International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:2, No:12, 2008

Change Detection and Non Stationary Signals Tracking by Adaptive Filtering

Mounira Rouaïnia, and Noureddine Doghmane

***Abstract*—**In this paper we consider the problem of change detection and non stationary signals tracking. Using parametric estimation of signals based on least square lattice adaptive filters we consider for change detection statistical parametric methods using likelihood ratio and hypothesis tests. In order to track signals dynamics, we introduce a compensation procedure in the adaptive estimation. This will improve the adaptive estimation performances and fasten it's convergence after changes detection.

***Keywords*—**Change detection, Hypothesis test, likelihood ratio least square lattice adaptive filters.

I. Introduction

CHANGE detection constitutes a crucial problem while analyzing non stationary signals. This is especially the case in adaptive filtering, biomedical diagnosis and industrial process control [2,3,5,7]. Our contribution in this domain is the establishment of complementary tools for non stationary signal adaptive processing. The resulting algorithms characterize eventual changes in processed dynamic signals and improve their tracking by the adaptive estimation.

The change detection methods we consider here are based on the analysis of the prediction error of the signal. In deed this parameter conveys important information on the real time estimation state. A signal variation causes nearly instantaneous changes in this parameter statistics. For this and because of the incertitude on the existence or not of changes, the probabilistic environment is imperative. Thus, the change detection remains to make the following hypothesis test:

Hypothesis H_0: There is no change

Hypothesis H_1: there is a change

Detection methods we consider are then based on statistical notions such as likelihood ratio and confidence interval. Two types of methods are addressed; the first is based on conditional probability function of the signal and on likelihood ratio. In the second, we construct Information Signals (IS) that must reflect changes information, the detection will be based then on confidence intervals established from these IS a statistical distribution.

Manuscript received September 30, 2006.

Mounira Rouaïnia is with The Skikdq university PB 26 Route El Hadaek Skikda 21000, Algeria (e-mail: rouainia_m@yahoo.fr).

Nouredine Doghmane is with Annaba university BP 12 Annaba, Algeria.

II. Signal Modelization

To estimate signals parameters we consider the autoregressive model. In deed, this model has proved efficiency for the change detection even if it doesn't really describe the studied signal. So the discrete process is described as follows:

$$y(k)= \sum_{i=1}^{p} a_i y(k-i) + e(k) \tag{1}$$

where a_i : signal parameter of i order

p: model order

(e(k)): white noise

Parameters estimation is made using adaptive lattice least square filter [1]. This numerically robust filter gives us reliable parameters estimation especially for the prediction error.

III. Likelihood Ratio based Change Detection

Based on conditional probability function and likelihood ratio [2, 4, 6], this approach generates detection tests by comparing a decisions function to a defined detection threshold.

Consider a discrete random observation y(k) with conditional probability function $P_\theta(y)$, where θ is the the set of signal parameters able to change. We suppose that the model before change is θ_0 and the model after change is θ_1.

The observations probability function conditioned by θ_0 is written $p_{\theta_0}(y)$. The probability function conditioned by θ_1 by $p_{\theta_1}(y)$.

Log likelihood ratio (LLR) of the two functions is:

$$s_i(y) = \mathrm{Ln}\frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)} \tag{2}$$

This parameter has an important role in the detection problem. In deed it's mean before and after change is less than zero and it's mean after change is positive. Thus:

$E[s(y)/H_0] < 0$ et $E[s(y)/H_1] > 0$

The two following detection algorithms are then based on this important property.

A. Cumulated Sums Algorithm (CUMSUM)

The decision function in this classical algorithm is the sum of all the LLR available at each sampling time. It is written as follows:

$$g(k) = \sum_{i=1}^{k} s_i(y) \tag{3}$$

Expression which is equivalent to:

$$g(k) = g(k\text{-}1) + s_k(y) \tag{4}$$

The detection time t_d will be the instant at which the decision function becomes upper than the detection threshold h: $t_d = \mathrm{Min}\{ t/ g(k)>h\}$

Consider for θ the set of the AR parameters vector and the variance σ^2 of the generating white noise ($\theta=[a_1^1, a_1^2, \ldots a_1^p, \sigma_1]^T$).

