Asociación Argentina
de Mecánica Computacional

Mecánica Computacional Vol XXIX, págs. 9389-9405 (artículo completo)
Eduardo Dvorkin, Marcela Goldschmit, Mario Storti (Eds.)
Buenos Aires, Argentina, 15-18 Noviembre 2010

THE USE OF MULTI-WAY ANALISYS IN THE CLASSIFICATION TASK OF PASSIVE SONAR[1] CONTACTS

Rubens L. Oliveira[a,b], Beatriz S. L. P. de Lima[a] and Nelson F. F. Ebecken[a]

[a] *Federal University of Rio de Janeiro, Pos Graduate School of Civil Engineering (COPPE/UFRJ), Rio de Janeiro – RJ – Brazil, http://www.coc.ufrj.br*

[b] *Brazilian Navy Research Institute (IPqM), Brazilian Navy, Rio de Janeiro – RJ – Brazil, http://www.ipqm.mar.mil.br*

Keywords: Multi-Way Analysis, Classification, SONAR Contacts.

Abstract. When a ship sails, it irradiates a particular noise that categorizes it in a specific class. The identification of the ship classes is a non trivial task and it is performed employing the passive SONAR systems. The information processed by the SONAR sensors is usually evaluated by an operator that identifies the contact. Recently, intelligent systems capable to extract important features from the acoustic signal irradiated by ships have been proposed to increase the reliability and speed up the process of decision making. In this work, it is employed a new methodology applied in the process of analyzing and mining databases based in tensor analysis. This methodology showed to be an interesting treatment for this type of problem overcoming other methods.

[1] SONAR = **SO**und **N**avigation **A**nd **R**anging.

1 INTRODUCTION

The SONAR (SOund Navigations and Ranging) is the main equipment used to capture noises in the sea. The passive SONAR only "hear" acoustic signal irradiated from ships. The particular noise irradiated by the sailing ships can categorize them due to the similarities in the acoustic signal in ships of the same class. The identification of the ship class through the analysis of its irradiated noise is a non trivial task and it is performed employing the passive SONAR systems. This noise is received by the SONAR sensors and processed to provide visual and auditory information to an operator especially trained to this task. The SONAR operator (SO) evaluates such information and identifies the contact. The efficiency of this method is directly related to the ability of the SO in isolating and identifying relevant characteristics of the received signal, both in terms of auditory information and in terms of the content of the frequency spectrum.

To increase reliability and speed up the process of decision making intelligent systems capable to extract important features from the acoustic signal irradiated by the ships have been proposed.-Most of the works about this subject perform analysis of the extracted features from decomposition of the acoustic signal, either in time domain or frequency domain, ignoring composition of these two dimensions. On the other hand, those datasets frequently contain values that represent combinations of different properties of the real world. Systems for modeling real noise produced by a ship must remove irrelevant components in order to obtain the true value of the data. Tensor analysis may be an interesting treatment for this type of problem due to its ability to deal with various components independently or in an integrated way. Tensor analysis adds two important aspects in the process of analyzing and mining databases. The first one is their ability to "dissect" the various procedures employed in the capture phase of the data (Data Cleaning). The second is segmentation of instances (or attributes) of a dataset, either directly by the matrix decomposition or from any standard method. The tensor decomposition also allows other types of analysis, for example, to verify the importance of an instance or critical attribute, allowing the representation of data in terms of a small number of substructures.

This work uses Multi-way analysis as a tool for modeling acoustic signature pattern, considering the variation of signal strength, simultaneously as a function of frequency and time, providing compact and robust treatment to background noise and consequently eliminating the need of a specialist. The parallel decomposition method CANDECOMP/PARAFAC is used to ensure a representative minimum structure of data instances, eliminating irrelevant information to class-ship mapping process. Model validation was performed with real dataset and good results were obtained.

2 CONTEXT

The complex decomposition of different noise sources originated by other ships and the self noise, in addition to the noise from local fauna make the operator task quite tiring. The use of automated methods to support the detection and classification task reduces pressure on the SO.

Datasets from many diverse areas as scientific, medical, engineering or social areas frequently contains values that represent combinations of different properties of the real world. For example, observing the noise captured from a ship by a SONAR system, the noise sources produce some intensity values of spectral power for a given frequency, nevertheless those observed values represent the sum of at least three different components: the true intensity of power in the observed frequency; the properties of the environment in which the ship has emitted the noise and the properties of the SONAR system itself. Techniques for modeling the real noise properties produced by a ship should remove most of the irrelevant components for obtaining the real value of the data.

