

ANÁLISIS ESPECTRAL PARA EL RECONOCIMIENTO DE HUELLAS ACÚSTICAS

SPECTRAL ANALYSIS FOR RECOGNITION OF ACOUSTIC FINGERPRINT



Eduardo E. Zurek¹, Margarita R. Gamarra A², José R. Escorcía G.²,
Carlos Gutierrez³, Henry Bayona³, Roxana Pérez³, Xavier García³

¹. Dpto. de Ing. Sistemas Universidad del Norte Barranquilla, Colombia

². Dpto. de Ing. Electrónica y Telecomunicaciones Universidad Autónoma del Caribe Grupo de investigación IET-UAC Barranquilla, Colombia

³. Escuela Naval de Suboficiales, ARC, Barranquilla. Grupo de investigación GITIN Barranquilla, Colombia

Recibido: 10 Febrero de 2015

Aceptado: 6 Febrero de 2016

*Correspondencia del autor: Eduardo E. Zurek. E-mail: ezurek@uninorte.edu.co

RESUMEN

Este artículo presenta los resultados del proceso de reconocimiento de huellas digitales acústicas utilizando características espectrales de la señal. El Análisis de Componentes Principales (PCA, por sus siglas en inglés) se aplica para reducir la dimensión de las características extraídas y luego se implementa un clasificador basado en el método de los k-vecinos más cercanos (KNN) para identificar el patrón de la señal de audio. Este clasificador se compara con una Red Neuronal Artificial (ANN, por sus siglas en inglés). Es necesario implementar un sistema de filtrado para las señales adquiridas para reducir el ruido de 60 Hz generadas por las imperfecciones en el sistema de adquisición. Los métodos descritos en este trabajo se utilizaron para el reconocimiento de embarcaciones marinas.

Palabras claves: Huella acústicas; PCA; espectrograma; FFT; KNN; ANN.

ABSTRACT

This article presents the results of the recognition process of acoustic fingerprints using the spectral characteristics of the signal. The Principal Component Analysis (PCA, for its acronym in English) is applied to reduce the size of the extracted features and then, based on the method of k-nearest neighbor (KNN), a classifier is implemented to identify the pattern of the audio signal. This classifier is compared with the Artificial Neural Network (ANN, for its acronym in English). It is necessary to implement a filtering system for the acquired signals in order to reduce the noise of 60 Hz generated by imperfections in the acquisition system. The methods described in this paper were used for recognition of marine vessels.

Keywords: Acoustic fingerprints; PCA; spectrogram; FFT; KNN; ANN.

I. INTRODUCCIÓN

La detección y el reconocimiento de los diferentes tipos de embarcaciones es una de las principales tareas realizadas por las unidades de submarinos. Por lo tanto la capacidad de reconocer qué tipo de embarcación se encuentra en la proximidad es importante con el fin de tomar medidas necesarias. Este reconocimiento puede ayudar a detectar, por ejemplo, barcos ilegales. En estas condiciones, es evidente que una detección incorrecta podría comprometer la integridad del submarino y los miembros de la tripulación, incluso el personal de apoyo solicitados para una acción específica, debido a la desinformación (1). Del mismo modo, la formación del personal responsable del reconocimiento acústico se lleva a cabo en el campo y el aprendiz tiene que esperar una operación real para poner en práctica sus conocimientos y no siempre es posible identificar todo tipo de embarcaciones.

El sonido producido por una embarcación es causada principalmente por la maquinaria de su impulsor y la cavitación causada por la hélice. Estos sonidos son únicos para cada tipo de embarcación y, por tanto, dicha información puede ser utilizada para fines de identificación y seguimiento. Esta señal se considera como un proceso estacionario dadas sus características de variación y su período (2–4), en consecuencia, las técnicas de análisis espectral de señales han demostrado ser útiles para diferenciar y clasificar los barcos.

