Non-Stationary Time-Series Segmentation Based on the Schur Prediction Error Analysis

Maciej Lopatka[1,2], *IEEE Student Member,* Christophe Laplanche[1], Olivier Adam[1],
Jean-François Motsch[1] and Jan Zarzycki[2]

[1)] LISSI - iSnS, Université Paris XII, 61 av de Gaulle, 94010 Créteil France
phone: (33) 1 45171774, fax: (33) 1 45171754
[2)] Signal Theory Section, Institute of Telecommunications, Teleinformatics and Acoustics,
Wroclaw University of Technology,
Build. C5 - 407 W.Wyspianskiego 27, 50370 Wroclaw Poland
phone: (48) 71 320 30 47, fax: (48) 71 320 31 89

Contact: maciej.lopatka@ieee.org

Abstract—this paper proposes a non-stationary time-series segmentation method based on the analysis of the forward prediction error issued from the adaptive Schur orthogonal signal parameterisation. There is no a priori information about the analysed signal thus this method can be easily adapted to a large family of different types of signals for which two different stochastic processes are present. In this paper we set out some of the advantages of the adaptive Schur filter in deducing the presence of different non-stationary transient or long-term events leading to the signal segmentation. For each sample, the adaptive Schur algorithm calculates the optimal second-order solution for the signal prediction resulting in a set of time-varying model parameters (inter alia forward prediction error). We define the likelihood ratio (LR) test based on the Schur forward prediction error that is evaluated at each sample, thus giving excellent time-reaction properties. The LR test allows us to effectively partition the analysed time-series into homogeneous segments by considering its second-order statistics which are tracked adaptively by the Schur filter. The results performed by applying the proposed method to simulated signals are shown to verify its high performance.

Index Terms— Detection, Segmentation, Non-Stationary Stochastic Sequences, Optimal Adaptive Schur Parameterisation

I. INTRODUCTION

SEGMENTATION of a time-series is of major interest in many fields, notably the data mining problem, speech recognition, automatic autonomous underwater tracking systems such as those used in water for animals' or other man-made targets' detection, identification and localisation. The problem consists of splitting a non-stationary time-series into segments within which the statistics do not change. In other words, this is a problem of identifying the exact instants at which the statistics of the observed signal change. Recently published papers propose the generalised SVD approach [1] or the evolutionary algorithm using pattern matching [2] to partition the analysed time-series into homogenous regions.

For the purpose of signal segmentation, we derived a new effective and robust method which is based on the adaptive Schur orthogonal parameterisation algorithm. The characteristics of the Schur approach i.e.: the excellent convergence behaviour, extremely fast start-up performance and the capability to quickly track parameters' changes make this method extremely well adapted to the signal segmentation task.

II. ADAPTIVE SCHUR FILTER

An extended description of the Schur approach is presented in [3][4][5]. Here, we present succinctly the most important characteristics of this algorithm, in signal processing literature known also as the innovations filter. The Schur filter - an optimal orthogonal filter calculates recursively at every time-step the solution for the signal prediction following from the orthogonal projection theorem (the least-square approach). If the considered time-series is stationary, the solution is time-invariant; in other cases the solution is time-dependent [6].

The ladder-form realisation of the Schur filter as applicable to the time-series is presented in fig.1 [7]. The filter consists of P equal sections where P denotes the order of the model. Each section is described by a set of double-recursive equations 1-3 (in time and in order), as follows:

$$\overbrace{\rho(n+1,t)=\rho(n+1,t-1)\left(1-e^{2}(n,t)\right)^{\frac{1}{2}}\left(1-r^{2}(n,t-1)\right)^{\frac{1}{2}}-e(n,t)r(n,t-1)}^{1}$$

$$\overbrace{e(n+1,t)=\left(1-\rho^{2}(n+1,t)\right)^{-\frac{1}{2}}\left(1-r^{2}(n,t-1)\right)^{-\frac{1}{2}}\left[e(n,t)+\rho(n+1,t)r(n,t-1)\right]}^{2}$$

$$\overbrace{r(n+1,t)=\left(1-\rho^{2}(n+1,t)\right)^{-\frac{1}{2}}\left(1-e^{2}(n,t)\right)^{-\frac{1}{2}}\left[\rho\ (n+1,t)e(n,t)+r(n,t-1)\right]}^{3}$$

where $\rho(n+1,t)$, $e(n+1,t)$ and $r(n+1,t)$ denote respectively, the reflection coefficient, the forward prediction error and the backward prediction error on the *(n+1)th* section at the time *t*.

