Simposio Latinoamericano y del Caribe en Geofísica
Segundo Congreso Latinoamericano de Sismología
Tercer Congreso Colombiano de Sismología

DISSIMILARITY-BASED CLASSIFICATION OF SEISMIC SIGNALS AT NEVADO DEL RUIZ VOLCANO

Orozco-Alzate Mauricio[1,3], García-Ocampo Marcelo Enrique[2], Duin Robert P.W.[3], and Castellanos-Domínguez César Germán[1]

[1]Grupo de Control y Procesamiento Digital de Señales, Campus La Nubia, km 7 vía al Magdalena, Universidad Nacional de Colombia Sede Manizales. *E-mail: morozcoa@unal.edu.co, gcastell@telesat.com.co*

[2]INGEOMINAS - Observatorio Vulcanológico y Sismológico de Manizales, Avenida 12 de Octubre No. 15-47, Colombia. *E-mail: megarcia@ingeominas.gov.co*

[3]Information and Communication Theory Group, Mekelweg 4, 2628 CD Delft, Delft University of Technology, The Netherlands. *E-mail: r.p.w.duin@tudelft.nl*

ABSTRACT

Automatic classification of seismic signals has been typically carried out on feature-based representations. Recent research works have shown that constructing classifiers on dissimilarity representations is a more practical and, sometimes, a more accurate solution for some pattern recognition problems. In this paper, we consider Bayesian classifiers constructed on dissimilarity representations. We show that such classifiers, based on dissimilarities, are a feasible and reliable alternative for automatic classification of seismic signals. Our experiments were conducted on a dataset containing seismic signals recorded by two selected stations of the Nevado del Ruiz monitoring network. Dissimilarity representations were constructed by calculating pairwise Euclidean distances and a non-Euclidean measure on the normalized spectra, which is based on the area difference between spectral curves. Results show that even though Euclidean dissimilarities have advantageous properties, non-Euclidean measures can be beneficial for matching spectra of seismic signals.

Key words: Classification, dissimilarity, Ruiz, seismic, signals, volcano.

RESUMEN

La clasificación automática de señales sísmicas se ha llevado a cabo típicamente sobre representaciones de características. Trabajos de investigación recientes han mostrado que construir clasificadores sobre representaciones de disimilitud es una solución más práctica y, algunas veces, más precisa para ciertos problemas de reconocimiento de patrones. En este artículo consideramos clasificadores bayesianos construidos sobre representaciones de disimilitud. Mostramos que tales clasificadores, basados en disimilitudes, son una alternativa viable y confiable para la clasificación automática de señales sísmicas. Nuestros experimentos fueron llevados a cabo sobre una base de datos que contiene señales sísmicas registradas por dos estaciones seleccionadas de la red de monitoreo del Volcán Nevado del Ruiz. Las representaciones de disimilitud fueron construidas mediante el cálculo de distancias euclidianas y de una medida no euclidiana sobre los espectros normalizados, ésta última está basada en la diferencia de área entre curvas espectrales. Los resultados muestran que aunque las disimilitudes euclidianas tienen propiedades ventajosas, las medidas no euclidianas pueden resultar benéficas para comparar espectros de señales sísmicas.

Palabras claves: Clasificación, disimilitud, Ruiz, sísmica, señales, volcán.

INTRODUCTION

Nevado del Ruiz volcano is capped by a large volume of snow and ice, forming a glacier which has a volume of about 1200~1500 million cubic meters. Nevado del Ruiz has three craters: Arenas —the active one—, and two parasite craters: Olleta and Piraña. Since seismic activity has been digitally recorded by the Volcanological and Seismological Observatory at Manizales (VSOM), a large and increasing amount of data has been produced by the monitoring networks; such a database is suitable for applying automatic classification/learning techniques.

