

CARACTERIZACIÓN DE SEÑALES SÍSMICAS UTILIZANDO MODELOS PARAMÉTRICOS Y TRANSFORMADA CEPSTRUM

HERNÁN HUMBERTO AGUDELO LÓPEZ¹

Resumen

En este trabajo se plantea una metodología para la extracción de características de señales sísmicas que permite identificar entre los diferentes tipos de sismos de origen volcánico que son estudiados por el Observatorio Vulcanológico y Sismológico de Manizales (OVSM).

Este proceso se lleva a cabo sobre una representación conjunta en los dominios del tiempo y la frecuencia conocida como superficie de ajuste. Se comparan los modelos paramétricos autorregresivos (AR, ARMA) y la transformada cepstrum en la identificación de procesos aleatorios. Se aplican diferentes criterios en la selección del orden del modelo paramétrico y se estudia el efecto de diferentes métodos de normalización sobre las características extraídas. Se emplea un clasificador bayesiano en la tarea de reconocimiento. Se obtienen porcentajes de reconocimiento del 99.8% entre dos clases de señales volcánicas.

Palabras clave

Superficie de ajuste, modelos paramétricos, transformada cepstrum.

¹ Estudiante de Ingeniería Electrónica. Pertenece al Grupo de Investigación y Desarrollo en Comunicaciones y Hardware Reconfigurable de la Universidad Tecnológica de Pereira. E-mail: hh@utp.edu.co

Abstract

This work proposes a methodology for the extraction of features from seismic signals that allows to identify different types of volcanic earthquakes that have been studied in the Volcanological and Seismological Observatory of Manizales (OVSM).

This process is made through a joint representation on time and frequency known as adjustment surface. Autoregressive parametric models (AR, ARMA) and the cepstrum transform are compared on the identification of random processes. Different criteria for the selection of the parametric models order are used and the effect of some normalization methods over the characteristics is studied. Recognition rates of 99.8% between two kinds of volcanic signals are obtained.

Key words

Adjustment surface, parametric models, cepstrum transform.

1. INTRODUCCIÓN

Los procesos geofísicos propios de las placas tectónicas a gran escala y la actividad del hombre a menor escala, generan variaciones en el movimiento y actividad de la superficie terrestre; esta actividad se traduce en ondas sísmicas que se propagan a través de la corteza terrestre (Kennett, 2001). Las señales sísmicas son estudiadas como procesos aleatorios no estacionarios (Benbrahim, Daoudi, Benjelloun, & Ibenbrahim, 2005), de origen tanto volcánico como tectónico; éstas se originan a partir del flujo de magma a través de la estructura interna de los volcanes, del desplazamiento de las placas internas de la corteza terrestre, o el desprendimiento de materiales como roca, hielo y lodo de las laderas de las montañas. El estudio de las ondas sísmicas, actividad de la que se encarga la sismología al recopilar, caracterizar y estudiar los eventos volcánicos y sísmicos, utilizando entre otros, las señales sísmicas adquiridas por instrumentos de medida especiales conocidos como sismógrafos, es fundamental para definir planes de prevención y cuidado en zonas de alto riesgo. Es de particular interés dentro de la sismología y la vulcanología poder definir la naturaleza de las señales sísmicas a fin de estudiar los fenómenos físicos asociados a la ocurrencia de las mismas (Londoño, 2002).

En la actualidad el proceso de caracterización y etiquetado de los eventos sísmicos observados en las estaciones pertenecientes a la Red Sismológica Nacional de Colombia (RSNC), es realizado de forma manual por expertos adscritos a los observatorios que componen la Red. Esta caracterización se lleva a cabo sobre la observación temporal de la señal o sobre la identificación de componentes frecuenciales fundamentales de los registros sismográficos (Londoño, 2002). La metodología propuesta en este trabajo propone realizar un análisis en los dominios conjuntos del tiempo y la frecuencia.

De otro lado, los modelos paramétricos han sido utilizados previamente para obtener características de señales sísmicas. En investigaciones realizadas anteriormente (Orozco, García,

Duin, & G. Castellanos, 2006) se han obtenido altos porcentajes de discriminación mediante el empleo de clasificadores robustos. En otros trabajos, han sido utilizadas técnicas de representación como la transformada wavelet y espectrogramas (Benbrahim, Daoudi, Benjelloun, & Ibenbrahim, 2005), coeficientes de predicción lineal LPC (Scarpetta, Giudicepietro, Petrosino, Pezzo, Martini, & Marinaro, 2005), y análisis del contenido espectral de las señales calculado mediante la función de autocorrelación, y medidas estadísticas sobre las señales mismas (Langer, Falsaperla, b, & Thompson, 2006).