If one inputs on the innovations filter a second-order signal, at the output one obtains a second-order white noise resulting in a set of time-varying model parameters called reflection or Schur coefficients that describe entirely the second-order random sequence (if the filter order P is sufficiently high).

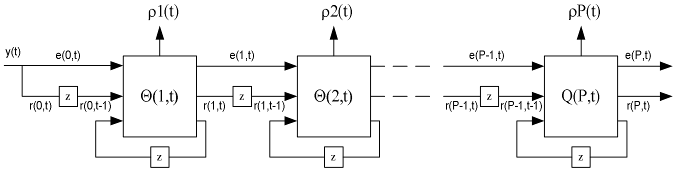

Fig.1 The ladder-form realisation of the adaptive Schur filter

In this paper we present a linear version of the Schur approach dedicated to second-order signals that can be easily implemented in a real-time system. In [8] the author presents the complete survey of the non-linear approach for the non-Gaussian stochastic signals which is much more complicated, and thus time-consuming. However, the same author in [9] proposes the staircase solution for complexity reduction of the non-linear approach.

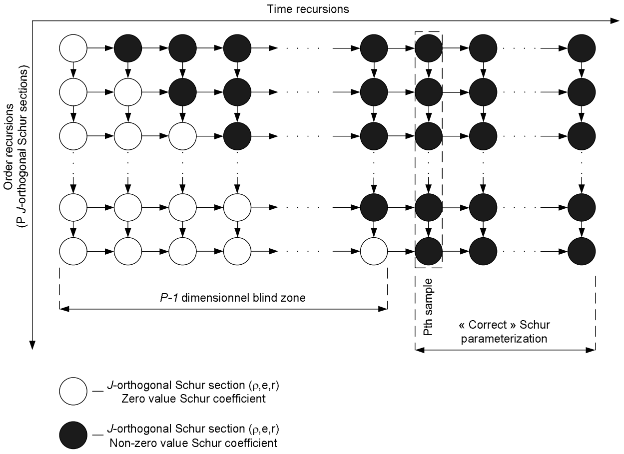

Fig.2 The scheme of time and order recursions for the adaptive Schur filter

The number of filter sections P is a function of the type of the analysed signal. The estimation quality depends on the decreasing rapidity of covariance function $c(k)$ where k denotes a time-lag:

- if the rapidity is high, estimation quality adjusts slowly and as a result we need to use the Schur filter of high order *P*,
- if the rapidity is low, estimation quality adjusts quickly – lower *P* order filters.

Evidently, for the white noise, $c(k)=0; k=1,2...$ presupposes zero quality improvement, which means that white noise prediction is futile.

The efficient tracking capabilities of the Schur algorithm presented in this paper are a consequence of continuously updated quantities (i.e. in our application: the forward prediction errors) being the exact solutions to the least-square problem for each sample in a time-series. The discussed algorithm is recursive both in time and order (fig.2).
The initialisation process of the adaptive Schur algorithm takes *P-1* samples, and this region we qualify as a blind zone.

In this paper we demonstrate the importance of the adaptive Schur orthogonal parameterisation to the segmentation of non-stationary events in a time-series.

III. NON-STATIONARY EVENTS DETECTION AND SEGMENTATION

The signal segmentation problem can be formulated as a problem of detection and localisation. By this we mean the detection and localisation of all the non-stationary events of which the signal is composed. Therefore, in practise the segmentation problem can be viewed as the detection and localisation of the signal's statistics changes (meaning the appearance or disappearance of non-stationary event). In our application we do not make any *a priori* presumptions about the arrival time of different events (mapped to new segments) and their frequency of appearing, thus we propose to use the Neyman-Pearson criterion in the evaluation of the detection threshold (like it is in radar applications). Moreover, we suppose that their amplitude and particular characteristics are not known, neither. In this paper we are considering two different stochastic processes of which the analysed signal is composed and the objective is to localise their beginning and end points.

The general problem of signal detection can be formulated as a simple exclusive binary test between null hypothesis H_0 and alternative hypothesis H_1 :

$$\begin{aligned} H_0 &: \underline{x}=\underline{n} \\ H_1 &: \underline{x}=\underline{n}+\underline{s} \end{aligned} \tag{4}$$

where x is an observation process (evidence variable) which is composed of noise alone (under hypothesis H_0) or noise and information signal (under hypothesis H_1). In our application this test can be defined as follows: the null hypothesis denotes no change in the signal structure (in the sense of its second-order statistics) and the alternative hypothesis denotes the change in the signal's statistics. Thus, by filtering the signal with the Schur orthogonal filter we partition the same the signal into segments.

