A Methodology for Detection and Localization of Dynamite Fishing

S. S. Bharadwaj[1], B.N. Sumukha[1], R. C. Kumar[1], S. Simha[2], H. S. Jamadagni[2], and K. George[1,3]
[1]Dept of Telecommunication Engineering, PES Institute of Technology, Bangalore, India
[2]Dept of Electronic Systems Engineering, Indian Institute of Science, Bangalore, India
[3]PES Centre for Intelligent Systems, Bangalore, India

Abstract— **The illegal practice of dynamite fishing upsets the balance of an ecosystem and can endanger human lives as well. Detection and localization of such activity is necessary from several viewpoints. A methodology for detection and localization of such dynamite blasts is presented in this paper. The sound signals are recorded using an array of sensors, filtered, and classified using an artificial neural network. The location of the sound-source is also simultaneously determined using these measured signals.**

Keywords — Detection, Localization, Dynamite Fishing

I. Introduction

Dynamite fishing is an illegal practice of killing a school of fish by blasting dynamites, which still remains widespread in countries like Indonesia, Philippines and Tanzania. It is also well documented in few other Southeast Asian countries, as well as in the Aegean Sea, and coastal Africa. This practice has also been spotted along the riverside of the river Kaveri in Karnataka, India. Homemade bombs that contain a mixture of powdered potassium nitrate and pebbles, or ammonium nitrate and kerosene, are employed to stun the fish for easy collection. The swim bladders of fish rupture creating a loss of buoyancy and make the fish float on the surface. In addition to the effect on fish, a large number of underwater organisms are killed ruthlessly by the use of these explosives. Moreover, coral reefs are damaged due to these explosions. Further, the explosions have been known to harm the physical ecosystem and the person using them as well. A methodology to detect and localize dynamite fishing is therefore necessary and important. However, there appears to be no prior work reported in the literature.

On the contrary, several researchers have dealt with impulsive sounds that are characterized by the sudden presence of a pressure wave. Examples include gunshots, thunder and screams. The approaches to detect such impulsive sounds are quite varied. For example, a correlator compares the autocorrelation against a threshold in [1] for the detection of gunshots. Six detection algorithms for gunshots are evaluated in [2]. It was suggested that a wavelet filter banks provides the best trade-off between detection efficiency and power requirements in a VLSI implementation. A combination of absolute threshold detector together with two exponentially weighted moving average detectors triggers local detection at a node of a wireless sensor network in [3]. A final voting process amongst a group of nodes is then used to detect volcanic eruptions. The variations of the sound energy are used as a measure to detect impulsive sounds in [4]. In contrast to these, the extracted features are used for both detection and classification in [5].

Once an impulsive sound has been detected, researchers have approached classification using a variety of tools. These include Hidden Markov Model (HMM) [1,4], Artificial Neural Networks (ANNs) [4], Gaussian mixture model (GMM) [4,5]. These classifiers require features to be extracted. Both temporal features such as correlation as well as perceptual features such as mel-frequency cepstral coefficients (MFCC) were utilized in [1]. A combination of Linear Predictive Coding (LPC), cepstral and MFCC coefficients are considered in [4]. A total of 49 features are initially considered in [4], and comprise temporal, spectral, perceptual, and correlation-based. Subsequently, the number of features is reduced by a procedure that involves both open-loop metrics and metrics that allow feedback from the classifier. A number of temporal and spectral features extracted from electromyography signals are described in [6].

The difference in the time of arrival of the signal at the sensors is used to localize the source in [5]. Here, the array of four sensors is T-shaped with the center microphone taken as the reference. The maximum likelihood generalized cross correlation method is used to estimate the time delays. In turn, this requires the computation of the minimum variance distortionless response spectra. Subsequently, the source is localized by solving a least-squares problem. Triangulation is a well-established technique for localization. For example, it is used to determine the location using a GPS receiver [8]. It is also used to track the location of a mobile phone using the signal strength to nearby antenna towers [9]. Triangulation is based on estimates of quantities such as time of arrival, time difference of arrival, or angle of arrival. (See, for example, [9]-[11].) In contrast, the Received Signal Strength Indicator (RSSI) is a measure of the strength of the signal received by a sensor from a source and can be indicative of the distance between the two. This measure has been used by several researchers. For example, this was used in [12] to determine the location of a mobile node in a wireless sensor network and applied to robot navigation in [13]. Here, an empirical formulation is done to obtain a relationship between received signal strength and distance.