Se desarrolla en este trabajo una metodología para la extracción de características en señales sísmicas que permite identificar entre dos tipos de sismos de origen volcánico. Se comparan los modelos paramétricos y transformadas tiempo-frecuencia en el procesamiento de las señales sísmicas. Este trabajo aporta nuevas características reveladas a partir de la metodología de caracterización para la identificación de tipos de sismos.

2. MATERIALES Y MÉTODOS

2.1 Datos de entrenamiento

La base de datos principal utilizada es una colección de registros sísmicos del Observatorio Vulcanológico y Sismológico de Manizales (OVSM). Consta de 622 datos clasificados como sismos Largo Periodo (LP), y 487 registros etiquetados como Vulcano-Tectónicos (VT), digitalizados con una frecuencia de muestreo de 100.16 Hz. Estos registros fueron medidos en la estación Olleta, sismógrafo ubicado aproximadamente a 4 km del cráter principal del Volcán Nevado del Ruiz, sobre 4.630 m.s.n.m.

Así mismo se utilizó una base de datos de prueba consistente en 294 señales sísmicas etiquetadas como LP y 150 como VT registradas a la misma frecuencia de la base de datos principal en la estación sismológica Cráter en el volcán Galeras. Estos registros pertenecen al Observatorio Vulcanológico y Sismológico de Pasto (OVSP).

2.2 Modelos paramétricos

Los modelos paramétricos permiten representar un sistema físico, en este caso un sistema geofísico, como la relación existente entre la entrada y la salida en forma de función de transferencia. Los diferentes modelos existentes determinan los elementos de dicha función, como los parámetros de entrada, de salida, o de entradas exógenas. Se utilizaron los modelos paramétricos Autorregresivos (AR) y Autorregresivos de Media Móvil (ARMA) (Ojeda, 2004).

En general, se puede definir un modelo paramétrico como:

$$\hat{y}_i = f(x_i; \theta) + \zeta_i \quad i = 1, 2, \dots, n,$$

donde \hat{y} es la salida del sistema, n es el número de parámetros del modelo y $f(x_i; \theta)$ es una función que depende de las condiciones x_i y del parámetro de interés θ . Es necesario aclarar que ζ_i representa la parte aleatoria del sistema. El modelo AR es un modelo de sólo polos que permite describir el valor actual de una serie de tiempo como una combinación lineal de los valores anteriores de dicha serie (Yao & Qiwei, 2005). Matemáticamente se puede describir como:

$$y_t = b_1 y_{t-1} + b_2 y_{t-2} + b_3 y_{t-3} + \dots + b_p y_{t-p} + \varepsilon_t \quad (1)$$

donde y_t es la salida actual del sistema en el tiempo t , $\{b_p, b_{p-1}, \dots, b_1\}$ son los coeficientes, ε_t es la parte aleatoria del sistema de distribución normal o gaussiana y p es el orden del modelo paramétrico.

Para describir el modelo ARMA es necesario hablar del modelo de Media Móvil (MA). Éste se caracteriza por ser un modelo de sólo ceros, el cual expresa la salida del sistema como la combinación lineal de las entradas actual y pasadas (Yao & Qiwei, 2005). Es así como el modelo ARMA es la unión de los dos modelos anteriores, lo que lleva a representar el sistema como una función de polos y ceros (Ver Ecu.).

$$y_t = b_1 y_{t-1} + b_2 y_{t-2} + b_3 y_{t-3} + \dots + b_p y_{t-p} + x_t + a_1 x_{t-1} + a_2 x_{t-2} + a_3 x_{t-3} + \dots + a_q x_{t-q} \quad (2)$$

En este caso $\{a_p, a_{p-1}, \dots, a_m\}$ son los parámetros asociados a las entradas anteriores y la entrada actual, $\{b_p, b_{p-1}, \dots, b_n\}$ son los parámetros asociados a los valores pasados de la serie de tiempo, p es el número de polos y q es el número de ceros del modelo. La parte aleatoria de este modelo la determinan las entradas de la ecuación x_{t-q} (Yao & Qiwei, 2005), que pueden ser de distribución normal o ruido blanco gaussiano con media cero y varianza σ^2 .

