

Underwater Acoustic Source Localization and Sounds Classification in Distributed Measurement Networks

Octavian Adrian Postolache^{1,2}, José Miguel Pereira^{1,2}
and Pedro Silva Girão¹
*¹Instituto de Telecomunicações,
²ESTSetúbal/IPS (LabIM)
Portugal*

1. Introduction

Underwater sound signals classification, localization and tracking of sound sources, are challenging tasks due to the multi-path nature of sound propagation, the mutual effects that exist between different sound signals and the large number of non-linear effects that reduces substantially the signal to noise ratio (SNR) of sound signals. In the region under observation, the Sado estuary, dolphins' sounds and anthropogenic noises are those that are mainly present. Referring to the dolphins' sounds, they can be classified in different types: narrow-band-frequency-modulated continuous tonal sounds, referred to as whistles, broadband sonar clicks and broadband burst pulse sounds.

The system used to acquire the underwater sound signals is based on a set of hydrophones. The hydrophones are usually associated with pre-amplifying blocks followed by data acquisition systems with data logging and advanced signal processing capabilities for sound recognition, underwater sound source localization and motion tracking. For the particular case of dolphin's sound recognition, dolphin localization and tracking, different practical approaches are reported in the literature that combine time-frequency representation and intelligent signal processing based on neural networks (Au et al., 2000; Wright, 2002; Carter, 1981).

This paper presents a distributed virtual system that includes a sound acquisition component expressed by 3 hydrophones array, a sound generation device, expressed by a sound projector, and two acquisition, data logging, data processing and data communication units, expressed by a laptop PC, a personal digital assistant (PDA) and a multifunction acquisition board. A water quality multiparameter measurement unit and two GPS devices are also included in the measurement system.

Several filtering blocks were designed and incorporated in the measurement system to improve the SNR ratio of the captured sound signals and a special attention was dedicated to present two techniques, one to locate sound signals' sources, based on triangulation, and other to identify and classify different signal types by using a wavelet packet based technique.

2. Main principles of acoustics' propagation

Sound is a mechanical oscillating pressure that causes particles of matter to vibrate as they transfer their energy from one to the next. These vibrations produce relatively small changes in pressure that are propagated through a material medium. Compared with the atmospheric pressure, those pressure variations are very small but can still be detected if their amplitudes are above the hearing threshold of the receiver that is about a few tenths of micro Pascal.

Sound is characterized by its amplitude (i.e., relative pressure level), intensity (the power of the wave transmitted in a particular direction in watts per square meter), frequency and propagation speed.

This section includes a short review of the basic sound propagation modes, namely, planar and spherical modes, and a few remarks about underwater sound propagation.

2.1 Plane sound waves

Considering an homogeneous medium and static conditions, i.e. a constant sound pressure over time, a stimulation force applied in YoZ plane, originates a plane sound wave traveling in the positive x direction whose pressure value, according to Hooke's law, is given by,

$$p(x) = -Y \cdot \varepsilon \quad (1)$$

where p represents the differential pressure caused the sound wave, Y represents the elastic modulus of the medium and ε represents the relative value of its mechanical deformation caused by sound pressure.

For time-varying conditions, there will be a differential pressure across an elementary volume, with a unitary transversal area and an elementary length dx, given by,

$$dp = \frac{\partial p(x,t)}{\partial x} \cdot dx \quad (2)$$

Using Newton's second law and the relationships (1) and (2), it is possible to obtain the relation between time pressure variation and the particle speed caused by the sound pressure,

$$\frac{\partial p}{\partial x} = -\rho \cdot \frac{\partial u(x,t)}{\partial t} \quad (3)$$

where ρ represents the density of the medium and $u(x,t)$ represents the particle speed at a given point (x) and a given time instant (t).

Considering expressions (1), (2) and (3), it is possible to obtain the differential equation of sound plane waves that is expressed by,

$$\frac{\partial^2 p}{\partial t^2} = \frac{Y}{\rho} \cdot \frac{\partial^2 p}{\partial x^2} \quad (4)$$

where Y represents the elastic modulus of the medium and ρ represents its density.

2.2 Spherical sound waves

This approximation still considers a homogeneous and lossless propagation medium but, in this case, it is assumed that the sound intensity decreases with the square value of the

distance from sound source ($1/r^2$), that means, the sound pressure is inversely proportional to that distance ($1/r$).

In this case, for static conditions, the spatial pressure variation is given by (Burdic, 1991),

$$\nabla p = \frac{\partial p}{\partial x} \cdot \hat{u}_x + \frac{\partial p}{\partial y} \cdot \hat{u}_y + \frac{\partial p}{\partial z} \cdot \hat{u}_z \quad (5)$$

where \hat{u}_x , \hat{u}_y and \hat{u}_z represent the Cartesian unit vectors and ∇ represents the gradient operator.

Using spherical polar coordinates, the sound pressure (p) depends only on the distance between a generic point in the space (r, θ, ϕ) and the sound source coordinates that is located in the origin of the coordinates' system. In this case, for time variable conditions, the incremental variation of pressure is given by,

$$\frac{1}{r} \cdot \frac{\partial^2 (r \cdot p)}{\partial t^2} = \frac{\partial^2 p}{\partial t^2} \quad (6)$$

where r represents the radial distance between a generic point and the sound source.

Concerning sound intensity, for spherical waves in homogeneous and lossless mediums, its value decreases with the square value of the distance (r) since the total acoustic power remains constant across spherical surfaces.

It is important to underline that this approximation is still valid for mediums with low power losses as long as the distance from the sound source is higher than ten times the sound wavelength ($r > 10 \cdot \lambda$).

2.3 Definition of some sound parameters

There are a very large number of sound parameters. However, according to the aim of the present chapter, only a few parameters and definitions will be reviewed, namely, the concepts of sound impedance, transmission and reflection coefficients and sound intensity.

The transmission of sound waves, through two different mediums, is determined by the sound impedance of each medium. The acoustic impedance of a medium represents the ratio between the sound pressure (p) and the particle velocity (u) and is given by,

$$Z_m = \rho \cdot c \quad (7)$$

where, as previously, ρ represents the density of the medium and c represents the propagation speed of the acoustic wave that is, by its turn, equal to the product of the acoustic wavelength by its frequency ($c = \lambda \cdot f$).

Sound propagation across two different mediums depends on the sound impedance of each one, namely, on the transmission and reflection coefficients. For the normal component of the acoustic wave, relatively to the separation plane of the mediums, the sound reflection and transmission coefficients are defined by,

$$\begin{aligned} \Gamma_R &= \frac{Z_{m1} - Z_{m2}}{Z_{m1} + Z_{m2}} \\ \Gamma_T &= \frac{2 \cdot Z_{m2}}{Z_{m1} + Z_{m2}} \end{aligned} \quad (8)$$

where Γ_R and Γ_T represent the reflection and transmission coefficients, and, Z_{m1} and Z_{m2} , represent the acoustic impedance of medium 1 and 2, respectively. For spherical waves, the acoustic intensity that represents the power of sound signals is defined by,

$$I = \frac{1}{r^2} \cdot \frac{(p^2)_{av}}{\rho \cdot c} \quad (9)$$

where $(p^2)_{av}$ is the mean square value of the acoustic pressure for $r=1$ m and the others variables have the meaning previously defined. The total acoustic power at a distance r , from the sound source, is obtained by multiplying the previous result by the area of a sphere with radius equal r . The results that is obtained is given by,

$$P = 4\pi \cdot \frac{(p^2)_{av}}{2\rho \cdot c} \quad (10)$$

This constant value of sound intensity was expected since it is assumed a sound propagation in a homogenous lossless propagation medium.

