Acoustic Fingerprint Recognition Using Artificial Neural Networks

Eduardo Zurek, Margarita R. Gamarra, José R. Escorcia, Carlos Gutierrez, Henry Bayona, Roxana P. Pérez, and Xavier García

Abstract—**This paper presents an implementation of Artificial Neural Networks (ANN) for acoustic fingerprints recognition, applied to the identification of marine vessels. In many cases, the vessel recognition process from an audible signal is performed by human operators, which could lead to failures in the identification process. Before entering the ANN classification process, the signal is filtered and its spectral characteristics are extracted. A comparison of the classification process between three types of neural networks is presented.**

Index Terms—**Acoustic fingerprint, FFT, PCA, ANN, feed-forward backpropagation, RBF, PNN.**

I. Introduction

THE detection of merchant vessels from an audible signal is one of the main tasks within the operations of underwater units. A human operator perform this process, which could lead to failures in the identification process due to factors such as physiological limitations, subjectivity in interpretation, emotional condition and noise present in the signal.

Therefore, poor identification cause misinformation and compromise the integrity of the submarine, the crew aboard it, and even the support staff that is requested for a specific action. Thus, the need for an automatic identification system to assist the operator is imperative to ensure safety and tactical coordination regarding specific operations.

The sound produced by a vessel is mainly caused by its impeller machinery and cavitation caused by the propeller. These sounds are unique to each type of vessel and therefore such information may be used for identification and tracking purposes. This signal is a stationary process given its characteristics of variation and its period [1–3], consequently, signal spectral analysis techniques have proven to be useful to differentiate and classify boats.

Section 2 presents related work. The proposed approach for the recognition of three types of vessels from its acoustic fingerprint is presented in Section 3. It involves the use of the Fast Fourier Transform (FFT) to extract the spectral characteristics of the signal, the Principal Component Analysis (PCA) method to reduce the dimension of extracted features and finally a classifier using ANN. Various tests were performed to evaluate the performance of each type of network in the process and these are presented in Section 4.

II. Related Works

Some techniques implemented for the automatic detection of boats are based on the extraction of features in the frequency domain. Some authors implement FFT, as in [4] to extract features. Using an omnidirectional hydrophone the power spectral density is extracted, for a moving object. With this information, a feed-forward neural network is trained.

Similarly, in [5] FFT is implemented to recognize acoustic fingerprint. This information is compared with a digital soundwaves database, which has been gathered for comparison and identification purposes.

In [6] equipment for the detection of UBA (Underwater Breathing Apparatus) is designed by acoustic signals. The acquired data were processed using Minimum Variance no Distortion Response (MVDR) techniques and the Multiple Signal Classification (MUSIC). Likewise, in [7] information from sonars is processed such that the amplitude spectrum is removed, then PCA is aimed to separate sets of boats, and the most significant components are used as input to a neural network applied for non-linear classification.

However, spectral analysis is not the only method used; autoregressive models are also used. In [8], authors use spectral information to design an autoregressive model or ARMA.

In [9], a method of treatment of the signal is developed to

Manuscript received on August 6, 2014; accepted for publication on October 2, 2014, published on November 15, 2014.

Eduardo Zurek (corresponding author) is with the Universidad del Norte, Department of Systems Engineering, Barranquilla, Colombia (e-mail: ezurek@uninorte.edu.co).

Margarita R. Gamarra is with the Universidad Autónoma del Caribe, Department of Electronic and Telecommunication Engineering, IET group, Barranquilla, Colombia (e-mail: margarita.gamarra@uac.edu.co).

José R. Escorcia is with the Universidad Autónoma del Caribe, Department of Electronic and Telecommunication Engineering, IET group, Barranquilla, Colombia, and with the Escuela Naval de Suboficiales A.R.C. Barranquilla, Colombia (e-mail: jescorcia@uac.edu.co).

Carlos Gutierrez is with the Escuela Naval de Suboficiales A.R.C. Barranquilla, GITIN group, Barranquilla, Colombia (e-mail: ing.cagm@ gmail.com).