Evidently, there appears to be no paper that suggests a technique for the detection and localization of dynamite blasts. In this paper, we propose a deployable system that can detect and localize a dynamite blast. The novel methodology incorporates pre-processing of signals, followed by feature extraction and classification using a trained artificial neural network for detection, and localization carried out using the measure RSSI. This is described in Section II, and the results presented in Section III.

II. The Methodology

The proposed methodology is depicted in Figures 1 and 2, and respectively describes detection and localization. The continuous-time sound signal $x(t)$ is recorded using a microphone and sampled at 8 kHz. The power spectrum of the resulting discrete-time signal $x(n)$ is analyzed and a band-pass filter $H(z)$ (with impulse response $h(n)$) is designed to filter out the noise. The features of the filtered output $y(n)$ are then extracted.

Three sets of features are considered here; these include parameters related to the statistics of the signal, those that are related to the spectrum of the signal, and those that are related to the compressed form of the spectral envelope of the signal. (The latter is obtained by using the tool linear predictive coding (LPC).) A trained artificial neural network is then used to classify whether or not the recorded signal had resulted from a dynamite blast.

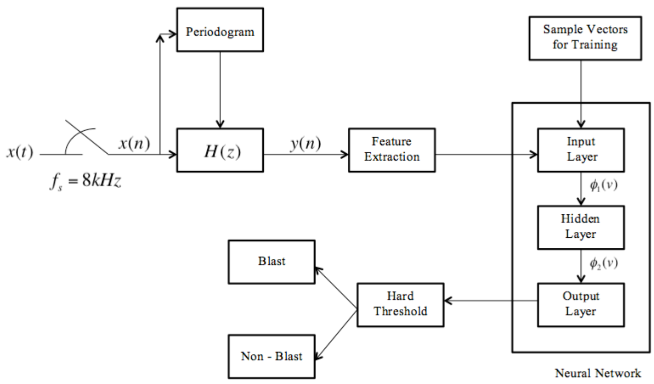

Fig. 1. Design flow of the proposed method: Detection.

Triangulation technique is incorporated to localize the blast. An arbitrary point is to be the source of the blast signal. The distance of the source from each vertex of the triangle is calculated using the received strength of the signal obtained at each vertex. (In this paper, we use the Received Signal Strength Indicator (RSSI) as a measure.) A unique solution to three equations of the circles formed with nodes as centers and distances between nodes and the point of blast as radii gives the coordinates of source of the blast.

In what follows, we describe the different steps involved:

A. *Pre-Processing*

For purposes of testing the methodology, sound signals from databases ([14] and [15]) are used. These databases include signals resulting from explosion of bombs and dynamites, thunder and gunshots. Whilst the former are referred to as "blast" signals in the sequel, the latter are referred to as "non-blast" signals. The database consists of 34 signals with 17 blast and 17 non-blast signals.

The sound signals from the database are sampled at 8 kHz and their spectrum estimated using the periodogram method. From the spectrum, it is observed that the frequency components of the signal lie in the range from 50 Hz to 450 Hz. Accordingly, a suitable band-pass filter is used to extract the signal. In this paper, an FIR filter with 566 coefficients is designed using a Hamming window.

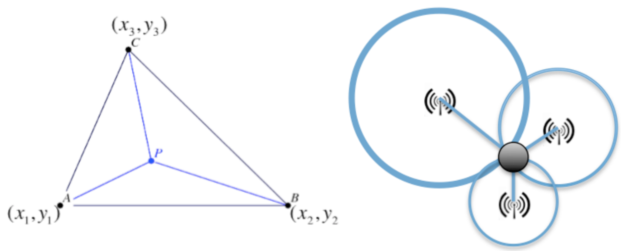

Fig. 2. Design flow of the proposed method: Localization.