Actualmente la clasificación de los buques se lleva a cabo por un operador, quien debe analizar las diversas señales que los sistemas de adquisición proporcionan para la detección y la identificación de embarcaciones. Por lo tanto, la probabilidad de que el operador identifique correctamente un barco puede verse afectada por el estrés o condiciones subjetivas. Por lo tanto, se hace necesario la implementación de un sistema de identificación automática que ayude al operador, para garantizar la seguridad y la coordinación táctica. El método propuesto para el reconocimiento de huellas acústicas de los barcos utiliza la Transformada Rápida de Fourier (FFT, por sus siglas en inglés) para extraer las características de las señales, Análisis de Componentes Principales (PCA) para reducir el número de características y un clasificador de los k-vecinos más cercanos (KNN) comparado con un clasificador basado en Redes Neuronales Artificiales (ANN).

2. ESTADO DEL ARTE

Inicialmente las técnicas aplicadas para la detección automática de los barcos se han basado en la extracción de características en el dominio de la frecuencia. Algunos autores implementan FFT, como en (5) para extraer características. Usando un hidrófono omnidireccional se extrae la densidad espectral de potencia, para un objeto en movimiento. Con esta información se entrena una red neuronal feed-forward.

Del mismo modo, en (6) se implementa FFT para reconocer la huella digital acústica. Esta información se compara con una base de datos de sonidos, que se han generado con el propósito de comparación e identificación.

De igual forma en (7) la información de un sonar es pre-procesado para extraer el espectro. Con el fin de reconocer el tipo de embarcación, se utiliza PCA. Esta información es la entrada a una red neural.

Sin embargo el análisis espectral no ha sido el único utilizado para reconocimiento de sonido, también se utilizan los modelos autorregresivos. En el caso de (8) la información espectral se utiliza para diseñar un modelo autorregresivo o ARMA.

En (9) se desarrolla un método de tratamiento de la señal para extraer automáticamente la estructura armónica del ruido radiado por pequeñas embarcaciones. Una herramienta para la extracción y análisis de armónicos (HEAT, por sus siglas en inglés) estima la frecuencia fundamental instantánea de los tonos armónicos, refina la estimación de la frecuencia fundamental usando un filtro de Kalman, y automáticamente extrae las amplitudes de los armónicos para generar una firma acústica para el barco.

En (10) no se utiliza la información espectral de la señal, en lugar de esto se construye un modelo autorregresivo o AR, del cual se extraen los polos periódicamente para diferenciar la fuente que emite la señal. Por último, usando clasificadores estadísticos se reconocen las diferentes señales.

Es de esperarse que si las investigaciones con transformada de Fourier arrojaron buenos resultados, la transformada wavelet (8,11) podría mejorar los resultados obtenidos. En (11) se usan dos algoritmos de pre-procesamiento: el primero hace separación de ventanas en dos pasos o TPSW por sus siglas en inglés, extrayendo posteriormente la información de la

densidad espectral de potencia promedio de los datos de entrada y el segundo utiliza la transformada wavelet de los mismos datos. Para la etapa de clasificación se usan 4 distintos clasificadores de forma que se realice un análisis comparativo entre ellos: PNN, MLP AKC y LVQ.

Así también a lo largo de las investigaciones anteriores se observa como las redes neuronales tienen un espacio particular e importante en la etapa de clasificación. En particular en (12) se usó una red neuronal de Kohonen con datos espectrales suavizados y las diferencias entre componentes de orden k como entradas a ella.

Algunos avances sobre la tecnología de sonares se ha llevado a cabo en los últimos años. Uno de estos desarrollos corresponde a SOBEK, una familia de tecnologías de sonares pasivos para la vigilancia, de-

sarrollado en (13). Las ventajas de sonar pasivo son múltiples: es de bajo costo, que permite la operación encubierta, no es dañino para la vida bajo el agua, y la reverberación, lo que puede plantear desafíos significativos para los sonares activos en aguas confinadas. Por otra parte, las firmas acústicas pasivas proporcionan una gran cantidad de información para la clasificación de destino.