In general, data mining techniques ignores the fact that datasets of real world represents combination of independent data and build specific models from them. If such datasets can be separated in independent components certainly the quality of data mining will be improved. A way to accomplish this separation is through the use of matrices decomposition methods, capable of “dissecting” the data, as it can consider the relationship between large of data collections and the relationship between its components to promote their separation.

There are two properties of complex datasets which makes data mining a nontrivial task (SKILLICORN, 2007):

- In general, each data instance doesn’t represent a single discrete property or direct action of the associated object, instead represents a fusion of values originated from different process that combined produce a single value captured in the database.
- The relations between attributes and between each attribute and target attribute are subtle, and more, in some databases some attributes are more significant for some data instances than other ones.

More efficient techniques are necessary to knowledge discovery from complex databases. Methods based on matrices decomposition provide an efficient way to data cleaning, allowing conventional techniques to be employed.

2.1 Objectives

In general, the databases are treated as two-dimensional arrays, where rows represent instances while columns represent different attributes. In this kind of structure each instance is described by an attribute vector. The application of matrix decomposition methods has been observed frequently in this type of structure, however, little is known about works using such techniques when databases are treated as multi-way arrays (tensors), where the instances are described by arrangements with two or more dimensions.

In the database employed in this work, the acoustic signal corresponding to noise radiated by ships depends on both time and frequency; therefore the use of methods of multi-way analysis is suggested in the database analysis.

Methods of multi-way analysis add two interesting aspects to the process of databases analysis:

- Are capable to "dissect" the different processes employed in the capture phase of data. So, the effects of processes that are irrelevant to the task of interest can be separated from data.
- Provides segmentation of instances (or attributes) of a dataset.

Tensor decomposition also provides other kind of analysis, for example, verify what is the importance of a critical instance/attribute, allowing the representation of data with a minimal substructure.

The objective of this work is to propose an implementation of a target classifier based on radiated noise. At different stages of development various strategies are employed:

- Spectral signal analysis corresponding to noise radiated from ships, taking into account the variation in signal strength, depending simultaneously from frequency and time.
- Tensor decomposition algorithms (Parallel Factor Analysis-PARAFAC) (HASHMAN, et al., 1994), for extracting meaningful information and filtering of interference, enabling the data cleaning process.
- Analysis of data model representative of signal characteristic of the ship.
- Classification using Multi-layer Perceptron (MLP).

3 CONVENTIONALS METHODS TO SONAR TARGET CLASSIFICATION

The operators of the first SONARs used only amplifiers and analog filters to select and emphasize the concerned frequency band to be extracted from the noise. From the acoustic signal, the SO identified the presence of noise that could be associated with the machines of the target. So, the classification task was accomplished based on previous knowledge of operator and annotation about ships that were expected to be encountered in specific missions.

The introduction of DEMON (DEMOdulated Noise) and LOFAR (LOw Frequency Analysis and Recording) analysis improved the comprehension about the noise extracted from the SONAR sensor. DEMON analysis provides an overview of cavitations' noise to obtain the rotation of the propeller shaft and the number of blades. LOFAR corresponds to a spectral analysis of the radiated noise, allowing a simultaneous view of several frequency ranges.

The LOFARGRAM represents the LOFAR analysis in time domain allowing the visualization of the spectrum variations over the time. The lines corresponding to the tones present in the noise can be viewed and subsequently be associated with the ship machines.

4 PASSIVE SONAR SYSTEM AND NOISE IRRADIATED FROM SHIPS

Passive SONAR uses acoustic energy propagated by ships to extract its characteristic. The captured noise is a combination of the SONAR system noise itself with environment noise, including the noise irradiated by other platforms (URICK, 1983).

A moving ship constitutes a great source of acoustic energy, for example, the sound of machine noise (main and auxiliary) and the propeller noise. Table 1 lists some source of noises for a diesel-electric propulsion system.

Machines in ships may be directly related to propulsion (main machines) or to maintenance of navigation conditions (auxiliary machines). Machines associated with propulsion modify their frequency as the ship speed increases, but auxiliary machines do not change. The frequency on which auxiliary machines generate sound and its stability can reveal useful information to identify ships (DAMAS, et al., 2006).

The noise generated by ship machinery can be considered as stationary in broad sense but in aquatic environment its composition changes, which makes the sound signal received by distant SONAR be considered no longer stationary, either by lost during noise propagation or by aggregation of the environment noise.

Noise type	**Noise source**
Machine noise	Main machine: diesel engines, gears and turbines. Auxiliary machine: generators, pumps and air conditioning equipment.
Propeller noise	Cavitations in propeller, excitation of the hull induced by propeller.
Hydrodynamic noise	Radiated flow noise, cavitations in structures.