Henry Bayona is with the Escuela Naval de Suboficiales A.R.C. Barranquilla, GITIN group, Barranquilla, Colombia (e-mail: elektronikesub@ gmail.com).

Roxana P. Pérez is with the Escuela Naval de Suboficiales A.R.C. Barranquilla, GITIN group, Barranquilla, Colombia (e-mail: roxp64@gmail. com).

Xavier García is with the Escuela Naval de Suboficiales A.R.C. Barranquilla, GITIN group, Barranquilla, Colombia (e-mail: xavier.garciah@ gmail.com).

automatically extract the harmonic structure of the noise radiated by small boats.

Artificial Neural Networks have a particular and important place in the classification stage of many of the strategies described in literature. In particular, [10] uses Kohonen neural network with smoothed spectral data and differences between components of order k as inputs.

In [11], a classification system based on neural networks for object recognition from a sound signal was used. The results suggest that the typical indoor objects can be distinguished on a range of distances with high accuracy based only on the information in the ultrasound echoes.

In [12], a method based on neural networks for detection and classification of underwater mines using sonar images is proposed. The cited articles show a successful classification rate higher than 90% using neural networks.

Likewise, in [13] a neural network scheme with supervised and unsupervised learning is used, so that the performances among these are compared, given the same spectral information in input.

The technological and scientific advances continue developing but despite the high value that acoustic fingerprints vessels have, publications regarding these applications are limited. The researches mentioned in this section present several applications using Neural Networks in the pattern recognition, but the use of techniques in the fingerprint acoustic recognition of vessel is limited. In this article we propose a vessel identification process using spectral features and ANN classifier, from acoustics signals generated by the own vessel.

III. Proposed Approach

In order to identify acoustic fingerprints the stages of the pattern recognition process is developed: filtering, extraction and selection of features and finally classification. In these stages, well-known audio signal processing techniques are implemented. Figure 1 shows the techniques used at each stage and the results.

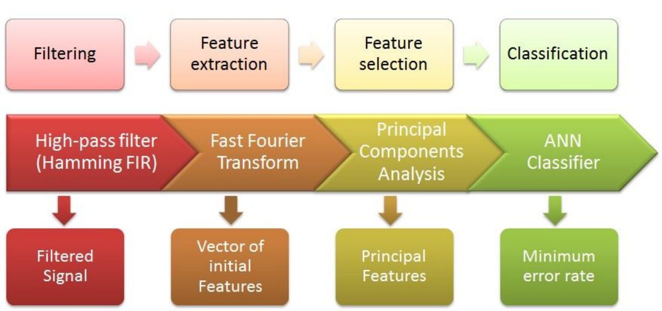

Fig. 1. Scheme of the proposed approach for recognition of acoustic fingerprint.

Initially, a database is generated with the reduced feature space, in order to perform the training of the ANN. Each audio sample is filtered, and then a vector of initial characteristics by applying the FFT to the filtered signal is obtained. The dimension of features is reduced using PCA and finally an ANN performs the classification process. This net is previously trained with the initial database. The next subsections present the techniques used in each stage.

A. *Filtering*

Due to acquisition hardware, acoustic signals have noise in frequency of 60 Hz and its harmonics. However, the representative frequencies generated by each boat are values higher than the 1000 Hz, and then it is necessary to apply a high pass filtering in order to eliminate the low frequency noise. This initial stage is suitable, because the next stages will work only with the representative frequencies of each signal, which facilitates the process of recognition.

In this case, the applied filter is FIR, taking into account the advantage presented in [14]: FIR systems are always stable, they are suited to multi-rate applications and they can design to yield a linear phase response. This filter use a window type Hamming, high-pass, order 100 and a cut-off frequency of 1 kHz. These parameters are selected by testing to the original acoustic signal such that a suitable filtered signal is obtained.

