PERFORMING DIMENSIONALITY REDUCTION OF NONSTATIONARY STOCHASTIC PROCESSES USING PERTURBATION ANALYSIS.

V. VIGNERON[1,2,3], H. MAAREF[1]

[1]LSC CNRS FRE 2494
40 rue du Pelvoux
91020 Evry, France
Tel: +33.1.69.47.75.24
vvigne,maaref@iup.univ-evry.fr

[2]LIS-INPG CNRS UMR 5083
46 av. Félix Viallet
38640 Grenoble, France
Tel: +33.4.76.57.43.50
vvigne@lis.inpg.fr

[3]MATISSE-SAMOS CNRS UMR 8595
90 rue de Tolbiac
75634 Paris Cedex 13, France
Tel: +33.1.44.07.89.35
vigneron@univ-paris1.fr

1. INTRODUCTION

This paper describes the performance characteristics of an adaptive filter, a digital filter based on the perturbation analysis of the eigenvalue equation. The algorithm find the singular value decomposition (SVD) of a general covariance matrix by stochastic approximation. The deviation in the eigenvalues through times is shown to be an index of the energy present in each channel of an sensors array The analysis of a simple nonstationary case is given.

2. SENSOR SELECTION STRATEGIES

While it may be appropriate to design networks that densely populate a region with microsensors in medical application for instance - such as electroencephalogram (EEG) or electromyogram (EMG)-, operation of the network may not require that all nodes be operating. Indeed, for efficient operation and efficient use of communication under restricted requirements, it may be desirable to select a subset of nodes to communicate at any fixed time since the selected subset can change over time.

Usual hypotheses are the following: (*i*) 2 sensors don't record exactly the same signal, (*ii*) all sources must be involved in the recording with a non-zero variance, (*iii*) electrodes providing irrelevant signals are rejected, (*vi*) the low power of the signal of interest can be improved, (*v*) there doesn't exist an optimal location of the sensors, constant in time as the "target" moves. These considerations, in addition to the low power of the signal, may explain why the locations of the electrodes can improve the signal extraction, while others can decrease its efficiency.

Note that fusion approaches range from simple rules of picking the best result to model-based techniques that consider how the signals are generated. Again, there is a tradeoff between performance and robustness [1]. Simple fusion rules are robust but suboptimal while more sophisticated and higher performance fusion rules may be sensitive to the underlying models [2]. The fusion algorithm should *recognize* the *dependency* in the information to be fused and avoid double counting.

We consider a digital signal processing sensor array system (Fig. 1) based on randomly distributed sensor nodes. In most array processing, the sensor array geometry is fixed and known. In this system, array calibration may be impractical due to unknown placement and orientation of the sensors with unknown frequency/spatial responses. We propose a "blind" selection technique, using only the measured sensor data, to form either a sample data or a sample correlation matrix.

Not so many articles have dealt with this topic, usually in the area of digital communications. The typical scenario involves narrowband sources of which some known characateristics are used for the pupose of detection. Among the features exploited are the cyclostationarity property [3], spectral coherence [4], the statistical difference between desired and undesired sources, including types of signal nonstationarity and higher order statistical parameters [5]. The latter class of problems has generated a wide variety of articles in which higher order cumulants have been effectively used to limit the effect of mesokurtic disturbances such as Gaussian noise [6].

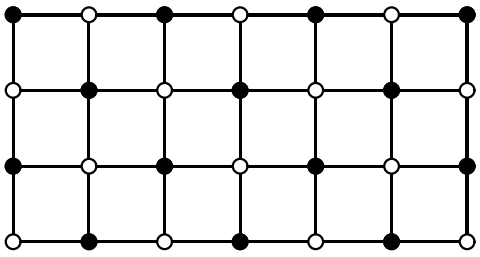

Figure 1: Networked sensors.

The aim of choosing electrodes is to enhance the quality of the output signal. This quality is strongly compromised by the fact that the signal of interest is overlayed by other signals with an equal or higher amplitude. Hence, to make the signal of interest observable, the sources have to be separated. This is not possible by conventional linear filtering as the signals overlap in time and frequency space, but if their mixture is linear, instantaneous and noise-free, the well-known method of independent component analysis (ICA) is able to recover the original sources, up to a scale factor and permutation [7][1]. As former investigations have shown, applying ICA to the signals of experimental found electrodes is not satisfying: by selecting the "best" electrodes the quality of the input signal of the ICA procedure will increase, and thus also the quality of the output. The reason for the fact that some electrodes are better than others is that the sources are not emitting homogenously in every direction but, for every source, there are regions of maximal and regions of minimal amplitude [8]. As a conclusion, the task is to find pairs of electrodes so that their *differential measurements* are optimized in order to separate the sources.