Table 1 - Sources of irradiated noises by diesel-electric propulsion ships.

When the irradiated noise by a ship reaches the SONAR sensor it presents quite distinct from that presented in its origin. These transformations vary over time depending on the medium, thus the received signal can no longer be considered stationary. However, these signals are treated as locally stationary (PFLUG, et al., 1997).

Besides the effects of scattering and attenuation due to environment, the noise received by SONAR sensor is also impregnated with the so-called background noise caused by external sources.

5 METODOLOGY

The methodology developed in this work involves the following development stages:

I. Characterization of an acoustic signature model using the concept of tensor.

II. The use of the CP[2] decomposition method on the model of acoustic signature characterized to eliminate noise.

[2] **C**anonical/**P**arallel factorization decomposition.

III. Compression of acoustic signature model "filtered" through the process of flattening and use of tensor operators, selecting features most relevant to the classification process.
IV. Classification using the acoustic signature model produced using an Artificial Neural Networks.

5.1 Characterization of an acoustic signature model as a third order tensor

From the available data was implemented an acoustic signature model that was capable of "improving" the process of association of each data instance with the corresponding ship.

In (DAMAS, et al., 2006) (MOURA, et al., 2007) (SANTOS, 2005) (SEIXAS, et al., 1999) (SOARES, 2001) is assumed that the noise captured from each ship has a stationary behavior. Thus the representation of each data instance is through a vector of values of spectral powers where the amplitude is only a function of frequency.

Despite of results presented in the above works it is observed that in treatment of databases is not taken into consideration possible interrelationship between variables: time, frequency and power spectrum.

In order to give a more integrated and compact representation of data instances we chose a structure able to encapsulate characteristics of frequency variation, time and spectral power. Each instance of the database is treated as a two-dimensional array where the spectral power is function of frequency and time.

Figure 1 shows a graphical representation of a data instance (corresponding to the matrix with the spectral power of each pair frequency vs. time).

Note that this graphical representation corresponds to a LOFARGRAM extracted from noise irradiated by ship. With this philosophy the set of representable instances of the database is done through a three dimensional tensor, where each plane represents a horizontal array of values of spectral power simultaneously as function of time and frequency.

Figure 2 shows the different layers of the third order tensor. Note that each horizontal layer A ((i ,:,:)) corresponds to an instance of the database, while the lateral and frontal layers A(:, i,:)) and A ((:,:, i)) refers to attributes set.

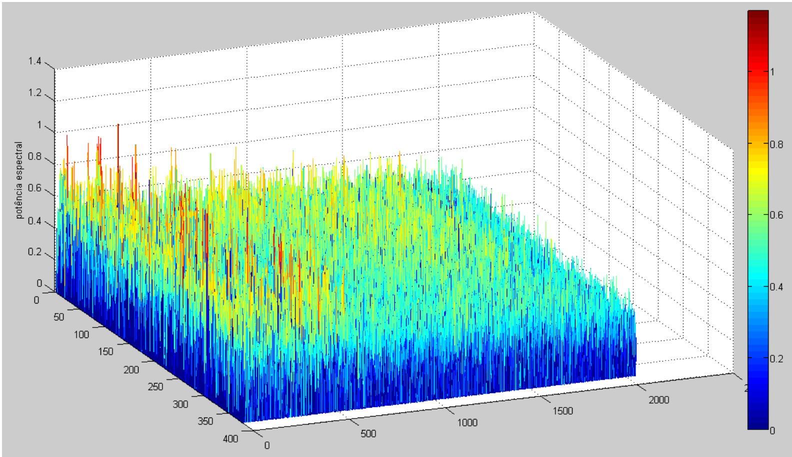


Figure 1 - Graphical representation of a data instance.

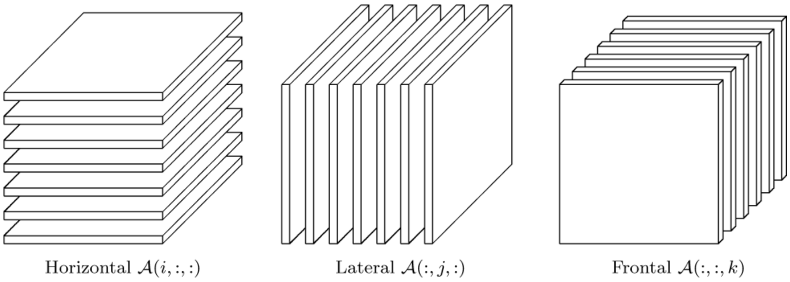


Figure 2 - Layer of the third order tensor.