2.3 Transformada cepstrum

Un sistema se puede representar como la convolución entre una entrada (fuente) y una función de transferencia (filtro). La transformada cepstrum es una transformación homomórfica que permite separar la fuente del filtro que representa el sistema. Esta transformación permite convertir la convolución

$$Y[n] = x[n] * h[n] \quad (3)$$

en una suma

$$\hat{Y}[n] = \hat{x}[n] + \hat{h}[n] \quad (4)$$

Hallando un N para el cual el cepstrum del filtro $h[N]$ sea cero, es posible recuperar aproximadamente la excitación $e[n]$ y viceversa.

En general el cepstrum de una señal es la transformada de Fourier del logaritmo (con fase instantánea o no envolvente) de la señal estudiada, razón por la cual en ocasiones se le conoce como el espectro del espectro. Existen dos tipos de transformadas cepstrum, la real y la compleja, las que comparten el concepto de la transformada en sí (Huang, Acero, & Hon, 2001). En la figura 1 se muestra algorítmicamente el procedimiento de cálculo de la transformada cepstrum real y compleja (Childers, Skinner, & Kemerait, 1977). Matemáticamente están expresadas así:

Transformada cepstrum real

$$c[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |X(e^{j\omega})| e^{j\omega n} d\omega \quad (5)$$

Transformada cepstrum compleja

$$\hat{x}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln X(e^{j\omega}) e^{j\omega n} d\omega \quad (6)$$

Donde es necesario utilizar el logaritmo complejo

$$\hat{X}(e^{j\omega}) = \ln X(e^{j\omega}) = \ln |X(e^{j\omega})| + j\theta(\omega) \quad (7)$$

En el cual se define

$$\theta(\omega) = \arg [X(e^{j\omega})] \quad (8)$$

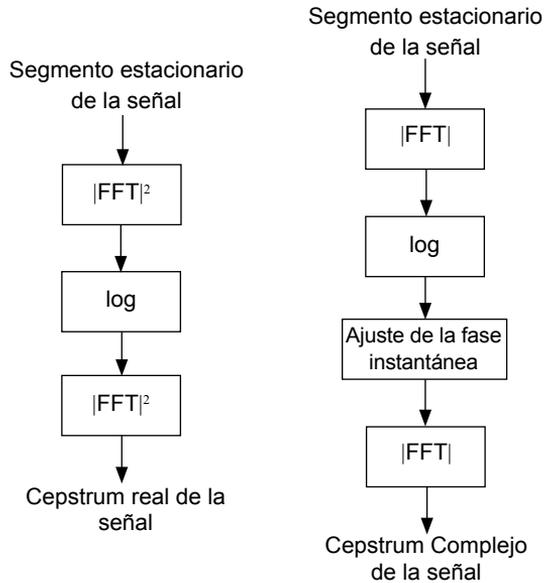


FIGURA 1. MÉTODOS DE CÁLCULO DE LAS TRANSFORMADAS CEPSTRUM REAL Y COMPLEJA

2.4 Selección de ventana de estacionariedad

Debido a su naturaleza, las señales sísmicas son procesos aleatorios y altamente no estacionarios. Es necesario, para el uso de modelos paramétricos y la transformada cepstrum, calcular un tamaño de ventana en la cual la señal cumpla con condiciones de estacionariedad, ya que estas técnicas de tratamiento de señales modelan adecuadamente series de tiempo estacionarias. Una alternativa es aplicar el test de estacionariedad *runtest*, que permite estimar segmentos de tiempo de las señales que cumplen con condiciones de cuasi estacionariedad (Shanmugan, 1988).

2.5 Selección de orden de los modelos paramétricos

Uno de los aspectos más importantes del uso de modelos paramétricos es la selección del número de parámetros. En general, si se selecciona un modelo con orden muy bajo, se obtiene un espectro altamente suavizado. Por otra parte, si el orden se selecciona muy alto, se corre el riesgo de introducir falsos picos de bajo nivel en el espectro. En el caso del modelo autorregresivo, el valor cuadrático medio del error residual constituye un indicador de desempeño. La característica de este error residual es que decrece a medida que el orden del modelo AR incrementa. Se puede monitorear la tasa de decremento y finalizar el proceso de actualización del orden cuando la tasa de decremento se vuelva relativamente lenta (Proakis & Manolakis, 2000).

