Title	Identification of non-Gaussian parametric model with time-varying coefficients using wavelet basis
Author(s)	Shen, MF; Zhang, YZ; Chan, FHY
Citation	International Conference on Neural Networks and Signal Processing Proceedings, Nanjing, China, 14-17 December 2003, v. 1, p. 659-662
Issued Date	2003
URL	http://hdl.handle.net/10722/46510
Rights	©2003 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

IDENTIFICATION OF NON-GAUSSIAN PARAMETRIC MODEL WITH TIME-VARYING COEFFICIENTS USING WAVELET BASIS

Minfen Shen*,   Yuzheng Zhang*,   Francis H. Y. Chan**

*Science Research Center, Shantou University, Guangdong 515063, China.

Email: mfshen@stu.edu.cn

** Department of Electrical and Electronics, Hong Kong University, Hong Kong, China.

ABSTRACT

Many time series in practice turn to be the time-varying (TV) non-Gaussian processes. In this paper, we address the problem of how to describe these non-stationary non-Gaussian time series. A non-Gaussian AR model with TV parameters is proposed to track the non-stationary non-Gaussian characteristics of the signal. Since wavelet has flexibility in capturing the signal's transient characteristics at different scales, a set of wavelet basis is employed so that the model parameters can effectively track the variations of TV signals and be used to estimate the corresponding TV bispectrum. The experiments results confirm the superior performance of the presented model over the previous method.

1. INTRODUCTION

To identify the systems and estimate the parameters, many of the existing methods exclusively assume that the signal is stationary. The stationarity requires that the underlying statistics and the model parameters that characterize the process are not dependent on time. However, in practice the time-varying characteristics of the systems are so important that they have close relations with physical accommodation. In spite of having been applied in many system, the stationary assumption is often not true for many physical signals such as speech signals, biomedical measurement, seismic signals and so on. To effectively identify the characteristics of the non-stationary signals, several kinds of methods were proposed for this purpose [1-3]. But up to now, there exists not any general mathematical framework for dealing with the time-varying process.

On the other hand, many signals encountered in practice are not only non-stationary, but also non-Gaussian and nonlinear [4,5]. The most common method used for the evaluation of the signals is based on the power spectral density (PSD) function which quantifies the power contents at different frequencies [7]. The PSD, however, suppresses the phase relationship of the frequency components of the signal. Dealing with non-Gaussian and nonlinear random process, higher-order statistics (HOS), like bispectrum, are of paramount importance in advanced statistical signal processing. The motivation behind the use of HOS for the non-Gaussian and nonlinear signals includes the following three aspects: (1) detect and characterize the non-Gaussianility and the nonlinearity of the underlying process, (2) extract the information of the phases relationship among the frequency components of the process, (3) significantly improve the signal-to-noise ratio (SNR) when dealing with the Gaussian additive noise.

As more evidences have mounted to confirm the existence of time-varying non-Gaussian and nonlinear behavior of the process, it is necessary and appropriate to adopt new method to address the non-stationarity, non-Gaussianility and nonlinearoty of the signal. For this purpose, a non-Gaussian autoregressive (AR) model with time-varying coefficients is employed in this contribution. By means of time-variant parametric bispectral analysis, the transient phase-locking relationship among the components of the signal can be identified.

2. THE PROPOSED SCHEME

Let the non-stationary time series $x(n)$ be modeled as the output of a time-varying non-Gaussian linear system which is expressed as the following AR(P) model:

$$x(n) = -\sum_{k=1}^{P} a_k(n)x(n-k) + e(n) \qquad (1)$$

where $e(k)$ denotes a stationary non-Gaussian, independent identically distribution (i.i.d.) input sequence with zero-mean and finite moments. $a_k(n)$ represent the time-varying model parameter to be estimated.

