15th Triennial World Congress, Barcelona, Spain

STOCHASTIC PROCESS ON MULTIWAVELET

Jaka Sembiring [*] Alireza S. Sabzevary [**]
Kageo Akizuki [***]

[*] *Waseda University, Tokyo, Japan*
[**] *Mashhad Institute of Technology, Mashhad, Iran*
[***] *Waseda University, Tokyo, Japan*

Abstract: This paper describes a multi-scale stochastic model defined on a multiwavelet structure. Previously such stochastic models are based either on a single or multiple binary tree as well as on an ordinary wavelet structure. The proposed model lies on a tree structure consists of several sets of data coefficients as a result of multiwavelet transformation. Each data sets is linked only through initial data set at root node and conditionally independent given this initial state. Multiwavelet possesses several interesting properties like simultaneously short support, symmetry and orthogonality. The effect of these properties on the proposed model is shown through simulation of smoothing process of a certain fractal signal. It will be demonstrated that several improvements over previously announced results are obtained.

Keywords: Stochastic systems, Stochastic modeling, Fractals, Estimation algorithms, Signal processing, Smoothing

1. INTRODUCTION

Research topic focusing on the descriptions and applications of the multi-scale stochastic process recently has gained momentum thanks to the nice properties of such model on dealing with fractal signal (Chou *et al.*, 1994*a*), (Chou *et al.*, 1994*b*), (Daniel and Willsky, 1997), (Fabre, 1996), (Sembiring and Akizuki, 2000*b*), (Sembiring and Akizuki, 2000*a*). Originally the model is defined on a single binary tree, see e.g. (Chou *et al.*, 1994*a*), (Chou *et al.*, 1994*b*), then it is expanded further with model on wavelet (Fabre, 1996). The authors gave another model from different point of view in (Sembiring and Akizuki, 2000*a*). On the other development, the wavelet theory gets another impulse driven by the introduction of multiwavelet, see e.g. (Strela *et al.*, 1999), (Strela *et al.*, 1995). The multiwavelet has been known to have several advantages over ordinary wavelet such as possesses simultaneously short support, symmetry and orthogonality properties.

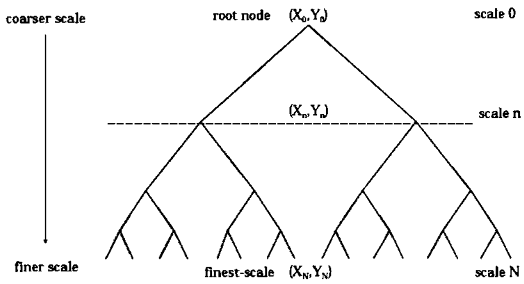

Fig. 1. Dyadic tree $\mathcal{T}$

The basic construction of a multi-scale stochastic model is a state space evolving on a dyadic tree $\mathcal{T}$, instead of $\mathbb{Z}$, given in Fig. 1, as a data structure. Upon going down the tree towards the finest scale N, the resolution doubles from scale to scale, and vice versa. The Haar wavelet decomposition or a pure decimator derived from the wavelet theory behaves just like the upward evolution, and the synthesis process is equivalent to the downward path. This paper expands the model further on multiwavelet structure by exploiting its strength.

2. MULTIWAVELET

2.1 Multi-resolution Analysis

This section contains a brief review of the multiwavelet theory, which is itself an addition to the body of wavelet theory. It falls on the realm of of multi-resolution analysis (MRA). But instead of assuming single scaling function $\phi(t)$, the multiwavelet extends the MRA by M scaling function $\Phi(t) = [\phi_1(t), \cdots, \phi_r(t), \cdots \ \phi_M(t)] \in \mathbb{L}^2(\mathbb{R})^M, 1 \leq r \leq M$, such that the following two-scale matrix dilation equation is satisfied

$$\Phi(t) = \sqrt{2} \sum_k G_k \Phi(2t - k), \text{ for } k \in \mathbb{Z} \quad (1)$$

where $G_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})^{M \times M}$ are $M \times M$ scaling coefficients. For orthonormal scaling function, it can be shown that

$$\sum_k G_k G_{2l+k}^T = I_M \delta_{0,l}, \text{ for } k, l \in \mathbb{Z} \quad (2)$$