We use the adaptive Schur algorithm as a transformation from one dimensional time space to P dimensional Schur model parameters' space. We define the likelihood ratio (LR) test based on the Schur forward prediction error as follows:

$$LR = \sum_{k=1}^{k=P} \left(\frac{\partial e_k}{\partial t} \right)^4 \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5}$$

where P is a filter order (number of the filter sections), e_k denotes the forward prediction error on k-th section and η denotes comparison threshold. The choice of the LR test as in (5) follows directly from the principle of the algorithm. At the moment when the non-stationary transient (new signal event) appears the margin for the forward prediction error increases. We sum the errors from all the filter sections to emphasize the effect. The value of the power in (5) was chosen empirically to obtain optimal performances of the detection process i.e. diminish the value of the LR test (make it close to zero) when signal is supposed to be "near stationary" and augment the LR value when the non-stationarities appear. By tracking changes in the LR value as a function of time, one is able to partition the signal with respect to its second-order statistics. At instants when the value of the LR test exceeds the value of the threshold one obtains a new segment of the signal.

The segmentation algorithm at each time-step performs the LR test. The value of $\underline{\eta} = [\eta_1, \eta_2, \eta_3, ...,]^T$ is calculated at each sample with respect to the Neyman-Pearson criterion for *a priori* defined value of the accepted false alarm rate and regarding the signal characteristics estimated adaptively by the Schur filter.

The formulation of the segmentation process consists of 3 steps (for *N*-length signal):

1. Filtering of the signal: calculation of P time-varying forward prediction error vectors; here, we introduce it as the forward prediction error matrix E (filter order increases from top to bottom in columns and time increases from left to right in rows):

$$E_{PxN} = \begin{bmatrix} e(1,1) & e(1,2) & \cdots & e(1,\mathrm{N}) \\ e(2,1) & e(2,2) & \cdots & e(2,\mathrm{N}) \\ \vdots & \vdots & \ddots & \vdots \\ e(\mathrm{P},1) & e(\mathrm{P},2) & \cdots & e(\mathrm{P},\mathrm{N}) \end{bmatrix}$$

2. Continuous estimation of the value of the threshold $\underline{\eta} = [\eta_1, \eta_2, ..., \eta_N]^T$;
3. Performing the Likelihood Ratio (LR) test every sample:
    - If $LR_i > \eta_i$ then there is a new segment
    - If $LR_i < \eta_i$ then there is the same segment,

    where *i* denotes the present signal sample.

At every sample in the time-series the adaptive Schur filter minimizes the error for the signal prediction. If the filter order is sufficient for the orthogonal parameterisation of the stochastic sequence than the resulted forward prediction error is close to zero. The algorithm tracks the signal *via* Schur coefficients describing the signal entirely. Each second-order signal's change influences the values of the model parameters: the reflection coefficients and the prediction errors. During periods where the signal is stationary or *quasi* stationary, the value of the forward prediction error is near zero and reflection coefficients have *quasi* constant values. However, in cases where there are changes in the signal structure (in the sense of its statistics) the margin for the prediction error increases and the reflection coefficients change their values. Therefore, tracking these two model parameters allows us to deduce the presence of new segments in the filtered signal.

We propose an adaptive estimation of the threshold η, to obtain the optimal performances of the Schur detector in the task of deducing the exact transition points of the signal in the second-order sense. The η estimation is calculated *via* the Neyman-Pearson criterion considering evidently signal characteristics which are estimated adaptively by the Schur filter. The LR test is performed at each-time step which gives very good time-reaction properties - the response of the detector to the presence of a new signal event is immediate (involving exact localisation of the new segment).

This algorithm may be simply applied in a real-time signal detection/segmentation system.

IV. RESULTS AND DISCUSSION

We illustrate our approach with simulated signals. In our analyses, several detection/segmentation simulations, each of 500000 Monte-Carlo runs was performed. The simulated signal consists of two different non-stationary stochastic processes. These processes are mixed that the signal is composed of only one or both of these processes at the same time (one of considered processes is embedded in the second one). The analyzed processes being non-stationary events are in fact Gaussian random processes filtered with a pre-defined B bandpass filter (Butterworth filter of the 12th order). By changing the frequency band of the process ($B \in \{B_1, B_2\}$), we simulate new signal segments. The exemplary realisation of the considered signal is presented in fig.3.

