Contribution to Passive Acoustic Ocean Tomography- Part IV: A Data Fusion Strategy For Blind Source Separation and Classification

S. Vallez, A. Martin, A. Mansour, C. Gervaise, A. Khenchaf

ENSIETA, E3I2 EA3876,
2 rue François Verny
29806 Brest Cedex 9, France
vallezsi@ensieta.fr, Arnaud.Martin@ensieta.fr, Ali.Mansour@ensieta.fr, Cedric.Gervaise@ensieta.fr, Ali.Khenchaf@ensieta.fr,

Abstract – **In this paper we applied Fusion algorithm in a passive Ocean Acoustic Tomography context. We applied several types of independent component analysis algorithms on a large number of realistic data and we propose to classify their respective outputs according to the principal properties of the opportunity sources. Finally, we apply a fusion's model*:* the Belief functions theory to outperform the Blind Source Separation/Classification approach.**

I. INTRODUCTION

Acoustic tomography initially developed by Munk et Wunsch [1], is a way to produce a fast, accurate and low cost monitoring of water mass. This monitoring requires an inversion procedure made with two steps. The first one is to estimate acoustic properties (such as the sound speed profile of the water column) from the measurement of a propagated known acoustic waveform between fixed sources and receivers. Then a second step consists in inferring some physical ocean parameters (temperature, bottom nature) from these previous estimated acoustic characteristics. Large scales deep water and small scales shallow water configurations were successfully studied and associated with matched delay, matched field and matched impulse response inversion processing.

Accurate estimates of acoustic properties demand the emission of powerful and recurrent signals in the adapted bandwidth and in agreement with the scale of the monitoring. But we would rather not send these hard active sounds through the water column in a potential (would-be) military underwater warfare context, or if mammal species health is considered. A recent solution has emerged in the community to tackle this problem with the passive tomography processing. Passive tomography processing consists in estimating acoustic properties by using opportunity sources present in the channel at the time of interest. Some experimentations were recently carried out using ships, marine mammals and surface noises [2] [3] [4] [5].

To perform at sea and real time passive tomography of an underwater channel and to take into account the loss of information (unknown emitted signal and location of sources) and difficulties such as several moving sources at the same time, a strategy for fusion approach and data analysis is proposed in this paper.

In the channel, a huge kind of opportunity sources can be found. Depending on those sources properties, the inversion step will be different. For instance, Jesus in [2] used narrow band boat noises to invert geoacoustic parameters and on an other hand, in [6] C.Gervaise used broad band marine mammals vocalizes to estimate the impulse response of the channel.

That is why we propose a strategy which consists in separating the different sources using ICA (Independent Component Analysis) algorithms and then classifying those sources according to their main properties (Bandwidth, stationnarity…). To do that, we propose to use some information fusion approaches in order to outperform the classification and ICA algorithms for Blind Source Separation (BSS). Indeed, Fusion approach, allows for a representation of both imprecision and uncertainty, that is why according to each ICA algorithm we will be able to model their performances and to increase the global performance of the method by combining several outputs. Finally, thanks to this approach we will be able to direct signals towards the adapted inversion stage which depend on their properties.

The figure 1 summarises the strategy we retained in this paper.

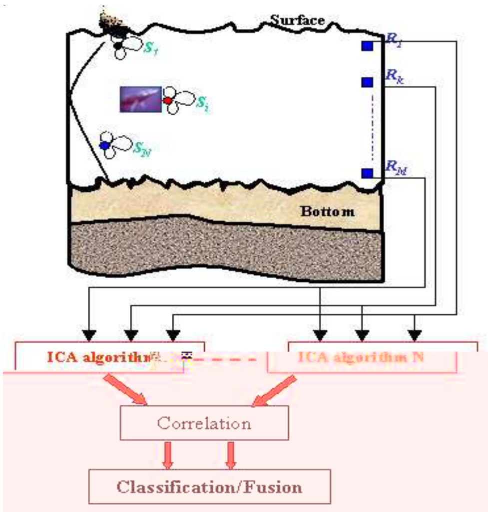

Fig. 1. Synoptic of our approach

In the first paragraph of this paper, we present the sources separation algorithms we used and related assumptions we made. The second paragraph is dedicated to the way we classify signals according to their main properties. The section III presents information Fusion method we chose to apply, this technique is based on the Belief function theory and the masses are calculated thanks to the distances method.