By writing the LLR in terms of the conditional probabilities $p_{\theta_0}(y)$ and $p_{\theta_1}(y)$, we obtain the following increment :

$$\mathbf{s_i} = \frac{1}{2}\mathrm{Ln}\frac{\sigma_0^2}{\sigma_1^2} - \frac{e_1^2(i)}{2\sigma_1^2} + \frac{e_0^2(i)}{2\sigma_0^2} \tag{5}$$

where $e_l(k)=y(k) - \sum_{i=1}^{p} a_l^i y(k-i) \qquad l=0,1$

The algorithm drives the difference between the normalized squared prediction errors under p_{θ_0} and p_{θ_1} .

B. Kullback Divergence Algorithm (DIVE)

This algorithm is based on a similar decision function to the precedent, but with an increment given by.

$$\tilde{\mathbf{s}}_i = \mathrm{Ln}\frac{p_{\theta_1}(y_i / Y_1^{i-1})}{p_{\theta_0}(y_i / Y_1^{i-1})} - E_{\theta_0}\left[\mathrm{Ln}\frac{p_{\theta_1}\left(y_i/Y_1^{i-1}\right)}{p_{\theta_0}\left(y_i / Y_1^{i-1}\right)}\right] \tag{6}$$

The difference is that the correction of the increment is made by subtraction of the conditional mean before change. This make the test be symmetric. In case of AR and Gaussian process the expression is:

$$\tilde{s}_i(y) = -\frac{e_i^0 e_i^1}{\sigma_1^2} + \frac{1}{2}\left(\frac{\sigma_0^2}{\sigma_1^2}+1\right)\frac{(e_i^0)^2}{\sigma_0^2} + \frac{1}{2}\left(\frac{\sigma_0^2}{\sigma_1^2}-1\right)\text{-v} \tag{7}$$

This algorithm drives the difference between the normalized squared prediction errors under p_{θ_0} and p_{θ_1}

C. Implementation and Application

As the set of parameters is not available, we must estimate it to run the tests. We have to estimate two models: M_0 (before change model and M_1 (after change model).

For M_0 we use the adaptive lattice least square algorithm. At each time all information is used to estimate this global model. And for M_1 we use a sliding window Burg algorithm.

We applied the two algorithms to non stationary signals presenting different kind of changes: abrupt and progressive, occurring in the model and in the signal energy.

Fig. 1 presents the multiple changes case and in Table II we present the different tests results.

We remark that:

-The tests don't require an optimal estimation since the model order changes are detected even if the order becomes upper to the order of the prediction algorithm (Test n° 3).

-The low variations are not detected

-The detection threshold influences the detection speed and the false alarm rate. A too high threshold gives an important detection delay but a small threshold may increase the false alarms risks.

TABLE I
CHANGE DETECTION TESTS

test	Changing parameter	t_c
1	Abrupt model change	500
2	Model order change	500
3	Progressive model change	t_c =350, 850
4	Abrupt change in variance	500
5	Progressive variance change	t_c 350, 850
6	Multiple changes	t_{c1}=500 t_{r2}= 1000

TABLE II
DETECTION RESULTS

test	$t_{d,CUMSUM}$	$t_{d,}$dive
1	535 (h=20)	588 (h=20)
2	No detection	No detection
3	629,914 (h=20)	589,849 (h=20)
4	515 (h=10)	522 (h=10)
5	577 (h=10)	573 (h=10)
6	517,1005 (h=10)	518, 1005 (h=10)

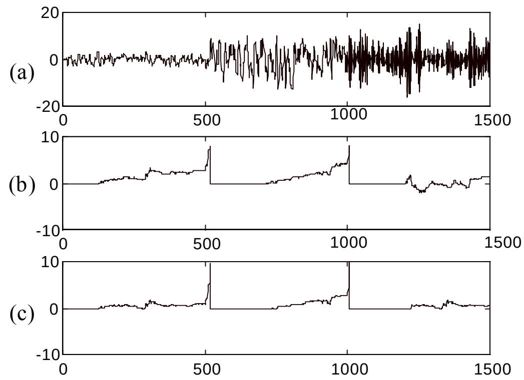

Fig. 1 Multiple changes detection
a: non stationary signal
b: decion function for CUMSOM
c: decision function for DIVE