A criterion or cost function for that optimization has to be defined. Two ways for finding optimal electrodes can be identified:

1. The first possibility is to get the measurements of all electrodes, to compute the cost function of every electrode pair and to select the pairs with the lowest value. This

[1]A necessary condition to recover the original sources is that the number of external sensors must be greater or equal to the number of original sources.

would mean the computation of all possible combinations of two electrodes.
2. The second way could be to select electrode pairs *randomly* and change their position iteratively by only using a subset of measurements.

Unaffected by the cost function used, several questions came up concerning the optimal electrodes found:

1. Are those sensors, for which the criterion is fulfilled best, also the best for another criterion, which is impossible to measure? One answer regarding an example criterion could be that the measurement with the lowest amplitude of the signal s_1 is the best for observing the signal s_2. But the signal s_2 may not be visible at all in this measurement.
2. What if the solution is not one single but a set of different measurements? An additional criterion has then to be applied to this set. *E.g.* measurements with minimal amplitude of s_1 are those orthogonal to s_1 but, among these, are also those orthogonal to the s_2. Thus, out of the sensors orthogonal to the signal s_1, those sensors have to be chosen, which are most parallel to the signal s_2.
3. Does there exist, among the latter set of different measurements, a trivial one? *E.g.* one measurement with minimal amplitude of s_1 could be received when using a first electrode as reference and a second electrode which is the same as, or as near as possible to that reference. In that case, this solution has to be excluded.

In the following, we attempt to formulate a "statistical theory of energy variation". This is a very difficult subject and the present work should be regarded as only a beginning.

3. FORMAL PROBLEM STATEMENT

We consider a sensor array system – a rectangular structure with base> height [9] (see Fig. 1). The sensors' relative positions are unknown and the sources have specific frequency characteristics that cannot be used to our advantage. In the following, every measurement of sensor signals is the *difference* between the potential at one sensor and the potential at a "reference" sensor, which is performed by a differential amplifier. Consider the situation in which R sensors are randomly distributed in a spatial region, which can be 1,2 or 3 dimensional. Denote the $R \times 1$ sensor data vector by

$$x_t = (x_1(t), \ldots, x_r(t), \ldots, x_R(t))^T. \tag{1}$$

x_t denote a random process. Let X_0 be the dataset collecting the samples $x_t, t = 1, \ldots, T$ and $C_0 = \mathbb{E}[x_t x_t^T]$, the autocorrelation function of x estimated by $\hat{C}_0 \cong \frac{1}{T} \sum_{t=1}^{T} x_t x_t^T$, $\quad_i$ and $\quad_i, i = 1, \ldots, n$ be the *eigenvectors* and *eigenvalues* of C_0. Note that $\hat{\quad}_i$ and $\hat{\quad}_i$ are *estimates* of $\quad_i$ and $\quad_i$. Assume that the $\quad_i$'s are disctincts. For an additive sample x_{T+1}, we have $C = (X_0 | x_{T+1}^T)^T \begin{pmatrix} X_0 \\ x_{T+1}^T \end{pmatrix} = X_0^T X_0 + x_{T+1} x_{T+1}^T = \hat{C}_0 + x_{T+1} x_{T+1}^T$.Let $\quad C = x_{T+1} x_{T+1}^T$ be a real symmetric perturbation matrix. We wish to obtain a first-order approximation of the eigenvectors and eigenvalues of C in terms of the $\quad_i$'s and $\quad_i$'s, where $C = C_0 + \quad C$. These may be obtained by retaining the terms of first order or lower of the equation:

$$(C_0 + \quad C)(\quad_i + \quad_i) = (\quad_i + \quad_i)(\quad_i + \quad_i), \tag{2}$$

where $C_0 \quad_i = \quad_i \quad_i$. The resulting equation is:

$$C_0 \quad_i + \quad C \quad_i \cong \quad_i \quad_i + \quad_i \quad_i. \tag{3}$$