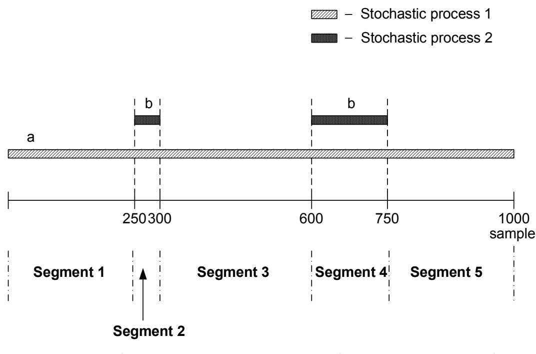

Fig.3 Time-domain structure of the investigated signal (example realisation)

The time-domain segmentation of the signal from the fig.3 is presented in fig.4. The analysed signal x consists of 1000 samples and is composed of 5 different segments (2 different random processes). The structure of the signal is as follows: $x_{1x1000} = [a_{1x250}, (a+b)_{1x50}, a_{1x300}, (a+b)_{1x150}, a_{1x250}]$ where a and b denote the two different Gaussian random processes of B_1 and B_2 bandwidths ($B_1 \neq B_2$) respectively.

We demonstrate the performances of our approach by introducing the Receiver-Operating Characteristics curves (fig.5). These curves describe the performances of the detection process (binary test) of the signal statistics changes. This is valid because our segmentation algorithm is based on the proposed LR test which indicates the presence of a new signal segment by deducing the signal statistics' change. Each ROC curve is described by the SNR value. This value defines the relation between two random processes to be split up (for $SNR = 0dB$ the 2-components signal is composed of two energy-equal processes; for $SNR < 0$ the second process is embedded in the first one with underlying energy ratio). In practice the idea is to detect the presence of the other random process in analyzed signal and localize it.

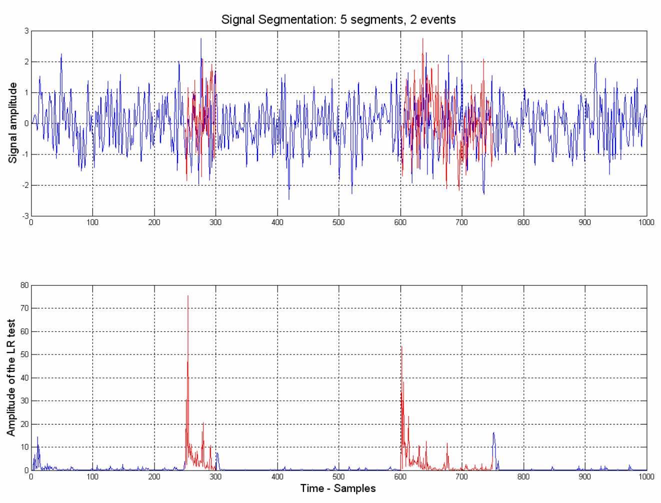

Fig.4 Example of the signal segmentation based on the adaptive Schur algorithm

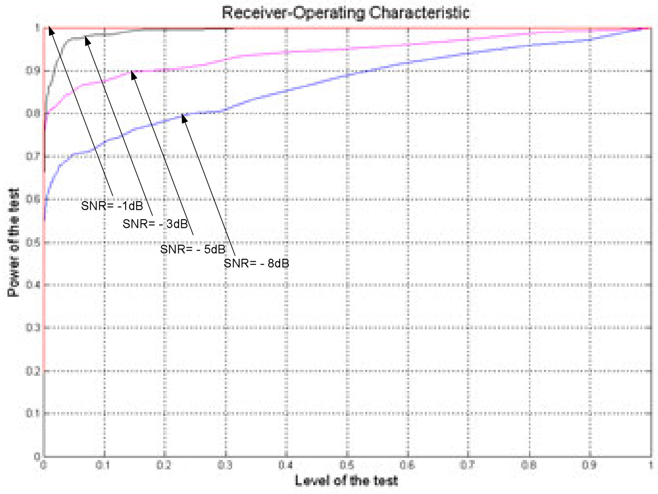

Fig.5 Receiver-Operating Characteristic of the adaptive Schur filter based detector

Each ROC curve was generated using 500 simulations and thereby we obtained a vector of 500 different SNR values. The mean value of this vector was taken into account in further analysis as the SNR value describing the particular ROC curves.