The multiwavelet theory stated that one can derive a multiwavelet $\Psi(t) = [\psi_1(t), \cdots, \phi_r(t), \cdots \ \psi_M(t)] \in \mathbb{L}^2(\mathbb{R})^M, 1 \leq r \leq M$, which also preserve the following two-scale matrix equation

$$\Psi(t) = \sqrt{2} \sum_k H_k \Phi(2t - k), \text{ for } k \in \mathbb{Z} \quad (3)$$

where $H_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})^{M \times M}$ are $M \times M$ wavelet coefficients. In orthogonal multi-resolution analysis the next relation

$$\sum_k G_k H_{2l+k}^T = 0_M, \text{ for } k, l \in \mathbb{Z} \quad (4)$$

is hold and the matrix H_k in this orthonormal system satisfies the condition below

$$\sum_k H_k H_{2l+k}^T = I_M \delta_{0,l}, \text{ for } k, l \in \mathbb{Z} \quad (5)$$

The pair of Eqs. (1) and (3) can be realized with matrix filter banks operating on M input data stream and producing $2M$ decimated by 2 output streams. If $x(t)$ denotes input signal, then multiwavelet decomposition of this signal can be obtained by recursively calculate the following

$$v_{n,k} = \sum_m G_{m-2k} v_{n-1,m} \text{ for } m, k \in \mathbb{Z} \quad (6a)$$

$$w_{n,k} = \sum_m H_{m-2k} v_{n-1,m} \text{ for } m, k \in \mathbb{Z} \quad (6b)$$

where $v_{n,k}, w_{n,k}$ are coefficients of multiwavelet transformation at scale $0 \leq n \leq N$ in $M \times 1$ column vectors. The original signal then can be reconstructed through synthesis equation

$$v_{n-1,k} = \sum_m G_{k-2m}^T v_{n,m} + \sum_m H_{k-2m}^T w_{n,m} \quad (7)$$

This paper will concentrate on two types of orthogonal multiwavelet, the Geronimo, Hardin and Massopust (GHM) and the Chui-Lian (CL) multiwavelet, see Table 1 for multiscaling and multiwavelet matrix filters of each system.

Table 1. Multiwavelet filters.

GHM multiwavelet	
$G_0 = \begin{bmatrix} \frac{3}{5\sqrt{2}} & \frac{4}{5} \\ -\frac{1}{20} & -\frac{3}{10\sqrt{2}} \end{bmatrix}$	$H_0 = \frac{1}{10} \begin{bmatrix} -\frac{1}{2} & -\frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 3 \end{bmatrix}$
$G_1 = \begin{bmatrix} \frac{3}{5\sqrt{2}} & 0 \\ \frac{9}{20} & \frac{1}{\sqrt{2}} \end{bmatrix}$	$H_1 = \frac{1}{10} \begin{bmatrix} \frac{9}{2} & -\frac{10}{\sqrt{2}} \\ -\frac{9}{\sqrt{2}} & 0 \end{bmatrix}$
$G_2 = \begin{bmatrix} 0 & 0 \\ \frac{9}{20} & -\frac{3}{10\sqrt{2}} \end{bmatrix}$	$H_2 = \frac{1}{10} \begin{bmatrix} \frac{9}{2} & -\frac{3}{\sqrt{2}} \\ \frac{9}{\sqrt{2}} & -3 \end{bmatrix}$
$G_3 = \begin{bmatrix} 0 & 0 \\ -\frac{1}{20} & 0 \end{bmatrix}$	$H_3 = \frac{1}{10} \begin{bmatrix} -\frac{1}{2} & 0 \\ -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}$
CL multiwavelet	
$G_0 = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & -1 \\ \frac{\sqrt{7}}{2} & -\frac{\sqrt{7}}{2} \end{bmatrix}$	$H_0 = \frac{1}{4\sqrt{2}} \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix}$
$G_1 = \frac{1}{2\sqrt{2}} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$	$H_1 = \frac{1}{4\sqrt{2}} \begin{bmatrix} -4 & 0 \\ 0 & 2\sqrt{7} \end{bmatrix}$
$G_2 = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -\frac{\sqrt{7}}{2} & -\frac{\sqrt{7}}{2} \end{bmatrix}$	$H_2 = \frac{1}{4\sqrt{2}} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$

2.2 Pre-filtering

As described in the previous subsection, the multiwavelet system needs M row of input data streams, and in particular choosing the GHM or CL wavelet results in $M = 2$ input data streams. There are several kinds of pre-filtering for this purpose, two of which are used in this paper, namely repeated row and critical sampling method described in (Strela *et al.*, 1999). Critical sampling pre-filtering combined with CL multiwavelet transformation produces data streams whose covariance matrix takes diagonal form, whereas repeated row processing with GHM multiwavelet produces almost diagonal covariance. Through simulation the effect of both pre-filtering methods will be given.