Diversos métodos se han descrito para la selección del orden del modelo autorregresivo, entre los cuales el más conocido es el criterio de información de Akaike (AIC). Este criterio se basa en seleccionar el orden que minimiza la expresión

$$AIC(p) = \ln(\hat{\sigma}_{\omega_p}^2) + \frac{2p}{N}, \quad (9)$$

donde $\hat{\sigma}_{\omega_p}^2$ es la varianza estimada del error lineal de predicción y p es el orden del modelo paramétrico. Este índice de desempeño está basado en la minimización del error cuadrático medio para un

predictor de un paso. Nótese que el término $\hat{\sigma}_{\omega p}^2$ decrece y por lo tanto $\ln(\hat{\sigma}_{\omega p}^2)$ decrece también a medida que el orden del modelo AR se incrementa. Sin embargo, $2p/N$ se incrementa en razón de p . Debido a esto, el mínimo valor es obtenido para algún valor p . Otro criterio muy utilizado a la hora de modelar sistemas físicos con modelos paramétricos es el criterio de información de Bayes (BIC). Este permite una mejor estimación del orden del modelo autorregresivo dado que está basado en seleccionar el orden que minimiza el largo de la descripción, donde BIC se define según:

$$BIC(p) = N \ln(\hat{\sigma}_{\omega p}^2) + p \ln(N) \quad (10)$$

Estos criterios también están definidos para el modelo autorregresivo de media móvil. En general se define la función:

$$\Psi(p, q) = \log(\hat{\sigma}_{\omega p}^2(p, q)) + (p + q)C(N)/N \quad (11)$$

donde $\hat{\sigma}_{\omega p}^2$ es la varianza estimada del error lineal de predicción, p es el número de polos y q es el número de ceros del modelo. $C(N)$ es una función de penalidad que para el caso de BIC es $\log(n)$ y para AIC es 2 (Potscher & Srinivasan, 1994).

Un método un poco más heurístico y computacionalmente más costoso se basa en hallar el orden para el cual el error de estimación del modelo sea mínimo. En este se busca variar el orden de los modelos paramétricos y estimar la señal, comparándola con la señal original punto a punto con el objetivo de minimizar el error cuadrático medio entre ambas.

2.6 Superficies de ajuste

Uno de los métodos utilizados para analizar señales aleatorias es utilizar representaciones en diferentes dimensiones, siendo las representaciones tiempo-frecuencia las más usadas, ya que permiten estudiar la dinámica frecuencial de la señal en el tiempo. Estas representaciones se obtienen al mapear, por medio del uso de kernels o núcleos, una señal en una dimensión (tiempo) en un

espacio de dos dimensiones (tiempo-frecuencia) (Chen & Ling, 2002).

Las superficies o matrices de ajuste son similares a las representaciones tiempo-frecuencia. Consisten en generar una representación tiempo-frecuencia, no sobre la señal misma, sino sobre una estimación de esta (Orozco & Castellanos, 2007). En la figura 2 se muestra el algoritmo de construcción de una superficie de ajuste, y en la figura 3 se muestran las superficies de ajuste de dos señales sísmicas de las clases estudiadas en este trabajo, calculadas a partir de la estimación de las señales obtenida de los