B. *Relevant Features*

The redundancy within a large number of samples of a signal is removed by extracting those features that retain the relevant information. The dimension is then naturally smaller and the resulting feature vector is easier to use for pattern classification compared to the actual signal. The choice of the relevant features depends on the application. A number of features can be extracted (see, e.g., [6]). For the purpose of classification of sound signals within the context of detection of dynamite blasts, the following features have been experimentally found to be relevant: These include statistical parameters like the average value, the variance and the related standard deviation, and the root mean square value. Other parameters include the average change of amplitude, the mean frequency, the peak frequency, the total power, and the mean and median powers. These ten features are summarized in Table 1. The last five parameters are computed using the power P_j at the jth frequency bin f_j.

In addition to these features, we also use Linear Predictive Coding (LPC) coefficients. LPC is a tool used primarily in audio and speech processing for representing the spectral envelope of the signal in compressed form, in the form of a linear predictive model. In this paper, 50 LPC coefficients are considered. These are computed using the Levinson method [7]. Therefore, in this paper, the feature vector consists of 60 elements.

Tab. 1. Some extracted features.

#	Feature	Formula
1	Average value (μ)	$\mu = \frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert$
2	Variance (σ^2)	$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N}(x(n)-\mu)^2$
3	Standard deviation (σ)	$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x(n)-\mu)^2}$
4	Root mean square (RMS)	$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N}x^2(n)}$
5	Average change in amplitude (AC)	$AC = \frac{1}{N}\sum_{n=1}^{N}\lvert x(n+1)-x(n)\rvert$
6	Total power (TP)	$TP = \sum_{j=1}^{M}P_j$
7	Mean power (MP)	$MP = \frac{1}{M}\sum_{j=1}^{M}P_j$
8	Median power (MDP)	$MDP = \frac{1}{2}\sum_{j=1}^{M}P_j$
9	Mean frequency (MF)	$MF = \frac{\sum_{j=1}^{M}f_j\,P_j}{\sum_{j=1}^{M}P_j}$
10	Maximum power (MXP)	$MXP = \max_{1\le j\le M}P_j$

C. The Neural Network for Pattern Classification

Pattern classification can be achieved using several approaches (see, for example, [17]). Although statistical methods such as Bayesian Decision Theory can be used, it has been our experience that Artificial Neural Networks (ANNs) are powerful tools that are easier to use. The principal reason for this is that ANNs are universal approximators in that a sufficiently smooth function can be approximated to any a priori specified accuracy by a three-layered feedforward neural network which has a sufficient number of neurons in the hidden layer. In the case of pattern classification, the separation of the classes (i.e., the discriminant function) is approximated well by the artificial neural network.

We consider a feedforward neural network that is trained in a supervised manner using the standard back propagation algorithm. The chosen network consists of one hidden layer. The input layer has 60 nodes corresponding to the number of elements of the feature vector, and the output consists of 2 neurons, as there are two classes of patterns – blast signals and non-blast signals. The hidden layer consists of 100 neurons. The hidden layer is a nonlinear layer with the sigmoidal activation function (1), and the output layer is also nonlinear with the *softmax* activation function (2).

$$\varphi_{1,k}(v_k) = \tanh(v_k), \quad 1 \le k \le 100 \tag{1}$$

$$\varphi_{2,l}(v_l) = \frac{e^{v_l}}{e^{v_1}+e^{v_2}}, \quad 1 \le l \le 2 \tag{2}$$

The softmax function is clearly a normalized exponential function, and it generalizes the logistic function. The resultant outputs lie in the interval [0,1], and they add to one. Therefore, they represent a categorical probability distribution. Accordingly, softmax classifiers are more suitable compared to binary classifiers when the classes are mutually exclusive.