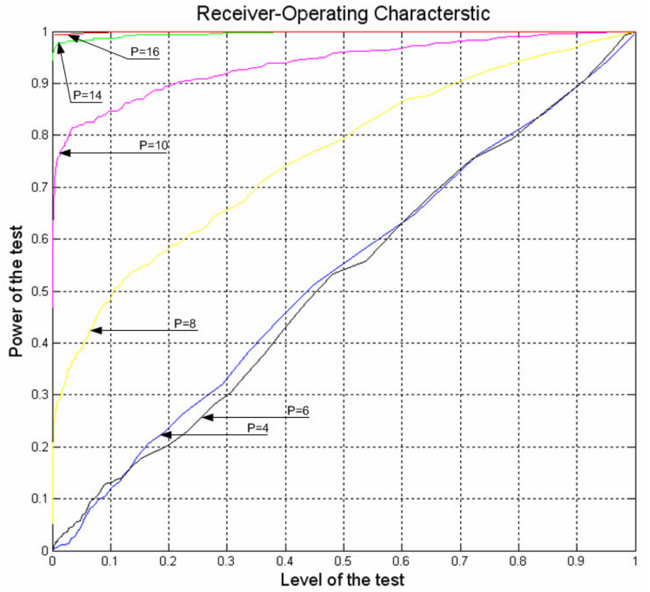

Fig.6 Exemplary Receiver-Operating Characteristics of the adaptive Schur filter based detector for different values of the filter order P and SNR= -13dB

In fig.6 we illustrated the role of the filter order P for the performances of the detector. We tested the influence of the filter order P for the detection process during the simulation with $SNR = -13dB$ (this is an exemplary value chosen to highlight the importance of the choice of the filter order for the performances of the detector, especially in low SNR conditions). The presented example shows that the number of the filter sections P has a great importance for the resulting detection scores. The filter order should be chosen as a function of the signal and noise types. If one wants to improve the performance of the detector, then one should increase the filter order (overestimating it). However, the drawback is the rise of the requisite calculation time (increased computational complexity of the algorithm).

V. CONCLUSION

In this paper, we presented arguments supporting the importance of the adaptive Schur algorithm in the two components signal segmentation task. The parameters of the model i.e. the Schur forward prediction errors calculated at each sample allow us to segment different parts in the analysed time-series with respect to the second-order statistics. We presented a short description of the Schur approach and discussed the influence of the order of the Schur filter as well as the influence of the signal's "conditions" on the segmentation performance ratio. The presented results we obtained on simulated signals in the segmentation problem task are promising for prospective real-world applications. The purpose of our research is justified by the fact that our laboratories work on the detection, identification, and real-time localisation of whales which spend most of the time under water. The objective is to propose the passive tracking system to detect, identify, and track certain species of marine mammals. The proposed approach for the segmentation of the time-series based on the analysis/tracking of the Schur model parameters is particularly important in that work.

REFERENCES

[1] Yadunandana N. Rao, Jose C. Principe, "Time Series Segmentation Using a Novel Adaptive Eigendecomposition Algorithm", Journal of VLSI Signal Processing vol. 32, pp. 7-17, 2002

[2] Fu-lai Chung, Tak-chung Fu, Robert Luk, Vincent Ng, “Evolutionary Time Series Segmentation for Stock Data Mining”, IEEE International Conference on Data Mining ICDM’02, Maebashi City, Japan, 2002

[3] I. Schur,”On Power Series Which Are Bounded in the Interior of the Unit Circle I”, in *I. Schur Methods in Operator Theory and Signal Processing*, Ed. I. Gohberg, Operator Theory: Advances and Applications, vol. 18, Birkhäuser-Verlag, pp. 31-60, 1986

[4] I. Schur,”On Power Series Which Are Bounded in the Interior of the Unit Circle II”, in *I. Schur Methods in Operator Theory and Signal Processing*, Ed. I. Gohberg, Operator Theory: Advances and Applications, vol. 18, Birkhäuser-Verlag, pp. 61-88, 1986

[5] T. Kailath, “A Theorem of I. Schur and Its impact on Modern Signal Processing”, in *I. Schur Methods in Operator Theory and Signal Processing*, Ed. I. Gohberg, Operator Theory: Advances and Applications, vol. 18, Birkhäuser-Verlag, pp. 9-30, 1986

[6] T. Kailath, “Time-Variant and Time-Invariant Lattive-Filters for Nonstationary Processes”, in *Outils et modèles mathématiques pour l’automatique, l’analyse de systèmes et le traitement du signal*, ed. CNRS, Paris, pp. 417-464, 1982

[7] D.T.L. Lee, M. Morf, B. Friedlander, “Recursive Least Squares Ladder Estimation Algorithms”, *IEEE Transactions on Circuits and Systems*, Vol. CAS-28, No. 6, June 1981

[8] J. Zarzycki, “Multidimensional Nonlinear Schur Parameterization of NonGaussian Stochastic Signals. Part Two: Generalised Schur Algorithm”, *MDSSP Journal*, 15, 243-275, 2004

[9] J. Zarzycki, “Multidimensional Nonlinear Schur Parameterization of NonGaussian Stochastic Signals. Part Three: Low-Complexity Staircase Solution”, *MDSSP Journal*, 15, 313-340, 2004