SUBSPACE DETECTORS FOR MULTICHANNEL SIGNALS

Salvador Olmos[1], *Juan Pablo Martínez*[1] *and Leif Sörnmo*[2]
[1]Dept. of Electronics and Communications Eng. University of Zaragoza.
[2]Dept. of Electroscience. Lund University, SWEDEN.
María de Luna, 3 - E50015 Zaragoza, Aragón, SPAIN.
e-mail: olmos@posta.unizar.es

ABSTRACT

In this paper a multichannel subspace detector is proposed based on a separable spatio-temporal linear model. The generalized likelihood ratio test (GLRT) for this model is derived for the case of coloured Gaussian noise and known temporal and covariance matrices. Simulation results support the theory and illustrate the benefits of the spatio-temporal detector versus channel-by-channel detectors for low values of SNR.

1 INTRODUCTION

The detection of a signal in noise is a frequently encountered problem in different signal processing domains. Usually, there is not a complete previous knowledge of the signal to detect. In these cases, the Generalized Likelihood Ratio Test (GLRT) is often used [1]. If the waveform of the signal is perfectly known (except for a proporcionality constant), the GLRT detector is the well-known matched filter. In several applications the signal waveform is not completely known, and a weaker assumption is often used where the signal lies in a subspace. The GLRT detector is then the matched subspace detector [2] of which the matched filter is a rank-one particular case.

In the field of biomedical signal processing, many bioelectrical signals, such as the electrocardiogram (ECG), electroencephalogram or body surface potential maps are recorded using a set of electrodes spatially distributed on the body surface. However, most signal processing methods rely on one-dimensional techniques, either in time or in space domain. In this paper, we propose a framework for detection problems in bioelectrical signals based on a multichannel spatio-temporal approach. In most of these applications, a wide variety of signal waveforms can be found (e.g. QRS morphologies in the ECG), and these are often modelled according to linear expansions.

The rest of the paper is organized as follows. A multichannel signal model is introduced first based on truncated spatio-temporal linear expansions. The GLRT for this model is considered in Section 3. A detection performance measure is given in Section 4. Some simulations are given in Sections 5 and 6.

This work was supported by grant TIC2001-2167-C02-02 from CICYT and P075/2001 from CONSID-DGA (Spain)

2 MULTICHANNEL SIGNAL MODEL

The information conveyed by a multichannel signal can be represented by a matrix $\mathbf{D} \in \mathbb{R}^{N \times L}$, N being the number of samples and L the number of sensors. The signal $\mathbf{D}$ can be decomposed as a linear combination of $N \times L$ linearly independent spatio-temporal functions (elementary matrices) $\mathbf{B}_{ij}$

$$\mathbf{D} = \sum_{i=1}^{N} \sum_{j=1}^{L} w_{ij} \, \mathbf{B}_{ij} \, . \tag{1}$$

The linear coefficients, w_{ij}, give information of the strength of the contribution of every function $\mathbf{B}_{ij}$ in the signal. Each $\mathbf{B}_{ij}$ carries spatial as well as temporal characteristics of the signal.

As these two characterstics are often decoupled, we can assume that the basis functions $\mathbf{B}_{ij}$ are separable (rank-one matrices)

$$\mathbf{B}_{ij} = \mathbf{t}_i \, \mathbf{s}_j^T \, , \tag{2}$$

where the temporal and spatial elementary vectors $\mathbf{t}_i$ and $\mathbf{s}_j$ are the i-th and j-th column of two matrices denoted respectively by $\mathbf{T}$ and $\mathbf{S}$. The only restriction for these matrices is that they must be full rank. The linear expansion (1) can then be written in matrix form as

$$\mathbf{D} = \mathbf{T} \, \mathbf{W} \, \mathbf{S}^T \, , \tag{3}$$

where the $N \times N$ matrix $\mathbf{T}$ contains the temporal information of the basis functions, $\mathbf{W}$ is the linear coefficient matrix formed by w_{ij}, and the $L \times L$ matrix $\mathbf{S}$ contains the spatial information of the basis functions. The particular case of channel-by-channel signal expansion can be obtained from (3) by setting $\mathbf{S} = \mathbf{I}$. Note that even if the spatio-temporal characteristics of the signal are not decoupled, eq. (3) still denotes a valid representation of any signal.