D. Localization

Localization is a process of reporting the origin of events, and determines either the physical position or the logical location. The main objective of localization is to determine the location of an unknown node, or an event, as accurately as possible from the information obtained from a set of nodes whose locations are predefined. In this paper, we use the Received Signal Strength Indicator (RSSI) as a parameter to determine the distance between the sensor nodes and the point of blast.

The RSSI is based on the fact that the signal strength is inversely proportional to the squared distance between the transmitting node and the receiving node. A known radio propagation model is used to convert the received signal strength into distance. Either empirical or theoretical models are used to translate signal strength into distance.

In the proposed methodology, the sensors that are used to record the sound signal are the nodes. For the purpose of this paper, we assume that these nodes are placed along lines parallel to each other and separated from each other by an a priori known fixed distance. The arrangement is shown in Fig. 3. Accordingly, the point of blast is inside a triangle with vertices formed by three sensors, with one sensor from one line of nodes, and two sensors from the other line. The described scenario is shown in Fig. 2.

Several methods can be used to determine the relationship between the power of the signal and the geographical distance. Two methods considered here are nonlinear regression analysis and an artificial neural network. Using the developed relationship, the coordinates of the blast can be determined by solving three circle equations. These circles are drawn with the nodes as centers and the distances between the nodes and the point of blast as radii. Such a triangulation process is shown in Fig. 2.

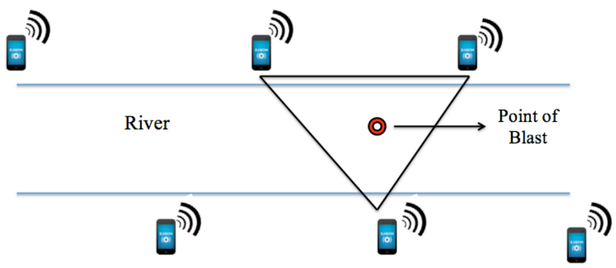

Fig. 3. Scenario for Localization

III. RESULTS AND DISCUSSIONS

The signals used in this paper are from the databases mentioned earlier. We first describe the training of the network using these signals. It may be recalled that, of the 34 available signals, 17 are blast signals and the remaining 17 are non-blast signals. Of these, 28 sound signals are used to train the network using the back-propagation algorithm, 10% of the signals (i.e., 3 signals) are used to test the trained network, and the remaining 10% used to validate the network. We use a network with 60 input nodes, 100 neurons in the hidden layer, and 2 neurons in the output layer. The activation functions for the hidden and output layers are respectively given in (1) and (2). The training, testing and validation accuracies are depicted as the confusion matrix in Fig. 4. Here, three matrices provide mismatch, false-positives, the accuracies of training, testing and validation and the fourth matrix provides the mismatch, false-positives and the accuracy of the trained neural network.

<u>Comments:</u> (i) If only the parameters depicted in Table 1 are considered for training, then the achieved classification accuracy is only 88%. On the contrary, if only LPC coefficients are used then the accuracy is 94%. This shows that both sets of parameters are required to maximize the accuracy. As mentioned earlier, the combination of these parameters is essential to yield 100% classification accuracy. (ii) One may argue that since variance and standard deviation are related, the features may contain redundancy. Whilst this is true in that using only either of the feature yields 100% classification accuracy, we use both for the following purpose: The match between the desired output and the actual output is near perfect when both the parameters are used thereby making the use of a hard-limit function redundant.

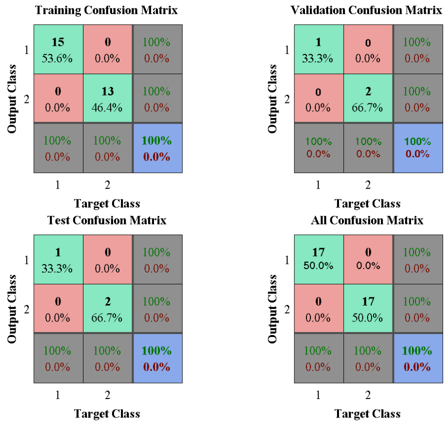

Fig. 4. Confusion matrix for the trained neural network.