Los avances tecnológicos y científicos siguen en desarrollo y dado el alto valor que tienen las huellas acústicas de las embarcaciones las publicaciones son limitadas respecto a este tipo de aplicaciones.

3. ENFOQUE PROPUESTO

Con la finalidad de identificar tres tipos de embarcaciones mediante su huella acústica se ha propuesto el esquema de reconocimiento planteado en la Figura. 1.



Figura 1. Esquema del enfoque propuesto para el reconocimiento de huellas acústicas.

Inicialmente se crea una base de datos con el espacio reducido de características. Las embarcaciones que ser identificados son: mercantes, buques y lanchas. El espectro de estas señales se presenta en la Figura. 2. La frecuencia de muestreo es de 22 kHz y se tomaron 500 muestras de audio para cada tipo de embarcación. Cada una de las muestras de audio se filtra y sus características se extraen mediante FFT; luego se seleccionan los componentes principales aplicando PCA. En el proceso de clasificación, se utiliza un clasificador KNN: la muestra para clasificar se compara con la base de datos de sonido y se asocia con una clase vecina más cercana. El desempeño de este clasificador se compara con el de una red neuronal tipo Feedforward. A continuación, se presentan cada una de las etapas y técnicas utilizadas.

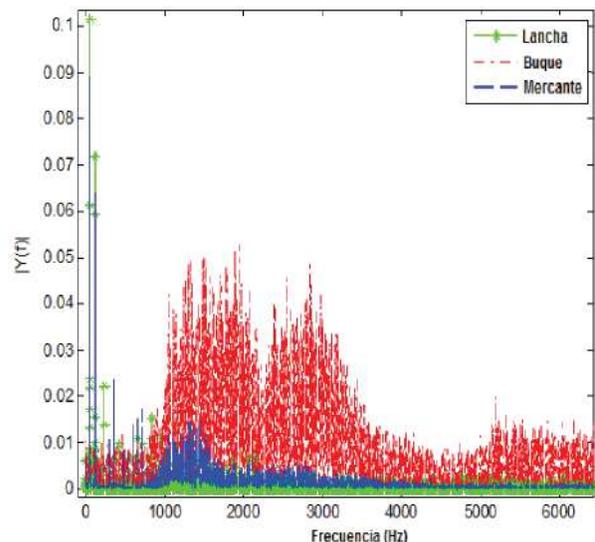


Figura 2. Espectro de las señales acústicas de las tres embarcaciones a identificar

A. Filtrado

Debido a los equipos de adquisición, las señales acústicas presentan ruido en la frecuencia de 60Hz y sus armónicos, como se puede ver en la Figura. 3. Las frecuencias características de cada embarcación se encuentran en valores superiores a los 1000Hz, por esta razón es necesario aplicar un filtrado tipo pasa-bajas con la finalidad de eliminar el ruido de baja frecuencia. Esta etapa inicial resulta adecuada, pues en las etapas posteriores se trabajará solamente con las frecuencias representativas de cada señal, lo que facilita el proceso de reconocimiento.

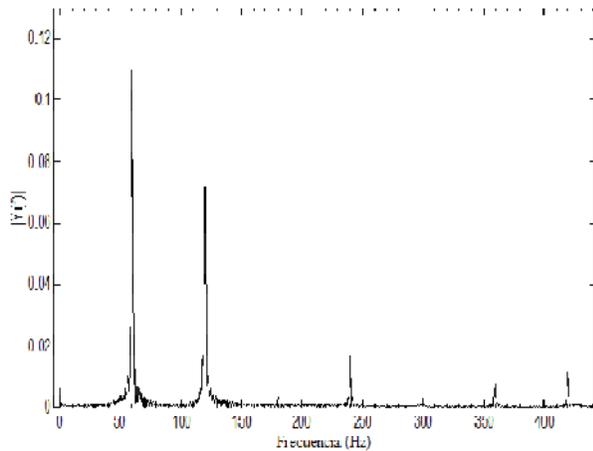


Figura 3. Espectro de la señal acústica de una lancha, mostrando solamente frecuencias bajas.