Truncated expansions are usually needed in several applications, such as data compression, feature extraction or filtering and can be interpreted as a restriction of the signal to a given subspace. Truncation in the linear model in eq. (3) is achieved by selecting $p<N$ basis functions from $\mathbf{T}$ and/or $q<L$ basis functions from $\mathbf{S}$ yielding the model $\mathbf{D}=\mathbf{T}_p\,\mathbf{W}\,\mathbf{S}_q^T$, where the columns of $\mathbf{T}_p$ and $\mathbf{S}_q$ are the truncated basis functions. For simplicity, in the rest of the paper, the subscripts p and q will be dropped.

3 GLRT FOR MULTICHANNEL SIGNAL MODEL

The separable linear model for a $N\times L$ multichannel signal $\mathbf{D}$ plus noise can be written as

$$\mathbf{X}=\mathbf{D}+\mathbf{N}=\mathbf{TWS}^T+\mathbf{N} \tag{4}$$

where $\mathbf{T}$ and $\mathbf{S}$ are temporal and spatial truncated basis function matrices with dimensions $N\times p$ and $L\times q$ respectively, and $\mathbf{N}$ is multichannel noise. The signal to detect is assumed to lie within the signal subspace spanned by $\{\mathbf{T},\,\mathbf{S}\}$.

The noise is assumed to be zero-mean Gaussian. Noise is usually considered uncorrelated, but in multichannel biomedical signals, there is a clear correlation of physiological noise (e.g. EMG noise or motion artifacts) between nearly placed electrodes. Accordingly, the case of Gaussian colored noise will be consider here. In general, the second-order moment of multichannel noise is described by a fourth-order tensor. A common simplification is to assume a separable autocovariance function

$$E\left\{n_{ij}\,n_{kl}\right\}=\sigma^2\,c_t\left(i,\,k\right)\,c_s\left(j,\,l\right)\,, \tag{5}$$

where σ^2 is the noise variance, $c_t\left(i,\,k\right)$ denotes the normalized temporal correlation between time instants i and k, and $c_s\left(j,\,l\right)$ the normalized spatial correlation between channels j and l. Then the equivalent $NL\times NL$ autocovariance matrix can be written as

$$\mathbf{C}=E\left\{\mathrm{vec}\left(\mathbf{N}\right)\mathrm{vec}\left(\mathbf{N}\right)^T\right\}=\sigma^2\left(\mathbf{C}_s\otimes\mathbf{C}_t\right) \tag{6}$$

where $\mathrm{vec}\left(\mathbf{N}\right)$ is the $NL\times 1$ column vector with all the samples from the noise matrix $\mathbf{N}$ ordered by columns, $\otimes$ denotes the Kronecker product, $\mathbf{C}_t$ is the $N\times N$ normalized temporal correlation matrix and $\mathbf{C}_s$ is the $L\times L$ normalized spatial correlation matrix.

The hypothesis testing problem is

$$\begin{array}{l}\mathcal{H}_0:\ \mathbf{W}=\mathbf{0}\\ \mathcal{H}_1:\ \mathbf{W}\neq\mathbf{0}\end{array}\,. \tag{7}$$

The GLRT decides between $\mathcal{H}_0$ and $\mathcal{H}_1$ depending on

$$L_G\left(\mathbf{X}\right)=\frac{p\left(\mathbf{X};\,\hat{\mathbf{W}}_1\right)}{p\left(\mathbf{X};\,\mathbf{W}=\mathbf{0}\right)}\underset{<}{\overset{>}{}}\gamma\,, \tag{8}$$

where $\hat{\mathbf{W}}_1$ is the maximum likelihood estimator of $\mathbf{W}$ under $\mathcal{H}_1$.

The p.d.f. of $\mathbf{X}$ can be written as

$$\begin{aligned}&p\left(\mathbf{X};\mathbf{W}\right)=p\left(\mathrm{vec}\left(\mathbf{X};\mathbf{W}\right)\right) \qquad (9)\\ &=\frac{\exp\left(\frac{-1}{2\sigma^2}\left[\mathrm{vec}\left(\mathbf{X}-\mathbf{TWS}^T\right)\right]^T\mathbf{C}^{-1}\left[\mathrm{vec}\left(\mathbf{X}-\mathbf{TWS}^T\right)\right]\right)}{\left(2\pi\sigma^2\right)^{NL/2}\det{}^{1/2}\,\mathbf{C}}\end{aligned}$$