In which concerns the sound pressure level, it is important to underline that this parameter represents, not acoustic energy per time unit, but acoustic strength per unit area. The sound pressure level (SPL) is defined by,

$$SPL = 20 \cdot \log_{10} (p/p_{ref}) \quad (11)$$

where the reference power (p_{ref}) is equal to 1 μ Pa for sound propagation in water or others liquids. Similarly, the logarithmic expression of sound intensity level (SIL) and sound power level (SL) are defined by,

$$\begin{aligned} I &= 10 \cdot \log_{10} (I/I_{ref}) \text{ dB(SIL)} \\ S_{WL} &= 10 \cdot \log_{10} (W / W_{ref}) \end{aligned} \quad (12)$$

where the reference values of intensity and power are given by $I_{ref}=10^{-12}$ W/m² and $W_{ref}=10^{-12}$ W, respectively.

2.4 A few remarks about underwater sound propagation

It should be noted that the speed of sound in water, particularly seawater, is not the same for all frequencies, but varies with aspects of the local marine environment such as density, temperature and salinity. Due mainly to the greater "stiffness" of seawater relative to air, sound travels approximately with a velocity (c) about 1500 m/s in seawater while in air it travels with a velocity about 340 m/s. In a simplified way it is possible to say that underwater sound propagation velocity is mainly affected by water temperature (T), depth (D) and salinity (S). A simple and empirical relationship that can be used to determine the sound velocity in salt water is given by (Hodges, 2010),

$$\begin{aligned} c(T, S, DP) &\cong A_1 + A_2 \cdot T + A_3 \cdot T^2 + A_4 \cdot T^3 + (B_1 - B_2 \cdot T) \cdot (S - C_1) + D_1 \cdot D \\ [A_1, A_2, A_3, A_4] &\cong [1449, 4.6, -0.055, 0.0003] \\ [B_1, B_2, C_1, D_1] &\cong [1.39, 0.012, 35, 0.017] \end{aligned} \quad (13)$$

where temperature is expressed in °C, salinity is expressed in parts per thousand and depth in m.

The sensitivity of sound velocity depends mainly on water temperature. However, the variation of temperature in low depth waters, that sometimes is lower than 2 m in river estuaries, is very small and salinity is the main parameter that affects sound velocity in estuarine salt waters. Moreover, salinity in estuarine zones depends strongly on tides and each sound monitoring measuring node must include at least a conductivity/salinity transducer to compensate underwater sound propagation velocity from its dependence on salinity (Mackenzi, 1981). As a summary it must be underlined that underwater sound transmission is a very complex issue, besides the effects previously referred, the ocean surface and bottom reflects, refracts and scatters the sound in a random fashion causing interference and attenuation that exhibit variations over time. Moreover, there are a large number of non-linear effects, namely temperature and salinity gradients, that causes complex time-variable and non-linear effects.

3. Spectral characterization of acoustic signals

Several MATLAB scripts were developed to identify and to classify acoustic signals. Using a given dolphin sound signal as a reference signal, different time to frequency conversion methods (TFCM) were applied to test the main characteristics of each one.

3.1 Dolphin sounds

In which concerns dolphin sounds (Evans, 1973; Podos et al., 2002), there are different types with different spectral characteristics. Between these different sound types we can refer whistles, clicks, bursts, pops and mews, between others.

Dolphin whistles, also called signature sounds, appear to be an identification sound since they are unique for each dolphin. The frequency range of these sounds is mainly contained in the interval between 200 Hz and 20 kHz (Reynolds et al., 1999). Clicks sounds are though to be used exclusively for echolocation (Evans, 1973). These sounds contains mainly high frequency spectral components and they require data acquisition systems with high analog to digital conversion rates. The frequency range for echolocation clicks includes the interval between 200 Hz and 150 kHz (Reynolds et al., 1999). Usually, low frequency clicks are used for long distance targets and high frequency clicks are used for short distance targets. When dolphins are closer to an object, they increase the frequency used for echolocation to obtain a more detailed information about the object characteristics, like shape, speed, moving direction, and object density, between others. For long distance objects low frequency acoustic signals are used because their attenuation is lower than the attenuation that is obtained with high frequency acoustic signals. By its turn, burst pulse sounds that include, mainly, pops, mews, chirps and barks, seem to be used when dolphins are angry or upset. These signals are frequency modulated and their frequency range includes the interval between 15 kHz and 150 kHz.

3.2 Time to frequency conversion methods

As previously referred, in order to compare the performance of different TFCM, that can be used to identify and classify dolphin sounds a dolphin whistle sound will be considered as reference. In which concerns signals' amplitudes, it makes only sense, for classification

purposes, to used normalized amplitudes. Sound signals' amplitudes depend on many factors, namely on the distance between sound sources and the measurement system, being this distance variable for moving objects, for example dolphins and ships. A data acquisition sample rate equal to 44.1 kS/s was used to digitize sound signals and the acquisition period was equal to 1 s. Figure 1 represents the time variation of the whistle sound signal under analysis.

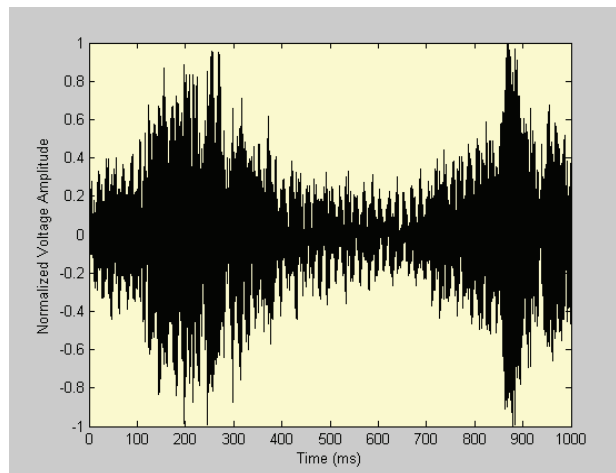


Fig. 1. Time variation of the dolphin whistle sound signal under analysis

Fourier time to frequency conversion method

The first TFCM that will be considered is the Fourier transform method (Körner, 1996). The complex version of this time to frequency operator is defined by,

$$X(f) = \int_{-\infty}^{+\infty} x(t) \cdot e^{-j2\pi \cdot f \cdot t} dt \quad (14)$$

where $x(t)$ and $X(f)$ represent the signal and its Fourier transform, respectively.

The results that are obtained with this FTFCM don't give any information about the frequency contents of the signal over time. However, some information about the signal bandwidth and its spectral energy distribution can be accessed. Figure 2 represents the power spectral density (PSD) of the sound signal represented in figure 1. As it is clearly shown, the PSD of the signal exhibits two peaks, one around 2.8 kHz and the other, with higher amplitude, is a spectral component whose frequency is approximately equal to 50 Hz. This spectral component is caused by the mains power supply and can be strongly attenuated, almost removed, by hardware or digital filtering.

It is important to underline that this FTFCM is not suitable for non-stationary signals, like the ones generated by dolphins.