Thus, we defined a model of acoustic signature as a third order tensor with dimensions: 253 x 929 x 90, where:

1. First dimension refers to the total number of instances of the database, corresponding to each instance of a ship.
2. Second dimension corresponds to the frequency bands, varying between the limits from zero to 2500Hz, and
3. Third dimension is associated with the record time of signals varying from zero to 33s.

5.2 Database Filtering from the PARAFAC model

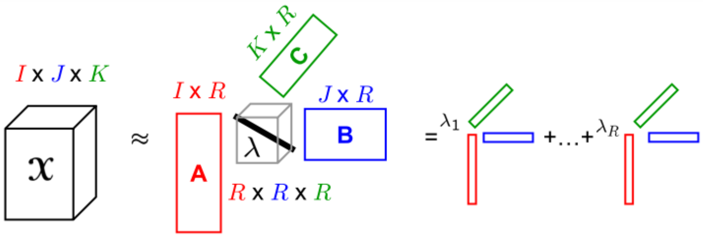
Second stage utilizes the PARAFAC[3] (HARSHMAN, et al., 1994) model for "filtering" acoustic signature model characterized in the previous stage. In the process of determining the number of components (rank-1 tensors) to obtain the best factors of decomposition was crucial. In this task we used the method CORCONDIA (CORE Consistency Diagnostics), proposed by (BRO, et al., 2003).The core consistency diagnostic was efficient for his performance presented in that study.

CORCONDIA method operates as follows: For a given number of components, adjust the PARAFAC model to data. Use the solution found for calculation of core consistency. If the PARAFAC model is validated, then the core consistency will be

[3] **PARA**llel **FAC**torization.

close to 100%. If the data cannot be approximately described by a model of third order or many components are needed the consistency of the core will be close to zero (or negative). If the consistency is close to 50%, the model is considered unstable. In practice the core consistency index grows slowly until reaching the optimum number of components and then decreases sharply. The number of components corresponding to value of greater consistency should be chosen.

As can be seen in Figure 3 the number of components in PARAFAC model is directly related to the number of the matrix factors columns (A, B and C).



$$\mathcal{X} \approx [\![\lambda\, ; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] = \sum_r \lambda_r\, \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$$

Figure 3 - PARAFAC model represents the sum of rank-1 tensor.

In practice determining the tensor rank involves determining the dimension R [core (λ) of model PARAFAC]. The smaller size of λ smaller the number of the matrix factors columns (A, B and C), providing greater reduction in size of the tensor *X*', obtained by reorganization *X*' ≈ *λ*', *A*', *B* ', *C*', where: *λ*', *A*', *B* 'and *C*' represent respectively elements: *λ*, *A*, *B* and *C*, decomposition of the original tensor *vs.* reduced tensor by elimination of elements "less significant" of the super-diagonal (original tensor rank).

The coefficients of λ are in descending key order and reflects the importance (explain) of each rank-1 tensor in the PARAFAC decomposition.

The application of PARAFAC decomposition provides identification of components more relevant to the ships characterization, allowing the elimination of those less relevant and thus eliminating spurious information (probably background noise), and allow the reduction of dimensionality of the original data.

5.3 Rearranging "filtered" signature model through the flattening process

In this stage factors decomposed in the previous stage are rearranged by the application of the operators to tensors and flattening process resulting in a compact structure. At this stage we search for a two-dimensional structure (already free from background noise) allowing its use in the classification task.

Firstly, is recomposed the three-dimensional tensor (reduced) multiplying core by first, second and third factors (obtained in the previous stage with the PARAFAC model). After that the tensor obtained is extended in first mode, getting a two-dimensional array with first dimension corresponding to data instances while second dimension corresponding to attributes (spectral power values related to both frequency and time).

5.4 Classification and performance analysis with "filtered" Signature Acoustic Model

At this stage to classification task is proceeding using an Artificial Neural Networks (ANN). Here, the main objective is to validate the model obtained.

The analysis of performance was made against works that have the purpose of ships classification from its irradiated noise and have employed the same type of database.

6 DATASET

The dataset used in this work was extracted from noise irradiated by ships. Each record represents the passage of a ship on a non-directional hydrophone positioned near the sea bottom. In each record the ship maintained constant course and speed. Each class was identified by letters (A, B, C e D) while each ship was identified by the letter of the class followed by a sequential number.

For each record, it was obtained the Spectrogram and DEMON analysis. Information related with tones (spectral amplitude as function of frequency and time) were extracted from the Spectrograms, while the rotational speed of main axis was obtained from the DEMON analysis.

The data used in the classification task are composed by information extracted from the tone and the rotation speed of main axis of the ship. Table 2 presents the number of records for each class.