B. *Feature Extraction* [15]

Aim to identify the boat is necessary to extract features of the signal which allows characterize each pattern and classify the boat. The original audio signal is reduced to 3000 samples, which is a representative data frame of the signal. Then convolution is performed with a Hamming window to reduce the undesirable effects of the rectangular window [16]. FFT is applied to the resulting signal; this transform yields real and complex values: the real values represent the distribution of the frequency components while the complex values carry information on the phase of the components [17].

This information is adequate to represent the acquired audio signal. In this stage 512 Fourier coefficients are extracted, but this amount is reduced aim to distinguish the three sources sound using only the discriminant features. Then, it is necessary to use a selecting process to define the most significant features, which allows discriminate the three input signals.

C. *Feature Selection* [15]

In order to reduce the dimensionality of the feature space produced from the FFT algorithm applied to the audio signals, Principal Component Analysis (PCA) is used. PCA is a data analysis technique that is used in order to extract the discriminant features from a large data set [18]. In this case the number of features used in classification is tested from 256 to 64 first principal components aim to select the adequate size.

There are different algorithms for calculating the principal components of a data set, however the one used in this case is based on eigenvalue decomposition.

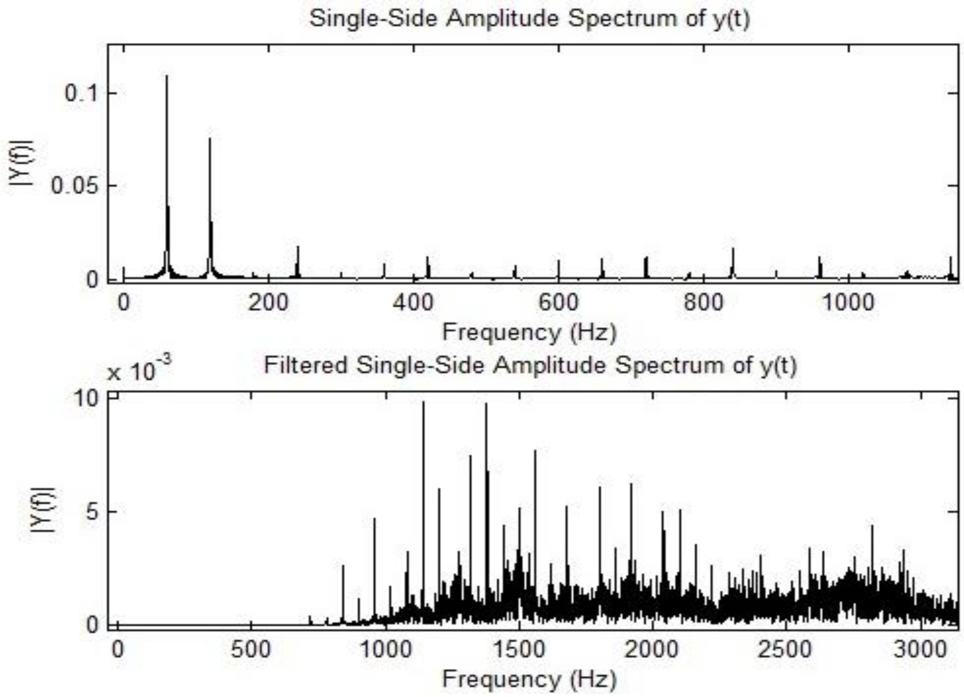

Fig. 2. Spectrum of the acoustic signal from a vessel. Top: original spectrum, below: filtered spectrum

IV. Classification

Once feature selection of all the database signals is completed, these features are used to train the ANN. After several re-training process, the ANN with the highest performance is selected [19].

Other set of acoustic signals are used to measure the performance of the neural network; this set is different to the training set. A comparison of hit rate between three different types of neural network classifier and the k-nearest neighbor (KNN) is performed. These neural networks are Feed-forward Backpropagation, Radial Basis Function (RBF) and Probabilistic Neural Network (PNN).

- Feed-forward networks have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors [20].

- The RBF is a feed forward network, consists of one input layer, hidden layer and the output layer. The RBF is different from the ordinary feed forward networks in calculating the activations of hidden neurons. The activations at the hidden neurons are computed by using the exponential of distance measures [21]. In this case, the net input to the transfer function is the vector distance between its weight vector w and the input vector p, multiplied by the bias *b*. The transfer function for a radial basis neuron is exponential.