Finally in the last part of this paper we present how we constitute our database of mixtures and some simulation results obtained with the synthetic realistic data we created.

II. SOURCES SEPARATION

In real world, several sources of opportunity can be found in the same channel. Up to now, most of the acoustic ocean tomography algorithms assumes the existence of just one source at same time (or at least the different sources don't have the same signature in the time-frequency plan). For these reason, we should firstly separate the different sources according to their properties (time-frequency or statistical properties). This situation is similar to BSS (Blind Source Separation) problem. To solve BSS problem, researchers use Independent Component Analysis (ICA) techniques [7].

Let us consider independent signals $_1( ),...., ( )$, in the following these signals are called sources, observed by using n sensors, and let $_1( ),...., ( )$ denotes the observed (mixed) signals, then for memory-less channel (instantaneous mixtures) one can write

$$\mathbf{x} \quad = \quad \mathbf{s}( ) \qquad (2.1)$$

The BSS problem consists in retrieving the sources by only using the observed data $\boldsymbol{x}( )$. In general case, three main assumptions are considered to solve the BSS problem:

- The sources are statistically independent
- The matrix is a full rank matrix.
- At most one of the sources can be Gaussian

Our experimental studies show that these assumptions can be applied for some acoustic signals and channel. Under water channel can be consider as multiple paths (. the signal received on the hydrophones can be seen as an attenuated and delayed version of the emitted one):

$$( ) = \sum_{=0} \quad (\, - \tau \,) + \, ( ) \qquad (2.2)$$

Where denotes the number of paths of the channel, a real gain and denotes an additive white Gaussian noise.

We can also rewrite this equation in a matrix form by introducing $\boldsymbol{H}$ which denotes the impulse response of the channel.

$$( ) = \sum_{=0} \; ( ) \; (\, - \,) + \; ( ) \qquad (2.3)$$

The previous equation is called a convolutive mixture model. At first, we considered a simplified model i.e. Instantaneous mixtures (memory-less channel). The instantaneous mixture model can be applied in the deep channel when the sources and the sensors are far enough (depending on the channel as well the waves properties) from the water surface as well as from the bottom of the channel.

Many ICA algorithms for instantaneous mixtures can be found in the literature. Most of them are dedicated to some specific applications. In [8], some ICA algorithms have been considered to deal with natural and artificial acoustic sounds. Among these algorithms, we focus on two particular criteria and algorithms which gave us the best performances: Jade [9] is a simultaneous diagonalization of the eigen matrices of a fourth order cross cumulant tensor. The second criteria are presented in FFPA [10] [11] (Fast Fixed point Algorithm), three versions are considered according to different nonlinearity used in FFPA. The main idea of FFPA consists on minimizing a cross normalized fourth order cumulants using a fixed point minimization algorithm.

Figure 2, shows the experimental results obtained by applying FFPA on underwater acoustic signals. In that figure, we present the spectrograms of the sources, the mixed signals and the separated ones. The two sources are: a marine mammal (Short-Finned Pilot whale vocalize) and an artificial sound (LFM).

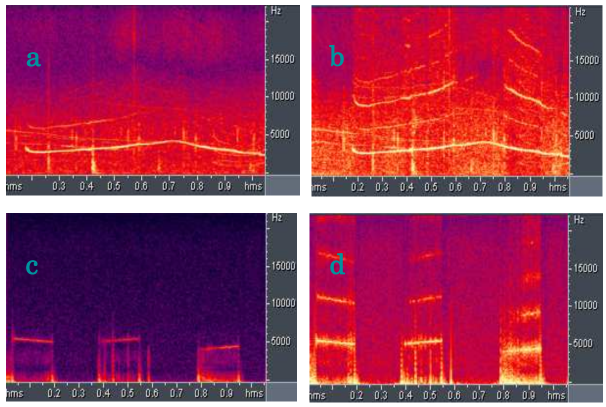

Fig. 2. (a) Initial spectrogram representation of a Short-Finned pilot whale vocalize (b) signal obtained after BSS (c) initial LFM (d) signal obtained after BSS

III CORRELATION/ FEATURE EXTRACTION

We mentioned earlier that the BSS literature is full, of many algorithms, many applications and divers possibilities. In our experimental studies, we couldn't found the best algorithm for all type of sources and/or channels.