En (14) se presenta una comparación entre filtros IIR y FIR. Esta información se presenta en la Tabla 1.

TABLA 1. COMPARACIÓN ENTRE FILTROS IIR Y FIR

ITEM	Filtros IIR	Filtros FIR
Fase	Difícil de controlar.	Siempre es posible una fase lineal.
Estabilidad	Puede ser inestable.	Siempre es estable.
Orden del filtro	Menor.	Mayor
Diseño análogo	Derivado de filtros análogos.	-

Teniendo en cuenta las ventajas presentadas en la Tabla I, el filtro diseñado es de tipo FIR, utilizando una ventana tipo Hamming, pasa-altas, de orden 100 y una frecuencia de corte de 1kHz. El filtro diseñado se muestra en la Figura. 4.

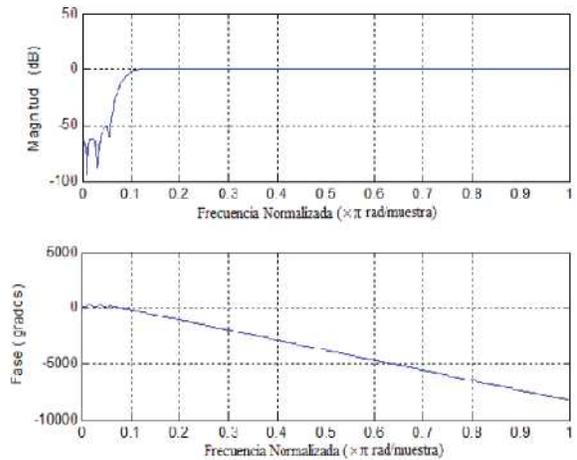


Figura 4. Filtro aplicado: magnitud y fase.

B. Extracción de Características

Con el objetivo de identificar la embarcación es necesario extraer características de la señal que permitan caracterizar y clasificar cada patrón de la embarcación. El audio original se reduce a 3000 muestras, que es una trama de datos representativa de la señal. Entonces se realiza una convolución con una ventana de Hamming para reducir los efectos indeseables de la ventana rectangular (15). Luego se aplica FFT a la señal resultante. La FFT produce valores reales y complejos: los valores reales representan la distribución de los componentes de frecuencia, mientras que los valores complejos llevan la información sobre la fase de los componentes (16).

Esta información es adecuada para representar la señal de audio adquirida. En esta etapa se extraen 512 coeficientes de Fourier. Sin embargo, es necesario utilizar un proceso de selección para definir las características más significativas que permiten discriminar las tres señales de entrada.

C. Selección de características

A fin de reducir la dimensionalidad del espacio de características producido a partir del algoritmo FFT aplicada a las señales de audio, se utiliza el Análisis de Componentes Principales (PCA). PCA es una técnica de análisis de datos que se puede utilizar con el fin de extraer las características "más importantes" de un gran conjunto de datos (17,18). En este caso se desarrolla un experimento donde el número de características utilizadas en la clasificación se varía de 256 a 64 componentes principales con el objetivo de seleccionar el tamaño adecuado.

Hay diversos algoritmos para el cálculo de los principales componentes de un conjunto de datos, sin embargo el utilizado en este caso se basa en la descomposición de valor propio (19).

D. Clasificación

Después de la selección de las características principales de las señales, este vector de características se aplicó a las señales de entrenamiento y prueba del clasificador KNN. La entrada para este clasificador consistió en valores señales de entrenamiento, correspondiente a la clase a la que pertenecen (20).