Classification of seismic signals is a crucial issue in order to discover the interaction between volcanic earthquakes and volcanic processes. In this study, we consider three classes of seismic signals originating from Nevado del Ruiz volcano: Volcano-Tectonic (VT) earthquakes, Long-

Period (LP) earthquakes and Icequakes (IC); of course, for every volcano, the seismologists use their own classification with more detailed description of every subtype of earthquakes (Zobin, 2003). VSOM staff currently classifies volcanic earthquakes by visual inspection; such a method supposes a great deal of workload for the seismic analysts. In consequence, an automatic classification tool dramatically reduces this arduous task and also turns classification reliable and objective, removing errors associated to tedious evaluations and changing of personnel.

Among the applications of pattern recognition techniques to seismic-volcanic signals, two recent works are remarkable: automatic classification of seismic signals at Mt. Vesuvius volcano, Italy (Scarpetta et al., 2005) and automatic classification of seismic events at Soufrière Hills volcano, Montserrat (Langer et al., 2006). Both of them propose the application of Artificial Neural Networks (ANN) to classify seismic events. In (Scarpetta et al., 2005), a multilayer perceptron (MLP) is used to distinguish between VT events and transient signals due to other sources such as underwater explosions, quarry blasts, and thunders; spectral features and amplitude parameters are used for characterization. In (Langer et al., 2006), an ANN is used to classify five fundamental classes of signals: VT events, regional (RE) events, LP events, hybrid (HB) events and Rockfalls (ROC); autocorrelation functions, high order statistical moments and amplitude ratios are introduced as features to the input nodes; a mismatch rate of 30% is reported, which was reduced up to 20% after a manual revision of the original a-priori classification. Typically, in the context of volcanic seismology, neural networks have been preferred rather than other classical statistical pattern recognition methods; they are still being used for discrimination of seismic signals, including modifications in the feature-based representation, e.g. the modified approach used in (Benbrahim et al., 2005). The popularity of neural networks models to solve pattern recognition problems has been primarily due to their seemingly low dependence on domain-specific knowledge and due to the availability of efficient learning algorithms (Jain et al., 2000).

Recently, a number of studies showed advantages of learning from dissimilarity representations instead of learning from feature-based representations (Duin et al., 1998; Pękalska et al., 2001; Pękalska and Duin, 2002; Paclík and Duin, 2003; Pękalska and Duin, 2005). A dissimilarity representation of objects, seismic events in our particular case, is based on pairwise comparisons and is expressed e.g. as an $N \times N$ dissimilarity matrix $D(T,T)$, where each entry corresponds to a dissimilarity between pairs of objects. Dissimilarity representations are more general than feature-based representations; in fact, the notion of dissimilarity is more fundamental than that of a feature (Pękalska and Duin, 2005). For dissimilarities the geometry is contained in the definition, giving the possibility to include physical background knowledge; in contrast, feature-based representations usually suppose a Euclidean geometry. This paper is devoted to explore dissimilarity representations to classify volcanic-seismic signals. Dealing with this particular problem, we advocate the dissimilarity-based classification of seismic signals as an advantageous and feasible alternative to the feature-based classification.