3. STOCHASTIC MODEL

Stochastic model based on wavelet decomposition has been covered in (Fabre, 1996). The extension of this model on multiwavelet structure, given apparent advantages of multiwavelet, is the main topic in this paper. The multiwavelet transformation consists of filtering $G_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})^{M\times M}$ and $H_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})^{M\times M}$ followed by a decimation by two. In this paper, the filtering process with G_k followed by such decimation will be referred as α, and for H_k will be denoted by β respectively. Thus the operators α and β work from $l^2(\mathbb{Z})$ to $l^2(2\mathbb{Z})$. For a given signal or vector X_n at scale n where $0 \le n \le N$, the αX_n is the coarse approximation at scale $n+1$, whereas βX_n yields the detail lost in this approximation. From Eq. (7), the *reverse* process, i.e. going from coarse to fine involves adjoint operators $\bar{\alpha}$ and $\bar{\beta}$ for filter G_k^T and H_k^T respectively. From Eqs. (2), (4), (5), and (7), it is clear that the following perfect reconstruction conditions are satisfied

$$\bar{\alpha}\alpha + \bar{\beta}\beta = I \quad (8a)$$
$$\bar{\alpha}\alpha = I = \bar{\beta}\beta \quad (8b)$$
$$\alpha\bar{\beta} = 0 = \beta\bar{\alpha} \quad (8c)$$

Upon close examination of the coefficients after applying operators α and β, the data set at scale n consists of M data rows either at approximation decomposition or at detail decomposition. In the subsequent parts when necessary, these rows of data are denoted by r where $1 \le r \le M$. In case of GHM and CL mutiwavelet, it is clear that there exists $M = 2$ rows of data at each scale n, see Table 1.

Now the relation of the transformation above with the stochastic modeling can be explained as follows. General multi-scale stochastic model can be written as

$$X_{n+1} = (\bar{\alpha}_n A_n + \bar{\beta}_n B_n)X_n + W_{n+1} \quad (9a)$$
$$Y_{n+1} = CX_n + V_n \quad (9b)$$

where n denotes scale index increasing towards finest scale N, recall Fig. 1. The $\bar{\alpha}, \bar{\beta}$ are multiwavelet transformation adjoint operators. X_n is the state vector with stationary white noise W_n, and Y_n is the observation vector also with stationary white noise V_n. These noises are independent each other, and independent to the data at initial root X_0. The variables A_n, B_n and C_n are matrices with appropriate dimensions. The model written in Eq. (9) can be associated with a synthesis process of multi wavelet transformation, see Eq. (7), such that the evolution of Eq. (9) is towards finer scale through adjoint operators $\bar{\alpha}$ and $\bar{\beta}$. If necessary the notation $(X_n)_{r_s}$ is used to denote the data at scale n, row r-th and node s-th.

Once transformation is carried out, the coefficients becomes the observation Y_n. Provided that several conditions are satisfied, a very efficient smoothing algorithm of (Chou *et al.*, 1994*a*) can be applied. These conditions are summarized in the following theorem.

Theorem 1. Let $\mathcal{X}Y_{r_s} \triangleq \{(w_nX_n)_{r_s}, (w_nY_n)_{r_s} | w_n = \{\alpha,\beta\}, 0 \le n \le N, 1 \le r \le M, 1 \le s \le n^2\}$ be the data set of multiwavelet transformation. Assume that the noises in Eq. (9) are white Gaussian with covariances P_n and Q_n respectively, constant over scale. The set $\mathcal{X}Y_{r_s}$ is linked through $(X_0)_{r_s}$ only, and conditionally independent given X_0 at root node.

PROOF. The proof mainly depends on the properties of noises W_n and V_n. On the G_k filter side, pre-multiply Eq. (9) with α, then using Eq. (8) one will have $w_nX_n = A.w_{n-1}X_{n-1} + w_nW_n$. Conditioning with respect to X_{n-1}, only $w_{n-1}X_{n-1}$ is needed to estimate X_{n-1}, so that w_nX_n is linked to the coarser scale through $w_{n-1}X_{n-1}$, and it is conditionally independent given X_{n-1}. Now using Eq. (2) the w_nX_n actually consists of $\{(w_nX_n)_{r_s} | 1 \le r \le M\}$ which is orthogonal each other. Noting that the covariance of the noises are constant over scale, it can be concluded that the $(w_nX_n)_{r_s}$ is uncorrelated. The same conclusion can be achieved at the H_k filter decomposition.

