecole Mohammadia d'Ingénieurs, Rabat, Morocco (corresponding author to provide phone: 212-63631938; fax: 212-55633878; e-mail: benbrahim10@yahoo.fr).

K. Benjelloun is with the Ecole Mohammadia d'Ingénieurs, Rabat, Morocco (e-mail: bkhalid@emi.ac.ma).

A. Ibenbrahim is with the Centre National pour la Recherche Scientifique et Technique Rabat, Morocco (e-mail: ibenbrahimr@cnr.ac.ma).

A. Daoudi is with the Régie Autonome de Distribution de l'eau et de l'électricité de Marrakech, Marrakech, Morocco (e-mail: adildaoudi@menara.ma).

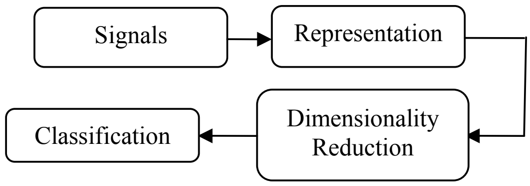


FIGURE 1: SYSTEM DIAGRAM

The remainder of this paper is structured as follows: first, we present the different representations for non-stationary signals and we propose a new complex wavelet. We then describe our algorithm for the part dimensionality reduction. The next paragraph is dedicated to classification by multilayer perceptron network. Finally, we discuss the results and the conclusions.

## II. NON STATIONARY SIGNALS REPRESENTATIONS

The choice of the representation is an important parameter that it must be chosen carefully in order to increase the classification performances. It consists to define a representation space permitting the extraction of the pertinent information.

For non-stationary signals, all previous works have highlighted that a representation space, where the power, the frequency and the time are .present, is an adequate space. We can replace the frequency parameter by the scale parameter that permits a multiresolution analysis of the signal.

### A. Time and Frequency Representations

The **time representation** is the natural form to represent a signal deriving from a given phenomena. It is not necessary to use any mathematical tool to perform it and we can have some information about signal from it. However, it is not adapted for the automatic classification.

The **frequency representation** is an alternative for the temporal representation of looking at a signal. It consists to represent the frequency content of the signal via the Fourier transform. In the non-stationary signals case, the frequency contents and all the statistic properties change with the time. Consequently, for the events with weak signal noise ratio, the classification based on the Fourier transform can give wrong results. Moreover, this representation limits the generalization of the automatic classification system for other classes where frequency content is similar.

### B. Time-Frequency Representations

In order to overcome the limits posed by the temporal and the frequency representations, the use of a time-frequency representation (**TFR**) provides localized information in time and frequency simultaneously. This representation gives a natural description for the non-stationary signals such as the seismic signals. Indeed, TFRs characterize signals over a time-frequency plane. They thus combine time-domain and frequency domain analyses to yield a potentially more revealing picture of the temporal localization of a signal's spectral components.

Several TFRs exist in the literature and they can be divided in three types [4]: linear, quadratic and non-linear non-quadratic representations. Moreover, when satisfying some properties, these representations can be divided in classes [5]. We cite the Cohen class ([6], [7]), the affine class ([4], [8], and [9]), the hyperbolic class ([10], [11]) and the power class ([10], [12]). However, which is the better representation for seismic signals between them?.

Because there is not any universal solution for all signals, we must necessary to do a choice based on the mathematical properties of the representation and its utility for a given signal. In the experimental part, we use the spectrogram to represent the seismic signals in the time-frequency space.

### C. Time-Scale Representations

The techniques based on windowed Fourier transform represent inaccurate and inefficient methods of time-frequency localization, as they impose a fix size of the analysis window.

A direct way to overcome the problems with a fixed window size is to use a time-scale representation (TSR). As the TFRs, the TSRs can be divided in linear and quadratic representations. For the linear case, we find the wavelets and for the quadratic case, the affine class is the most important class of the covariant TFRs ([4], [8], and [9]).

For the continuous wavelet transform, there are two popular functions: the Mexican hat wavelet and the Morlet wavelet [13]. The first is a real function and because it is the second derivative of the Gaussian function, it is most adapted to detect discontinuities in signals. The second wavelet is complex valued, enabling one to extract information about the amplitude and phase of the signal being analyzed [14].