Short time Fourier transform method

Short time Fourier transform (STFT) is a TFCM that can be used to access the variation of the spectral components of a non-stationary signal over time. This TFCM is defined by,

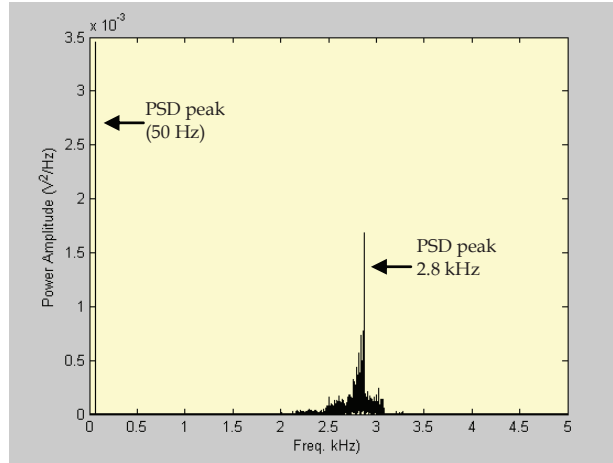


Fig. 2. Power spectral density of the dolphin whistle sound signal

$$X(t,f) = \int_{-\infty}^{+\infty} x(t) \cdot w(t-\tau) \cdot e^{-j2\pi f \cdot \tau} d\tau \quad t \in \mathfrak{R} \quad (15)$$

where $x(t)$ and $X(t,f)$ presents the signal and its STFT, respectively, and $w(t)$ represents the time window function that is used to evaluate the STFT. With this TFCM it is possible to obtain the variation of the frequency contents of the signal over time. Figure 3 represents the spectrogram of the whistle sound signal when the STFT method is used. The spectrogram considers a window length of 1024 samples, an overlap length of 128 samples and a number of points that are used for FFT evaluation, in each time window, equal to 1024.

However, the STFT of a given signal depends significantly on the parameters that are used for its evaluation. Confirming this statement, figure 4 represents the spectrogram of the whistle signal obtained with a different window length, in this case equal to 64 samples, an overlap length equal to 16 samples and a number of points used for FFT evaluation, in each time interval, equal to 64. In this case, it is clearly shown that different time and frequency resolutions are obtained. The STFT parameters previously referred, namely, time window length, number of overlapping points, and the number of points used for FFT evaluation in each time window, together with the time window function, affect the time and frequency resolution that are obtained. Essentially, if a large time window is used, spectral resolution is improved but time resolution gets worse. This is the main drawback of the STFT method, there is a compromise between time and frequency resolutions. It is possible to demonstrate (Allen & Rabiner, 1997; Flandrin, 1984) that the constraint between time and frequency resolutions is given by,

$$\Delta f \geq \frac{1}{4\pi \cdot \Delta t} \quad (16)$$

where Δf and Δt represent the frequency and time resolutions, respectively.

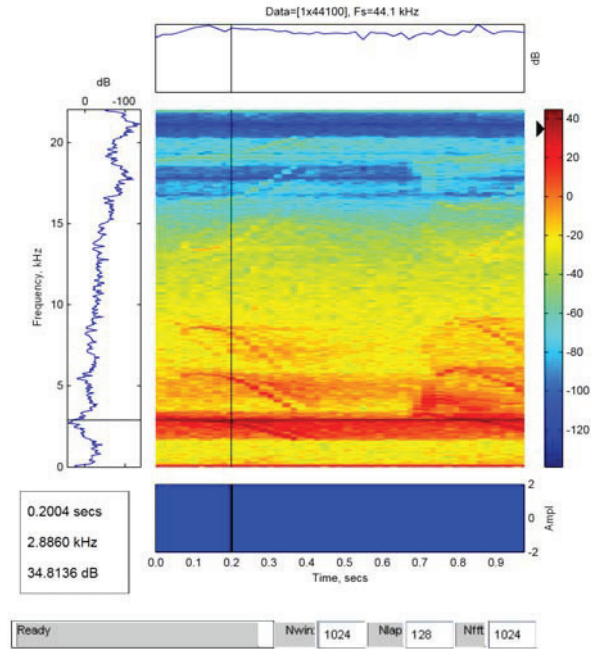


Fig. 3. Spectrogram of the whistle sound signal (window length equal to 1024 samples, overlap length equal to 128 samples and a number of points used for FFT evaluation equal to 1024)

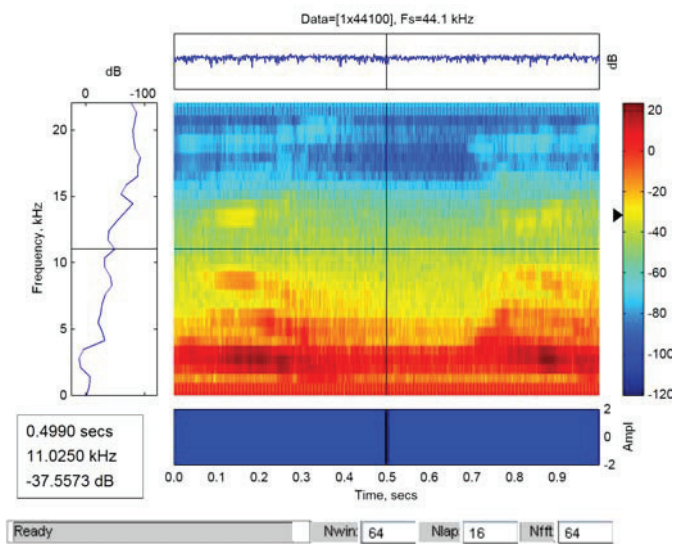


Fig. 4. Spectrogram of the whistle sound signal (window length equal to 64 samples, overlap length equal to 16 samples and a number of points used for FFT evaluation equal to 64)

Time to frequency conversion methods based on time-frequency distributions

When the signal exhibits slow variations in time, and there is no hard requirements of time and frequency resolutions, the STFT, previously described, gives acceptable results. Otherwise, time-frequency distributions can be used to obtain a better spectral power characterization of the signal over time (Claasen & Mecklenbrauker, 1980; Choi & Williams, 1989). A well know case of these methods is the Choi-Williams time to frequency transform that is defined by,

$$X(t,f) = \int_{-\infty}^{+\infty} e^{-j2\pi 2\pi} \int_{-\infty}^{+\infty} \sqrt{\sigma/4\pi \cdot \tau^2} \cdot e^{-\frac{\sigma(\mu-t)^2}{4\tau^2}} \cdot x(\mu + \tau/2) \cdot x^*(\mu - \tau/2) \cdot d\mu \cdot d\tau \quad (17)$$

where $x(\mu+\tau/2)$ represents the signal amplitude for a generic time t equal to $\mu+\tau/2$ and the exponential term is the distribution kernel function that depends on the value of σ coefficient. The Wigner-Ville distribution (WVD) time to frequency transform is a particular case of the Choi-Williams TFCM that is obtained when $\sigma \rightarrow \infty$, and its time to frequency transform operator is defined by,

$$X(t,f) = \int_{-\infty}^{+\infty} e^{-j2\pi 2\pi} \cdot x(\mu + \tau/2) \cdot x^*(\mu - \tau/2) d\tau \quad (18)$$

These TFCM could give better results in which concerns the evaluation of the main spectral components of non-stationary signals. They can minimize the spectral interference between adjacent frequency components as long as the distributions kernel function parameters' are properly selected. These TFCM provide a joint function of time and frequency that describes the energy density of the signal simultaneously in time and frequency. However, Choi-Williams and WVD TCFM based on time-frequency distributions depends on non-linear quadratic terms that introduce cross-terms in the time-frequency plane. It is even possible to obtain non-sense results, namely, negative values of the energy of the signal in some regions of the time-frequency plane. Figure 5 represents the spectrogram of the whistle sound signal calculated using the Choi-Williams distribution. The graphical representation considers a time window of 1 s, a unitary default Kernel coefficient ($\sigma=1$), a time smoothing window (Lg) equal to 17, a smoothing width (Lh) equal to 43 and a representation threshold equal to 5 %.