DATASET RECORD							
						TOTAL	%
CLASS A							
SHIP	A1	A2	A3	A4	A5		
# record	16	12	24	37	--	89	35,2
CLASS C							
SHIP	B1	B2	B3	B4	b5		
# record	15	39	06	08	14	82	32,4
CLASS F							
SHIP	C1	C2	C3	C4	C5		
# record	16	14	07	21	--	58	22,9
CLASS J							
SHIP	D1	D2	D3	D4	D5		
# record	12	--	--	12	--	24	9,5
TOTAL OF RECORDS						253	100

Table 2 - Number of records per class.

A preprocessing step was performed in the signal in order to emphasize useful details and remove unnecessary information to classification step.

Spectral analysis of acoustic signal captured by SONAR is a good representation of the signal for subsequent classification task (OPPENHEIM, et al. 1975). In this work the spectral analysis is used to obtain the power spectra. Much of the information used to discriminate the classes of ships is derived from noise tones generated by auxiliary engine on board.

6.1 Spectrograms definition

Methods for selection of optimal resolution in the signal frequency range are not the focus of this work. In (SOARES, 2001), it was presented a detailed study suggesting several experimental procedures to select more appropriate methods and parameters. We emphasize that the expert's knowledge is crucial for an appropriate parameterization.

Using the same procedure adopted in (DAMAS et al., 2006), the analysis was done in Spectrograms extracted in the frequency range between 0 and 5500 Hz. Figure 4 (DAMAS, et al., 2006) presents a diagram of the performed signal processing to extract the corresponding Spectrogram.

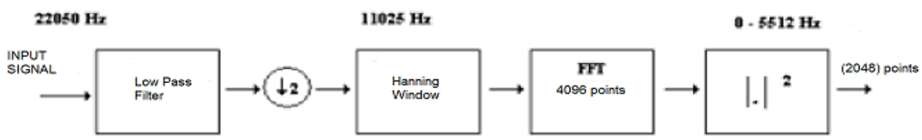


Figure 4 - Diagram of processing for obtaining signal spectrum.

From the sound signal (LOFARGRAMs) it was observed that the most significant tones (high amplitude) were within the frequency range from 0 to 2500 Hz, therefore this range was chosen to be focused with the purpose of showing in detail the noise behavior.

Figure 5 presents a LOFARGRAM restrict to this frequency range.

6.2 Correction of spectrum using an estimate of background noise

The spectral of irradiated noise consists of tones associated with machines in operation within the ship. The SONAR capture these tones superimposed on a continuous spectrum noise caused by local background noise. In most cases, the estimated background noises are used to correct the spectrum itself, emphasizing information about amplitude picks in the spectral (SOARES, 2001).

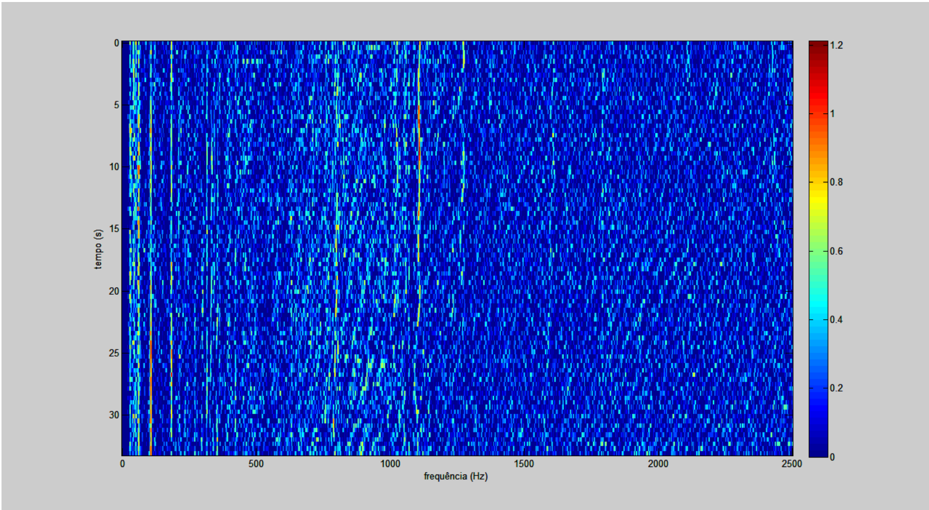


Figure 5 - LOFARGRAM corresponding to a frequency range from 0 t0 2500 Hz.