Using the properties of the inverse and determinant of a Kronecker product of matrices,

$$\mathbf{C}^{-1}=\mathbf{C}_s^{-1}\otimes\mathbf{C}_t^{-1}\ \text{ and }\ \det\left(\mathbf{C}\right)=\det\left(\mathbf{C}_s\right)^N\det\left(\mathbf{C}_t\right)^L\,, \tag{10}$$

the p.d.f. (9) can be written as

$$p\left(\mathbf{X};\mathbf{W}\right)=\frac{\exp\left(\frac{-1}{2\sigma^2}\mathrm{tr}\left\{\left(\mathbf{X}-\mathbf{TWS}^T\right)\mathbf{C}_s^{-1}\left(\mathbf{X}-\mathbf{TWS}^T\right)^T\mathbf{C}_t^{-1}\right\}\right)}{\left(2\pi\sigma^2\right)^{NL/2}\det\left(\mathbf{C}_s\right)^{N/2}\,\det\left(\mathbf{C}_t\right)^{L/2}}. \tag{11}$$

The maximum likelihood (ML) estimator of the coefficient matrix under $\mathcal{H}_1$ is obtained as

$$\begin{aligned}\hat{\mathbf{W}}_1&=\arg\max_W\left\{p\left(\mathbf{X};\mathbf{W}\right)\right\}\\ &=\left(\mathbf{T}^T\mathbf{C}_t^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{C}_t^{-1}\mathbf{X}\mathbf{C}_s^{-1}\mathbf{S}\left(\mathbf{S}^T\mathbf{C}_s^{-1}\mathbf{S}\right)^{-1}\,. \qquad (12)\end{aligned}$$

Thus, $\hat{\mathbf{D}}_1=\mathbf{T}\hat{\mathbf{W}}_1\mathbf{S}^T$ is the ML estimator of the signal under $\mathcal{H}_1$, and it is an oblique projection of $\mathbf{X}$ into the subspace defined by $\mathbf{T}$ and $\mathbf{S}$. But using the metric defined by the weighted inner product

$$\left\langle\mathbf{A},\mathbf{B}\right\rangle_{\mathbf{C}_s,\mathbf{C}_t}=\mathrm{tr}\left\{\mathbf{A}\,\mathbf{C}_s^{-1}\mathbf{B}^T\,\mathbf{C}_t^{-1}\right\}\,, \tag{13}$$

and its induced norm, $\hat{\mathbf{D}}_1$ is the orthogonal projection of $\mathbf{X}$.

From (8), (11), (12) and (13), the GLRT can be written as

$$\begin{aligned}T(\mathbf{X})\ &=\ 2\log L_G\left(\mathbf{X}\right)\\ &=\frac{\left\langle\mathbf{X},\mathbf{X}^T\right\rangle_{\mathbf{C}_s,\mathbf{C}_t}-\left\langle\mathbf{X}-\hat{\mathbf{D}}_1\,,\,\mathbf{X}-\hat{\mathbf{D}}_1\right\rangle_{\mathbf{C}_s,\mathbf{C}_t}}{\sigma^2}\\ &=\frac{\left\langle\hat{\mathbf{D}}_1,\mathbf{X}\right\rangle_{\mathbf{C}_s,\mathbf{C}_t}}{\sigma^2}=\frac{\left\langle\hat{\mathbf{D}}_1,\hat{\mathbf{D}}_1\right\rangle_{\mathbf{C}_s,\mathbf{C}_t}}{\sigma^2}\,. \qquad (14)\end{aligned}$$

Two alternative interpretations can be given of the test statistic in (14): an estimator-correlator detector, but using the metric defined in (13) (see Fig. 1), and the ratio of the norm of the projected signal on the subspace spanned by $\{\mathbf{T},\,\mathbf{S}\}$ to the mean noise power. Table 1 gives the expression of the GLRT statistic $T(\mathbf{X})$ for some particular cases, including uncorrelated/correlated noise and unitary/non-unitary transformations.