- The PNN consists of input layer, two hidden layers, and an output layer. The process based classification that differentiates PNN and RBF is that PNN works on the estimation of probability density function [21]. When an input is presented, the first layer computes distances from the input vector to the training input vectors, and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a competitive transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 for that class and a 0 for the other classes [22].

V. Results

All stages proposed in Figure 1 were performed in Matlab 7.11. In Figure 2, the spectrum of the signal without filtering is shown, where the harmonics of the frequency of 60Hz are predominant and below, the filtered signal is shown where the 60Hz harmonics are eliminated. The used filter was described in subsection III.A.

Although higher frequencies of 1kH, 60Hz harmonics are also present, the dominant signal is typical of each vessel, facilitating the subsequent classification process.

In the training process of the ANN, 1500 samples, 500 of each type of vessel used to identify. To determine the performance of the network created 1500 different samples were used and tested in the classification process.

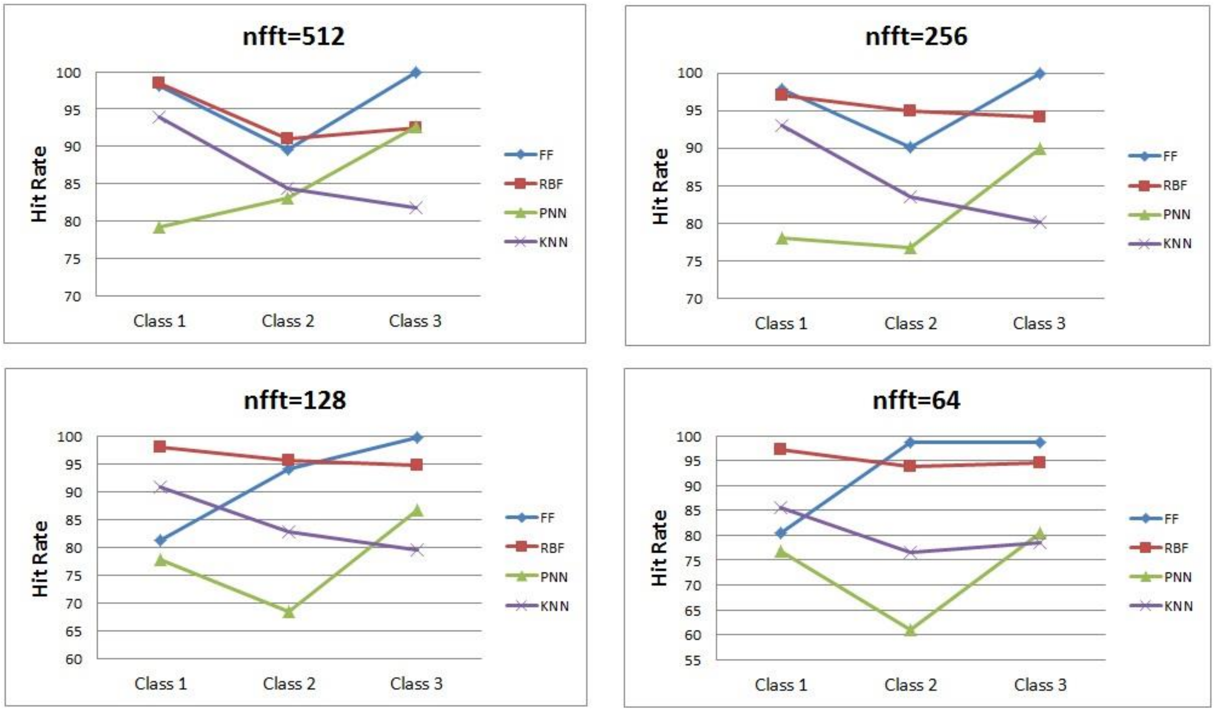

Fig. 3. Graphics of Hit Rate for each classifier

The propagation constant for RBF and PNN was set at 1. The Feed-Forward Backpropagation neural network was set with the following parameters:

– 1 hidden layer sigmoid.
– 20 neurons in the hidden layer.
– Training function: Leveberg-Marquardt optimization.
– Transfer function: Hyperbolic tangent sigmoid.