Wavelets time to scale conversion method

Conversely to others TFCM that are based on Fourier transforms, in this case, the signal is decomposed is multiple components that are obtained by using different scales and time shifts of a base function, usually known as the mother wavelet function. The time to scale wavelet operator is defined by,

$$X(\tau(\alpha)) = \int_{-\infty}^{+\infty} x(t) \cdot \alpha^{0.5} \cdot \psi(\alpha(t-\tau)) dt \quad (19)$$

where ψ is the mother wavelet function, α and τ are the wavelet scaling and time shift coefficients, respectively.

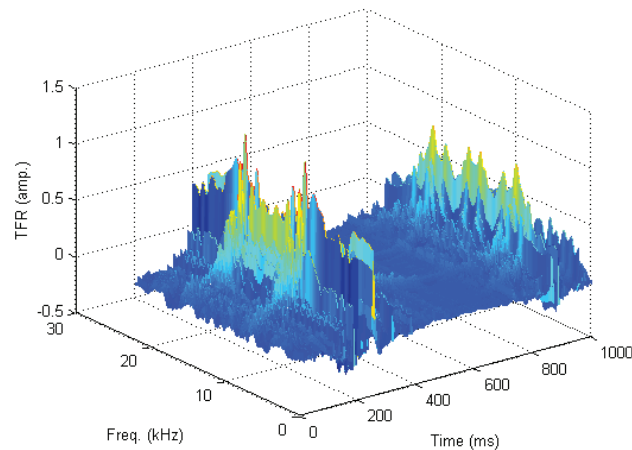


Fig. 5. Spectrogram of the whistle sound signal using the Choi-Williams distribution (time window=1 s, a unitary default Kernel coefficient, time smoothing window=17, a smoothing width=43)

It is important to underline that the frequency contents of the signal is not directly obtained from its wavelet transform (WT). However, as the scale of the mother wavelet gets lower, a lower number of signal's samples are contained in each scaled mother wavelet, and there the WT gives an increased knowledge of the high frequency components of the signal.

In this case, there is no compromise between time and frequency resolutions. Moreover, wavelets are particularly interesting to detect signals' trends, breakdowns and sharp peaks variations, and also to perform signals' compressing and de-noising with minimal distortion.

Figure 6 represents the scalogram of the whistle sound signal when a Morlet mother wavelet with a bandwidth parameter equal to 10 is used (Cristi, 2004; Donoho & Johnstone, 1994). The contour plot uses time and frequency linear scales and a logarithmic scale, with a dynamic range equal to 60 dB, to represent scalogram values. The scalogram was evaluated with 132 scales values, 90 scales between 1 and 45.5 with 0,5 units' increments, an 42 scales between 46 and 128 with 2 units' increments.

The scalogram shows clearly that the main frequency components of the whistle sound signal are centered on the amplitude peaks of the signal, confirming the results previously obtained with the Fourier based TFCM.

3.3 Anthropogenic sound signals

In which concerns underwater sound analysis it is important to analyze anthropogenic sound signals because they can disturb deeply the sounds generated by dolphins' sounds.

Anthropogenic noises are ubiquitous, they exist everywhere there is human activities. The powerful anthropogenic power sources come from sonars, ships and seismic survey pulses. Particularly in estuarine zones, noises from ships, ferries, winches and motorbikes, interfere with marine life in many ways (Holt et al., 2009).

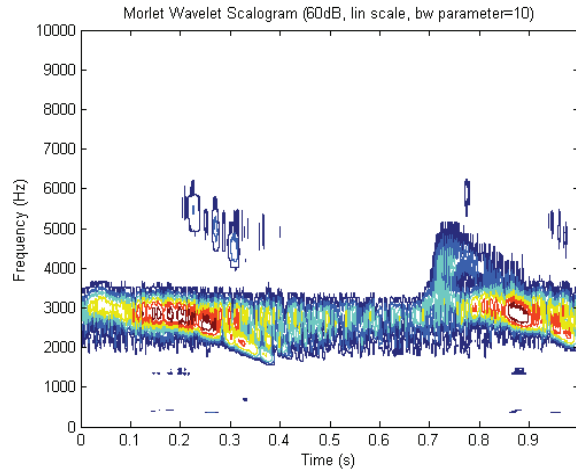


Fig. 6. Scalogram of the whistle sound signal when a Morlet mother wavelet with a bandwidth parameter equal to 10 is used

Since the communication between dolphins is based on underwater sounds, anthropogenic noises can originate an increase of dolphin sounds' amplitudes, durations and repetition rates. These negative effects happen, particularly, whenever anthropogenic noises frequencies overlap the frequency bandwidth of the acoustic signals used by dolphins. It is generally accepted that anthropogenic noises can affect dolphins' survival, reproduction and also divert them from their original habitat (NRC, 2003; Oherveger & Goller, 2001). Assuming equal amplitudes of dolphin and anthropogenic sounds, it is important to know their spectral components. Two examples of the time variations and scalograms of anthropogenic sounds signals will be presented. Figures 7 and 8 represent the time variations and the scalograms of a ship-harbor and submarine sonar sound signals, respectively. As it is clearly shown, both signals contain spectral components that overlap the frequency bandwidth of dolphin sound signals, thus, affecting dolphins' communication and sound signals' analysis.

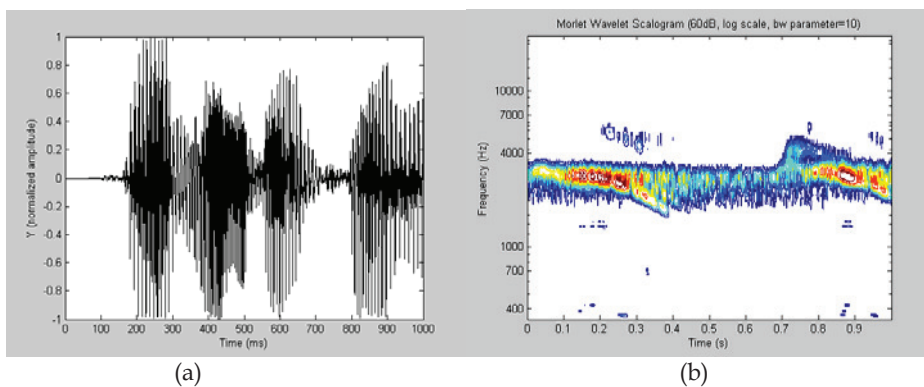


Fig. 7. Ship-harbour signal: (a) time variation and (b) scalogram

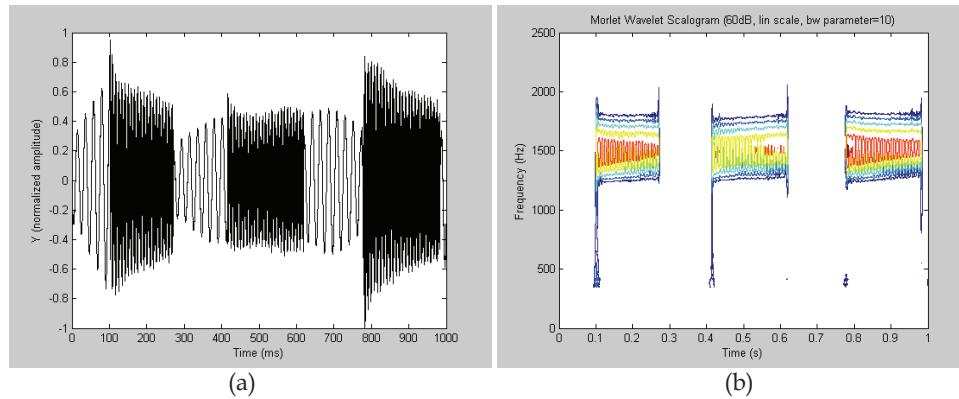


Fig. 8. Submarine sonar signal: (a) time variation and (b) scalogram

4. Measurement system

The measurement system includes several measurement units that can, by its turn, be integrated in a distributed measurement network, with wireless communication capabilities (Postolache et al., 2006). Each measurement unit, whose description will be considered in the present section, includes the acoustic devices that establish the interface between the electrical devices and the underwater medium, a water quality measurement unit that is used for environmental assessment purposes, and the signal conditioning, data acquisition and data processing units.