Furthermore $(w_nY_n)_r$ is linked to $(X_n)_r$ through $(w_nX_n)_r$, and it is conditionally independent given $(X_n)_r$. This conclusion can be achieved from $w_nY_n = C.w_nX_n + w_nV_n$ and the similar arguments above.

Combining the two facts completes the theorem.

Based on the theorem, the fast algorithm derived in (Chou *et al.*, 1994*a*) can be applied on each set $\mathcal{X}Y_{r_s}$ where the smoothing process is independent. Different from the model on the wavelet structure where the number of data sets is equivalent to the length of filter taps, in multi wavelet decomposition the number is M times as much.

4. SIMULATION

The subject of this section is the smoothing procedure of a certain fractal signal using the model described in the previous section. Assume an observation signal corrupted by white Gaussian noise (SNR $= \sqrt{2}$) depicted as dashed line in Figs. 2, 3, 4 and 5. The sample path of the true signal of this noisy observation is generated through multi-scale system elaborated in (Chou *et al.*, 1994*a*), (Chou *et al.*, 1994*b*) and given by the dotted line in those figures. The smoothing results

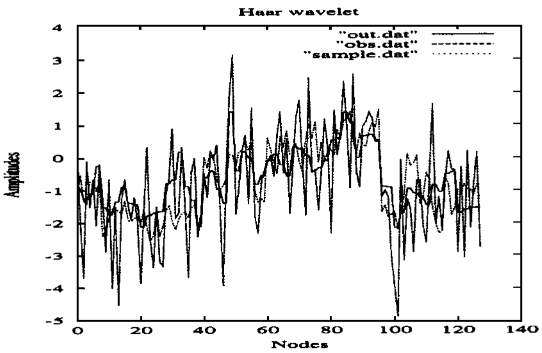

Fig. 2. Smoothing on Haar wavelet

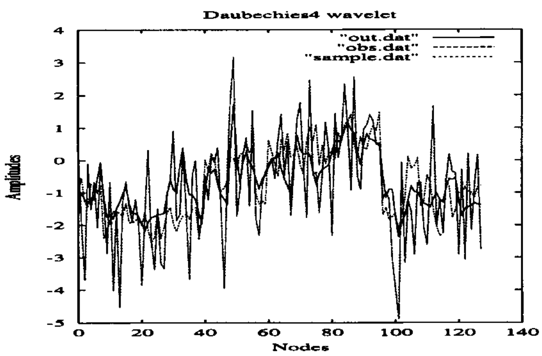

Fig. 3. Smoothing on Daubechies4 wavelet

on a variety of wavelets are given by solid line. On the close examination of the graphs, a smoothing based on Haar wavelet reveals some blocky nature mentioned in (Chou *et al.*, 1994*a*). A better result can be achieved through Daubechies4 wavelet, see Fig. 3, refer also (Fabre, 1996). Finally Figs. 4 and 5 depicted the smoothing of the signal on the GHM with repeated row pre-filtering and CL multiwavelet structure with critical sampling pre-filtering, respectively, which shows further improvement over previously announced method, especially in capturing the high frequency variation of the sample signal.

The qualities of the proposed smoothing result based on GHM and CL multiwavelet are given numerically in Table 2, as well as the result using Haar wavelet in (Chou *et al.*, 1994*a*) and Daubechies4 wavelet in (Fabre, 1996) for comparison. The MSE value of each process shows that the proposed method gives better performance. Other important quality measurement is the Bhattacharrya distance which is given to show the closeness of two random vector. It is based on the following formula

$$B_d(R_1, R_2) = \frac{1}{2}\ln\left[\left|\frac{1}{2}(R_1 + R_2)\right|\right] - \frac{1}{4}\ln\left[\left|(R_1 R_2)\right|\right] \qquad (10)$$

where R_1, R_2 denote the covariance matrices of both random vectors, and $|R|$ is the determinant