The observed noisy time series is described as

$$y(k) = x(k) + n(k) \qquad (2)$$

where the noise n(k) is zero-mean, additive Gaussian noise which is statistically independent of x(k). Since the lth -order cumulants $C_l(m,n)$ of Gaussian process will vanish for $l \geq 3$, we have the third-order cumulants relation as $C_{3z}(m,n) = C_{3y}(m,n)$. Obviously, the additive Gaussian noise is significantly suppressed due to the HOS. The time-varying coefficients $a_i(t)$ are expressed as wavelet basis functions [6,8]

$$a_i(t) = \sum_k a_{i,k} \phi_{i,k}(t) \qquad (3)$$

Where $\phi_{i,k}(t)$ are a number of orthonormal basis functions. The TVAR coefficients are projected on the space of the basis functions and this projection, symbolized by $a_{i,k}$, does not change with time. Once $\phi_{i,k}(t)$ were decided, we can get x(t) only if we have known the constant coefficients as

$$a_i(t) = \sum_m a_{i,m} \phi_{i,m}(t)$$

$$a_i(t)\phi_{i,k}(t) = \sum_m a_{i,m} \phi_{i,m}(t)\phi_{i,k}(t)$$

where the $\phi_k(t)$ are defined as an orthonormal set

$$\int \phi_k(t)\phi_m(t)dt = \delta(k-m) = \begin{cases} 1 & k=m \\ 0 & k \neq m \end{cases}$$

so we have

$$\int a_i(t)\phi_{i,k}(t)dt = a_{i,k}$$

$$a_{i,k} = \int a_i(t)\phi_{i,k}(t)dt \qquad (4)$$

Thus the TV signal can be obtained

$$x(t) = -\sum_{i=1}^P a_i(t)x(t-i) + e(t)$$
$$= -\sum_{i=1}^P [\sum_k a_{i,k}\phi_{i,k}(t)]x(t-i) + e(t) \qquad (5)$$

or

$$X(t) = -\sum_{i=1}^P A(i).X(t-i) + E(t) \qquad (6)$$

where $X(t),E(t)$ denote the projections of the x(t) and e(t) on the basis function space:

$$X(t) = [\phi_0(t)x(t), \phi_1(t)x(t), \cdots\cdots, \phi_q(t)x(t)]^T$$
$$E(t) = [\phi_0(t)e(t), \phi_1(t)e(t), \cdots\cdots, \phi_q(t)e(t)]^T$$
$$A(k) = [\phi_0(t)a_k, \phi_1(t)a_k, \cdots\cdots, \phi_q(t)a_k]^T$$

where $a_k = [a_{k0}, a_{k1}, \cdots, a_{kq}]^T$. Then the expression (6) can be evolved into a system of q+1 equations:

$$\phi_0(t)x(t) = -\sum_{k=1}^P \phi_0(t)a_k \cdot X(t-k) + \phi_0(t)e(t)$$
$$\phi_1(t)x(t) = -\sum_{k=1}^P \phi_1(t)a_k \cdot X(t-k) + \phi_1(t)e(t)$$
$$\vdots$$
$$\phi_q(t)x(t) = -\sum_{k=1}^P \phi_q(t)a_k \cdot X(t-k) + \phi_q(t)e(t)$$

We have also known that the mth order cumulant sequence of X(t) satisfies the following recursive equation [5]:

$$\sum_{k=0}^P A(k)C_{m,X}(\tau_1, \cdots; \tau_{m-2}, \tau-k) = 0, \quad \tau > 0 \qquad (7)$$

When m=3, (7) can be changed to :

$$\sum_{k=1}^P \begin{bmatrix} \phi_0(t)a_k \\ \phi_1(t)a_k \\ \vdots \\ \phi_q(t)a_k \end{bmatrix} \cdot \begin{bmatrix} c_{j,1,1}, \cdots c_{j,q+1,1} \\ c_{j,1,2}, \cdots c_{j,q+1,2} \\ \vdots \quad \vdots \\ c_{j,1,q+1}, \cdots c_{j,q+1,q+1} \end{bmatrix} \equiv 0 \qquad (8)$$

where $c_{j,1,1} = E[x_j(t)x_1(t+\tau_1)x_1(t+\tau-k)]$ and $x_j(t)$ is the jth element of X(t).