IV. Confidence Interval based Change Detection

In opposition with the precedent approach, the method we consider in this paragraph is independent from a detection threshold choice. The generated detection tests is function of confidence intervals constructed from the statistical distribution of specific information signals (IS)[7]. These signals must reflect the change information, that's why we consider the prediction error and it's variance as IS. We develop three tests based on χ^2, Fisher and Student distributions. The test will be applied in three steps:

1-Information signal construction.

2-Choice of the confidence degree α necessary parameter for confidence interval construction.

3-Choice of n the freedom degrees of the considered statistical distribution. This parameter is in relation with the length of the quasi stationary slices of the signal [7], [8].

A. χ^2 Distribution based Test

This test is based on the comparison of an estimated value $\hat{\chi}^2$ with two threshold values delimiting the confidence interval [8]. The χ^2 **variable** generated by the prediction error e(k) is given by :

$$\beta_n(k)=\sum_{i=k-n+1}^{k} e_n(k)^2 \tag{8}$$

where $e_n(k)=\dfrac{e(k)}{\sigma(k)} \in N(0,1)$

And the detection test will be:

If $\chi^2_{n,1-\alpha/2} \le \beta_n(k) \le \chi^2_{n,\alpha/2}$:

$\mathbf{H_0}$ true, no change

Else $\mathbf{H_1}$ true, there is a change

B. Fisher Distribution based Test

The Fisher variable is the ratio of two χ^2 variables divided by their respective freedom degree number [8].

By considering the prediction error variances s_1^2 et s_2^2 estimated from two different size samples of the signal, and normalized by their respective freedom degree, we obtain the following Fisher variable:

$$F=\frac{s_1^2}{s_2^2} \tag{9}$$

The detection test will be then:

If $\mathbf{F<F_{n1,n2;\alpha}}$ **and** $\mathbf{F^{-1}<F_{n2,n1;\alpha}}$ **:**

$\mathbf{H_0}$ true, no change

Else $\mathbf{H_1}$ true, there is a change

C. Student Distribution based Test

The Student variable results from the ration of two variables, the first normally distributed and the second is a Fisher variable. By considering the prediction error and its estimated variance we obtain the following Student variable:

$$T=\frac{e(k)-m_e(k)}{\dfrac{\sigma_e(k)}{s_e(k)}} \tag{10}$$

The detection test will be:

If $\mathbf{-T_{n,\alpha/2} \le T(k) \le T_{n,\alpha/2}}$ $\mathbf{H_0}$ true, no change
else $\mathbf{H_1}$ true, there is a change

D. Application and Results Discussion

For the confidence degree α, we choose the value of 5%. This means that the estimation will have a confidence rate of 95%.

Fig. 2 illustrates the three tests applied on the signal (a) of Fig. 1 graphs in full line correspond to Information Signals and doted ones to the confidence interval.

In Table III, we present the detection results for Table I signals we can remark that:

- The change detection is faster than with the precedent approach.
-The false detection rate is significant in some cases (Tests 4 and 3).

But we must say that this false detection problem is resolved if we choose a lower value for the parameter α, a value of 1% for example. But this may cause non detections especially when the change is not important.

TABLE III
χ^2, Fisher and Student. Tests Results

test	$t_{d,\chi 2}$	$t_{d,Fish}$	$t_{d,Stud}$
1	502	501,701	317, 534
2	512	502,702	519,723
3	629,914	589,849	575,997
4	511	503	280, 501
5	524	524,724	317
6	511,1005	512 1075	399,609, 1001

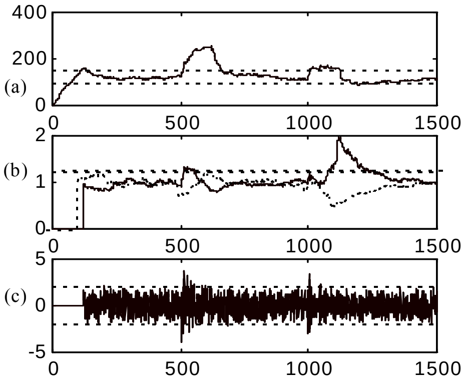

Fig. 2 Multiple changes detection
a: χ^2 test, b: Fisher test, c: Student test

V. CHANGE COMPENSATION PROCEDURE

The delay considered between eventual successive change detections is T= 200 samples which is an important value. This is caused by the long temporization that we consider in order to ensure the global estimation convergence after any change detection.