For the non-stationary signals, in order to profit of the intrinsic properties of the Mexican hat wavelet and Morlet wavelet, we consider the new complex wavelet called the **Ben wavelet** [15] (see Fig. 2 for the time representation of the three wavelets):

$$\psi(t) = \frac{2}{\sqrt{3}}\pi^{-1/4}\left(1-t^2\right)e^{-t^2/2}\,e^{i\omega_0 t} \qquad (1)$$

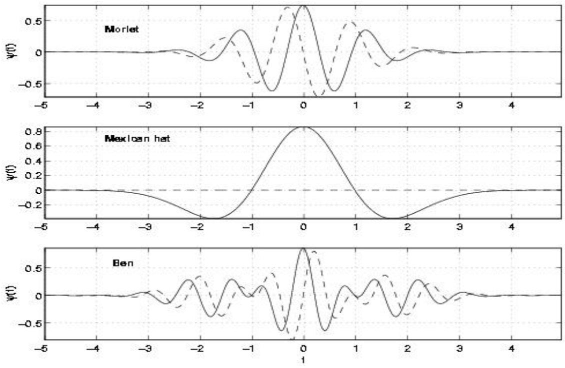


FIGURE 2: REAL PART (SOLID LINES) AND IMAGINARY PART (DASHED LINES) OF THE MORLET W AVELET, THE MEXICAN HAT WAVELET AND THE BEN WAVELT RESPECTIVELY

For representing the non-stationary signals in time-scale space, we use the scalogram of this wavelet. The scalogram is defined as the squared magnitude of the wavelet transform of the signal considered (see the Fig. 3 for the scalogram of an earthquake using the three wavelets).

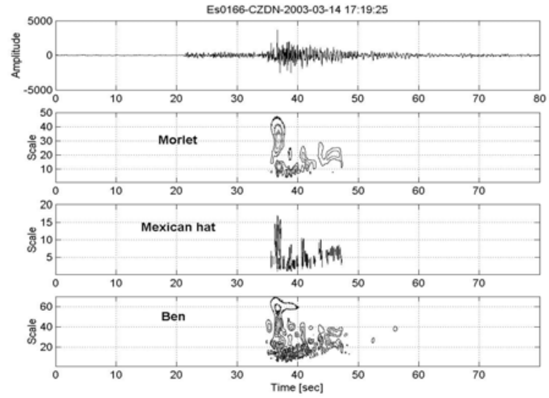


FIGURE 3: TIME REPRESENTATION OF AN EARTHQUAKE AND THE CORRESPONDING T SCALOGRAMS FOR MORLET, MEXICAN HAT AND BEN WAVELETS RESPECTIVELY

## III. DIMENSIONALITY REDUCTION

The bidimensional representations of seismic signals by TFRs and TSRs give high dimensional images. For a system of automatic classification, in order to eliminate the problems due to high dimensional data such as the curse of dimensionality [16], the dimensionality stage must be integrated in the system.

For the case of seismic signals, the dimension of the bidimensional representations is variable in the temporal axe because the length of time of the seismic events is variable. In the precedents works, in order to obtain images with the same length, the method used consists to clip the signal. But it is possible to loose the pertinent information in the ignored part. To overcome this problem, we use a new algorithm based on the combination of the random projection (**RP**) and principal component analysis (**PCA**).

### A. Random Projection

In RP, the original d-dimensional data is projected to a k-dimensional (k<<d) subspace through the origin, using a random k×d matrix *R* whose columns have unit lengths. Using matrix notations where $X_{d\times N}$ is the original set of N d-dimensional observations,

$$X^{RP}_{k\times N} = R_{k\times d} X_{d\times N} \quad (2)$$

is the projection of the data onto a lower k-dimensional subspace. The key idea of random mapping arises from the Johnson-Lindenstrauss lemma [17]: If points in a vector space are projected onto a randomly selected subspace of suitable high dimension, then the distances between the points are approximately preserved.

RP is computationally very simple: forming the random matrix *R* and projecting the d×N data matrix into k dimensions is of order *O(dkN).*

### B. Principal Component Analysis

PCA is a widely used dimensionality reduction technique in data analysis. Its popularity comes from two important properties. First, it is the optimal (in terms of mean squared error) linear scheme for compressing a set of high dimensional vectors into a set of lower dimensional vectors and reconstructing ([18], [19]). Second, the model parameters can be computed directly from the data. Indeed, dimensionality reduction by PCA consists of projecting data onto a subspace spanned by the most important eigenvectors:

$$X^{PCA}_{k\times N} = E^{T}_{d\times k} X_{d\times N} \quad (3)$$

where the $E^{T}_{d\times k}$ is the transpose of the d×k matrix $E_{d\times k}$ that contains the k eigenvectors corresponding to the k largest eigenvalues of the data covariance matrix.