4.1 Hardware

Figure 9 represents the intelligent distributed virtual measurement system that was implemented for underwater sound monitoring and sound source localization. The system includes two main units: a base unit, where the acoustic signals are detected and digitized, and a remote unit that generates testing underwater acoustic signals used to validate the implemented algorithms for time delay measurement (Carter, 1981; Chan & Ho, 1994), acoustic signal classification and underwater acoustic source localization.

A set of three hydrophones (Sensor Technology, model SS03) are mounted on a 20m structure with 6 buoys that assure a linear distribution of the hydrophones. The number and the linear distribution of the hydrophones permit to implement a hyperbolic algorithm (Mattos & Grant, 2004; Glegg et al., 2001) for underwater acoustic source localization, and also to perform underwater sound monitoring tasks including sound detection and classification. The main characteristics of the hydrophones includes a frequency range between 200 Hz and 20 kHz, a sensitivity of -169 dB relatively to 1 V/ μ Pa and a maximum operating depth of 100 m.

The azimuth angle (φ) obtained from hydrophone array structure, together with the information obtained from the GPS1 (Garmin 75GPSMAP) device, installed on the base unit, and the information obtained from the fluxgate compass (SIMRAD RFC35NS) device, are used to calculate the absolute position of the remote underwater acoustic source. After the estimation of the underwater acoustic source localization, a comparison with the values

given by the GPS2 is carried out to validate the performance of the algorithms that are used for sound source localization.

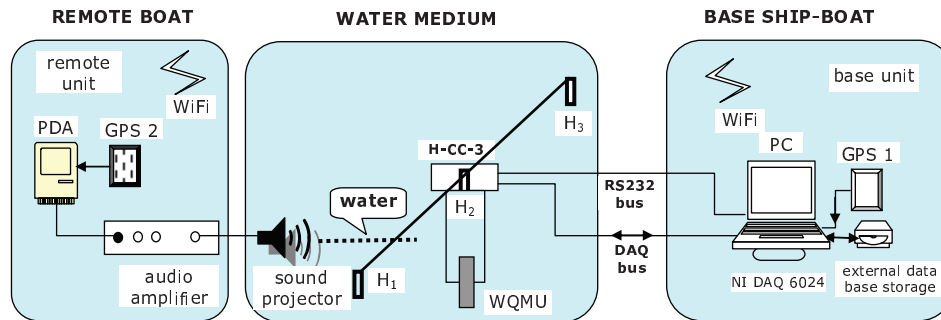


Fig. 9. The architecture of the distributed virtual system for underwater acoustic signal acquisition, underwater sound source localization and sound analysis (H1,H2,H3-hydrophones, H-CC-3- channels' conditioning circuits, WQMU- water quality measurement unit, NI DAQCard-6024 - multifunction DAQCard, GPS1 and GPS2- remote and base GPS units).

As it is presented in figure 9, a three-channel hydrophones' conditioning circuit (H-CC) provides the analog voltage signals associated with the captured sounds. These signals are acquired using three analog channels' inputs (ACH0, ACH1 and ACH2) of the DAQCard using a data acquisition rate equal to 44.1S/s. The azimuth angle information, expressed by $V \cdot \sin\phi$ and $V \cdot \cos\phi$ voltages delivered by the electronic compass, is acquired using ACH3 and ACH4 channels of the DAQCard.

The water quality parameters, temperature and salinity, are acquired using a multiparameter Quanta Hydrolab unit (Eco Environmental Inc.) that is controlled by the laptop PC through a RS232 connection. During system's testing phase, acoustic signals generation is triggered through a Wi-Fi communication link that exists between the PC and the PDA, or by a start-up table that is stored in the PDA and in the PC. Thus, at pre-defined time instants, a specific sound signal is generated by the sound projector (Lubell LL9816) and it is acquired by the hydrophones. The acquisition time delays are then evaluated and localization algorithms, based on the time difference of arrivals (TDOA), are used to locate sound sources. The main characteristics of the sound projector includes a frequency range (± 3 dB) between 200 Hz and 20 kHz, a maximum SPL of 180 dB/ μ Pa/m at a frequency equal to 1 kHz and a maximum cable voltage-to-current ratio equal to 20 $V_{rms}/3$ A.

Temperature and salinity measurements, obtained for the WQMU (Postolache et al, 2002; Postolache et al., 2006; Postolache et al., 2010) are used to compensate sound source localization errors caused by underwater sound velocity variations (13).

4.2 Software

System's software includes two main parts. One is related with dolphin sounds classification and the other is related with the GIS (Postolache et al., 2007). Both software parts are integrated in a common application that simultaneously identify sound sources and locate them in the geographical area under assessment. In this way, it is possible to

locate and pursue the trajectory of moving sound sources, particularly dolphins in a river estuary.

4.2.1 Dolphin sounds classification based on wavelets packets

This software part performs basically the following tasks: hydrophone channel voltage acquisition and processing, fluxgate compass voltage data acquisition and processing, noise filtering, using wavelet threshold denoising (Mallat, 1999; Guo et al., 2000), digital filtering, detection and classification of sound signals. Additional software routines were developed to perform data logging of the acquired signals, to implement the GIS and to perform geographic coordinates' analysis based on historical data. The laptop PC software was developed in LabVIEW (National Instruments) that, by its turn, includes some embedded MATLAB scripts.

The generation of the acoustic signals, at the remote unit, is controlled by the distributed LabVIEW software (laptop PC software and PDA software). The laptop software component triggers the sound generation by sending to the PDA a command using the TCP/IP client-server communication's protocol. The sound type (e.g. dolphin's whistle) and its time duration are defined using a specific command code.

In which concerns the underwater acoustic analysis, the hydrophones' data is processed in order to extract the information about the type of underwater sound source by using a wavelet packet (WP) signal decomposition and a set of neural network processing blocks.

Features' extraction of sound signals is performed using the root mean square (RMS) values of the coefficients that are obtained after WP decomposition (Chang & Wang, 2003). Based on the WP decomposition it is possible to obtain a reduced set of features parameters that characterize the main type of underwater sounds detected in the monitored area.

It is important to underline that conversely to the traditional wavelet decomposition method, where only the approximation portion of the original signal is split into successive approximation and details, the proposed WP decomposition method extends the capabilities of the traditional wavelet decomposition method by decomposing the detail part as well. The complete decomposition tree for a three level WP decomposition is represented in Fig. 10.

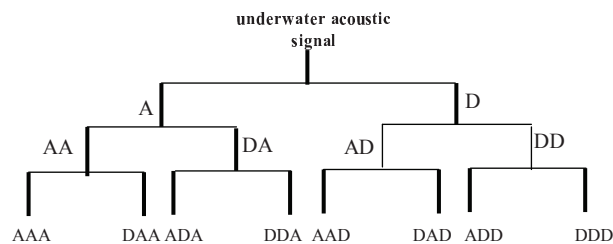


Fig. 10. Decomposition tree for a three level WP decomposition (D-details associated with the high-pass decimation filter, A-approximations associated with the low pass decimation filter)

4.2.2 Geographic Information System: their application to locate sound sources

This software part implements the GIS and provides a flexible solution to locate and pursue moving sound sources. The main components, included in this software part, are the

hyperbolic bearing angle and range algorithms, both related with the estimation of sound sources' localizations.