To minimize this parameter, the solution we propose is to compensate the change in the pertinent estimation algorithm parameters in order to improve its tracking performances.

In lattice filters, the most important parameters that influence the estimation convergence are the partial correlation coefficients (PARCOR), we opt then for them compensation.

As we use two adaptive estimations, a global one and a sliding window based one, we have the idea to compensate directly the global estimation PARCOR (M_0 model). This is done by replacing them by those of the sliding window estimation (M_1 model). There are two possibilities to do that:

1-Compensate when a change is detected

2-Compensate when the distance between the two estimation algorithms PARCORs is higher then a defined threshold s.

We choose the second method because in practice, it produces fast change detections. In addition, the tracking performances of the adaptive estimation are improved. The result we obtain is to reduce the time delay T from 200 to 100.

Fig. 3 presents change detection with and without compensation and using CUMSOM and DIVE methods (Test 1). With compensation we obtain the following detection times:

$t_{d,CUMSOM}$ =504

$t_{d,Dive}$ =505

The detection is really fastened and the decision functions are more smooth which decrease the false detection risks.

Using methods of paragraph 4, we obtain similarly good results. For the example of Test 1, we have:

$t_{d,\chi 2}$ = 502

$t_{d,Fish}$ = 501

$t_{d,Stud}$ = 534

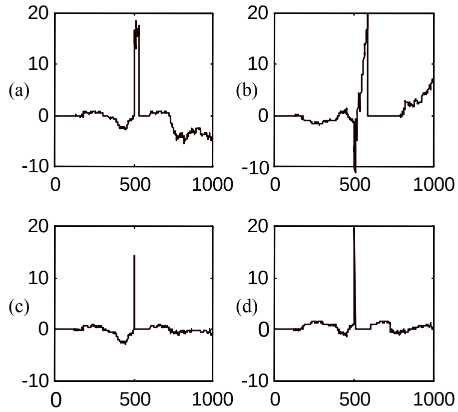

Fig. 3 Change detection with and without compensation
CUMSUM decision functions (a,b), DIVE (c,d)

VI. CONCLUSION

In this work we considered statistical change detection applied with adaptive filtering. We implemented simple and performing tools that combined with adaptive estimation gave satisfying change detection. We also introduce a new compensation step which improved the tracking performances of the adaptive estimation and produced fastener change detection.

REFERENCES

[1] P. Diniz, Adaptive filtering: Algorithms and practical implementation. Kluwer academic pub, 2002.
[2] M.Basseville, I.Nikiforov. Detection of abrupt changes: Theory and application. Prentice Hall, 1997.
[3] M. Basseville. Detecting changes in signals and systems – A survey-. Automatica, V-24, 1988. p 309-325.
[4] I. Nikiforov, Two strategies in the problem of change detection and isolation. IEEE Trans on information theory. V43(2) 1997 p 770-776.
[5] M.Seck, R.Blouet, F.Bimbot. Comparaison de critères de segmentation par détection de rupture sur un signal sonore. 17ème colloque GRETSI sur le traitement du signal et des imagesV4(1)1999 p 989-992.
[6] G. V. Moustakides. Optimal stopping times for detecting changes in distribution. The annals of statistics. V 14.1986 p 1479-1487.
[7] Ramdani. Détection de rupture: outil de diagnostic d'adaptativité dans le cas non stationnaire. ICSS Alger, 1994. p III45-III49.
[8] D. Siegmund, Sequential analysis tests and confidence intervals. Series in statistics, Springer, 1985.