Se calculó el error de clasificación como el porcentaje de las clases clasificadas incorrectamente. Del mismo modo, se obtuvieron indicadores de rendimiento. Se tomó en cuenta un balance entre el tiempo de ejecución y la tasa de aciertos para seleccionar los parámetros adecuados del filtro y número de coeficientes extraídos con la FFT. Se realizó una comparación entre KNN y ANN, en términos de la tasa de aciertos.

4. RESULTADOS

Los experimentos con el sistema de reconocimiento se realizaron en Matlab® 7.11. Con la finalidad de entrenar el clasificador KNN, se utilizaron 1500 muestras, 500 de cada embarcación. Para la validación se utilizaron 1500 muestras diferentes. Las muestras de entrenamiento y validación se tomaron de algunos tramos de la señal de audio adquirida. Se utilizaron los 5 vecinos más cercanos para asociar la muestra a una clase. Este valor se determinó de manera experimental.

De igual forma, se entrenó un clasificador basado en ANN. El tipo de red utilizada fue Feed Forward Backpropagation, con una capa oculta tipo sigmoide con 20 neuronas. Esta arquitectura de red es una de las más utilizadas en reconocimiento de patrones (21) y su caracterización se determinó experimentalmente. Para la validación, se utilizaron las mismas muestras que en clasificador KNN.

La Tabla 2 presenta una relación entre el número de coeficientes de Fourier extraídos, *nfft*, la tasa de aciertos y el tiempo de ejecución del algoritmo de ex-

tracción. Inicialmente se extrajeron 512 coeficientes, pero usando PCA se reduce la dimensión a 256, 128 o 64. El tiempo de ejecución se calcula para el procesamiento de 1500 muestras y utilizando solamente el clasificador KNN.

TABLA 2. TASA DE ACIERTOS Y TIEMPO DE EJECUCIÓN PARA DIFERENTES COEFICIENTES

<i>nfft</i>	Tasa de aciertos (ANN)	Tasa de aciertos (KNN)	Tiempo de ejecución (s)
512	95.9%	86.73%	14.031
256	96.0%	85.60%	5.862
128	91.6%	84.33%	3.373
64	90.7%	80.27%	1.949

Los resultados presentados en la Tabla 2 muestran que el clasificador con redes neuronales tiene una mayor tasa de aciertos que el clasificador KNN. Si se utilizan 256 coeficientes se presenta un desempeño similar que utilizando 512 coeficientes; este comportamiento se puede explicar debido a que algunas características (coeficientes) no son discriminantes y esto puede conducir a una clasificación incorrecta.

Métricas comúnmente utilizadas para medir la calidad de los métodos de clasificación son: la tasa de aciertos y la Matriz de Confusión, la cual contiene información relativa a esta tasa (22). Los siguientes indicadores de desempeño se calculan con base en la Matriz de Confusión:

- a) Proporción de Verdaderos Positivos (Sensitividad): $PVP=VP/(VP+FN)$
- b) Proporción de Verdaderos Negativos (Especificidad): $PVN=VN/(VN+FP)$
- c) Valor Predictivo Positivo: $VPP = VP/(VP + FP)$.
- d) Valor Predictivo Negativo: $VPN = VN/(VN + FN)$.

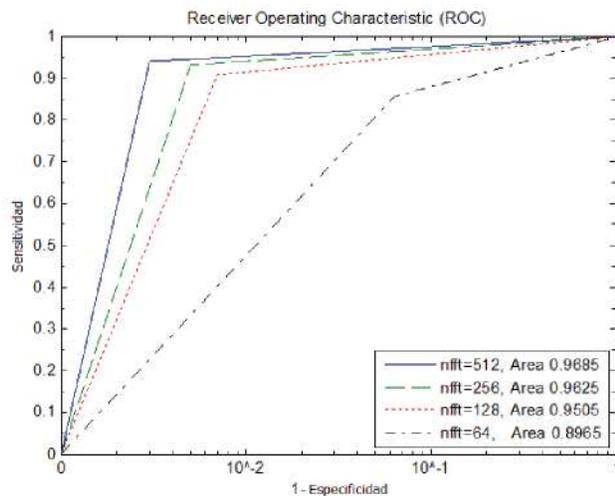
Después que se realizaron las pruebas, se elaboró una matriz de confusión para determinar los indicadores de desempeño para cada valor de *nfft*, como se muestra en la Tabla 3. Estos indicadores se pueden interpretar mejor utilizando una curva de desempeño.