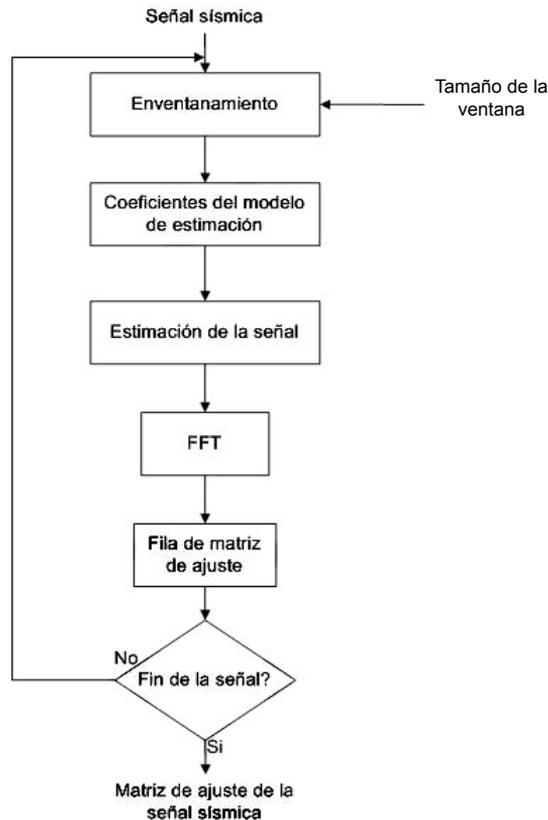


FIGURA 2. ALGORITMO DE CONSTRUCCIÓN DE UNA SUPERFICIE DE AJUSTE

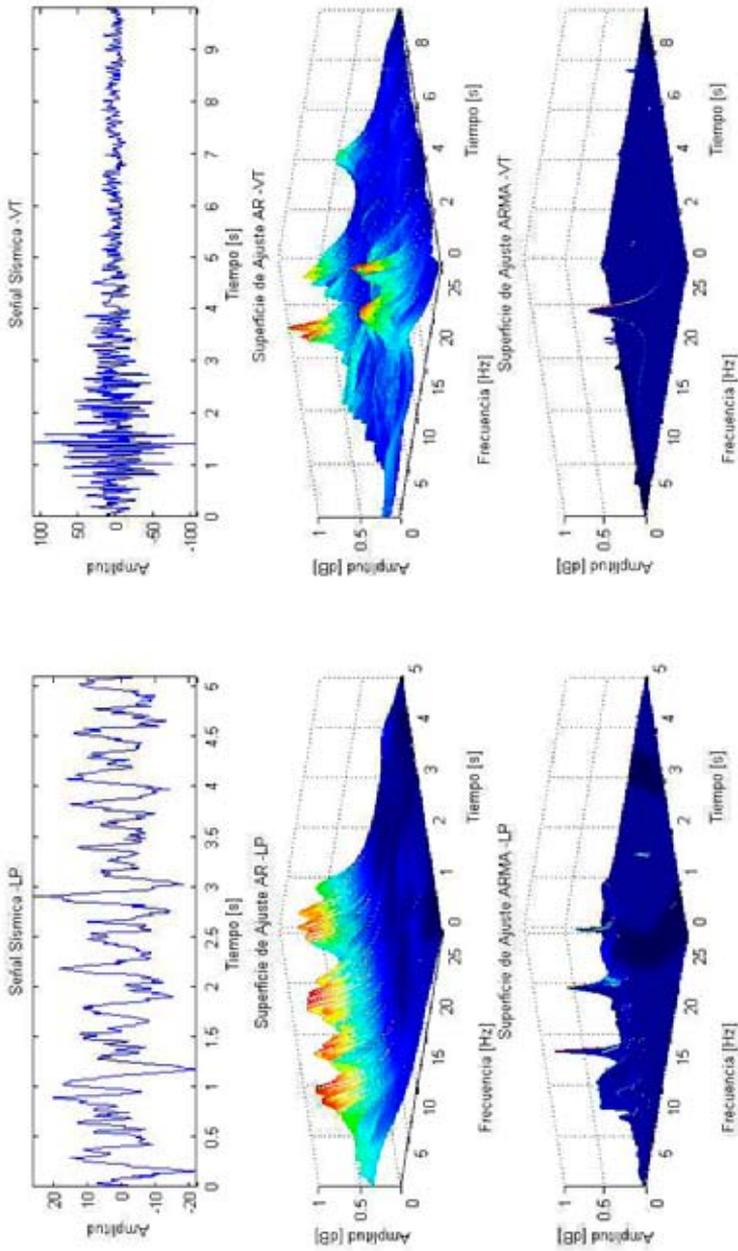


FIGURA 3. SUPERFICIES DE AJUSTE DE SEÑALES SÍSMICAS

modelos paramétricos AR y ARMA de cada segmento estacionario por medio del uso de la transformada de Fourier enventanada. En dicha figura se puede observar que ambas representaciones permiten diferenciar las dos señales, es decir, que son discriminantes.