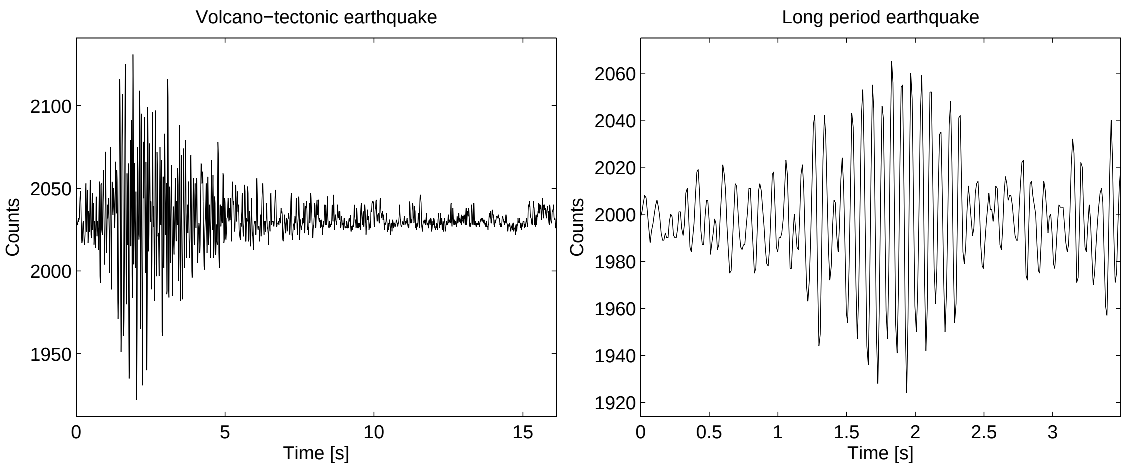
DATASET

Nevado del Ruiz data set contains signals, which were selected from data collected by the monitoring network of the VSOM monitoring system. Stations of the monitoring network are located at strategic places; for instance near to the glacier and craters. Signals from two stations

(Olleta crater station and Glacier station) have been selected for the experiments because, according to the VSOM staff experience, they are a reference for the volcanic and ice-related events. Stations are located 4.08 km and 1.8 km from the active crater respectively. Signals were digitized at 100.16 Hz sampling frequency by using a 12 bits analog to digital converter. A description of the Nevado del Ruiz data set is provided in Table 1. Typical waveforms are shown in Figure 1.

TABLE 1. Composition of the data set (number of events per class). VT and LP events were recorded at the Olleta crater station. IC events were registered at the glacier station.

VT	LP	IC
483	580	782



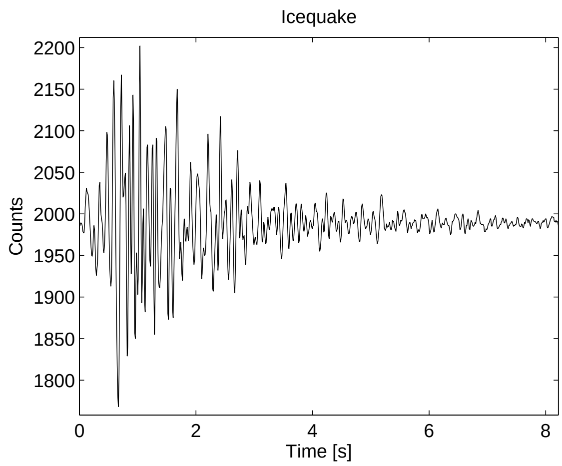


Figure 1. Typical waveforms of the three classes considered: VT, LP and IC events.

VT and LP events are the most frequent earthquakes registered at Nevado del Ruiz volcano. In our experiments, the events VT and LP (the *Ruiz-LP,VT* two-class problem) are used as well as all the classes (the multi-class *Ruiz-all* problem).

DISSIMILARITY REPRESENTATIONS AND CLASSIFIERS

Dissimilarity representations can be derived in many ways, e.g. from raw (sensor) measurements such as images, histograms or spectra or from an initial representation by features, strings or graphs (Pękalska et al., 2006); nonetheless, the particular way dissimilarities are computed is crucial and relies on the additional knowledge experts —volcanic seismologists— have about the problem.

The spectra of seismic records are commonly used for classification and monitoring of seismic activity. Since differences in spectral content allow a visual discriminating of different types of volcanic earthquakes (Zobin, 2003), we have calculated the spectrum for each record by using two different approaches: (i) N-point Fast Fourier Transform (FFT) and, (ii) parametric estimation of the power spectral density (PSD). In such a way, we explore the difference between deriving dissimilarities from a data-based spectral estimation and from a model-based spectral estimation such as the Yule-Walker AR method. DC bias was removed before computing the spectra; in addition, when spectra are to be directly compared, they are required to be normalized and to have the same length. In consequence, considering the length of the shortest event and a length-resolution trade-off, we calculated 128-point spectra.

Two different dissimilarities measures have been computed between spectra: (i) pointwise Euclidean distance and (ii) area difference: the area of non-overlapping parts (L_1 -norm) as shown in Fig. 2.