To calculate $\quad_i$, we left-multiply (3) by $\quad_i^T$ and, since $\quad_i^T C_0 = \quad_i \quad_i^T$, we have:

$$\quad_i \cong \quad_i^T \quad C \quad_i. \tag{4}$$

Since the $\quad_j, i = 1, \ldots, n$ form a set of basis vector, *i.e.* $\quad_i^T \quad_j = \quad_{ij}, \forall i, j$, we can write $\quad\quad_i$ as a linear combination of the $\quad_j$'s as follows

$$\quad_i = \sum_{j=1}^{n} b_{ij} \quad_j, \tag{5}$$

where $b_{ij} = \quad_j^T \quad\quad_i$. To deal with nonstationary processes in discrete time, it is reasonnable to expect that the derivative of the autocorrelation function for x can be computed: $E[x_t^{(i)} x_t^{(j)^T}] = \frac{d^{(i+j)}}{dt} C_0$ (see Box and Miller [10]). In particular, we suppose that when we pass to continuous time, many of the definitions generalize in a straigthforward way, although some new problems arise in connexion with limiting operations, and then: $\frac{d}{dt} C_0 \approx E[\dot{x}_t \dot{x}_t^T]$.

From the eigenvalue equation, $C_0 = \quad\quad^T$, with $\quad = (\quad_1 \ldots \quad_R)$ and supposing we can manipulate derivatives of this equation freely, exactly as differentiating ordinary functions, we have

$$\dot{C}_0 = \dot{\quad\quad}^T + \dot{\quad}\dot{\quad}^T + \quad\quad\dot{}^T, \tag{6}$$

Note that $\dot{C}_0$ is simply obtained by performing $x_t - x_{t-1}$ for any t. Hence, by identifying $\dot{\quad}_i \approx \quad_i$ and $\dot{phi}_i \approx \quad_i$, we have

$$\dot{C}_0 = \sum_i \quad\quad_i \quad_i \quad_i^T + \sum_i \quad\quad_i \quad_i^T + \sum_i \quad_i \quad_i \quad\quad_i^T, \tag{7}$$

By right and left-multiplying by $\quad_i$ and rearrange using (7), we have:

$$\quad_i = \frac{(\dot{C}_0 - \quad C) \quad_i}{2 \quad_i}. \tag{8}$$

At this stage, the R-measurements have been combined together providing R more informative signals, as it can be seen in the following example.

Example 1 *Consider the simple scenario where $R = 50 \times 50$ diaphragm electrodes, located around a patient thoracic cage (see Fig. 2), each of them providing uncertain measurements, to be combined in order to be able to extract the d sources in real time.*

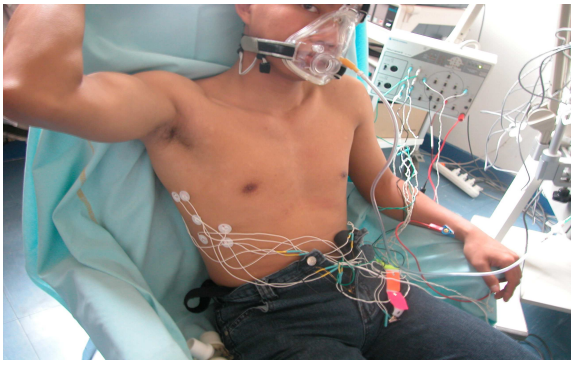

Figure 2: Multi-electrodes surface data acquisition.

A subset $\mathcal{X}^*$ of $n \ll R$ signals recorded by selected sensors will be processed by a Blind Source Separation (BSS) algorithm, discarding all other electrodes (Fig. 3.a).

The quality of the extracted electromyogram diaphragm using only $\mathcal{X}^*$ is improved by comparison to the performances reached if the whole set of signals ($n = 100$) is used in the extraction process.

A set of electrode pairs can be easily identified (the K sensors for which the amplitude of the of the $_i, i = 1,\ldots,K$ is maximized) based on the variation of information bring at each sample.

3.1 Finding Further Pairs...

Once a set of pairs of sensors have been found, further pairs have to be found from the measurements $x_1 x_2 \ldots x_n$, where $x_i = (x_{i1},\ldots,x_{iT})^T$ is the data vector of sensor i and ^{T} is the transpose operator. More electrode pairs can be found by minimizing the correlation between an additional pair and a pair already found. This could be achieved by examining the following correlation difference:

$$D. = \mathbb{E}[x_d x_r] - \mathbb{E}[x_o x_r], \tag{9}$$

where x_r denotes the signal of the already found pair, x_o denotes the signal of the additional pair at the actual position and x_d is the signal of the additional pair.
When computing the correlation by equation (9), it is only possible to receive the correlation between two sensor pairs. But it is necessary to have one single criterion which reflects the correlation between one new signal and an arbitrary number of already found signals. It was proposed to use the angle between the axis of the new signal and the eigenvector of the new signal. By eigenvector and eigenvalue is meant the eigenvector and eigenvalue of the full correlation matrix of the signals $\mathbb{E}[X^T X]$. The explanation for this uses statistical matters [11] which will be outlined in the following, while the first and second signals are always the signals of the already found sensor pairs and the third signal is always the signal of a new sensor pair whose placement has to be optimized. The problem is located near the problem of principle component analysis, which is discussed in detail in [12].