2.7 Extracción de características

De la representación tiempo-frecuencia conseguida a partir de las superficies de ajuste es necesario extraer un conjunto de características discriminantes. Fueron extraídas 67 características, entre las que se encuentran la media, varianza, desviación estándar, kurtosis y asimetría, todas extraídas sobre los marginales del tiempo, de la frecuencia, y sobre los coeficientes de los modelos, su primera y segunda derivada (Echeverry, 2006). Se calculó de igual forma una medida de la energía contenida en la superficie utilizando los criterios de Teager (Zhou, Hansen, & Kaiser) y Shannon, de donde se extrajeron las mismas medidas estadísticas antes mencionadas. Así mismo se calculó el parámetro de perturbación sobre la media por filas de la superficie.

2.8 Normalización

Es sabido que uno de los procesos más utilizados en el tratamiento de señales y en la caracterización de las mismas es la normalización. En ocasiones, este proceso genera pérdida de información importante para los procesos de discriminación (Orozco A., 2008), razón por la cual se probó la metodología sin normalizar y normalizando las características obtenidas. El criterio de normalización utilizado fue el Min-Max definido como una transformación lineal del rango de entrada original en un nuevo rango de datos especificado (Álvarez López, 2004). Dicha transformación se realiza mediante

$$y' = \left(\frac{y - \min_1}{\max_1 - \min_1} \right) (\max_2 - \min_2) + \min_2 \quad (12)$$

donde min_1 , max_1 son los valores mínimo y máximo de la serie original, min_2 , max_2 son los nuevos valores mínimo y máximo (típicamente 0 y 1), y es el valor original y y' el nuevo valor.

2.9 Metodología

En la figura 4 se muestra la metodología que se plantea en este trabajo.

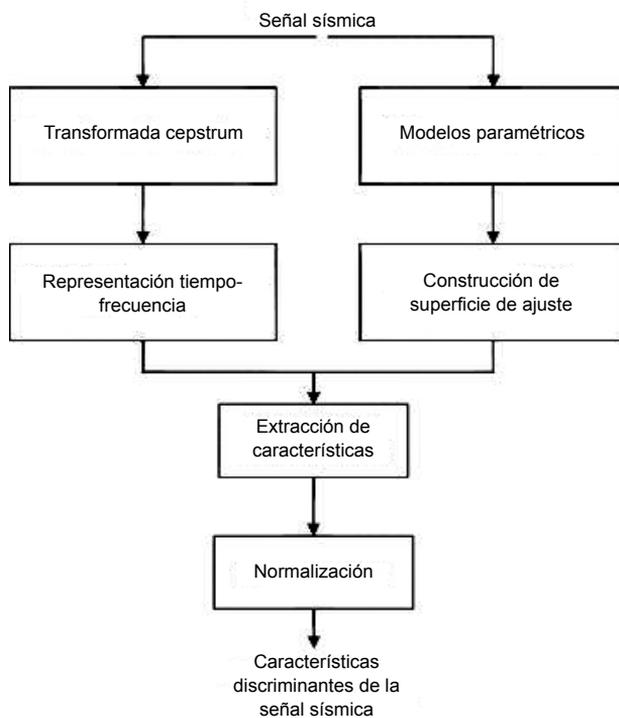


FIGURA 4. METODOLOGÍA DE CARACTERIZACIÓN PLANTEADA

3. RESULTADOS

Inicialmente se calcularon los tamaños de las ventanas en las cuales las diferentes clases estudiadas cumplen con condiciones de estacionariedad usando el método *runtest* (Shanmugan, 1988). Estos intervalos son de 0.7962 segundos (equivalentes a 80 muestras) para la clase LP y de 0.6957 segundos (equivalentes a 70 muestras) para la clase VT. Con el fin de estandarizar el tamaño de la ventana de análisis, de manera que fuera independiente del tipo de señal, de los intervalos encontrados se utilizó el de la clase LP dado que por ser el más extenso, permite incorporar uno o más segmentos estacionarios de la otra clase de análisis. Posteriormente se determinó el orden de los modelos AR y ARMA. Se utilizaron los criterios de Akaike (AIC), Bayes (BIC) y el criterio de mínimo error de estimación. Para el modelo AR se obtuvo un orden óptimo de 7 polos por medio del criterio de Bayes y el de mínimo error de estimación. En el modelo ARMA se utilizaron los criterios de Bayes y el criterio de mínimo error de estimación. Se determinó que 5 polos y 4 ceros constituían el orden óptimo para este modelo.

