

Multi-scale Representation of Stochastic Processes using Subband Coder as Modeling Filter

Brejesh Lall

Dept of Elect Engg
IIT Delhi
brejesh@ee.iitd.ac.in

Shiv Dutt Joshi

Dept of Elect Engg
IIT Delhi
sdjoshi@ee.iitd.ac.in

RKP Bhatt

Dept of Elect Engg
IIT Delhi
rkpb@ee.iitd.ac.in

Abstract

Multi-scale modeling and analysis of signals has aroused a lot of interest over the last few years. Multiresolution signal analysis has been used as a tool for a variety of applications and a fundamental attribute for these techniques is the ability to represent a wide class of signals. In this paper, we propose a modeling scheme which models evolution of a stochastic process across scale. The scheme consists of multiple stages where each stage models the generation mechanism of the process as it evolves in scale. It is further shown that for specific choice of filters appearing in the model, the proposed structure is equivalent to a non-causal Auto Regressive (AR). Thus the non causal Auto Regressive processes will be one specific class of processes that will be represented by the proposed scheme. Algorithms for the estimation of the model parameters from the statistics of the process as well as in the least squares sense are also presented here. Finally, simulation results are presented which validate the model.

1. Introduction

The problem of modeling a stochastic process is of fundamental concern in a variety of applications which include, among others, speech and image processing, analysis of econometric time-series, bio-medical signal processing, seismic data processing etc.. In the past a number of researchers [8]-[10] have focused on representation of signals (deterministic) in the multiscale framework. Willsky and coworkers [1]-[6] have considered multi-scale modeling in the stochastic framework. In [4], multiscale stochastic models are applied to texture segmentation. A multiscale AR modeling scheme, which typically can be used to model isotropic processes, has been proposed by Willsky et al [1], [2]. They also provide Levinson-Durbin type of recursions for the computation of parameters for the proposed model. Here, in our work we propose a generation mechanism for a stochastic process in a multiscale setting. The model can be used for the generation and analysis of cyclostationary processes in general and this paper illustrates the scheme by using a particular case of stationary random processes. The proposed scheme models a stochastic process as evolving across scale. The information from coarser scale is used to obtain an estimate of the process at the finer scale. This estimate is refined by adding information which is novel to

this scale transition. This paper proposes the modeling scheme and then applies it to represent stationary processes.

2. Multi-scale Model

2.1. Introduction

The proposed scheme, as shown in Fig.1, models the stochastic process in the following manner: It is assumed that the process is evolving in scale, so each stage models the evolution of process in scale. We assume that the finer scale of the process has two components: one which can be extracted from the coarser scale (i.e. that part of the information that lies in the space spanned by coarse scale information) and second which is the novel information that needs to be added to the coarser scale to obtain the finer scale. The i th scale approximation of the stochastic process can be obtained by passing the $(i+1)$ th scale version through

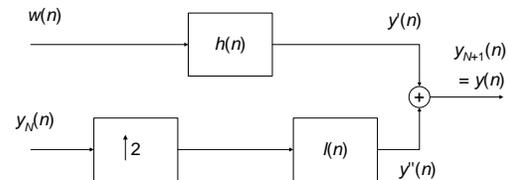


Figure 1: The filter that generates the $(N+1)$ th scale realization of the stochastic process, given the N th scale realization of the process. $h(n)$ is causal IIR in form and $l(n)$ is a two sided FIR

a blurring filter, which in the present study is assumed to be known. This decimation operation should be looked at as a fundamental operation (similar to the delay operation in Linear Time Invariant Filter Framework) on the signal. To regenerate the signal at the $(i+1)$ th scale, the i th scale version is upsampled and passed through an interpolating filter, to generate that part of the $(i+1)$ th scale version which can be extracted from the i th scale. Added to this is the required new information modeled as AR (to be able to represent a large number of colouring types). Let us denote the i th scale approximation of the process as $y_i(n)$. Then the above modeling procedure can be viewed as follows: $y_{i+1}(n)$ has two components: one is that part that can be extracted from $y_i(n)$ and the other is new information that is added at the $(i+1)$ th scale, and which is orthogonal to $y_i(n)$.