An alternative interpretation of the detector arises if $\mathbf{C}_t^{-1}$ and $\mathbf{C}_s^{-1}$ are decomposed as:

$$\mathbf{C}_t^{-1}=\mathbf{H}_t^T\mathbf{H}_t\quad\mathbf{C}_s^{-1}=\mathbf{H}_s^T\mathbf{H}_s\,. \tag{15}$$

A modified observed signal $\mathbf{X}'$ can be defined using $\mathbf{H}_t$ and $\mathbf{H}_s$ as filtering matrices

$$\mathbf{X}'=\mathbf{H}_t\,\mathbf{X}\,\mathbf{H}_s^T=\mathbf{T}'\mathbf{W}\mathbf{S}'^T+\mathbf{N}'\,, \tag{16}$$

	Unitary $\mathbf{T}$ and $\mathbf{S}$	Non-unitary $\mathbf{T}$ and $\mathbf{S}$
Colored Noise	$T(\mathbf{X}) = \dfrac{\mathrm{tr}\left\{\mathbf{T}\left(\mathbf{T}^T\mathbf{C}_t^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{C}_t^{-1}\mathbf{X}\,\mathbf{C}_s^{-1}\mathbf{S}\left(\mathbf{S}^T\mathbf{C}_s^{-1}\mathbf{S}\right)^{-1}\mathbf{S}^T\mathbf{C}_s^{-1}\mathbf{X}^T\mathbf{C}_t^{-1}\right\}}{\sigma^2}$	
Uncorrelated Noise	$T(\mathbf{X}) = \dfrac{\mathrm{tr}\left\{\mathbf{T}\mathbf{T}^T\mathbf{X}\,\mathbf{S}\mathbf{S}^T\mathbf{X}^T\right\}}{\sigma^2}$	$T(\mathbf{X}) = \dfrac{\mathrm{tr}\left\{\mathbf{T}\left(\mathbf{T}^T\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{X}\,\mathbf{S}\left(\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{S}^T\mathbf{X}^T\right\}}{\sigma^2}$

Table 1: GLR Tests for different characteristics of the noise and the elementary matrices.

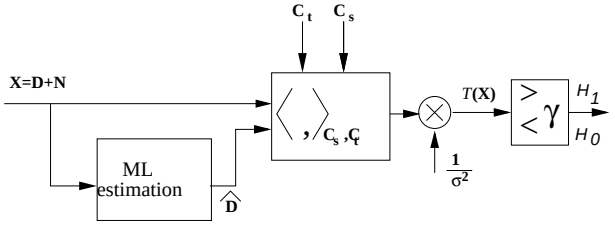

Figure 1: Block diagram of the GLRT as an estimator-correlator detector.

where $\mathbf{T}' = \mathbf{H}_t\mathbf{T}$ and $\mathbf{S}' = \mathbf{H}_s\mathbf{S}$. It is easy to show that $\mathbf{N}' = \mathbf{H}_t\,\mathbf{N}\,\mathbf{H}_s^T$ is uncorrelated ($E\{n'_{ik}\,n'_{jl}\} = \sigma^2\,\delta_{ij}\,\delta_{kl}$). After noise prewhitening, the detector structure can be notably simplified because the noise is uncorrelated (see Table 1). The related realization is shown in Fig. 2.

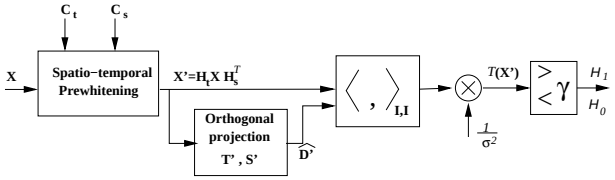

Figure 2: Block diagram of the GLRT as an estimator-correlator detector of prewhitened data.

The test statistic from (14) for the prewhitened data is

$$T\left(\mathbf{X}'\right) = \frac{\left\langle \widehat{\mathbf{D}'_1}, \mathbf{X}' \right\rangle_{\mathbf{I},\mathbf{I}}}{\sigma^2} = \frac{\left\langle \widehat{\mathbf{D}'_1}, \widehat{\mathbf{D}'_1} \right\rangle_{\mathbf{I},\mathbf{I}}}{\sigma^2}\,, \qquad (17)$$

which can also be interpreted in two different ways: as an estimator-correlator detector for whitened data and the ratio between the energy of the projected signal and the mean noise power.