Con estos elementos se procedió a probar la metodología planteada con las bases de datos de entrenamiento descritas en la sección 2.1. Se traslaparon los intervalos de estacionariedad en 6%. Esta metodología fue validada utilizando un clasificador Bayesiano y validación *leave one out*, la cual permite una estimación no sesgada del error del clasificador.

En la tabla 1 se muestran los resultados de clasificación de obtenidos con las señales del Volcán Nevado del Ruiz para todas las metodologías y clases.

CLASE	MODELOS PARAMÉTRICOS Y SUPERFICIES DE AJUSTE		TRANSFORMADA CEPSTRUM	
	AR	ARMA	Real	Compleja
VT	99.9%	99.9%	99,4%	99,6%
LP	99,2%	98.6%	98,9%	99.9%
Total	99,6%	99.3%	99,1%	99,8%

TABLA 1. RESULTADOS DE VALIDACIÓN DE LA METODOLOGÍA PLANTEADA CON LAS SEÑALES DEL VOLCÁN NEVADO DEL RUIZ

En la tabla 2 se muestran los resultados de clasificación con las señales del Volcán Galeras descritas en la sección 2.1.

CLASE	MODELOS PARAMÉTRICOS Y SUPERFICIES DE AJUSTE		TRANSFORMADA CEPSTRUM	
	AR	ARMA	Real	Compleja
VT	44,7%	99,9%	99,3%	98,7%
LP	98,3%	95,9%	98,6%	99,9%
Total	71,5%	98,0%	99,0%	99,3%

TABLA 2. RESULTADOS DE VALIDACIÓN DE LA METODOLOGÍA PLANTEADA CON LAS SEÑALES DEL VOLCÁN GALERAS

En la tabla 3 se muestra un consolidado de los resultados que permite comparar entre los diferentes métodos de procesamiento en ambas bases de datos.

Base de datos	MÉTODO	PARAMÉTRICOS		CEPSTRUM	
	Normalización	AR	ARMA	Real	Complejo
Volcán Nevado del Ruiz	Ninguna	88,8%	97,5%	95,0%	96,2%
	Min-max	99,6%	99,3%	99,1%	99,8%
Volcán galeras	Ninguna	90,0%	97,8%	93,8%	94,5%
	Min-max	71,5%	98,0%	99,0%	99,3%

TABLA 3. RESULTADOS DE VALIDACIÓN CONSOLIDADOS

Los mejores resultados fueron obtenidos utilizando la transformada cepstrum compleja.

4. DISCUSIÓN

En el estado del arte sobre clasificación de señales sísmicas se ha dado mayor relevancia al uso de clasificadores robustos y no a la tarea de definir una metodología clara de caracterización de dichas

señales (Álvarez & Henao, 2007) (Benbrahim, Daoudi, Benjelloun, & Ibenbrahim, 2005) (Langer, Falsaperla, b, & Thompson, 2006). Este trabajo se centra en este tema. Se determinó la utilidad del modelo paramétrico autorregresivo de media móvil y de la transformada cepstrum para reconstruir las señales sísmicas de las clases objeto de estudio. Aunque el modelo autorregresivo ha sido utilizado para calcular la estimación de la densidad espectral de potencia, lo que ha permitido extraer características que luego fueron clasificadas usando procesos gaussianos (Álvarez & Henao, 2007) para obtener resultados acierto del 94%, no se tuvo en cuenta la no estacionariedad de la señal y la necesidad de determinar intervalos que cumplan con esta condición para poder aplicar modelos paramétricos lineales.

Para trabajos futuros, se utilizarán clasificadores de mayor desempeño como máquinas de soporte vectorial o procesos gaussianos, buscando mejores resultados de clasificación. Así mismo se busca estudiar otras representaciones tiempo-frecuencia como las transformadas Wavelet y Cohen y se hace necesario explorar procesos de selección de características como PCA o LDA.

5. CONCLUSIONES

- La representación tiempo-frecuencia que se obtiene a partir de las superficies de ajuste permiten estudiar de forma adecuada la dinámica de las señales sísmicas en el tiempo y la frecuencia de forma conjunta.
- Las características más discriminantes entre las clases de sismos estudiadas fueron las obtenidas a partir de la representación tiempo-frecuencia generada con la transformada cepstrum.
- Los resultados obtenidos con los modelos paramétricos autorregresivo y autorregresivo de media móvil son muy cercanos a los obtenidos con la transformada cepstrum, probando que permiten también obtener características altamente discriminantes entre clases.