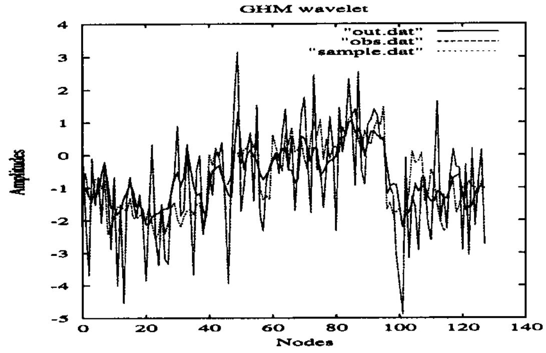

Fig. 4. Smoothing on GHM multiwavelet, repeated row pre-filtering

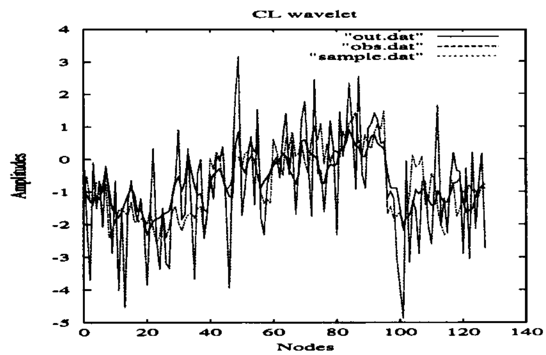

Fig. 5. Smoothing on Chui-Lian multiwavelet, critical sampling pre-filtering

of the matrix R. Given two random vectors with probability of obtaining a sample from either vectors is $1/2$, the probability of error of deciding from which process the sample was is bounded by the following

$$P(\text{error}) \leq \frac{1}{2}\exp^{-B_d(R_1,R_2)} \qquad (11)$$

Bhattacharrya distance and the bound of the probability of error confirm that the smoothing based on the CL multiwavelet outperforms ordinary wavelet base. The CL multiwavelet in particular produces better smoothing results compare to the GHM as seen in Figs. 4, 5 and Table 2. This is due to the input vectors of the CL multiwavelet which has diagonal covariance matrix as a result of critical sampling pre-filtering employed. In addition, the computation of critical sampling pre-filtering followed by CL multiwavelet is faster than repeated row pre-filtering plus GHM multiwavelet. There are two reasons for this facts, first the size of each wavelet itself, 2×6 for CL multiwavelet compare to 2×8 for GHM. The second is that the critical sampling pre-filtering produce a half length data vectors for system input.

Table 2. Quality of smoothing.

Wavelet	MSE	B_d	P(error)
Haar	1.319928	5.501938	0.002039
Daubechies4	1.333280	3.806540	0.011112
GHM	1.313699	5.779562	0.001545
Chui-Lian	1.301453	3.586756	0.013844

5. CONCLUSION

This paper describes an attempt of modeling a stochastic system evolving on multi wavelet structure. Simulation results show that based on the derived model, in particular a smoothing procedure of a certain fractal signal gives several improvement over previously published construction.

REFERENCES

Chou, Kenneth C., Alan S. Willsky and Albert Benveniste (1994*a*). Multiscale recursive estimation, data fusion, and regularization. *IEEE Trans. on Automatic Control* **39**(3), 464–478.

Chou, Kenneth C., Alan S. Willsky and Ramine Nikoukhah (1994*b*). Multiscale systems, Kalman filters, and Riccati equations. *IEEE Trans. on Automatic Control* **39**(3), 479–492.

Daniel, M. and Alan S. Willsky (1997). A multiresolution methodology for signal-level fusion and data assimilation with application in remote sensing. *Proc. of IEEE* **85**(1), 164–180.

Fabre, Eric (1996). New fast smoothers for multiscale systems. *IEEE Trans. on Signal Processing* **44**(8), 1893–1911.

Sembiring, Jaka and Kageo Akizuki (2000*a*). Application of multiple tree stochastic theory on estimating signal over network. In: *Proc. of IFAC SYSID 2000.* Santa Barbara, USA.

Sembiring, Jaka and Kageo Akizuki (2000*b*). Realization of multi-scale stochastic process on multiple tree. *Trans. of IEE of Japan* **120-C**(2), 194–200.

Strela, Vasily et al. (1995). Application of multiwavelets to signal compression and denoising. In: *Proc. UK Symposium on Applications of Time-Frequency and Time-Scale Methods.*

Strela, Vasily et al. (1999). The application of multiwavelet filter banks to signal and image processing. *IEEE Trans. on IP* **8**(2), 548–563.