TABLA 3. INDICADORES DE DESEMPEÑO PARA DIFERENTES CASOS DE *nfft*

Indicador	512	256	128	64
Sensitividad	0.9400	0.9300	0.9080	0.8560
Especificidad	0.9970	0.9950	0.9930	0.9370
VPP	0.9937	0.9894	0.9848	0.8717
VPN	0.9708	0.9660	0.9557	0.9286

La curva de desempeño ROC (Receiver Operating Characteristic) se define como un espacio de coordenadas usado para mostrar el desempeño del clasificador. Las curvas ROC muestran claramente la relación entre Sensitividad y Especificidad del clasificador; un incremento en la Sensitividad es acompañado por un decremento en la Especificidad (23).

Una curva más cerca de la esquina superior izquierda en el espacio ROC, representa un método de identificación más preciso, asociado a una mayor tasa de aciertos. Una curva cerca de la diagonal de 45° en el espacio ROC, representa un clasificador con menos éxito. En la Figura. 5 se grafica la curva ROC mostrando los resultados obtenidos en la evaluación del clasificador KNN.

**Figura 5.** Curva ROC para diferentes *nfft*

La gráfica en la Figura 5 muestra que el área bajo la curva para un *nfft*=512 y *nfft*=256 son muy similares. Teniendo en cuenta estos resultados y la tasa de aciertos presentada en la Tabla 2, un *nfft* igual a 256 ofrece un desempeño adecuado tanto para el clasificador KNN como para ANN, utilizando un tiempo de ejecución bajo. De acuerdo a los resultados experimentales obtenidos, este valor de *nfft* es adecuado para clasificar una embarcación perteneciente a una de las tres clases evaluadas.

5. CONCLUSIONES

En la investigación desarrollada se utiliza una técnica de análisis espectral para el reconocimiento de embarcaciones utilizando su huella digital acústica. Se necesita un proceso previo de filtrado con el fin de reducir el ruido de 60 Hz en la señal de audio.

El número de puntos de muestreo (*nfft*) utilizados para calcular la FFT se varían para determinar el valor adecuado que aporte un mayor porcentaje de aciertos y menor tiempo de ejecución. A partir de pruebas realizadas este valor se fijó en *nfft* = 256, de acuerdo con los resultados presentados en la Tabla 2 y en la Figura 5.

En la Tabla 2 se presenta una comparación entre los clasificadores KNN y ANN: el clasificador utilizando Redes Neuronales tuvo un mejor rendimiento que el clasificador de los k-vecinos, para un *nfft* dado.

Con el uso de un filtro pasa-altas FIR, 256 coeficientes extraídos aplicando FFT y un clasificador ANN es posible obtener un buen rendimiento en el proceso de reconocimiento de embarcaciones.

Resulta relevante implementar el proceso de reconocimiento propuesto en un sistema embebido, lo que reduciría el tiempo de reconocimiento y una identificación más precisa, comparada con un operador humano.

RECONOCIMIENTOS

Este trabajo ha sido financiado por Colciencias, mediante la convocatoria número 632, la Armada Nacional y el Ministerio de Defensa de Colombia, bajo el proyecto: “Desarrollo de un Sistema de Reconocimiento de Huellas Acústicas de Embarcaciones Marinas para la Flotilla de Submarinos de la Armada Nacional Utilizando Técnicas de Inteligencia Artificial”.