Using (13), (15) and (16) it can be shown that both test statistics (14) and (17) are identical

$$\begin{aligned} T\left(\mathbf{X}\right) &= \frac{\mathrm{tr}\left\{\hat{\mathbf{D}}_1\mathbf{C}_s^{-1}\mathbf{X}^T\mathbf{C}_t^{-1}\right\}}{\sigma^2} \\ &= \frac{\mathrm{tr}\left\{\mathbf{H}_t\hat{\mathbf{D}}_1\mathbf{H}_s^T\mathbf{H}_s\mathbf{X}^T\mathbf{H}_t^T\right\}}{\sigma^2} = T\left(\mathbf{X}'\right)\,. \end{aligned} \qquad (18)$$

4 DETECTION PERFORMANCE

Given a certain signal $\mathbf{D}$, the p.d.f. of the test statistic $T(\mathbf{X})$ under assumption $\mathcal{H}_0$, $T\left(\mathbf{X};\,\mathcal{H}_0\right)$, is a χ^2_{pq} distribution since $T(\mathbf{X})$ can be expressed as the summation of $p{\times}q$ independent normalized zero-mean Gaussian random variables. In a similar way, the p.d.f. of $T\left(\mathbf{X};\,\mathcal{H}_1\right)$ is that of a $\chi'^2_{pq}\left(\lambda\right)$ distribution where the centrality parameter is

$$\lambda = \frac{\langle \mathbf{D}, \mathbf{D}\rangle_{\mathbf{C}_s,\mathbf{C}_t}}{\sigma^2} = \frac{\langle \mathbf{D}', \mathbf{D}'\rangle_{\mathbf{I},\mathbf{I}}}{\sigma^2}\,. \qquad (19)$$

From the distributions of the GLRT under both hypotheses, it is straightforward to obtain the detection and false alarm probabilities (P_D and P_{FA}) for different threshold values γ, i.e. the Receiver Operating Characteristic (ROC) curve.

An often used criterion for the evaluation of a detector is the distance measure between distributions

$$d = \frac{\left|E\left\{T(\mathbf{X});\,\mathcal{H}_1\right\} - E\left\{T(\mathbf{X});\,\mathcal{H}_0\right\}\right|}{\sqrt{0.5\left[\mathrm{var}\left(T(\mathbf{X});\,\mathcal{H}_0\right) + \mathrm{var}\left(T(\mathbf{X});\,\mathcal{H}_1\right)\right]}} \qquad (20)$$

which can be calculated using the properties of the previously given p.d.f.'s

$$d = \frac{p\,q + \lambda - p\,q}{\sqrt{0.5\left[2\,p\,q + 2\,p\,q\, + 4\,\lambda\right]}} = \frac{\lambda}{\sqrt{2\,p\,q\, + 2\,\lambda}}\,. \qquad (21)$$

For a given subspace dimensionality, $\{p, q\}$, the performance depends on λ which is the energy to noise ratio once the signal has been prewhitened. Note that signals with equal SNR may have different energy after prewhitening and thus different detection performance.

5 SIMULATION SETUP

To support the above theory and to evaluate the performance of the spatio-temporal subspace detector we used a Monte Carlo simulation. The signals used in the simulation were 9-channel 1000 Hz sampled ECG signals (specifically, QRS complexes). The spatio-temporal separable subspaces were obtained calculating the temporal and spatial Karhunen-Loève Transform (KLT) bases from a total of 37000 heartbeats from 114 patients. The KLT has the property of optimal energy packing. Thus, it seems a reasonable choice for subspace detection. The signal subspace $\{\mathbf{T}, \mathbf{S}\}$, was chosen by selecting the first $p = 5$ and $q = 2$ KL basis functions.

Clean signals matching the model were simulated considering the elements of $\mathbf{W}$, w_{ij} as independent Gaussian random variables with zero-mean and variance $\sigma_{w_{ij}} = \lambda_i\pi_j$, being λ_i and π_j respectively, the eigenvalues associated with $\mathbf{t}_i$ and $\mathbf{s}_j$. For each realization of $\mathbf{W}$, a signal $\mathbf{D} = \mathbf{TWS}^T$ was computed.

Spatial and temporal noise autocovariance matrices $\mathbf{C}_s$ and $\mathbf{C}_t$ were estimated from isoelectric intervals of the database. A colored noise realization was computed as $\mathbf{N} = \mathbf{H}_t^{-1}\mathbf{V}{\mathbf{H}_s^{-1}}^T$, where $\mathbf{V}$ is uncorrelated Gaussian noise.

Four different detectors were compared: matched filter (MFD), as a clairvoyant detector, spatio-temporal subspace detector (ST-SD) in (14), channel-by-channel temporal subspace detector (T-SD) and energy detector (ED). The last two detectors can be understood as particular cases of the ST-SD, making $\mathbf{S} = \mathbf{I}$ for T-SD, and $\mathbf{S} = \mathbf{I}$, $\mathbf{T} = \mathbf{I}$ for ED.