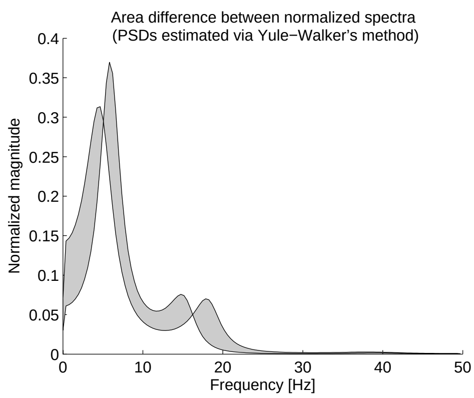


Figure 2. Dissimilarity measure as the difference between normalized spectra.

The *k* Nearest Neighbor classifier (*k*-NN)

k-NN is considered a direct approach for dissimilarity-based classification. This rule classifies a new object by assigning it the class label most frequently represented among the k nearest prototypes; i.e., by finding the k neighbors with the minimum distances between the new object and all the prototypes. For $k=1$, the rule is called 1-NN. Even though *k*-NN is asymptotically optimal in the Bayes sense, it is sensitive to noise and erroneously labelled prototypes.

Linear and Quadratic normal density based classifiers

Previous studies (Pękalska et al, 2001; Pękalska and Duin, 2002; Paclík and Duin, 2003) have shown that Bayesian (normal density based) classifiers, particularly the linear (LDC) and quadratic (QDC) normal based classifiers, perform well in dissimilarity spaces. For a 2-class problem, the LDC based on the representation set R is given by

$$f(D(x,R)) = \left[D(x,R) - \frac{1}{2}\left(\mathbf{m}_{(1)} + \mathbf{m}_{(2)}\right)\right]^T \times C^{-1}\left(\mathbf{m}_{(1)} - \mathbf{m}_{(2)}\right) + \log\frac{P_{(1)}}{P_{(2)}} \tag{1}$$

and the QDC is derived as

$$f\left(D(x,R)\right) = \sum_{i=1}^{2}(-1)^i\left(D(x,R) - \mathbf{m}_{(i)}\right)^T C_{(i)}^{-1}\left(D(x,R) - \mathbf{m}_{(i)}\right) + 2\log\frac{P_{(1)}}{P_{(2)}} + \log\frac{\left|C_{(2)}\right|}{\left|C_{(1)}\right|} \tag{2}$$

where C is the sample covariance matrix, $C_{(1)}$ and $C_{(2)}$ are the estimated class covariance matrices, and $\mathbf{m}_{(1)}$ and $\mathbf{m}_{(2)}$ are the mean vectors, computed in the dissimilarity space $D(T,R)$. $P_{(1)}$ and $P_{(2)}$ are the class prior probabilities. If C is singular, a regularized version must be used. In this study, the following regularization is used:

$$C_{reg}^{\lambda} = (1-\lambda)C + \lambda\, diag(C). \tag{3}$$

We have fixed λ to be 0.01. Nonetheless, regularization parameter should be optimized in order to obtain the best possible results for the normal density based classifiers.

EXPERIMENTAL RESULTS

Experiments were conducted to compare the results of the *k*-NN rule and the LDC and QDC classifiers built on the dissimilarity representations described above. Experiments were performed 25 times for randomly chosen training and test sets. Since in this study we are particularly interested in recognition accuracy rather than in computational complexity and storage requirements, the entire training set T has been used as the representation set R. Nonetheless, R may be properly reduced by prototype selection procedures (Pękalska et al., 2006). Training and testing sets were generated by selecting equal partitions for the classes.

The results of our experiments are shown in Fig. 3 and 4. They present the generalization errors as a function of the number of training objects randomly chosen. Fig. 3 presents the results for four dissimilarity representations of the *Ruiz-VT,LP* subset; similarly, the results for the *Ruiz-all* subset are shown in Fig. 4. Standard deviations for averaged test error decrease rapidly, varying around 0.02 after at least 10 training objects per class are available; for clarity reasons, standard deviations are not presented in Fig. 3 and 4.