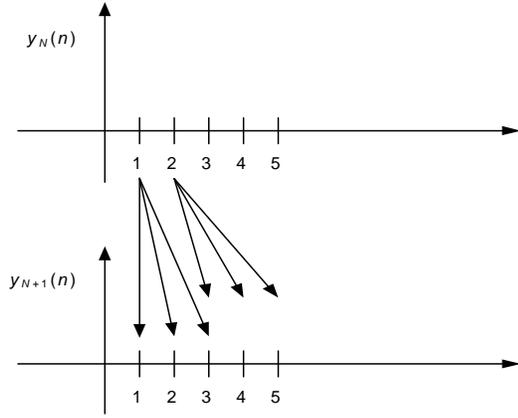


Figure 2: Samples of coarser scale version that contribute to the finer scale, based on blurring model (0.5, 1.0, 0.5)

2.2. Introduction

The model for obtaining finer scale version from the coarser scale version is schematically shown in Fig. 1. In the following discussion $y(n)$ is used to denote the $(N+1)^{\text{th}}$ scale realization (i.e. $y_{N+1}(n)$) for notational convenience. Here we can write

$$y_{N+1}(n) \equiv y(n) = y'(n) + y''(n) \quad (1)$$

Here $y''(n)$ represents that part of the information of $y(n)$ which is contained in $y_N(n)$ and $y'(n)$ represents that part of the new information about $y(n)$ which is added at the $(N+1)^{\text{th}}$ scale, and assumed to be orthogonal to the information at the N^{th} scale. Hence it represents the innovation as we move from scale N to $N+1$.

To obtain form for the filter $l(n)$, consider the blurring operation in the analysis filter. The decimator blurs the information of the finer scale by performing a low pass filtering operation (average of neighbouring pixels). Therefore, in the estimation of the finer scale information (inverse operation) from the coarser scale information, those pixels of the coarser scale are chosen to which this finer scale pixel contributed to in the blurring operation. Figure 2 illustrates this by listing the coarser scale pixels used for filtering in the inverse operation.

Given the forms for the filters $l(n)$ and $h(n)$, the two signals $y'(n)$ and $y''(n)$ are given by the following equations:

$$y'(n) = \sum_{i=1}^p a_i y'(n-i) + w(n) \quad (2)$$

and

$$y''(n) = \sum_{i=-r}^s b_i y_N\left(\frac{n-i}{2}\right) \quad (3)$$

Clearly, here $y''(n)$ has two different forms for two different time instances, one each for even and odd time instances. In eqn. (3) the index i takes even values if n is even and takes odd values if n is odd. Thus, filter coefficients with even index contribute to $y''(n)$ for even instances of time (n even) and filter coefficients with odd index contribute to $y''(n)$ for

odd instances of time (n odd). Substituting (2) and (3) in eqn. (1), we get:

$$y(n) = \sum_{i=1}^p a_i y'(n-i) + \sum_{k=-r}^s b_k y_N\left(\frac{n-k}{2}\right) + w(n) \quad (4)$$

where k is even for n even and k is odd for n odd. This equation can be modified to obtain a representation of $y(n)$ entirely in terms of the samples of the sequence at its N and $N+1$ scale and the innovations sequence $w(n)$. The following proposition specifies this representation.

Proposition 1 - The stochastic process generated by the filter shown in Fig. 1 admits the following representation for each scale

$$y(n) = \sum_{i=1}^p a_i y(n-i) - \sum_{k=-r}^s \theta_k y_N\left(\frac{n-k}{2}\right) + w(n) \quad (5)$$

where k is even for n even and k is odd for n odd and θ_k is a function of the filter coefficients a_i 's and b_i 's and the mapping is defined later in this section.