6 RESULTS

First, a Monte Carlo simulation analysis was performed using a simulated signal $\mathbf{D}$ with 3000 noise realizations. The computed ROC curves for different SNR and methods are shown in Figure 3.

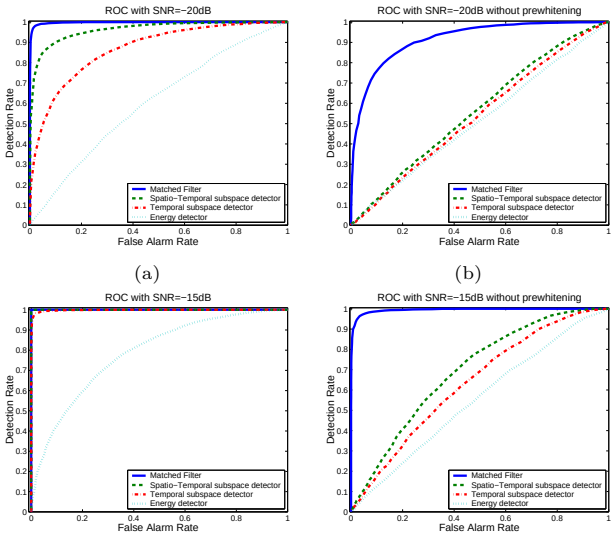

Figure 3: ROC curves for different detectors.

We appreciate in Figure 3(a) that, for SNR=-20 dB, the ST-SD clearly outperforms the T-SD, and is near the upper bound defined by the non-realizable MFD . For SNR =-15 dB in pannel (c), all detectors improve their performance. On the other hand, comparing pannels (b) and (d) with (a) and (c), a clear degradation is found for all detectors when they are applied assuming that the noise is uncorrelated (without any previous whitening). These results highlight the relevance of including the noise autocovariance to achieve a good detection performance.

Figure 4 shows the theoretical and simulated distance measure d for the studied detectors. The experimental results fit clearly the theoretical ones. When SNR is high enough, the performance of ST-SD, T-SD and ED converge, being otherwise always poorer than the unreachable MFD.

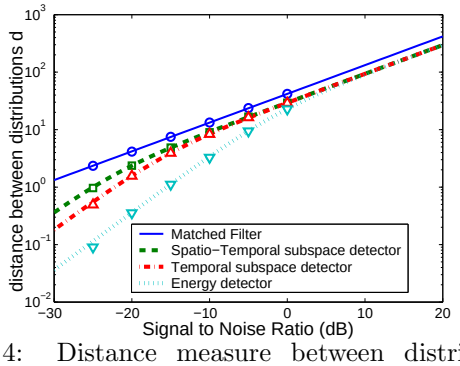

Figure 4: Distance measure between distributions against SNR.

A second simulation considered 50 different signal morphologies in order to independize the results from a particular waveform. 500 noise realizations were added to each morphology. The ROC curves computed from the total of simulations are given in Figure 5. They show a similar behaviour as the ones in Figure 3.

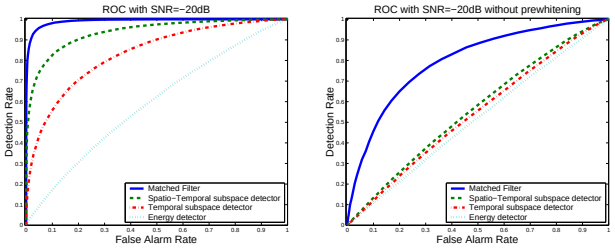

Figure 5: ROC using 50 different waveforms.

7 CONCLUSIONS

A subspace detector has been analyzed for multichannel signals in noise with known temporal and spatial covariance matrices. It has been shown that for low SNR, it outperforms the classical channel-by-channel temporal detector, nearly achieving the clairvoyant matched filter, which cannot be used in some applications with a wide range of waveforms.

When the noise is correlated and the detector assumes that it is uncorrelated, the detection performance degrades notably making the detectors useless. In real applications, covariance matrices are unknown and need to be estimated. The effect of the estimation error in detection performance should be evaluated in a further work.

Selection of the subspace and its dimensionality $\{p, q\}$ were not analyzed in this paper. Future works should consider the effect of model mismatching.

References

[1] S. Kay, *Fundamentals of statistical signal processing*, vol. II-Detection theory. Prentice-Hall, 1998.

[2] L. L. Scharf, "Matched subspace detectors," *IEEE Trans. Sig. Proc.*, vol. 42, pp. 2146–2157, August 1994.