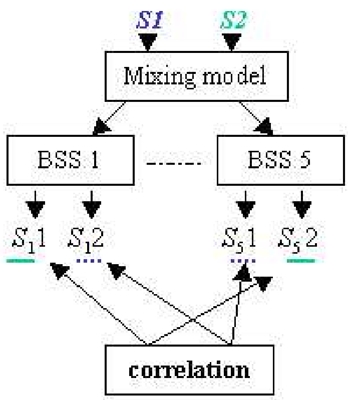

Fig. 3. Correlation principle

For this reason, we consider hereinafter the following strategy, see figure 3: By considering the output signals of many ICA algorithms and by using some classification and fusion techniques, can we improve the total performances? To answer this question, one should be able to consider an estimated signal as the most "similar" (according to some criteria) output signals of the different ICA algorithms. On the other hand, it is well known that the separated signals are the sources up to a scale factor and a permutation [7]. As the sources are independent form each other, we considered a "similarity" criteria based on a normalized cross-correlation.

In the second section of this paper, we present the source separation algorithms we used, moreover, as we mentioned in the introduction, our goal consists in directing the different signals towards the inversion schemes which are related to their respective properties. To perform this task, we need a feature extraction stage.

By studying the different kind of signals existing in the channel, we chose to fix 4 possible classes of signals depending on their principal properties:

- Wide band stationary signals
- Narrow band stationary signals
- Wide band nonstationary signals (frequency modulated signal, marine mammals vocalizes for instance)
- Narrow band nonstationary signals (Transitory impulsionnal signals, snapping shrimps for instance)

We chose to perform the classification at the same time than fusion by using the belief functions theory with some mass functions calculated by a distances approach. This idea is developed in the fourth section of this paper.

By choosing this way to perform the classification, we have to extract from our signals relevant features which could be used to calculate some distances between a signal's features and an other signal contained in a learning database ones.

To extract relevant features, we used two approaches. The first one is a simple signal processing approach and the second one is classic in speech recognition applications.

Actually, to evaluate if a signal is broad or narrow band, we simply calculated its bandwidth by the equation (3.1) and (3.2):

$$" \quad = \frac{\int_0^\infty " .\gamma(") \, "}{\int_0^\infty \gamma(") \, "} \qquad (3.1)$$

$$= \sqrt{\frac{\int_0^\infty (" - " \quad)^2 .\gamma(") \, "}{\int_0^\infty \gamma(") \, "}} \qquad (3.2)$$

Moreover, to evaluate if a signal is stationary or not, we calculated its central frequency in a moving window of fixed larger. Then, we were able to observe the evolution of the central frequency which allowed us to see if a signal is stationary or not.

Finally, this simple signal processing approach provides us two relevant features.

We also extract other features, to provide comparison with features currently used in speech recognition applications, we provide results obtained by using 13 Mel-frequency cepstral coefficients (MFCCs). We do not detailed the way to obtain those features in this paper, interested readers would see [12] for more details. Those cepstral coefficients obtained with a Mel-frequency scale filter bank which has been defined according to human perception specificity are well adapted to speech recognition application but not necessarily to underwater acoustics signals classification.

The figure n°4 presents the synoptic to obtain the MFCC's:

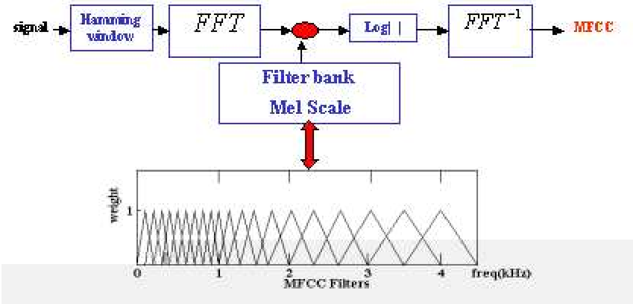

Fig. 4. MFCC's extraction synoptic

IV FUSION BY BELIEF FUNCTIONS THEORY

In this paragraph, we briefly describe the theory of the fusion model we apply in this paper.