Proof: Using eqn. (1) we first express $y'(n-i)$ in terms of $y(n-i)$ and $y''(n-i)$. Further, using eqn. (3), we can write $y''(n-i)$ in the following form

$$y''(n-i) = \sum_{j=-r}^s b_j y_N\left(\frac{n-i-j}{2}\right) \quad (6)$$

where j is even for $(n-i)$ even and odd for $(n-i)$ odd. Using (1) and (6), we obtain the following representation for $y'(n-i)$

$$y'(n-i) = y(n-i) - \sum_{j=-r}^s b_j y_N\left(\frac{n-i-j}{2}\right) \quad (7)$$

where j is even for $(n-i)$ even and odd for $(n-i)$ odd. Substituting (7) in (4) we obtain the following representation for $y(n)$

$$y(n) = \sum_{i=1}^p a_i y(n-i) - \sum_{i=1}^p a_i \sum_{j=-r}^s b_j y_N\left(\frac{n-i-j}{2}\right) + w(n) \quad (8)$$

where j is even for $(n-i)$ even and odd for $(n-i)$ odd. The two summations that appear in the second term in the RHS of (8) can be combined to obtain the representation for $y(n)$ specified by (5), where θ_k is a function of the filter coefficients and is given by $\theta_k = \sum_{i,j} a_i b_j$, such that $i+j=k$.

Also, where k is even for n even and k is odd for n odd. This implies that even index θ_k contribute to generation of even samples of $y(n)$ and odd index θ_k contribute to generation of odd samples of $y(n)$.

Q.E.D.

This representation can be further modified to obtain a representation for $y(n)$ entirely in terms of the past and present sample of $y(n)$ at scale $N+1$, giving a non causal AR form to the representation of $y(n)$. The following proposition formally states this result.

Proposition 2 - The stochastic process generated by the filter shown in Fig. 1 admits the following representation for the signal at each scale

$$y(n) = \sum_{i=-(r+1)}^{p+s+1} e_i y(n-i) + w(n) \quad (9)$$

where e_i 's are functions of the filter parameters a_i 's and b_i 's and the mapping is defined later in this section.

Proof: The signal $y(n)$ in the coarser scale is a blurred version of the signal in the finer scale and this dependence can be represented by the following equation:

$$y_N(n) = \sum_{i=-t}^u c_i y_{N+1}(2n-i) \quad (10)$$

The blurring function a mapping from the finer scale to the coarser one. Choosing the following form for this mapping, helps fix a particular blurring operation. Choose $c_0=1$, $c_{-1}=1/2$ and $c_1=1/2$. Using these and (10) the following representation for $y_N(n-k/2)$ can be obtained

$$y_N\left(\frac{n-k}{2}\right) = \frac{1}{2} \{y(n-k-1) + 2y(n-k) + y(n-k+1)\} \quad (11)$$

In general, the choice of blurring model will impact the amount of information being passed on to the coarser scale version and will, therefore, impact the filter parameters of the innovations filter. Also, the blurring function choice will impact the performance of the model. This work demonstrates the ability of the model with an assumed blurring model. Substituting (11) in (5), we obtain

$$y(n) = \sum_{i=1}^p a_i y(n-i) - \frac{1}{2} \sum_{k=-r}^s \theta_k \{y(n-k-1) + 2y(n-k) + y(n-k+1)\} + w(n) \quad (12)$$

where k is even if n is even and k is odd if n is odd. Combining terms with same shift i in the RHS of (12) the following representation for $y(n)$ is obtained.

$$y(n) = \sum_{i=-(r+1)}^{p+s+1} e_i y(n-i) + w(n) \quad (13)$$

where e_i 's are given by

$$e_i = a_i + \frac{1}{2}(\theta_{i-1} + 2\theta_i + \theta_{i+1}) \quad (14)$$

where it is given that $\theta_k = \sum_{i,j} a_i b_j$, such that $i+j=k$.

Q.E.D.

Thus we have a representation scheme for modeling stochastic processes in the multi-scale framework. The representation scheme has a generation model similar to one stage of the synthesis filter of the subband coder. The algorithms for obtaining the filter coefficients from the given statistics or from a given realization are now presented.