In general the background noise is estimated using the algorithm Two-Pass Split Window (TPSW) (NIELSEN, 1991). This algorithm estimates an average value of amplitudes. Figure 2 illustrates the processing (SANTOS, 2005). Initially, it is estimated a local average which each point represents an average of its closest neighbors. This local average is multiplied by a factor defining a threshold of detection (Figure 6a). The points of spectrum that exceed this threshold are replaced by the local mean at that point (Figure 6b). A second convolution of this new spectrum with a new window produces a final estimate of the local average (Figure 6c). This estimated local mean corresponds to the spectrum background noise estimate. Figure 6d shows the spectrum after the subtraction of the background noise estimate.

One way to correct the spectrum by background noise is using Equation 1.

$$y^k(n) = log\left(x^k(n)\right) - TPSW\left(log\left(x^k(n)\right)\right)$$

Equation 1

Where $x^k(n)$ represents the n^{th} spectrum of class k and $y^k(n)$ the corrected spectrum. Equation 1 permits both correction and normalization of spectrum.

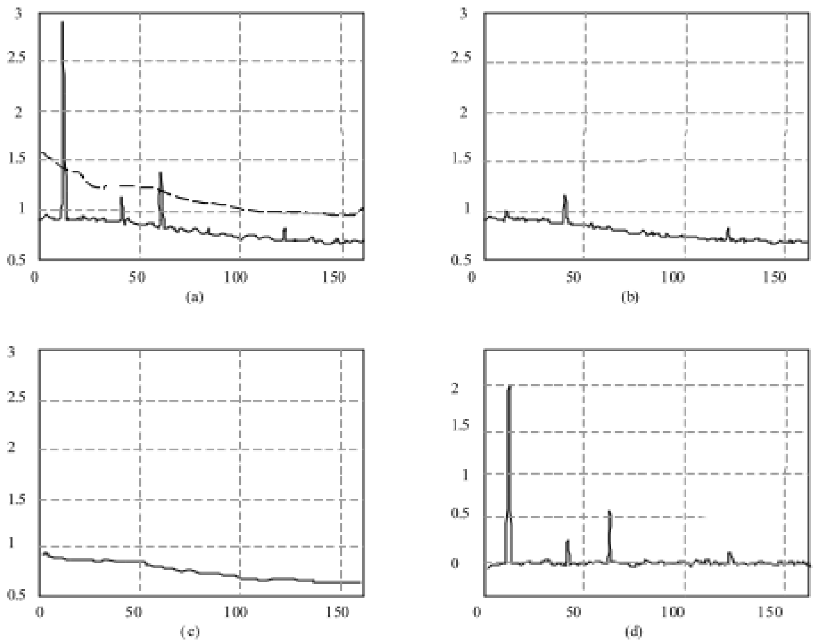


Figure 6 - Sequence of TPSW algorithm to estimate background noise (SOARES, 2001).

7 EXPERIMENTS, RESULTS AND COMPARATIVE ANALYSIS

Here we describe a series of experiments using the acoustic signature model proposed with methodology described in the previous chapter.

7.1 Data Analysis

In this section data analysis was processed identifying the entity (function) to be learned, data instances (rows) and attributes (variables) of each instance in addition to compression and data standardization.

The function to be learned must map the acoustic noise spectrum irradiated by ship with its corresponding class. Thus each data instance representing the noise spectrum and the attributes correspond to the spectral amplitudes values associated with each point of LOFARGRAM (intersection point frequency vs. time) included in a considered range. Each attribute is represented by a real value associated with the appropriate spectral amplitude.

From the determinations of the frequency bands, time, and their respective resolutions we constructed a representation of LOFARGRAM as two-dimensional array 929 x 90 dots, where each point represents spectral amplitude in frequency domain vs. time.

Thinking in reducing the processing time of learning program we applied a process of data compression. This procedure allowed reduced representation of the original data (volume reduction). Thus it was decided to consider the average of spectral amplitudes taken at three points of the frequencies axis and not change the time axis, which in practice can be seen as a reduction in the resolution of the frequency, obtaining a new two-dimensional array 309 x 90 as a representation of each LOFARGRAM.

As each LOFARGRAM is treated as a two-dimensional array 309 x 90, and we have 253 data instances (LOFARGRAMs) we generated a third-order tensor with dimensions 253 x 309 x 90 points.

7.2 Data base filtering

In the process of database filtering we use the PARAFAC decomposition model. The core consistency diagnostic (Figure 7) indicated that the most appropriate number of factors 115. With this decomposition, the tensor rank-1 (result of decomposition of the original tensor) should form a basis capable of representing the original tensor purged of interference caused by factors that don't compose the original signal irradiated by the ship (background noise).