In order to transform the relative position coordinates, determined by the system of hydrophones (Hurell & Duck, 2000), into absolute position coordinates, it is necessary to transform the GPS data, obtained from Garmin GPSMAP76, into a cartographic representation system. The mapping scale used to represent the geographical data is equal to 1/25000. This scale value was selected taking into account the accuracy of the GPS device that was used for testing purposes. The conversion from relative to absolute coordinates is performed in three steps: Molodensky (Clynch, 2006) three-dimensional transformation, Gauss-Krüger (Grafarend & Ardalan, 1993) projection and, finally, absolute positioning calculation. In the last step, a polar to Cartesian coordinates' conversion is performed considering the water surface as a reference plane (XoY), and defining the direction of X-axis by using the data provided by the electronic compass.

Figure 11 is a pictorial representation of the geometrical parameters that are used to locate underwater acoustic sources.

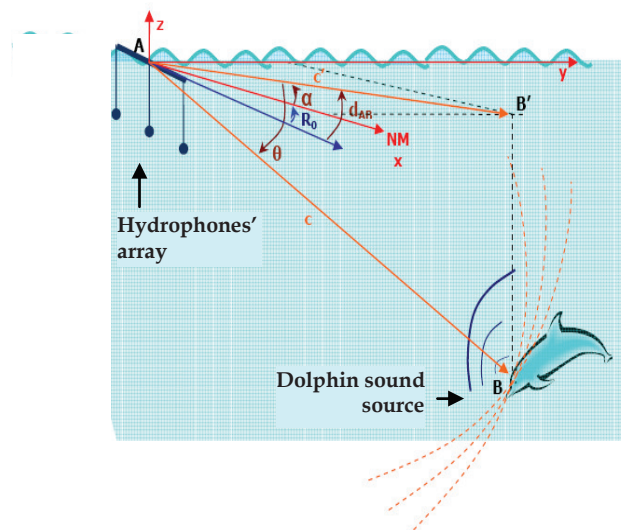


Fig. 11. Geometrical parameters that are used to locate underwater acoustic sources

The main software tasks performed by the measurement system are represented in figure 12.

Finally, it is also important to refer that the underwater sound source localization is calculated and displayed on the user interface together with some water quality parameters, namely, temperature, turbidity and conductivity, that are provided by the WQMU.

Future software developments can also provide improved localization accuracy by profiling the coverage area into regions where multiple measurements results of reference sound sources are stored in an allocation database. The best match between localization measurement data and the data that is stored in the localization database is determined and then interpolation can be used to improve sound source localization accuracy.

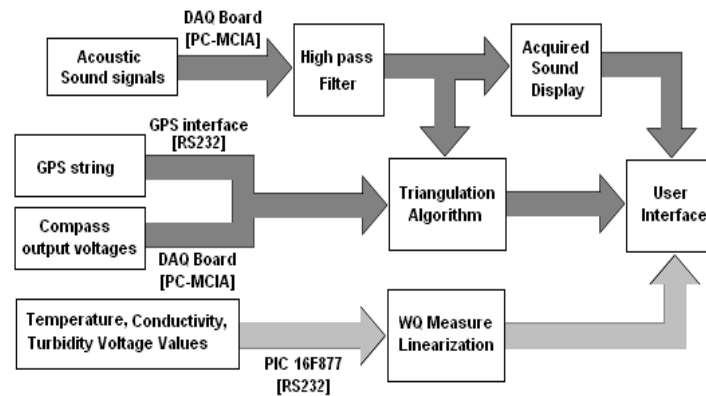


Fig. 12. Measurement system software block diagram

5. Experimental results

To evaluate the performance of the proposed measurements system two experimental results will be presented. The first one is related with the capabilities provided by the sound source localization algorithms, and second one is related with the capabilities of wavelet based techniques to detect and classify dolphin sounds.

5.1 Sound source localization

Several laboratory experiments were done to test the different measuring units including the WQMU. Field tests, similar to the ones previously performed in laboratory, were performed in Sado estuary. During field test of the measurement system, dolphins were sighted but none produced a clear sound signal that could be acquired or traced to the source. In order to fill this gap, a number of experiments took place involving pre-recorded dolphin sounds. For sound reproduction an underwater sound projector was used, which allowed the testing of the sound source localization algorithms. The sound projector, installed in a second boat (remote boat), away from the base ship-boat, was moved way from the hydrophones' structure and several pre-recorded sounds were played. This methodology was used to test the performance of the hydrophones' array structure and the TDOA algorithm to measure the localization of the sound source for different values of distances and azimuth angles. Using the time delay values between the sounds captured by the hydrophones and the data obtained from the electronic compass (zero heading), sounds can be traced to their sources. Table I represents the localization errors that are expected to estimate the localization of the sound source as a function of the angle resolution, that can be defined by the electronic compass, and by the distance between the hydrophones' and the position of the sound source. The data contained in the table considers that the distance from the hydrophones to sound source is always lower than 500 m.

As it can be verified, in order to obtain the desired precision, characteristic of a 1/25000 scale representation, the resolution in the zero heading acquisition angle, for distances lower than 500 m, cannot be greater than $\pm 1/2$ degree. Experimental results that were performed using GPS₁ and GPS₂ units gave an absolute error lower 10 m. This value is in accordance

with the SimRad RFC35NS electronic compass characteristics whose datasheet specifies an accuracy better than 1° and repeatability equal to $\pm 0.5^\circ$.

		Distance (m)					
		10	50	100	150	300	500
Angle ($^\circ$)	0,25	0,04	0,22	0,44	0,65	1,31	2,18
	0,5	0,09	0,44	0,87	1,31	2,62	4,36
	1	0,17	0,87	1,75	2,62	5,24	8,73
	2,5	0,44	2,18	4,36	6,54	13,09	21,81
	5	0,87	4,36	8,72	13,07	26,15	43,58
	10	1,74	8,68	17,36	26,05	52,09	86,82

Table 1. Localization errors as a function of the zero heading angle accuracy and of the distance between the hydrophones and the sound source

5.2 Wavelet based classification of dolphin sounds

The WP based method that was used to identify and classify sound signals enables a large flexibility to choose the best combination of the WP features that are used for detection and classification purposes.

During the design of the features extraction algorithm, different levels of decomposition, varying between 2 and five, and different sounds' periods, varying between 30 ms and 1000 ms, were used. Referring the wavelet packet decomposition, used for underwater sound features' extraction, a practical approach concerning the choice of the best level of decomposition and mother wavelet function was carried out. Thus, the capabilities of Daubechies, Symlets and Coiflets functions as mother wavelets were tested.