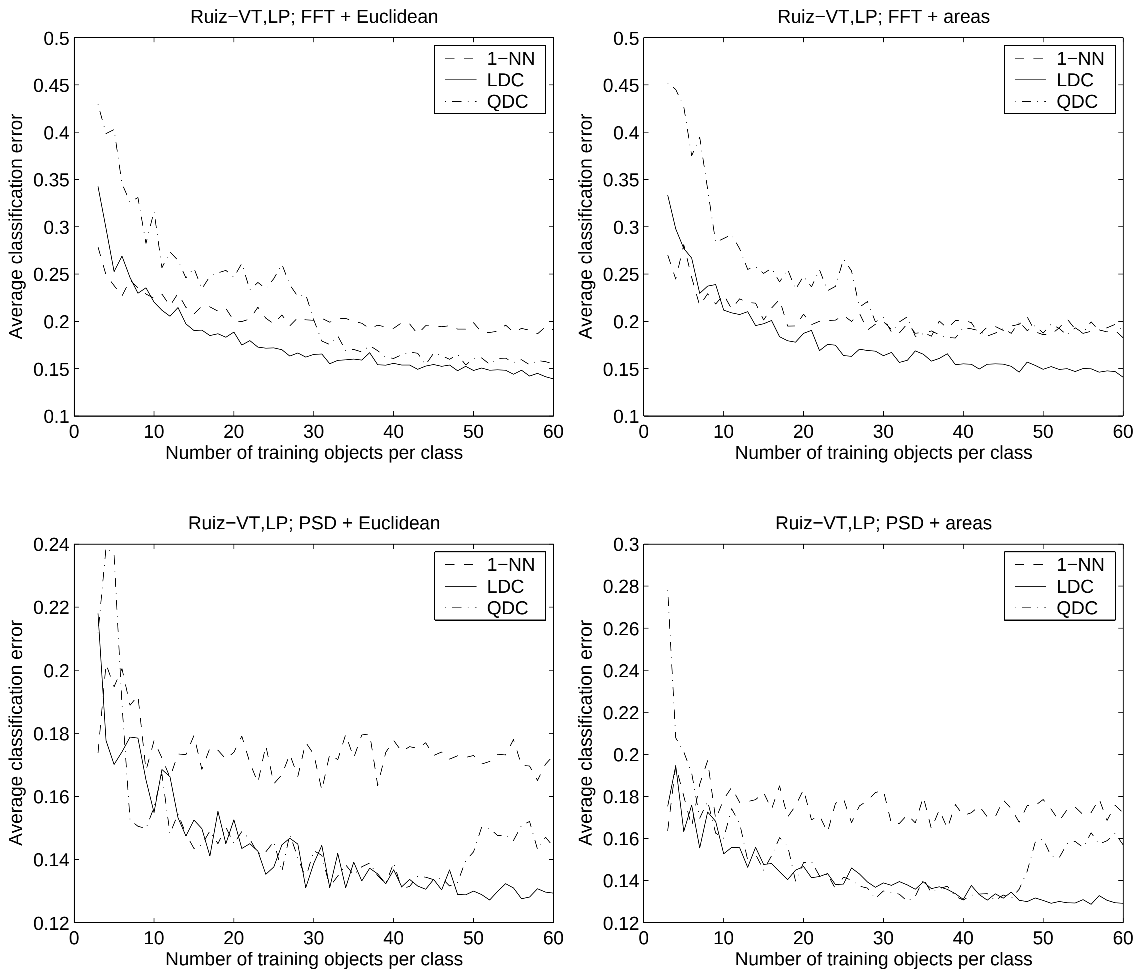


Figure 3. *Ruiz-VT, LP* data. Average classification error of the Bayesian classifiers and the 1-NN classifier in different dissimilarity spaces (FFT+Euclidean, FFT+areas, PSD+Euclidean and PSD+areas) as a function of the number of prototypes per class.