BIBLIOGRAFÍA

1. Catalfamo, Amelia. *Redes Neuronales Aplicadas a Señales Sonoras Obtenidas de un Sonar*. Buenos Aires, Argentina: ITBA; 2006.
2. Urick, R.J. *Principles of Underwater Sound*. New York: McGraw-Hill; 1983.
3. Regazonni, C; Tesei, A; Tacconi, G. *A Comparison Between Spectral and Bispectral Analysis for Ship Detection From Acoustical Time Series*. Adelaide; 1994.
4. Pflug, L A; Loup, G E; Jackson, P. *Variability in Higher Order Statistics of Measured Shallow-water Shipping Noise*. Banff; 1997.
5. Baran, R. H.; Coughlin, J. P. *A Neural Network for Target Classification Using Passive Sonar*. 1991.
6. Hernandez, A. *Implementación de un Algoritmo de Generación de Huella Digital Acústica para el Reconocimiento de Sonidos*. Universidad de la Sierra Juárez; 2013.
7. Soares-Filho, W.; Seixas, J. M.; Caloba, L. P. *Enlarging Neural Class Detection Capacity in Passive Sonar Systems*. Scottsdale; 2002.
8. Eom, K.; Wellman, M.; Srour, N.; Hillis, D.; Chellappa, R. *Acoustic Target Classification Using Multiscale Methods*. 1997.
9. Ogden, G.; Zurk, M.; Jones, M.; Peterson, M. *Extraction of Small Boat Harmonic Signatures from Passive Sonar*. *J Acoust Soc Am*. 2011;
10. Huang, J.; Zhao, J.; Xie, Y. *Source Classification Using Pole Method of AR Model*. *IEEE Int Conf Acoust Speech Signal Process*. 1997;1:567–70.
11. Chen, C.; Lee, J.; Lin, M. *Classification of Underwater Signal Using Wavelet Transforms and Neural Networks*. *Math Comput Model*. 1998;27(2):47–60.
12. Meister, J. *A Neural Network Harmonic Family Classifier*. *J Acoust Soc Am*. 1993;93(3):1488–95.
13. Hunter, A.; Fillinger, L.; Zampolli, M.; Clarijs, M. *Recent Developments in SOBEK Passive Sonar Technology*. Alicante, Spain; 2012.
14. *Lecture 8. DT Filter Design: IIR Filters*. *Discrete-Time Signal Processing*. Massachusetts Institute of Technology; 2006.
15. Sanabria, A.C.; Predraza, C.; Vitola, J. *Quick Algorithm for Searching Audio by Content on GPUs*. *Intekhnia*. 2011;6(1):35–43.
16. Mitrovic, Dalibor; Zeppelzauer, Matthias; Breiteneder, Christian. *Features for Content-Based Audio Retrieval*. *Adv Comput*. 2010;78:71–150.
17. Burka, Zak. *Perceptual Audio Classification Using Principal Component Analysis*. Rochester: Golisano College of Computing and Information Sciences; 2010.
18. Bishop, Christopher. *Pattern Recognition and Machine Learning*. New York, USA: Springer; 2006.
19. *Digital Image Processing Using MATLAB*. Prentice Hall; 2004.
20. Gamarra, Margarita Rosa; Quintero, Christian. *Using genetic algorithm feature selection in neural classification systems for image pattern recognition*. *Ing E Investig*. 2013;33(1):52–8.
21. Ripley, B.D. *Pattern Recognition and Neural Networks*. Cambridge, United Kingdom: Cambridge University Press; 1996.
22. Molinero, L. *Métodos Estadísticos De Clasificación*. *Asoc Soc Esp Hipertens*. 2002;
23. Prati, R.C; Batista, G; Monard, M.C. *Evaluating Classifier Using ROC Curves*. *IEEE Latin America Transactions*. 2008;6(2):215–22.