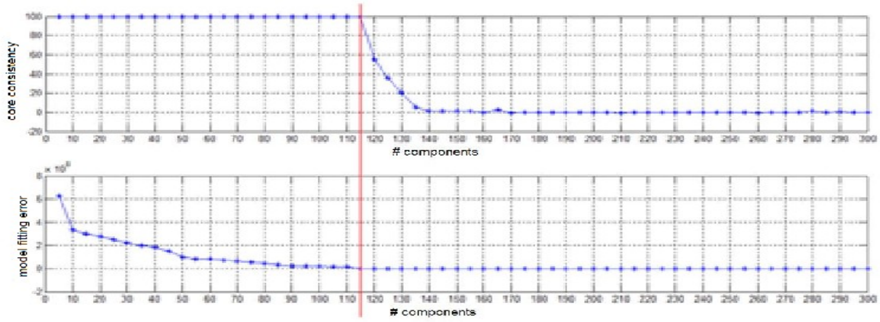


Figure 7 - Consistency index of core tensor and model fitting error to the number of components considered in decomposition.

7.3 Data array decomposition

The original acoustic signature model generated is represented by a third order tensor. By applying PARAFAC decomposition with 115 factors we obtained a new tensor X ' ≈ λ '; A', B', C ' where:
λ' = supercube (115 x 115 x 115);
A' = array (253 x 115);
B' = array (309 x 115) and
C' = array (1990 x 115).

Representing, respectively, the elements: λ, A, B and C, decomposition of the original tensor X.

The process of re-composition of the data matrix involved the following steps:

1. The third order tensor (reduced) was recomposed by multiplication of the core (λ') the first (A'), second (B ') and third (C') factors (obtained with the model PARAFAC), resulting tensor is a 253 x 309 x 90;
2. The tensor was extended (flattening operation) along the first mode, so we obtained a two-dimensional array (253 x 27810), with the first dimension corresponding to data instances and the second to attributes (values of powers associated both frequency and time);
3. All columns beyond the 115 was purged from the matrix (only first 115

columns was weighted by the matrix λ'), so a new matrix 253 x 115 was obtained;

4. A column corresponding to rotation speed of the ship axis was added, resulting in a 253 x 116 matrix.
5. The columns were normalized using the z-score method.

7.4 Experiments

In this work, a Multilayer Perceptron (MLP) - Artificial Neural Network (ANN) is used as a tool for training and classification of ships from the produced acoustic signature model. ANN with MLP was chosen as the classifier due to the cost-benefit ratio between its processing time and learning ability.

Five experiments were performed for classification with the produced acoustic signature model as following:

1. Classification without applying any filter: TPSW or the application of PARAFAC decomposition[4];
2. Classification with application of the TPSW filter and without applying PARAFAC decomposition[5];
3. Classification without the application of the TPSW filter and by applying PARAFAC decomposition with 115 decomposition factors[6];
4. Classification with application of the two filters: TPSW and PARAFAC decomposition with 115 decomposition factors;
5. Classification with application of the two filters: TPSW and PARAFAC decomposition with 125 decomposition factors[7];

7.5 Results

Table 3 presents a comparative table with the classification performances obtained in each of the experiments listed above.

# Experiment	**Instance correctly classified**	
	absolute	**Percentage (%)**

[4] In practice this experiment is equivalent to not applying the filter TPSW and application of PARAFAC decomposition with an arbitrary large number of factors (1000).

[5] In practice this experiment is equivalent to applying the filter TPSW and application of PARAFAC decomposition with an arbitrary large number of factors (1000).

[6] To the number of decomposition factors was obtained a structure for training and classification matrix 253 x 116.

[7] In this experiment due to have applied filtering algorithm TPSW to original data model was decided by a new evaluation of optimum number of components for application of PARAFAC decomposition by applying CORCONDIA algorithm we adopted the optimal factor number equal to 125.

1	174	70.75
2	216	85.38
3	243	96.05
4	226	89.33
5	233	93.10

Table 3 - Comparison Table Performance Classification of Experiments Performed.

The best performance classification was obtained in experiment three, when we used the produced acoustic signature model with the background noise filtering performed by PARAFAC with decomposition of 115 factors.

The second best result was obtained in experiment five, with the filtering of background noise using both TPSW algorithms and PARAFAC decomposition with 125 factors. Note that the performance obtained is lower than that obtained in experiment three and can be associated with a sub-optimal parameterization of the TPSW algorithm by the specialist. Another interesting aspect is the increase in the optimal number of factors for the PARAFAC decomposition (from 115 in the third experiment, to 125 in last one). This fact probably is related to the increased complexity of the original data model related to the parameterization in filtering the background noise introduced by the expert.