The belief functions theory allows for a representation of both imprecision and uncertainty through two functions: plausibility and belief [13, 14]. Both functions are derived from a mass function defined on each subset of the space of discernment $=\{C_1, \ldots, C\}$ (2) onto [0,1], such that:

$$\sum_{\subseteq} \quad (\quad) = 1 \qquad (4.1)$$

The first difficulty is the choice of a mass function. There are several approaches in the literature, but in this paper we apply the Denoeux Distances Method. Indeed this method use a classification by the -nearest neighbors scheme which allow us to avoid using a specific classifier. Actually, in [13], Denoeux defined the mass functions by:

$$\begin{cases} {}_{!}(\{ \quad \}/ \quad ^{()})(\quad) = \alpha_{!}\varphi(\quad ^{()}) \\ {}_{!}(\quad / \quad ^{()})(\quad) = 1 - \alpha_{!}\varphi(\quad ^{()}) \end{cases} \qquad (4.2)$$

where $^{()}$ corresponds to a learning vector, $^{()} = (\; ^{()}, \;)$ is a distance (to be fixed) between and $^{()}$ and is the class of $^{()}$; φ is a distance function which verifies:

$$\begin{cases} \varphi(0) = 1 \\ \lim \lim_{\to +\infty} \varphi(\quad) = 0 \end{cases} \qquad (4.3)$$

Many distance function can be used, we will use the function proposed in [13] by Denoeux:

$$\varphi(\quad) = \exp(-\nu \quad ^2) \qquad (4.4)$$

where V corresponds to a positive parameter according to the class . This function is well adapted to the Euclidean distance. If the training database is important, the distances calculation can take time, in our application we will focus on the nearest neighbors to limit this calculation time.

The second step of fusion is the combination. There are several conjunctive rules proposed in the literature to perform this combination. In this article, we used the Smets rule one [14] which is nowadays the most widespread one.

$$\begin{cases} (\ ) = \sum_{1\cap\ldots\ =\ \neq\phi} \prod_{!=1} {}_!(\ {}_!) \\ (\phi) = \sum_{1\cap\ldots\ =\phi} \prod_{!=1} {}_!(\ {}_!) \end{cases} \quad (4.5)$$

In the distance approach, this combination can be written by:

$$\begin{cases} {}_!(\{\ \})(\ ) = \frac{1}{\#}\left(1-\prod_{\in\$}(1-\alpha_!\varphi(\ ^{()}))\right). \\ \prod_{\neq}\prod_{\in\$}(1-\alpha_!\varphi(\ ^{()})) \\ {}_!(\ )(\ ) = \frac{1}{\#}\prod_{=1}\prod_{\in\$}(1-\alpha_!\varphi(\ ^{()})) \end{cases} \quad (4.6)$$

With **#** a normalized constant and **$** the set of neighbors of in the class .

The last step of fusion is the decision. In the belief functions theory, we can use the maximum of plausibility, maximum of belief or maximum of pignistic probability. In this article, we use the maximum of pignistic probability proposed by Smets [14] which constitutes a good compromise between the plausibility and belief maximum.

V SIMULATION RESULTS

To improve our algorithms, we constructed a database of realistic synthetic mixtures.

We used several signals of marine mammals, ships and ambient noise founded on the web and in two oceanic campaigns: the first one, TINA2001 is an active Tomographic campaign performed by the SHOM (**%** **&** **'(** **)** ***** **&** **'(**) and the other one, was performed by ISMER (**$**) from the University of Quebec at Rimousky (UQAR) in order to locate Belugas in the Saint Laurentian Channel. The descriptions of those experimentations can be found respectively in [15] and in [16].

We decided to use a shallow water environment with a simplified celerity profile and a sediment layer corresponding to the TINA campaign. The figure 5 is a representation of the model we retained.

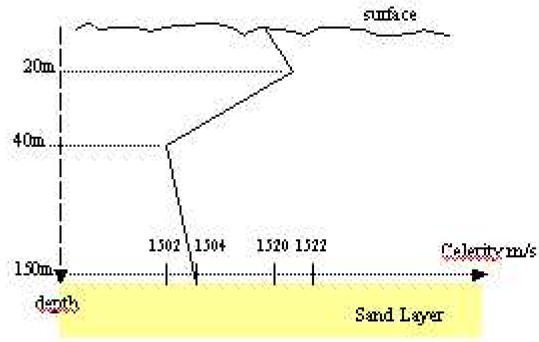

Fig. 5. Model of Channel used to construct our database

To construct realistic mixture matrices, we used the software Bellhop developed to simulate the rays propagation model. We used the retained channel described in the figure 5, thus, thanks to this software, we were able to consider a source in the channel and to obtain the Transmission Loss due to propagation between the opportunity source and the receivers. On the figure 6, we present an example of the transmission loss obtained with a 10m-depth source.