The experiments in the laboratory are conducted as follows: One of the signals corresponding to blast is played, and the resulting signals recorded using a Cyanogen OnePlus cellular phone instrument. This instrument has three in-built microphones for better noise cancellation and audio enhancement. The end result is a single recorded signal. The simulated blast is recorded at spatial intervals of 50 cms.

The time-domain and the frequency-domain plots of the measured signals are respectively shown in Fig. 5 (a) and (c), and of the filtered signals respectively in Fig. 5 (b) and (d). It may be recalled that the filter is band-pass and is based on a Hamming window. It can be observed from this figure that the noise in the measured signal is effectively removed. The filtered signal characterized by its feature vector is then presented to the trained network, which correctly classifies it as a blast signal.

In order to localize the blast, we first obtain the distance versus power plot; this is shown in Fig. 6. The measured data is shown as circles in the figure. A curve is fit to the measured data using nonlinear regression analysis; the resulting curve is shown as a solid line in Fig. 6. The resulting nonlinear function is given in (3).

$$d(p) = 12.47e^{-52.46p} + 1.66e^{-4.581p} \qquad (3)$$

Using this relationship, the three circles are drawn using equations (4), and the location of the blast is determined. The accuracy of localization is found to be 94%.

$$
\begin{aligned}
(x-x_1)^2+(y-y_1)^2&=d_1^2\\
(x-x_2)^2+(y-y_2)^2&=d_2^2\\
(x-x_3)^2+(y-y_3)^2&=d_3^2
\end{aligned}
\tag{4}
$$

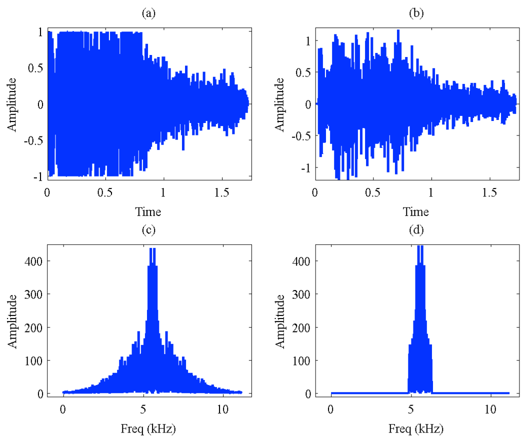

Fig. 5. Blast signal: (a) Time-domain plot of the raw signal, (b) time-domain plot of the filtered signal, (c) frequency-domain plot of the raw signal, and, (d) frequency-domain plot of the filtered signal.

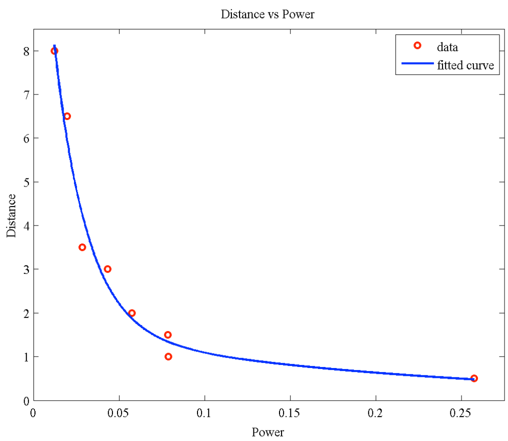

Fig. 6. Curve fitting using non-linear regression analysis

It is possible to exploit the function approximation property of an artificial neural network to obtain the relationship between power and distance. The advantages of using a network are that it can learn the patterns in a given data, and fit a curve. It has been our experience that for the data depicted in Fig. 6, the localization accuracy obtained using a neural network is comparable to the results obtained using nonlinear regression analysis. However, neural networks have an additional advantage in that it can handle larger number of data samples as opposed to regression analysis. The resulting localization accuracy improves with a larger number of measurements.