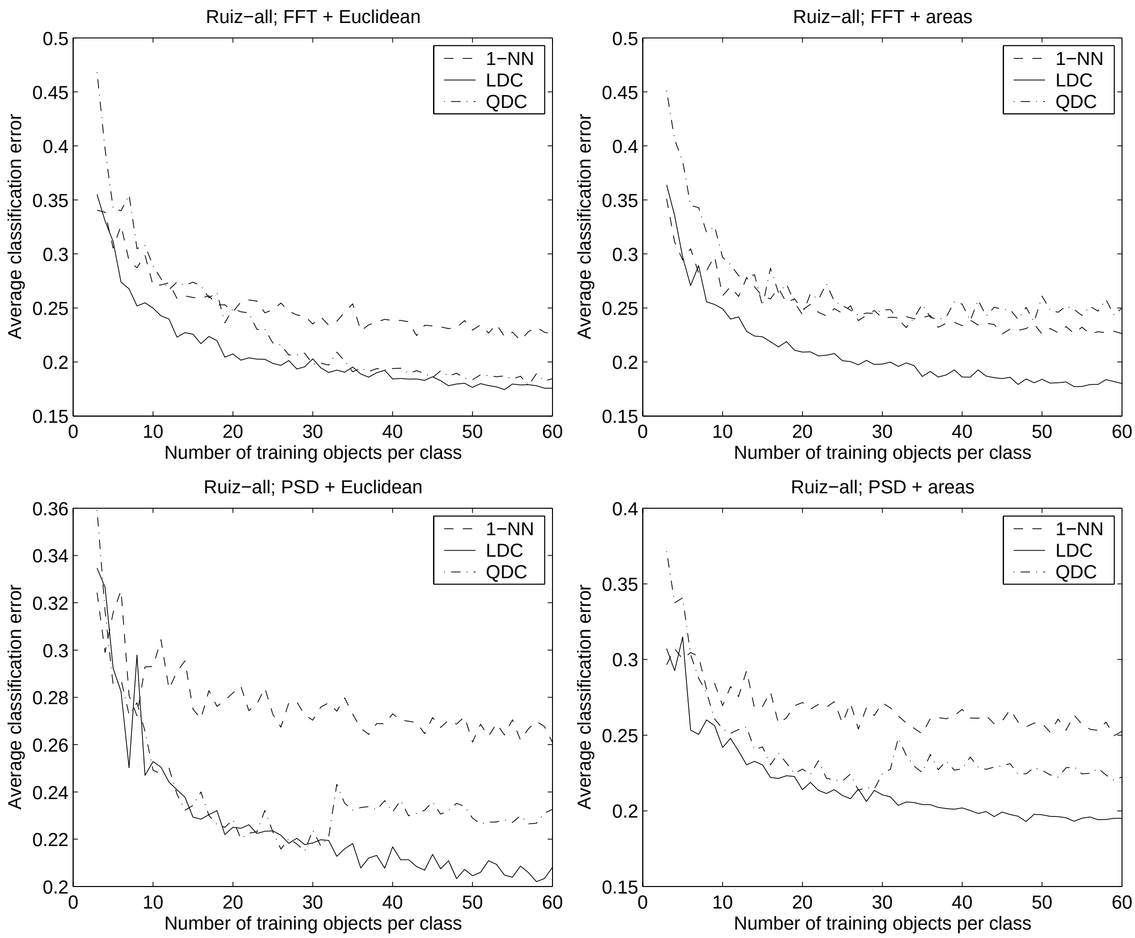


Figure 4. *Ruiz-all* data. Average classification error of the Bayesian classifiers and the 1-NN classifier in different dissimilarity spaces (FFT+Euclidean, FFT+areas, PSD+Euclidean and PSD+areas) as a function of the number of prototypes per class.

DISCUSSION AND CONCLUSIONS

We have explored and tested a dissimilarity-based strategy for classifying three different classes of seismic signals recorded by the monitoring network at Nevado del Ruiz volcano. Two classification problems were conducted: a two-class problem including VT events and LP events and a multi-class problem including ice-related events. Several dissimilarity representations were derived by combining two different approaches for spectral estimation: N-point FFT and parametric PSD estimation, as well as two dissimilarity measures: Euclidean distance and area difference between spectral curves. These dissimilarity representations $D(T,T)$ allow us for using traditional statistical decision rules, particularly normal density based classifiers. The 1-NN rule was employed as a reference for performance comparison.