3. Estimation of Model Parameters

Given a stochastic process $y(n)$, it is modeled as evolving across scale. The finer version is obtained from coarser scale by adding to it the new information appearing at the transition from the coarser scale version to the finer scale version. The innovations are orthogonal to the coarser scale

version of the stochastic process. This orthogonality is exploited in this section to obtain the filter coefficients in terms of the second order statistics of the given stochastic process. Equation (4) gives a representation for $y(n)$ in terms of the $y'(n-i)$'s and the samples of the coarser scale version. Now since $w(n)$ is orthogonal to $y_N(n)$, therefore, $y'(n)$ is also orthogonal to $y_N(n)$. If we now obtain the inner product of both sides of eqn. (4) with $y_N((n-j)/2)$, the following eqn results (using the principle of orthogonality):

$$E\left[y_N\left(\frac{n-j}{2}\right) y_{N+1}(n)\right] = \sum_{i=-r}^s b_i R_N\left(\frac{j-i}{2}\right) \quad (15)$$

where i is even for n even and i is odd for n odd. Using the blurring function on LHS of (15) we obtain.

$$\sum_{k=-t}^u c_k R_{N+1}(j+k) - \sum_{i=-r}^s b_i R_N\left(\frac{j-i}{2}\right) = 0 \quad (16)$$

Here j is even for n even and odd for n odd. Since the correlations at the finest scale are assumed to be available (and, further, the process is assumed to be ergodic) and the immediate coarser scale version is easily obtainable from the known, and assumed blurring function, all the required statistics for obtaining b_i 's are available. This set of equations can be put in a convenient matrix form. The first summation in (16) is a function of j and k but since the blurring function is known (i.e. c_k 's are available) it is essentially a function of j . We can define the following mapping to facilitate the evaluation of b_i 's

$$R'_{N+1}(j) = \sum_{k=-t}^u c_k R_{N+1}(j+k) \quad (17)$$

Eqn (17) can be used in (16) to obtain Eqn (18) as the matrix form for the solution of b_i 's

$$\begin{bmatrix} b_{-s+1} \\ b_{-s+3} \\ \vdots \\ b_1 \\ \vdots \\ b_{r-1} \end{bmatrix} = \begin{bmatrix} R_N(0) & \dots & R_N(s/2) & \dots & R_N((r+s-2)/2) \\ R_N(1) & \dots & R_N((s-2)/2) & \dots & R_N((r+s-4)/2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ R_N(s/2) & \dots & R_N(0) & \dots & R_N((r-2)/2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ R_N((r+s-2)/2) & \dots & R_N(r/2) & \dots & R_N(0) \end{bmatrix}^{-1} \begin{bmatrix} R'_{N+1}(-s+1) \\ R'_{N+1}(-s+3) \\ \vdots \\ R'_{N+1}(1) \\ \vdots \\ R'_{N+1}(r-1) \end{bmatrix} \quad (18)$$

Another form for (16) where the b_i 's are represented entirely by the statistics of the finest scale version of the process can be easily obtained from (16). This form involves a triple summation but gives the expression for the estimation of parameters based entirely on the statistics of the finest scale version in a linear fashion. The following equation represents the filter parameters b_i 's entirely in terms of the finest scale version statistics.

$$\sum_{l=-t}^u c_k R_{N+1}(j+l) - \sum_{i=-r}^s \sum_{k=-t}^u \sum_{l=-t}^u b_i c_k R_{N+1}(l+j-i-k) = 0 \quad (19)$$

where i is even for n even and i is odd for n odd.

Once the values of b_i 's have been obtained the statistics of $y'(n)$ can be obtained in terms of the statistics of the finest scale version and the values b_i 's and c_i 's. Using (6) and (10) in (7), with $i = 0$, we obtain

$$y'(n) = y_{N+1}(n) - \sum_{k=-u}^t \sum_{i=-r}^s b_i c_k y_{N+1}(n-i-k) \quad (20)$$

where i is even for n even and i is odd for n odd. The signal $y'(n)$ is Wide Sense Cyclostationary with period 2 (as it is the difference a stationary signal and a signal obtained by upsampling a stationary signal by 2) and can, therefore, be represented in terms of two correlation sequences [11]. These correlation sequences of $y'(n)$ is given by

$$\begin{aligned} R_{y'}^p(m) &= R_{N+1}(m) - \sum_{i,k} b_i c_k R_{N+1}(m-i-k) \\ &\quad - \sum_{j,l} b_j c_l R_{N+1}(m+l+j) \\ &\quad + \sum_{i,j,k,l} b_i b_j c_k c_l R_{N+1}(m+l+j-i-k) \end{aligned} \quad (21)$$