IV. Conclusions

In this paper we developed a methodology for detection and localization of dynamite blast fishing. The achieved accuracy of pattern classification is 100% by the particular choice of the feature vector. The detection included statistical parameters; parameters derived from the power spectrum of the signal, and LPC coefficients. The location of the dynamite blast is determined using triangulation with RSSI as a measure, and nonlinear regression analysis of the measured data. The location prediction accuracy achieved is 94%. Work is currently in progress for real-time deployment in field trials. Towards this, the developed methodology has been successfully implemented on Intel Galileo Gen 2.

Acknowledgment

We thank Sesha Raghav, Department of DESE, IISc Bangalore, for providing us with invaluable and constructive suggestions.

References

[1] I. L. Freire and J. A. Apolinário Jr., "Gunshot detection in noisy environment," in *Proceeding of 7th International Telecommunications Symposium (ITS2010)*, Manaus, Brazil, 2010.

[2] A. C. Rodríguez, L. Castro, P. Alvarad, P. Juliá and N. Hernánde, "Evaluation of gunshot detection algorithms", *IEEE Transactions on Circuits and Systems — I,* vol. 58, no. 2, pp. 363-373, February 2011.

[3] G. Warner-Allen, J. Johnson, M. Ruiz, J. Lees, and M. Welsh, "Monitoring volcanic eruptions with a wireless sensor network," in *Proceedings of the Second European Workshop on Wireless Sensor Networks (EWSN'05),* Istanbul, Turkey, January-February 2005, pp.108-120.

[4] M. A. Dufaux, "Detection and Recognition of Impulsive Sound Signals," Ph.D. Thesis, Institut de Microtechnique, Universite de Neuchatel, Switzerland, 2001.

[5] G. Valenzise, L. Gerosa, M. Tagliasacchi, F. Antonacci and A. Sarti, "Scream and gunshot detection and localization for audio-surveillance systems," in *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS 2007),* London, UK, September 2007, pp. 21-26.

[6] A. Phinyomark, A. Nuidod, P. Phukpattaranont, C. Limsakul, "Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification," *Elektronika ir Electrotechnika,* vol. 122, no. 6, 2012, pp. 27-32.

[7] N. S. Behbahan, S. Azari and H. Bahadori, "An introduction to the features extracted from the audio signal," *International Journal of Advanced Research in IT and Engineering*, vol. 2, no. 8, August 2013, pp. 39-46.

[8] A. El-Rabbany, *Introduction to GPS: The Global Positioning System,* Norwood, MA, USA: Artech House, 2002.

[9] M. Pourhomayoun and M. Fowler, "Improving WLAN-based indoor mobile positioning using sparsity," *Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Pacific Grove, CA, USA, November 2012, pp. 1393-1396.

[10] S. R. Leelavathy and S. Sophia, "Providing localization using triangulation method in wireless sensor networks," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* vol. 1, no. 6, November 2014.

[11] A. Varshavsky, M.Y. Chen, E. de Lara, J. Froehlich, D. Haehnel, J. Hightower, A. LaMarca, F. Potter, T. Sohn, K. Tang, and I. Smith, “Are GSM phones THE solution for localization?” *Seventh IEEE Workshop on Mobile Computing Sysetms and Applications (WMCSA’06),* Orcas Island, WA, USA, August 2005, pp. 34-42.

[12] N. Zhou, X. Zhao and M. Tan, “Deployment and routing method for fast localization based on RSSI in hierarchical wireless sensor network,” in *Proceedings of the IEEE 10th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS),* Hangzhou, China, October 2013, pp. 614-619.

[13] N. Zhou, X. Zhao and M. Tan, “RSSI-based mobile robot navigation in grid-pattern wireless sensor network,” in *Proceedings of the Chinese Automation Congress (CAC),* Changsa, China, November 2013, pp. 497-501.

[14] Sound database
http://soundbible.com/

[15] Sound database
https://www.audioblocks.com

[16] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification,* 2nd ed., New York, NY, USA: John Wiley & Sons, 2001.

[17] S. Haykin, *Neural Networks – A Comprehensive Foundation,* 2nd ed., Englewood Cliffs, New Jersey, USA: Pearson Education, 2002.