PCA computationally very expensive compared to RP. The computational complexity of estimating the PCA is $O(d^2N)+O(d^3)$.

### C. Algorithm

- Step1: Normalization of the TFR or TSR images,
- Step2: Reduction of dimensionality of each image by random projection,
- Step3: Calculus of the power mean for each frequency or scale level,
- Step4: Feature extraction by principal component analysis.

Because the PCA is computationally expensive, applying RP before PCA in this algorithm, permit to overcome the problems posed by the PCA.

## IV. Classification

Multilayer Perceptron networks (MLP) are artificial neural networks formed of cells simulating the low level functions of neurons. MLP networks are very useful for classification of input signals where the signals cannot be defined mathematically. Further, MLP networks have redundant networking and are very robust, providing a mathematical flexibility not available to algorithms based classifiers.

Seismic signals can mathematically be defined as chaotic signals and therefore suited for artificial neural networks classifiers. MLP network responds to an input by producing an output. This is a result of the transmission of the input through the network of neurons linked by weights. The output of the MLP network is a combination of outputs of each of the neurons in the output stage of the MLP.

Before the MLP can be used for classification, it has to be trained during the time when it learns of the input/output relationship for training vector set. During the training cycle, MLP is given sets of input patterns and corresponding target outputs representing the training vector.

## V. Experimental results

In order to demonstrate the performances of our system, we consider two applications: the first concerns the discrimination between local earthquakes and chemical explosions and the second application concerns the classification of leaks in water distribution pipes.

TABLE I
Experimental Results

| Representation | Percentage of correct classification |
|---|---|
| Scalogram of Mexican hat wavelet | 82.50% |
| Scalogram of Morlet wavelet | 86.25% |
| Scalogram of modified Mexican hat wavelet | 88.75% |
| Spectrogram | 85.% |

### A. Classification of seismic signals

To discriminate between local earthquakes and chemical explosions, we used a data set of 90 events (45 local earthquakes and 45 explosions) detected by the same station and transmitted to the geophysics laboratory of the National Center for Scientific and technical research in Morocco. These data are used to perform the training set (30 local earthquake and 30 explosions) and the test set (15 local earthquakes and 15 explosions).

For the artificial neural network, we used a MLP network with architecture of 20-7-2 (i.e. 20 input nodes, 7 hidden nodes and 2 output nodes). This network was trained with the scaled conjugate gradient algorithm. The learning rate, after a series of trial and error processes, was set to η=0.001.

TABLE II
Leak signal classification set up

| Cycle | Iron Pipe | PVC Pipe | *Total* |
|---|---|---|---|
| Learning | 13 | 21 | 34 |
| Testing | 8 | 13 | 21 |
| *Total* | 21 | 34 | **55** |

We obtain using this model in average [20]:

TABLE II
LEAK SIGNAL CLASSIFICATION SET UP

| Cycle | Iron Pipe | PVC Pipe | *Total* |
|---|---|---|---|
| Learning | 13 | 21 | 34 |
| Testing | 8 | 13 | 21 |
| *Total* | 21 | 34 | **55** |

The network architecture was chosen after a series of trial and error processes. This means that the obtained performances are not optimal in classification. Indeed, for a network with a weak number of hidden layers, it is not possible to reach the optimal performances and for a network with high number of hidden layers, it has not performances in generalization. For this application, with architecture of 20-1-2 or 20-50-2, we obtain bad results. For solving this problem, we can use incremental neural networks or to combine several systems with different architectures.

TABLE III
RECOGNITION PERFORMANCES

| Test | Iron Pipe | PVC Pipe | *Total* |
|---|---|---|---|
| Tested | 8 | 13 | 21 |
| Correct | 7 | 9 | 16 |
| *Percentage* | 87,5% | 69% | **76%** |

### B. *Classification of leaks signals*

For leaks signals, we consider a data set of 34 signals for the training set and 21 for the test as explained in the table 2:

And for the MLP architecture we use architecture of 30-5-2 (mean 30 input nodes, 5 hidden nodes and 2 output nodes). The learning rate was set to $\eta$=0.001.

We obtain using this model in average using the Ben wavelet in the representation stage [21]:

We note the low recognition rate for plastic pipes compared to iron pipes. It can be due to following reasons:

- leaks signal set for iron pipes is insufficient to build reliable pattern,
- several leaks are happened in confusion zones,
- Recognition rate is lower for plastic pipes; it is true, because of its characteristics concerning acoustic wave propagation: signal attenuation, low speed of sound, etc.
- External noisy sources, generally leak seeker have to choose convenient time to listen to pipes.