For the studied cases, the RMS of the WP coefficients for different bands of interests were evaluated. As an example, figure 13 represents the features' values obtained with a three level decomposition tree and a db1 mother wavelet, when different sound types are

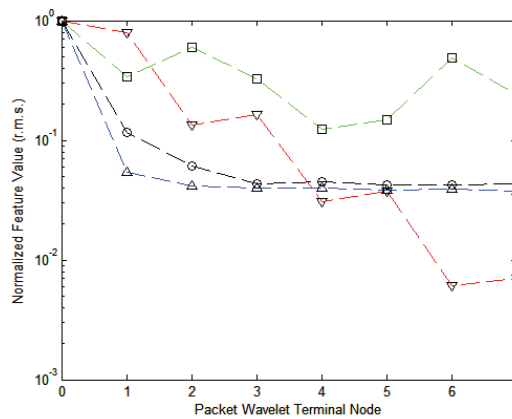


Fig. 13. Wavelet based feature extraction using a three level decomposition tree (green line- dolphin chirp, red line- dolphin whistle, blue line- motorbike, black line- ping sonar).

considered, namely, a dolphin chirp, a dolphin whistle, and two anthropogenic noises, in this case, a water motorbike and a ping sonar sound. As it can be easily verified, the features' values are significantly different for a third level WP decomposition.

Figure 14 represents the WP coefficients for a dolphin whistle when it is used a six level WP decomposition tree. In this case, the number of terminal nodes is higher, 64 instead of 8, but the features' variation profile over packet wavelet terminal nodes have a similar pattern and the sound classification performance is better. As expected, there is always a compromise between the data processing load and the sounds' classification performance.

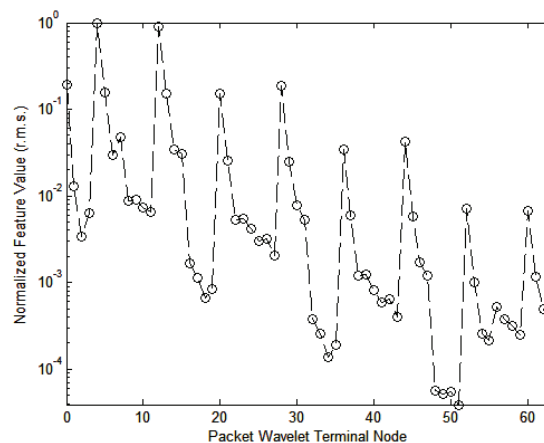


Fig. 14. Wavelet based feature extraction of a dolphin whistle using a six level decomposition tree

Neural Network Classifier

The calculated features for different types of sound signals, real or artificially generated by the sound projector located in the remote unit, are used to train a neural network sound classifier (NN-SC) characterized by a Multilayer Perceptron architecture with 8 neurons in the input layer, a set of 10 to 20 neurons in the hidden layer and one or more neurons (n_{out}) in the output layer (Haykin, 1994). Each input neuron collects one of the 8 features obtained from the WP decomposition. During the training phase of NN-SC the target vector elements are defined according the different sound types used for training purposes. For a similar sound type, for example dolphin whistles, the target vectors' values are within a pre-defined interval. It is important to underline that all values that are used in NN-SC are normalized to its maximum amplitude in order to improve sound identification performance of the neural network.

The number of output neurons (n_{out}) of the NN-SC depends on the number of different signal types to identify. Thus, in the simplified case of the detection of dolphins sounds, independently of their types, the NN-SC uses a single neuron in the output layer with two separated features' range intervals that corresponds to "dolphin sound" and "no dolphin sound" detected, respectively. In order to identify different sound sources and types, it is required more than two features' ranges. This is the case when it is required to classify different sound types, like dolphins bursts, whistles and clicks, or other anthropogenic

sounds, like water motorbikes, ship sound or other underwater noise sounds. Figure 15 represents the NN-SC features' range amplitudes that are used for sound classification.

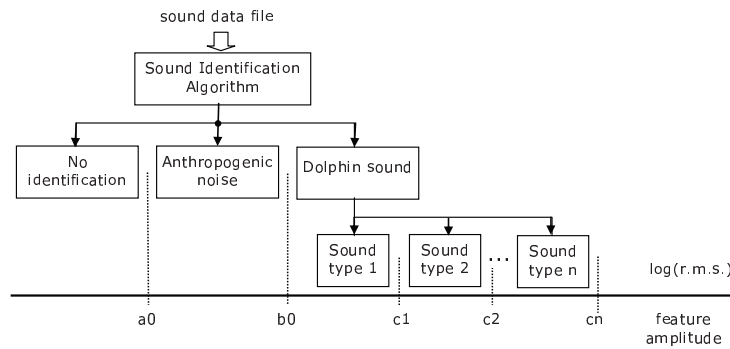


Fig. 15. NN-SC features' amplitudes that are used for sound classification

To test the performance of the sound classification algorithm the following features' amplitudes were considered: between 0.1 and 0.3 for anthropogenic noise sounds; between 0.5 and 0.7 for dolphin whistles and between 0.9 and 1.1 for dolphin chirps. When features' amplitudes are outside the previous intervals there is no sound identification. This happens if the NN-SC gives an erroneous output or if the training set is reduced in which concerns the number of different sound types to be identified.

Figure 16 represents NN-SC normalized output values that are obtained for a third level wavelet packet decomposition, a training and validation sets with 16 elements (sound signals), each one, a root-mean-square training parameter goal equal to 10^{-5} , a number of hidden layer neurons equal to eight and when the Levenberg-Marquardt minimization algorithm is used to evaluate the weights and bias of each ANN neuron.

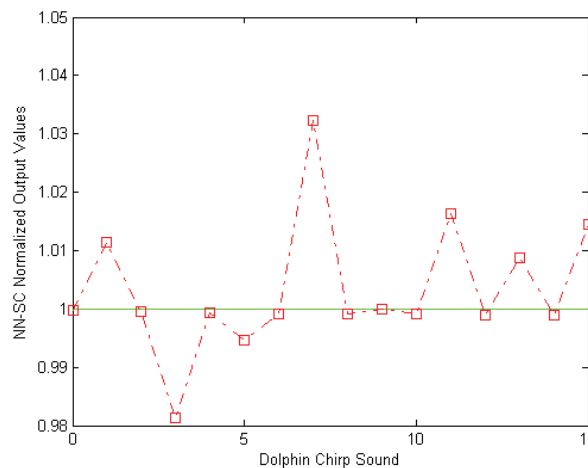


Fig. 16. NN-SC normalized output values that were obtain using a third level wavelet packet decomposition ANN classifiers

The results that were obtained present a maximum relative error almost equal to 3.23 % and the standard deviation of the errors values is almost equal to 1.14 %. Since, in this example, all dolphin sounds' features are in the range between 0.9 and 1.1, we conclude that there is no classification error. Different tests were performed with others sound types and it was verified a very good performance, higher than 95 % of right classifications, as long as there is no NN-SC features' identification ranges overlapping.

6. Conclusions

This chapter includes a review of sound propagation principles, the presentation of different TFCM that can be used to represent frequency contents of non-stationary signals and, finally, the presentation of a measurement system to acquire and process measurement data. In which concerns sound propagation principles, a particular attention was dedicated to the characterization of plane and spherical sound propagation modes and to the definition of power related acoustic parameters. Some details about underwater sound propagation are presented, particularly the ones that affect sound propagation speed. Variations of sound propagation speed in estuarine waters, where salinity can exhibit large variations, must be accounted to minimize measurement errors of sound sources' localizations.

Referring to TFCM a particular attention was dedicated to short time to frequency transforms and wavelet characterization of underwater sounds. Several examples of the application of these methods to characterize dolphin sounds and anthropogenic noises were presented.

Several field tests of the measurement system were performed to evaluate its performance for sound signals' detection and classification, and to test its capability to locate underwater sound sources. To validate triangulation algorithms an underwater sound projector and an array of hydrophones were used to obtain a large set of measurement data. Using the GPS coordinates of the sound projector, located in a remote boat, and the GPS coordinates of the base ship-boat, where the hydrophones' array and data acquisition units are located, validation of relative and absolute localization of sound sources were performed for distances lower than 500 m and for a frequency range between 200 Hz and 20 kHz. The measurement system also includes data logging and GIS capabilities. The first capability is important to evaluate changes over time in dolphin's habitats and the second one to locate and pursue the trajectory of moving sound sources, particularly dolphins in a river estuary.