3. WAVELET BASIS FUNCTION

In this contribution, we choose orthonormal wavelet basis with the form:

$$B = \{2^{1/2}\varphi(2^j t - k), j, k \in Z\}$$

where the function $\varphi(t)$ is called wavelet basis.

According to the model proposed above, the wavelet series can be expressed

$$a_i(t) = \sum_k a_{i,k} \varphi_{i,k}(t)$$

$$a_{i,k} = \int a_i(t)\varphi_{i,k}(t)dt \qquad (9)$$

where $a_{i,k}$ are constant coefficients of wavelet functions, which reflect the type of features presented in the time series. The use of wavelet basis is advocated in our work, since wavelets can well localize in time and frequency.

For time series with sharp jumps or steps, one would choose a boxcar-like wavelet function such as the Harr and Daubechies basis. While for smoothly varying time series, one would choose a smooth function such as Morlet wavelet.

For Harr basis functions, they are represented as

$$\varphi_{m,n}(t) = 2^{-m/2}\varphi(2^{-m}t - n)$$

Fig.1 shows the Harr basis functions across scales. It is noted that Harr basis has interesting behavior as being capable of capturing the global as well as the local behavior of the TV coefficients. It consists of scaled and translated versions of a single function. To obtain the parametric estimation, Harr basis is used in (9) and provides information about $a_i(t)$ at different resolutions while the traditional Fourier basis cannot provide satisfied time and frequency localization.

4. SIMULATIONS AND DISCUSSION

The first simulation is considered with a piecewise AR(2) model

$$x(t) = a_1(t)x(k-1) + a_2(t)x(k-1) + e(t)$$

where $\{a_1(k), a_2(k)\}$ have the values as

$$a_1(k) = \begin{cases} -1.5 & k \in [1, N/4] \cup [N/2+1, 3N/4] \\ -0.9 & k \in [N/4+1, N/2] \cup [3N/4+1, N] \end{cases}$$

$$a_2(k) = \begin{cases} 0.8 & k \in [1, N/4] \cup [N/2+1, 3N/4] \\ 0.2 & k \in [N/4+1, N/2] \cup [3N/4+1, N] \end{cases}$$

We investigate the data set with N=6000 data points. Fig.2 shows the estimation of $a_1(k)$ via Harr basis. The result using Fourier basis was also given as a comparison with dotted line.

The second simulation considered a non-Gaussian AR(2) model with the following non-piecewise TV parameter

$$a_1(k) = 0.6*\sin(12*(k./6000))$$

$$a_2(k) = -0.7*\cos(24*(k./6000))$$

Morlet wavelet was used in this example with m=7. Fig.3 shows the estimation result of $a_1(k)$ with the dotted line.

It is important to note that the estimation based on wavelet basis responds rapidly and precisely to the parameters' time-variances. This finding highlights the capacity of wavelets to identify time-variances that may have physiologic relevance and retains parametric identity even though model coefficients vary continuously over the data segment studied. We can also get the same results by analyzing the error criterion, which is estimated by the following definition [3]:

$$d_0(x_1, x_2, \cdots, x_N; a_1, a_2 \cdots a_P)$$

$$= \sum_{i=P+1}^{N} [X_i - \sum_{j=1}^{P} \hat{a}_j \hat{X}_{i-j}]^2 / N \qquad (10)$$

The error of wavelet basis in the first simulation with $d_{01} = 1.7332$ is less than the error of Fourier basis with $d_{02} = 2.0256$. Moreover, it is remarkable from the results that the wavelet basis is more suitable than Fourier basis to the estimation of time-variance of AR parameters. As a novel method, non-Gaussian AR model with wavelet basis provides us a new way to describe the TV bispectral structures of the non-stationary non-Gaussian process.

5. CONCLUSION

To cope with the problem of the identification of the time-varying linear systems, this paper proposes a non-Gaussian AR model with TV parameters. In this approach, model parameters that characterize the time-varying system are functions of time and were approximated by a family of wavelet basis functions. A comparison between wavelet basis and Fourier basis of cumulant-based method was also demonstrated. Simulation results show the applicability and the effectiveness of this new procedure, while more signal processing techniques are needed to be applied to minimize the estimated error.