## VI. CONCLUSION

In this paper, we tested the performance of a new system well adapted to automatic discrimination of non-stationary signals and can be used for other similar situations. The principal advantage of this system is its modularity that permits to use different methods for each module.

Two new results were proposed. The first is a new wavelet well adapted to oscillatory signals with possible discontinuities. The second is a new algorithm for dimensionality reduction for high dimensional signals with different lengths.

## REFERENCES

[1] S. Abeysekera and B. Boashash, "Methods of signal classification using the images produced by Wigner distribution" Pattern Recognition Letters, 12, pp. 717-729, 1991.

[2] L. Atlas, J. Droppo and J. McLaughlin, "Optimizing time-frequency distributions for automatic classification" In the International Society for Optical Engineering, 1997.

[3] M. Davy, C. Doncarli, and G. F. Boudreaux-Bartels, "Improved optimization of time-frequency based signal classifiers" IEEE Sig. Proc. Let, 8, pp. 52-57, 2001.

[4] F. Hlawasch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations" IEEE Sig. Proc. Mag, 9, pp. 21-67, 1992.

[5] A. Papandreou-Suppappola, F. Hlawasch and G. Boudreaux-Bartels, "Quadratic time-frequency representations with scale covariance and generalized time-shift covariance: a unified framework for the affine, hyperbolic and power classes" Digital signal Processing, 8, pp.3-48, 1998.

[6] L. Cohen, "Generalized phase-space distribution functions" J. Math. Phys, 7, pp. 781-786, 1966.

[7] L. Cohen, "Time-frequency analysis" Prentice Hall, 1995.

[8] O. Rioul and P. Flandrin, "Time-scale energy distributions : a general class extending wavelet transforms" IEEE Trans on Signal Processing, 40, pp. 1746-1757, 1992.

[9] P. Flandrin, "Temps-fréquence", Academic Press, 1998.

[10] F. Hlawasch, A. F. Papandreou-Suppappola and G. Boudreaux-Bartels, "The power classes of quadratic time-frequency representations : a generalization of the hyperbolic and affine classes" In 27th Asilomar Conf on Signals, Systems and computers, Pacific Grove, CA, 1265-1270, 1993.

[11] F. Hlawasch, A. F. Papandreou-Suppappola and G. Boudreaux-Bartels, "The hyperbolic class of quadratic time-frequency representations. Part II: Subclasses, intersection with affine and power classes, regularity unitarity" IEEE Trans on Signal Processing, 45, pp. 303-315, 1997.

[12] A. Papandreou-Suppappola, F. Hlawasch and G. Boudreaux-Bartels, "Power class time-frequency representations: interference geometry, smoothing and implementation" In IEEE Symposium on Time-Frequency and Time-Scale Analysis, Paris, pp.193-196, 1996.

[13] I. Daubechies, "Ten lectures on wavelets", SIAM, Philadelphia, Pa, 1992.

[14] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis" Bull. Amer. Meteor. Soc, 79, pp. 61-78, 1998.

[15] M. Benbrahim, K. Benjelloun and A. Ibenbrahim, "Discrimination des signaux sismiques par réseaux de neurones artificiels" In Proc of 3èmes journées nationales sur les systèmes intelligents: théorie et applications, Rabat, Morocco, pp. 62-66; 2004.

[16] R. Bellman, "Adaptive control processes: A guided tour" Princeton University Press, Princeton, 1961.

[17] W.B. Johnson and J. Lindenstrauss, "Extensions of Lipshitz mapping into Hilbert space" In Conference in modern analysis and probability, volume 26 of Contemporary Mathematics, Amer. Math. Soc, pp. 189-206, 1984.

[18] I. T. Jolliffe, "Principal component analysis", Springer-Verlag, 1986.

[19] J. E. Jackson, "A user's guide to principal components", John Wiley, New York, 1991.
[20] M. Benbrahim, A. Daoudi, K. Benjelloun and A. Ibenbrahim, "Discrimination of seismic signals using artificial neural networks" In Proc of the second world enformatika congress, WEC'05, Istanbul, Turkey, pp 4-7, 2005.
[21] , A. Daoudi, M. Benbrahim and K. Benjelloun, "An intelligent system to classify leaks in water distribution pipes" In Proc of the second world enformatika congress, WEC'05, Istanbul, Turkey, pp 1-3, 2005.