In the following, all signals are assumed to have zero mean, which could be easily achieved by subtracting the mean from the signals. The correlation between two signals could also be seen as in figure 4 where the lines indicate the eigenvectors of the signals. When one signal is already recorded and a second one has to be found, the two signals are uncorrelated if the eigenvector of the second signal is lying parallel to the axis of the second signal (figure 4.b). The two signals can be correlated if the eigenvalues are rotated as in figure 4.a.
Thus the angle between the eigenvector and the axis of the new signal can be used as a scalar criterion for the uncorrelatedness of the signals. The angle between two vectors a and b is calculated by equation (10):

$$= \arccos\left(\frac{a \cdot b}{|a||b|}\right) \tag{10}$$

This calculation works also for higher dimensions and is used as the quality criterion. Figure 5 shows the eigenvectors of the signals before (a) and after (b) the iterative optimization. One problem of this optimization was, that the differential signal of two sensors very near to each other is uncorrelated with any other signal but sensors like that are not useful for a measurement. The approach of the two sensors to each other is accompanied by a significant decrease of the eigenvalue of the new measurement. Thus if the eigenvalue is beneath a certain threshold, the sensors are replaced randomly.

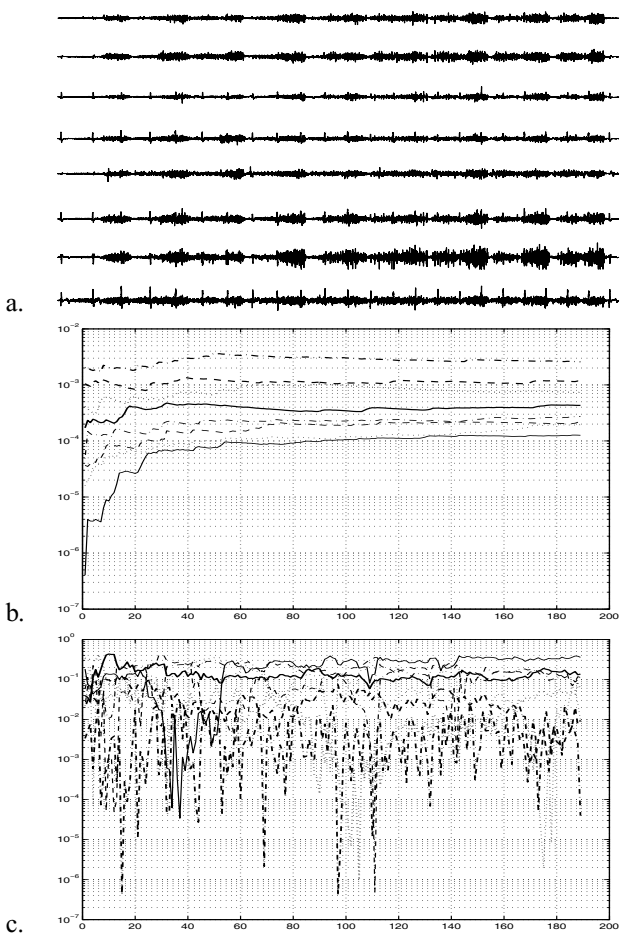

Figure 3: (a). 8 channnels electromyographic signals, (b). Eigenvalues of signals with time, (c). Eigendeviation with times.