The two-class *Ruiz-VT,LP* problem seems the easiest because it contains signals recorded and identified at the same station (Olleta crater station); in consequence, it is expected that sensor and noise conditions are the same, influencing the subsequent steps for representation and classification. In addition, it is well know that, in general, multi-class problems are more difficult.

For the two-class problem, experiments based on parametric PSD estimation outperform those based on the FFT. It makes sense because event lengths are, in general, short and, consequently, a parametric spectral estimation yields a higher resolution; in addition, the autoregressive methods (AR) tend to adequately describe spectra of peaky data, which is precisely the nature of seismic volcanic signals. In contrast, for the multi-class problem, the FFT yields to better results but, in these particular cases, differences are not significant.

Experiments confirm that Bayesian classifiers outperform the 1-NN classifier, when a sufficient number of prototypes is provided. The LDC constructed on the different dissimilarity representations, for both *Ruiz-VT,LP* and *Ruiz-all* problems, always outperforms the 1-NN rule.

LDC accuracies for the *Ruiz-VT,LP* problem vary between 85% and 87% when the average classification error curve reaches a steady state. Similarly, classification accuracies for the *Ruiz-all* problem vary between 81% and 84%.

QDC shows a loss of accuracy when certain number of prototypes is provided. Therefore, a further study on a proper regularization for the QDC should be conducted in order to obtain an improvement of this classifier. LDC accuracies could be an intrinsic limit of our classification problem; however, a further study on other dissimilarity-based classifiers is needed as well as a re-analysis of the original a-priori classification, in order to find more suitable classifiers and to confirm the labels assigned by the experts.

REFERENCES

- Benbrahim, M., A. Daoudi, K. Benjelloun, and A. Ibenbrahim (2005). Discrimination of Seismic Signals Using Artificial Neural Networks, *2nd World Enformatika Conference*, Istanbul, Turkey, February, 4-7.
- Duin, R.P.W., D. de Ridder, and D.M.J Tax (1998). Featureless pattern classification, *Kybernetika*. **34**, no. 4, 399-404.
- Jain, A.K., R.P.W. Duin, J. Mao (2000). Statistical pattern recognition: a review, *IEEE Trans. Pattern Anal. Machine Intell.* **22**, no. 1, 4-37.

- Langer, H., S. Falsaperla, T. Powell, and G. Thompson (2006). Automatic classification and a-posteriori analysis of seismic event identification at Soufrière Hills volcano, Montserrat, *Journal of volcanology and geothermal research*. **153**, 1-10.
- Paclík, P., and R.P.W. Duin (2003). Dissimilarity-based classification of spectra: computational issues, *Real Time Imaging*. **9**, 237-244.
- Pękalska, E., P. Paclík, and R.P.W. Duin (2001). A generalized Kernel approach to dissimilarity based classification, *J. Mach. Learn. Res.* **5**, 175-211.
- Pękalska, E., and R.P.W. Duin (2002). Dissimilarity representations allow for building good classifiers, *Pattern Recognition Lett.* **23**, 943-956.
- Pękalska, E., and R.P.W. Duin (2005). *The Dissimilarity Representation for Pattern Recognition: Foundations and Applications*. World Scientific, Singapore.
- Pękalska, E., R.P.W. Duin, and P. Paclík (2006). Prototype selection for dissimilarity-based classifiers, *Pattern Recognition*. **39**, 189-208.
- Scarpetta, S., F. Giudicepietro, E.C. Ezin, S. Petrosino, E. Del Pezzo, M. Martini, and M. Marinaro (2005). Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks, *Bulletin of the Seismological Society of America*. **95**, no. 1, 185-196.
- Zobin, V. (2003). *Introduction to volcanic seismology*. Elsevier, The Netherlands.

ACKNOWLEDGEMENTS

This work was partially supported by the research project: “Técnicas de computación de alto rendimiento en la interpretación automatizada de imagines médicas y bioseñales”. We thank the VSOM staff for providing the raw data set and Prof. Ricardo Henao for his valuable suggestions.