- Se obtuvo una tasa de discriminación del 99.8% utilizando la transformada cepstrum y de 99.6% para el modelo paramétrico AR.

6. AGRADECIMIENTOS

El autor agradece al Observatorio Sismológico y Vulcanológico de Manizales (INGEOMINAS) por la base de datos facilitada y la colaboración prestada a lo largo del proyecto de investigación. Este trabajo se desarrolló en el Grupo de Investigación y Desarrollo en Comunicaciones y Hardware Reconfigurable en el marco del proyecto de investigación “CLASIFICACIÓN AUTOMATIZADA DE EVENTOS SÍSMICOS”, financiado por la Universidad Tecnológica de Pereira con código 9-07-04.

7. BIBLIOGRAFÍA

- Álvarez López, M. A. (2004). Reconocimiento de voz sobre diccionarios reducidos usando Modelos Ocultos de Markov. *Tesis de grado*. Universidad Nacional de Colombia Sede Manizales.
- Álvarez, M., & Henao, R. (2007). Clasificación De Eventos Sísmicos Empleando Procesos Gaussianos. *Scientia et Technica*.
- Benbrahim, M., Daoudi, A., Benjelloun, K., & Ibenbrahim, A. A. (2005). Discrimination of Seismic Signals Using Artificial Neural Networks. *Transactions On Engineering, Computing And Technology*.
- Chen, V. C., & Ling, H. (2002). *Time-frequency transforms for radar imaging and signal analysis*. Boston: Artech House.
- Childers, D. G., Skinner, D. P., & Kemerait, R. C. (1977). The Cepstrum - A Guide to Processing. *Proceedings of the IEEE, Vol. 65*.
- Echeverry, J. D. (2006). *Caracterización acústica de bioseñales empleando transformadas tiempo frecuencia y modelado paramétrico*. Pereira: Universidad Tecnológica de Pereira.
- Fan, J., & Yao, Q. (2005). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer.
- Huang, X., Acero, A., & Hon, H.-W. (2001). *Spoken language processing*. Prentice Hall.

- Kennett, B. L. (2001). *The Seismic Wavefield*. Cambridge University Press.
- Langer, H., Falsaperla, S., b, T. P., & Thompson, G. (2006). Automatic classification and a-posteriori analysis of seismic event identification at Soufrie're Hills volcano, Montserrat. *Journal of Volcanology and Geothermal Research*.
- Londoño, J. M. (2002). *Manual de Sismología y Volcánica Básica*. Observatorio Vulcanológico de Manizales.
- Ojeda, M. M. (2004). La Modelación Estadística. *Foro de Matemáticas del Sureste*, (p. 76).
- Orozco, A. (2008). *Clasificación de patrones provenientes de señales de actividad biológica no-estacionaria. Aplicación a la cirugía de la enfermedad de Parkinson*. Valencia, España: Universidad Politécnica de Valencia.
- Orozco, A., & Castellanos, G. (2007). *Análisis de señales y sistemas*. Pereira: Universidad Tecnológica de Pereira.
- Orozco, M., García, M. E., Duin, R. P., & G. Castellanos, C. (2006). Dissimilarity-Based Classification Of Seismic Signals At Nevado Del Ruiz Volcano. *Earth Sciences Research Journal*.
- Potscher, B. M., & Srinivasan, S. (1994). A comparison of order estimation procedures for ARMA models. *Statistica Sinica* 4, 29-50.
- Proakis, J., & Manolakis, D. (2000). *Tratamiento Digital de Señales*. Prentice Hall.
- Scarpetta, S., Giudicepietro, F., Petrosino, E. C., Pezzo, E. D., Martini, M., & Marinaro, M. (2005). Automatic Classification of Seismic Signals at Mt. Vesuvius Volcano, Italy, Using Neural Networks. *Bulletin of the Seismological Society of America*.
- Shanmugan, K. S. (1988). *Random Signals: Detection Estimation and Data Analysis*. Wiley.
- Stoica, Petre, & L., M. R. (1997). *Introduction to Spectral Analysis*. Prentice Hall.
- Yao, J., & Qiwei, F. (2005). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer.
- Zhou, G., Hansen, J. H., & Kaiser, J. F. (n.d.). A new nonlinear feature for stress classification. *Robust speech processing laboratory - Duke University*.