6. ACKNOWLEDGEMENTS

and the Grant of Natural Science Research of Guangdong Education Bureau (No. Z02036).

7. REFERENCES

[1] J. J. Rajan , Peter J. W. Rayner "Generalized Feature Extraction For Time-varying Autoregressive Models", IEEE Trans. on Signal Processing. vol. 44, No. 10, October 1996.

[2] D. Aboutajdine "Fast Adaptive Algorithms for AR Parameters Estimation Using Higher Order Statistics" IEEE Trans. on Signal Processing. Vol. 44, No.8, Aug. 1996.

[3] Marc Lavielle "Optimal Segmentation of Random Processes," IEEE Trans. on Signal Processing. vol. 46, No.5, May 1998

[4] D. Geman,"Random fields and inverse problems in imaging," in Lecture Notes in Mathematics. New York: Springer -Verlag,1990

[5] Lee M. Garth "A comparison of optimized higher-order spectral detection techniques for non-Gaussian signals," IEEE Trans. on Signal Processing. vol. 44, No.5, May 1996.

[6] M. K. Tsatsanis and G.. B Giannakis. "Time-varying System Identification and Model Validation Using Wavelets." IEEE Trans. on Signal Processing, 41(12): 3512-3523, 1993.

[7] M. Martone "Non-Gaussian Multivariate Adaptive AR Estimation Using the Super Exponential Algorithm," IEEE Trans. on Signal Processing, vol. 44, No.10, Oct. 1996.

[8] G. Thonet, "Stationarity Assessment with Time-varying Autoregressive Modeling" Signal Processing Laborator, pp. 3721-3724, 1997.
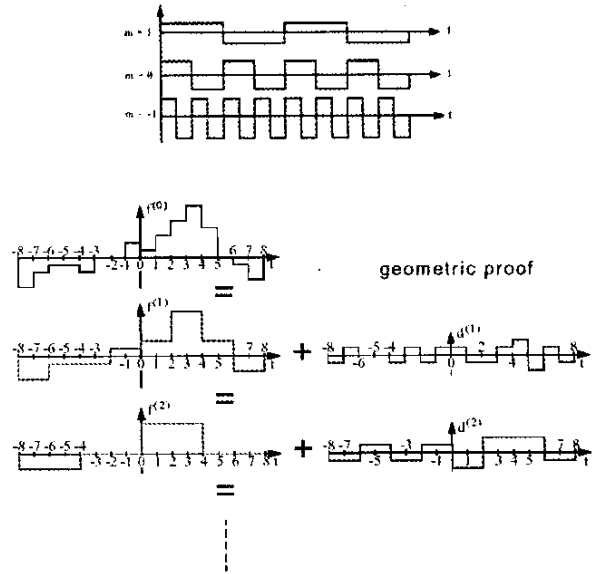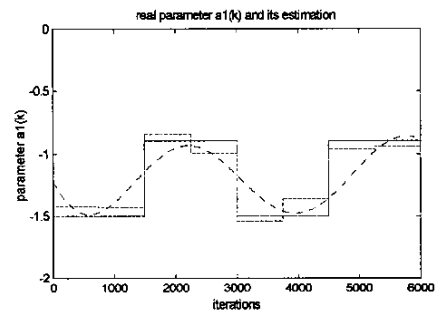
Fig. 1. Harr basis functions across scales

Fig. 2 The estimation of $a_1(k)$ in which the dotted line denotes the estimation based on the Fourier basis, and the solid line is the estimation based on Harr basis with m=9.
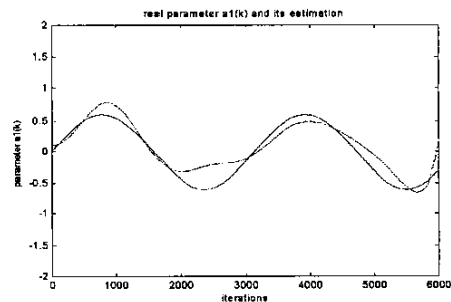

Fig. 3. The estimation of the TV $a_1(k)$ via Morlet wavelet with m=7.