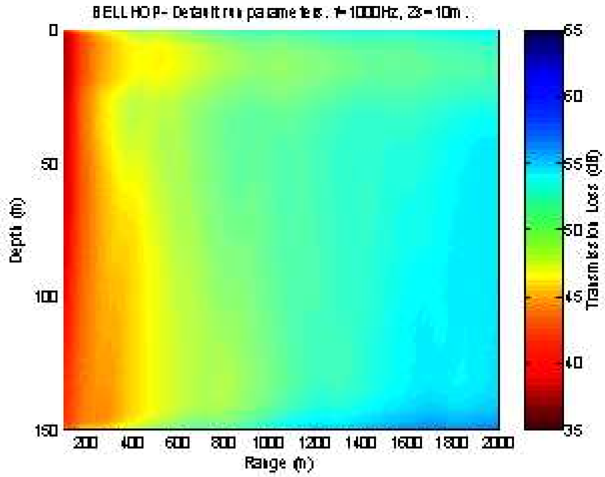

Fig. 6. Transmission Losses (dB) obtained with Bellhop for a configuration corresponding to the figure 5

Moreover, to construct our different mixture matrix, we need the knowledge of the sources of opportunity's emission levels. We found those levels in the literature [17]. We resumed it in the table 1.

We put three receivers in the channel in order to sample it. The first one at 5m depth, the second at 50m depth and the third one at 120m depth. And finally, by combining the source emission levels, the transmission losses and by using several sources proofs and ranges we were able to obtained a large number of realistic synthetic mixture matrix.

Actually, we created above 200 mixtures by considering 2, 3 or 4 sources moving in the channel. We created a 80 signals learning database and a 120 signals test database.

TABLE 1: OPPORTUNITY SOURCES EMISSION LEVELS

Opportunity Sources	Frequency (Hz)	Source Level (dB re 1µPa-m)
Trawlers	100	158
Tanker (s)	430/60/33/8	169/180/181/185
Twin diesels ships	630	159
Active Tomography	Chirp 300-1000	177
Dolphins	0.8-24 Echolocation,: 23-67 kHz	125-173 180
Sperm Whales	Clcks: 2-4Khz; 10-16 kHz	160-180
Beluga	Vocalizations 0.5-169 kHz Echolocation: 40-60;100-120 kHz	160-180 206-225
Dwarf minke whale	Grunts: 60-140	155-175
Killer Whale	12-25KHz	180
False Killer Whale	25-30; 95-130 kHz	220-228
Humpback whales	Song :30-8000 Pulse trains:25-1250	144-174 179-181
Short-finned pilot whale	30-60 kHz	180
Harbour Porpoise	Clicks 0.04 –12 Echolocation 110-150	120-148 135-177
Snapping shrimps	2 to 9 KHz	189

+

In this section, we present some results obtained by applying our classification/fusion approach on the outputs of the four ICA's algorithms described in the section II. The figure 7 et 8 present a comparison between rates of good classification obtained with ICA's algorithms with the nearest neighbours and rates of good classification obtained with the fusion approach. The figure 7 corresponds to the case of two parameters and the figure 8 corresponds to the results obtained with the 13 MCFF's.

We repeat the training 20th times in order to achieve a good estimator of the classification rate. The strait lines in the figure 7 and 8 correspond to the mean rates of good classification which are also explained respectively in the table 2 and 3.