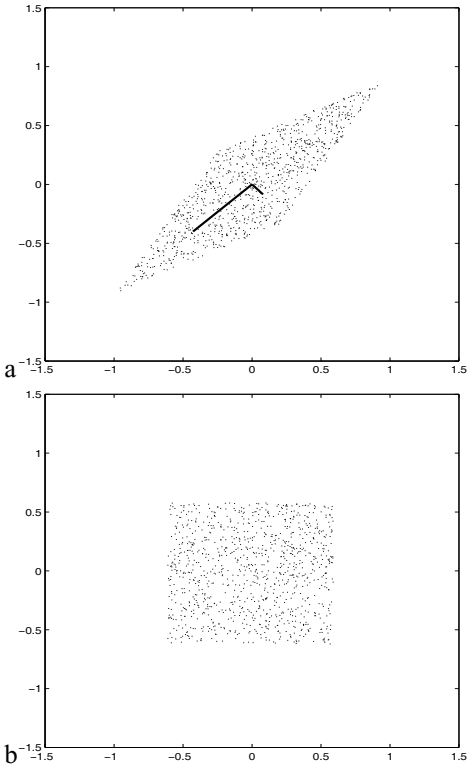

Figure 4: Distributions of Variables

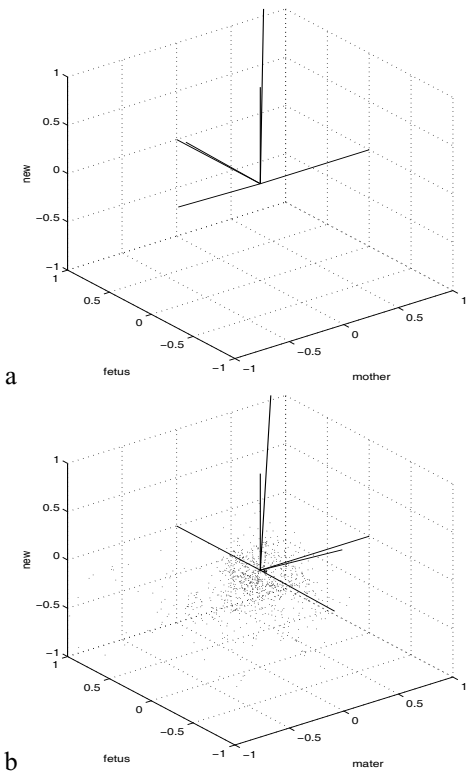

Figure 5: Eigenvectors for three dimensions

4. CONCLUSION

In this paper, we propose strategies based on sensor selection. We have shown that sensor selection is a *viable* approach in the absence of reliable and detailed prior information by simply using the autocorrelation of the signals and its derivative.

REFERENCES

[1] J. Denzler and C. Brown. Information theoretic sensor data selection for active object recognition and state estimation. *IEEE trans. on Pat. Anal. and Mach. Intel.*, 24(2):145–156, February 2002.

[2] K. Yao, R.E. Hudson, Chen D., and F. Lorenzelli. Blind beamforming on a randomly distributed sensor array system. *IEEE Journal on selected areas in communication*, 16(8):1555–1566, October 1998.

[3] D.T.M. Slock. Blinf fractionally spaces equalization, perfect reconstruction filter banks and multicannel linear prediction. *Proc. IEEE ICASSP*, pages 585–588, 1994.

[4] B.G. Agee, S.V. Schell, and W.A. Gardner. Spectral self-coherence restoral : a new appraach to blind adaptive signal extraction using antenna arrays. *Proc. IEEE*, 78:753–767, 1999.

[5] A. Souloumiac. Blind source detection and separation using second order nonstationarity. *Proc. ICASSP*, pages 1912–1915, 1995.

[6] V. Vigneron and C. Jutten. *Independent Component Analysis and Blind Signal Separation*, volume LNCS 3195 of *Lectures Notes in Computer Science*, chapter Fisher Information in Source Separation Problems, pages 168–176. Springer, 2004.

[7] P. Comon, C. Jutten, and Herault J. blind separation of sources, part 2: Problems statement. *Signal Processing*, 24:11–20, 1991.

[8] V. Vigneron, A. Azancot, C. Hérail, M. Schmidt, O. Sibony, and C. Jutten. Sensor array for foetal ECG. part 1: Simulations. In *proceeding IEEE CIMED'2005*, 2005.

[9] V. Vigneron, A. Azancot, A. Sibony, C. Hérail, and C. Jutten. Dispositif matériel et logiciel d'extraction d'ECG foetal. enveloppe Soleau, INPI, 2002.

[10] D.R. Box and H.D. Miller. *The theory of stochastic processes*. Chapman and Hall, 1965.

[11] A. Borovkov. *Statistique mathématique*. Editions Mir, Moscou, 1987.

[12] A. Hyvärinen, J. Karhunen, and E. Oja Oja. *Indepentent Component Analysis*. John Wiley & Sons, Chichester, 2001.