Table 1 presents a comparison between the hit percentages for each type of neural network classifier and KNN in each case varying the number of Fourier coefficients used (*nfft*).

TABLE 1.
HIT RATE WITH DIFFERENT NFFT

nfft	Feed-Fordward Backpropagation	Radial Basis	Probabilístic	KNN
512	95.9%	93.9%	84.9%	86.73%
256	96.0%	95.4%	81.6%	85.60%
128	91.6%	96.1%	77.6%	84.33%
64	90.7%	95.2%	72.7%	80.27%

Using 256 coefficients Feed-Forward and RBF neural networks presented a higher performance that using 512; this can be explained due to some features (coefficients) are not distinctive and it can conduce to a misclassification.

An adequate threshold for *nfft* can be considered 128 value and upper, but a value higher than 512 could have a lower or equal performance that this value, with more computational cost.

Traditional metrics for the quality of the classification method are the hit rate and the confusion matrix which conveys information relative to this rate [23]. Taking into account the values of the confusion matrix, the graphics of the Figure 3 were obtained. The hit rate of the classifiers for each class of vessel to identify is drawn for different *nfft* values.

This graphics shows that the Feed-forward and Radial Basis classifiers have a higher hit rate that the PNN classifier. For Class 1 and 2, KNN classifier presented better performance that the PNN classifier. In all *nfft* cases, the Feed-forward network had the highest hit rate for the identification of the vessel Class 3.

Then, for this particular recognition case (vessel identification from acoustic fingerprint) using 512 or 256 coefficients, the implementation of Feed-forward and RBF neural networks is suggested; due to these classifiers had a higher hit rate (greater than 90%) than the other ones tested.

VI. CONCLUSION

A spectral analysis was used for boat recognition using their acoustic fingerprint. A previous filtering process is needed in order to reduce the 60Hz noise in the audio signal.

The FFT was calculated using four different numbers of sampling points (nfft) in order to set the value that gets higher performance. The hit rate for four each nfft number was obtained using three different ANN architectures and KNN classifiers.

According to results presented in Table 1 and Figure 3 the Feed-forward and RBF neural networks classifiers have a higher hit rate for three vessel classes, compared with PNN and KNN classifiers.

Using a FIR high-pass filter, 256 coefficients extracted from FFT application and an ANN classifier (Feed-forward or RBF) is possible obtain a good performance in vessel recognition process using the acoustic fingerprint.

It is important to implement the identification proposed approach in an embedded system, which would greatly reduce the processing time and improve the vessel identification process.

ACKNOWLEDGMENT

The work was supported by the Escuela Naval de Suboficiales ARC Barranquilla and Colciencias. The work has been financed by a research grant from Colciencias under call 632, Armada Nacional and Ministerio de Defensa de Colombia: "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".

REFERENCES

[1] R. J. Urick, *Principles of Underwater Sound*. New York: McGraw-Hill, 1983.

[2] C. Regazonni, A. Tesei, and G. Tacconi, "A Comparison Between Spectral and Bispectral Analysis for Ship Detection From Acoustical Timese R. R. Raories," in *IEEE International Conference on Acoustics, Speech and Signl Processing*, Adelaide, 1994.

[3] L. A. Pflug; G. E. Loup, and P. Jackson, "Variability in Higher Order Statistics of Measured Shallow-water Shippin Noise," in *IEEE Signal Processing Workshop on Higher-Order Statistics*, Banff, 1997.

[4] R. H. Baran and J. P. Coughlin, "A Neural Network for Target Classificacion Using Passive Sonar," in *Proceedings of the Conf. on Analysis of Neural Networks Application*, Assoc. Comput. Mach., 1991.