The best performance classification reached in the third experiment is probably related to the capacity to rewrite the original database using a minimal representation model, by indication of the minimum number of factors for the PARAFAC decomposition through the algorithm CORCONDIA. This approach allows a reduction of both the problem of under-parameterization, ensuring that the subscription model noise produced is capable of representing all the instances associated and the problem of over-parameterization, preventing the model of storing useless information, as the background noise.

The produced acoustic signature model eliminates the need of other algorithms for attribute selection, both to reduce the dataset size and for elimination of uncorrelated attributes with the target classes.

7.6 Comparison with methods using the same data set

MOURA, et al. (2007) implemented a neural classifier for passive SONAR using a database similar to that employed in this work. The average performance rating obtained was 89%. Besides the classification task, the system from the information specialist enables the platform detection noise itself when it is under certain conditions. This procedure provides better filtering of the echo SONAR signal, allowing better targets classification.

In (DAMAS, et al., 2006) with the same database used in this work was presented a study on the extracted features of tone in the ship classification task. In this work, several features extracted of the tones generated from ships were considered. Its emphasized the importance of specialist in the indication of frequency bands associated with the tones of interest and in the identification process of what characteristics to be considered in each tone selected (average frequency of the tones, trend, and etc). With this approach was obtained accuracy around 85%.

In both systems mentioned above is clear the need for a specialist. The acoustic signature model proposed in this work requiring no expert knowledge, which reduces the time to data preparation beyond to allow the compression one depending on the complexity of the data itself, allowing time reducing to training.

In terms of performance classification, with the acoustic signature model proposed was obtained better result than that presented in (DAMAS, et al., 2006) using the same database (96% vs. 85%).

8 CONCLUSIONS

The classification performance obtained utilizing acoustic signature model proposed in this work (96.05%) showed considerably higher in performance compared with both when no filter has been applied (70.75%) and when it only was applied the filter TPSW (85.38%).

Another important aspect is the fact that the approach proposed here dispenses *a priori* knowledge of the expert on the environment and tactical scenario. The conditions of propagation of the acoustic signal radiated by a ship vary depending on the environment and this varies according to several aspects which makes it almost impossible for the OS the complete knowledge about this domain.

The need of expert on a heterogeneous scenario would increase time of data pre-processing, since their considerations would be essential to the filters parameterization process. With the acoustic signature model proposed in this work, not only this pre-processing stage can be suppressed but also is provided a more compact data representation allowing time reduction on data preprocessing and training.

The good performance classification presented with adoption of acoustic signature model proposed can be attributed to the robustness of its own data model, guaranteed by a minimal representative data structure expunging information not relevant to process class-ship mapping.

9 REFERENCES

Bro, R., Henk, A. e Kliers, L. 2003. A new efficient method for determining the number of components in PARAFAC model. Journal of Chemometrics, vol. 17, pp. 274-286. Published on line in Wiley InterScience (www.interscience.wiley.com), 2003.

Damas, M. D. 2006. Study of tons and their characteristics for use in a ship classifier based on neural networks. Federal University of Rio de Janeiro - COPPE-PEE. Master dissertation, 2006.

Freitas, A. A. 2001. Understanding the Crucial Role of Attribute Interaction in Data Mining. Artificial Intelligence Review. Vol. 16, pp. 177-199, 2001.

Harshman, A. R., Lundy, E. M. Parallel Factor Analysis. Computational Statistics &

Data Analysis, Vol. 18, pp. 39-72, 1994.

Moura, N. N., Grael, F. F. e Seixas, J. M. Detection and classification of passive SONAR signal using neural networks. Rio de Janeiro: s.n., VII Encontro de Tecnologia em Acústica Submarina, 2007.

Nielsen, R. O. Sonar Signal Processing. s.l.: Artech House, 1991.

Oppenheim, A. E., Schafer, R. Digital Signal Processing. [ed.] Englewood Cliffs. s.l.: Prentice Hall, 1975.

Pflug, A. L. Variability in higher order statistics of measured shallow-water shipping noise. IEEE, Signal Processing Workshop on Higher-Order Statistics, pp. 400-404, 1997.

Santos, F. L. Aid the classification of passive SONAR with characteristics tones. Federal University of Rio de Janeiro - COPPE-PEO, master dissertation, 2005.

Seixas, M. J., et al. A compact online neural system for classifying passive SONAR signals. International Conference on Signal Processing Application and Technology. 1999.

Skillicorn, D. B. Understanding Complex Datasets. s.l.: CRC Press, pp. 17-18, 2007.

Soares, W. F., Seixas, M. J., Calôba, L. P. Preprocessing Passive SONAR for Neural Classification. Journal of Acoustical Society of America, 2001.

Urick, R. J. Principles of Underwater Sound. 1983.