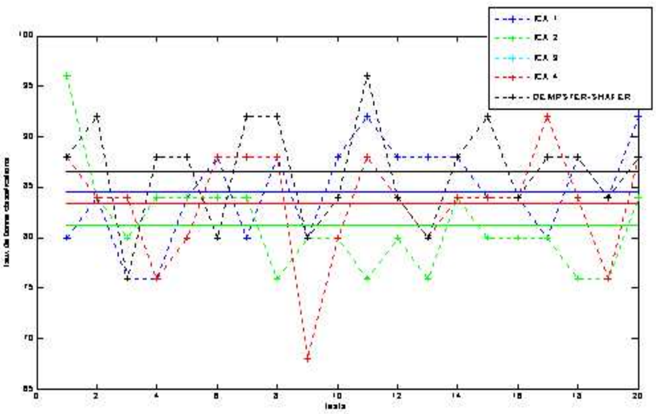

Fig. 7. Classification rate obtained with band and stationary parameters

TABLE 2: CLASSIFICATION MEAN RATE WITH TWO PARAMETERS

ICA 1	ICA 2	ICA 3	ICA 4	Belief functions theory
74.1	73.6	72.08	72.3	76.4

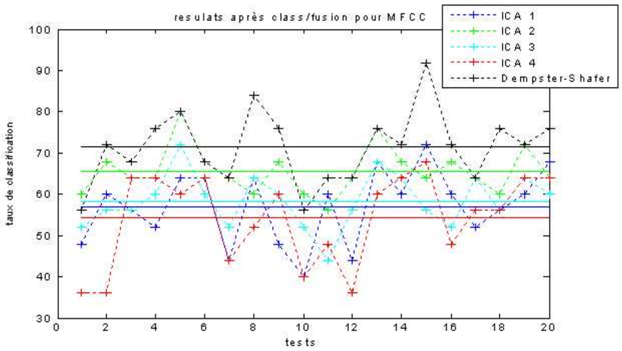

Fig. 8. Classification rate obtained with 13 MFCC's

TABLE 3: CLASSIFICATION MEAN RATE WITH 13 MFCC's (%).

ICA 1	ICA 2	ICA 3	ICA 4	Belief functions theory
55.2	57.09	59.12	67.41	71.6

You can note on those two figures that the fusion model give better results in mean than every individual ICA/k-nearest neighbors algorithm. The improvement of the Belief functions theory is thus statistically significant (between 4 and approximately 10 percents depending on the features we retained), but not still really determinant. We hope to obtain better results with a better knowledge of the truth, i.e. with better learning and test data bases.

VI CONCLUSION AND FUTURE WORKS

We have used four ICA algorithms based on different principles or on different parameterizations. These algorithms allow quite good performances, but those performances are not similar according to the type of sources. So, we have used a fusion approach: Belief functions theory based on the distance method. This approach outperform the results of each ICA/classifier approach separately.

We mentioned, that the ICA algorithms are currently being adapted to convolutive mixtures, thus we will soon be able to perform simulation much closer from reality. Moreover, our Tomography team is currently working on other approaches to separate sources (Time-Frequency and Time-Frequency-Space approaches). So, as future work, we would like to apply our Classification/Fusion approach to several source separation methods.

Acknowledgments

The authors want to thank the S.H.O.M. (% & '( )* & '( ) and the University of Quebec at Rimousky for providing us real data.

REFERENCES

[1] W. Munk and C. Wunsch, Ocean acoustic tomography: rays and modes, Reviews of geophysics and space physics, vol. 21, n°4, p 777-793, May 1983.

[2] S.M. Jesus, C. Soares, J. Onofre, P. Picco, & ' , ' $ " -.. , ECUA 2002.

[3] C. Park, W. Seong, P. Gerstoft, D. Battle, P. Nielsen, / & % & ' , OCEANS 2003.

[4] A.M. Thode, G.L. D'Spain, W.A. Kuperman, ' 0" &1& % 1 & % " 2' % 3 , JASA, vol. 107, n° 3, March 2000.

[5] N.M.Carbone, G.B. Deane, M.J. Buckingham, & ' ' 2 % " ' 2 2 " ' % ' " ' 2 , J. Acoust. Soc. Am. 103 (2), Feb 1998.

[6] C.Gervaise & al, Contributions to passive acoustic oceanic tomography-Inversion Algorithms based on Time Frequency Space Representation, PSIP 2005.

[7] P. Comon. $ 1 2 4 & &, Vol.36, n°3, pp287-314, April 1994.

[8] A. Mansour, C. Gervaise, $ % & ' WSEAS, Issue2, Vol. 1, April 2004.

[9] J.F Cardoso, &'0 " $, Neural Computation, Vol. 11, pp157-192, 1999.

[10] J.F. Cardoso and A. Soulamic, " &" 0& & . IEEE Proceedings-F, 140(6):362, December 1993.

[11] A. Hyvaerinen and E. Oja. " " & ' " . Neural computation, Vol.9 pp 1483-1492, 1997.

[12] B. Gold, N. Morgan 1 ' & 5 & , 5 & 5 " ' 1 John Willey and Sons, U.K., 1999.

[13] T. Denoeux, 0 &' " 0 ' " / ' , IEEE Trans. on Systems, Man and Cybernetics, Vol.25 n°5, pp 804-813, 1995.

[14] P. Smets, / ' " ' % ' / " " , IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.12 n°5, pp 447-458, 1990.

[15] + * & /$6 7..81 EPSHOM/CMO/GEMOS/NP

[16] http://www.ismer.ca/

[17] R.J. Urick, 5 " 9 2 , McGraw-Hill, New-York, 3rd ed., 1983.