[5] A. Hernandez, *Implementación de un Algoritmo de Generación de Huella Digital Acústica para el Reconocimeinto de Sonidos*, Universidad de la Sierra Juarez, 2013.

[6] L. Fillinger, A. Hunter, M. Zampolli and M. Clarijs, "Passive acoustic detection of closed-circuit underwater breathing apparatus in an operational port environment," *J. Acoust. Soc. Am.*, 2012.

[7] W. Soares-Filho, J. M. Seixas, and L. P. Caloba, "Enlarging Neural Class Deetection Capacity in Passive Sonar Systems," in *IEEE Int. Symp. on Circuits ans Systems*, Scottsdale, 2002.

[8] K. Eom, M. Wellman, N. Srour, D. Hillis, and R. Chellappa, "Acoustic Target Calssification Using Multscale Methods," in *Sensors and Electron Devices Symp*., 1997.

[9] G. Ogden, M. Zurk, M. Jones, M. Peterson, "Extraction of Small Boat Harmonic Signatures from Passive Sonar," *J. Acoust. Soc. Am.*, 2011.

[10] J. Meister, "A Neural Network Harmonic Family Classifier," *J. Acoust. Soc Am*, vol. 93, no. 3, pp. 1488–1495, 1993.

[11] W. Streilein, P. Gaudiano, and G. Carpenter, "A neural network for object recognition through sonar on a mobile robot," in *Proceedings of Intelligent Control (ISIC)*, held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA), Intelligent Systems and Semiotics (ISAS), 1998.

[12] S. Geethalakshmi, P. Subashini, and S. Ramya, "A Study on Detection and Classification of Underwater Mines Using Neural Networks," *Int. J. Soft Comput. Eng. IJSCE*, vol. 1, no. 5, pp. 150–157, 2006.

[13] B. Howell, S. Wood, and S. Koksal, "Passive sonar Recognition and analysis using hubrid neural networks," *IEEE Oceans*, vol. 4, pp. 1917–1924, 2003.

[14] "Lecture 8. DT Filter Design: IIR Filters," in *Discrete-Time Signal Processing*, Massachusetts Institute of Technology, 2006.

[15] E. Zurek, M. Gamarra, J. Escorcia, C. Gutierrez, R. Pérez, X. García, H.. Bayona, "Spectral Analysis Techniques for Acoustic Fingerprints Recognition," in *2014 XIX Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, IEEE, 2014.

[16] A. C. Sanabria, C. Predraza, J. Vitola, "Quick Algorithm for Searching Audio by Content on GPUs," *Intekhnia*, vol. 6, no. 1, pp. 35–43, 2011.

[17] D. Mitrovic, M. Zeppelzauer, C. Breiteneder, "Features for Contet-Based Audio Retrieval," *Adv. Comput.*, vol. 78, pp. 71–150, 2010.

[18] Z. Burka, *Perceptual Audio Calssification Using Principal Component Analysis*, Golisano College of Computing and Information Sciences, Rochester, 2010.

[19] M. R. Gamarra, C. Quintero, "Using genetic algorithm feature selection in neural classification systems for image pattern recognition," *Ing. E Investig.*, vol. 33, no. 1, pp. 52–58, 2013.

[20] "Neural Networks Toolbox," *Czech Technical University*. [Online]. Available: http://radio.feld.cvut.cz/matlab/toolbox/nnet/backpr52.html.

[21] T. Sitamahalakshmi, A. Vinay, M. Jagadeesh, and K. Chandra, "Performance Comparison of Radial Basis Function Networks and Probabilistic Neural Networks for Telugu Character Recognition," *Glob. J. Comput. Sci. Technol.*, vol. 11, no. 4, 2011.

[22] *Radial Basis Networks*. [Online]. Available: http://www.inf.ufrgs.br/~engel/data/media/file/cmp121/rbf-matlab.pdf.

[23] R. C. Prati, G. Batista, M. C. Monard, "Evaluating Classifier Using ROC Curves," *IEEE Latin America Transactions*, vol. 6, no. 2, pp. 215–222, 2008.