In which concerns the detection and classification of underwater acoustic sounds, a wavelet packet technique, based on a third level decomposition and on a RMS features' extraction of the terminal nodes' coefficients, followed by an ANN classification method, is proposed. The classification results that were obtained present a maximum relative error almost equal to 3.23 % and a standard deviation, of the error values, almost equal to 1.14 %.

Further tests are required to evaluate the sound detection and classification algorithms when different sound sources interfere mutually, particularly when dolphin sounds are mixed with anthropogenic noises.

7. References

- Allen, J. & Rabiner, L. (1977). "A Unified Approach to Short-Time Fourier Analysis and Synthesis" Proc. IEEE, Vol. 65, No. 11, pp. 1558-64, 1977

- Au, W.; Popper, A.N & Fay, R.F. (2000). "Hearing by Whales and Dolphins", New York: Springer-Verlag
- Burdic, W.S. (1991). "Underwater Acoustic System Analysis", 2nd edition, Prentice Hall, Inc., Peninsula Publishing, California, U.S.A., 1991
- Carter (1981). "Time Delay Estimation for Passive Signal Processing", IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-29, No.3, pp. 463-470, 1981
- Chang, S.H. & Wang, F. (2003). "Underwater Sound Detection based on Hilbert Transform Pair of Wavelet Bases", Proceedings OCEANS'2003, pp.1680-1684, San Diego, USA, 2003
- Choi, H. & Williams, W.J. (1989). "Improved Time-Frequency Representation of Multicomponent Signals Using Exponential Kernels," IEEE Trans. ASSP, Vol. 37, No. 6, pp. 862-871, June 1989
- Claasen, T. & Mecklenbrauker, W. (1980). "The Wigner Distribution - A Tool for Time-Frequency Signal Analysis" 3 parts Philips J. Res., Vol. 35, No. 3, 4/5, 6, pp. 217-250, 276-300, 372-389, 1980
- Clynch, J.R. (2006). "Datums - Map Coordinate Reference Frames Part 2 - Datum Transformations", Feb. 2006 (available at http://www.gmat.unsw.edu.au/snap/gps/clynch_pdfs/Datum_ii.pdf).
- Cristi, R. (2004). "Modern Digital Signal Processing", Thompson Learning Inc., Brooks/Cole, 2004
- Donoho, D.L & Johnstone, I.M. (1994). "Ideal Spatial Adaptation by Wavelet Shrinkage", Biometrika, Vol 81, pp. 425-455, 1994.
- Evans, W.E. (1973). "Echolocation by marine dauphines and one species of fresh-water dolphin", J. Acoust. Soc. Am. 54:191-199, 1973
- Eco Environmental, "Hydrolab Quanta" (available at http://www.ecoenvironmental.com.au/eco/water/hydrolab_quantag.htm)
- Flandrin, P. (1984). "Some Features of Time-Frequency Representations of Multi-Component Signals" IEEE Int. Conf. on Acoust. Speech and Signal Proc., pp. 41.B.4.1-41.B.4.4, San Diego (CA), 1984
- Glegg, S.; Olivieri, M.; Coulson, R. & Smith, S.(2001). "A Passive Sonar System Based on an Autonomous Underwater Vehicle", IEEE Journal of Oceanic Engineering, Vol. 26, No. 4, pp. 700-710, October 2001.
- Grafarend, E. & Ardalan, A. (1993). "World Geodetic Datum 2000", Journal of Geodesy 73, pp. 611-623, 1993
- Guo, D.; Zhu, W.; Gao, Z. & Zhang, J. (2000). "A study of Wavelet Thresholding Denoising", Proceedings of IEEE ICSP2000, pp. 329-332, 2000.
- Haykin (1994), "Neural Networks", Prentice Hall, New Jersey, USA, 1999.
- Chan, Y. & Ho, k. (1994). "A Simple and Efficient Estimator for Hyperbolic Location", IEEE Transactions on Signal Processing, Vol.42, No. 8, pp. 1905-1915, Aug. 1994
- Hodges, R.P. (2010). "Underwater Acoustics: analysis, design and performance of SONAR", John Wiley & Sons, Ltd, 2010
- Oherveger, K. & Goller, F. (2001). "The Matabolic Cost of Bird Song roduction", Journal Exp. Biol., (204), pp. 3379-3388, 2001.
- Holt, M.M.; Noren, D.P; Veirs, V.;Emmons, C.K. & Veirs, S. (2009). "Speaking Up: Killer Whales, Increase their Call Capability in Response to Vessel Noise", Journal of Acoustics Society of America, 125(1), 2009

- Hurell, A. & Duck, F. (2000). "A two-dimensional hydrophone array using piezoelectric PVDF", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 47, Issue 6, pp.1345-1353, Nov. 2000
- Körner, T.W. (1996). "Fourier Analysis", Cambridge, Cambridge University Press, United Kingdom, 1996
- Mallat, S. (1999). "A Wavelet Tour of Signal Processing", Elsevier, 1999
- Mattos, L. & Grant, E. (2004). "Passive Sonar Applications: Target Tracking and Navigation of an Autonomous Robot", *Proceedings of IEEE International Conference on Robotics and Automation*, pp.4265-4270, New Orleans, 2004
- Mackenzi, K.V. (1981). "Discussion of Sea Water Sound-Speed Determination", *Journal of the Acoustic Society of America*, (70), pp. 801-806, 1981
- National Instruments (2005). "LabVIEW Advanced Signal Processing Toolbox", Nat. Instr. Press, 2005.
- NRC (National Research Council) (2003). "Ocean Noise and Marine Animals", National Academies Press, Washington DC, 2003
- Podos, J.; Silva, V.F. & Rossi-Santos, M. (2002). "Vocalizations of Amazon River Dolphins, *Inia geoffrensis*: Insights into the Evolutionary Origins of Delphinid Whistles", *Ethology*, Blackwell Verlag Berlin, 108, pp. 601-612, 2002
- Postolache, O.; Pereira, J.M.D. & Girão, P.S. (2002). "An Intelligent Turbidity and Temperature Sensing Unit for Water Quality Assessment", *IEEE Canadian Conference on Electrical & Computer Engineering, CCECE 2002*, pp. 494-499, Manitoba, Canada, May 2002
- Postolache, O.; Girão, P.S.; Pereira, M.D. & Figueiredo, M. (2006). "Distributed Virtual System for Dolphins' Sound Acquisition and Time-Frequency Analysis", *IMEKO XVIII World Congress*, Rio de Janeiro, Brasil, Sept.
- Postolache, O.; Girão, P.S. & Pereira, J.M.D. (2007). "Intelligent Distributed Virtual System for Underwater Acoustic Source Localization and Sounds Classification", *Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2007)*, pp. 132-135, Dortmund, Germany, September 2007
- Postolache, O.; Girão, P. & Pereira, M. (2010). "Smart Sensors and Intelligent Signal Processing in Water Quality Monitoring Context", *4th International Conference on Sensing Technology (ICST'2010)*, Lecce, Itália, June 2010
- Reynolds, J.E.; Odell D.H. & Rommel, A. (1999). "Biology of Marine Mammals", edited by John E. Reynolds III and Sentiel A. Rommel, Melbourne University Press, Australia, 1999
- Wright, D. (2001). "Undersea with Geographical Information Systems", ESRI Press, USA