where p is either 0 or 1 (corresponding to the 2 correlation sequences). For $p = 0$, i is even and j is even for m even and j is odd for m odd. For $p = 1$, i is odd and j is odd for m even and j is even for m odd. Once the information regarding the second order statistics of $y'(n)$ is available obtaining the values of a_i 's is straightforward, since $y'(n)$ is the output of filter $h(n)$ for white noise as input. This amounts to AR parameterization which is widely used. A point worth noticing here is that the white noise $w(n)$ is wide sense cyclostationary with period 2, and therefore can be represented by two correlation sequences, both of which are impulses but of different strengths. The signal $y'(n)$ has two correlation sequences one each for $p = 0$ and $p = 1$ in (21).

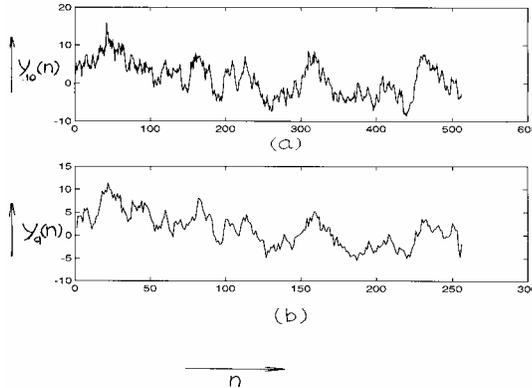


Figure 3: A typical process that can be represented and generated by the proposed modeling scheme for $p = 2$ and $r = s = 1$, for the following filter coefficient values: $a_1 = 0.6$, $a_2 = 0.3$, $b_1 = 0.9$, $b_0 = 1.0$ and $b_1 = 0.3$. a) represents the process at the 10th scale and b) represents the process at the 9th scale.

4. Least Squares Estimation

The previous section contained the procedure for finding the filter parameters given the statistics of the process. Here we present the equations which can be used for the computation of filter parameters from a given realization of the process. As seen in Section II a representation of the given stochastic process is given by (5). The equation has the following form.

$$\underline{y} = \mathbf{H} \underline{\theta} + \underline{w} \quad (22)$$

Since the vector \underline{w} represents the error in estimation, it consists of variables that are uncorrelated to each other, hence, the usual least squares estimation can be applied to the above set of equations to obtain the values of the coefficients. Thus the coefficients in the vectors $\underline{\theta}$ can be obtained using the following equation

$$\underline{\theta} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \underline{y} \quad (23)$$

Once these parameters are known the filter coefficients can be calculated from them, as a mapping, from the filter coefficients to the parameters estimated, exists. The mapping is as given in section II. Some simulations have been done to show the kinds of processes that can be generated using the generation mechanism of the proposed model. Figures 3 and 4 show two such processes. The orders of the model used for the generation of the two processes are taken to be: $p = 2$ and $s = r = 1$.

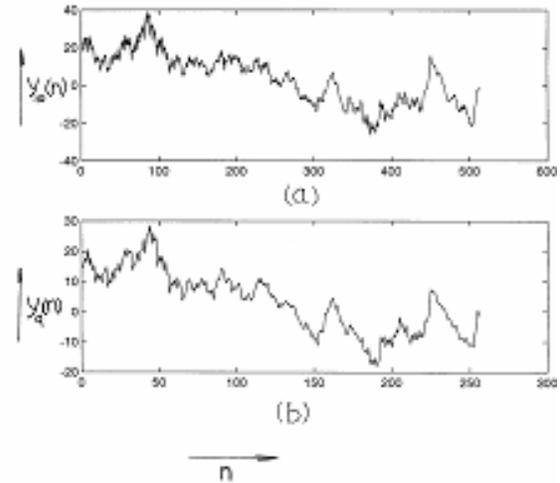


Figure 4: A typical process that can be represented and generated by the proposed modeling scheme for $p = 2$ and $r = s = 1$, for the following filter coefficient values: $a_1 = 0.2$, $a_2 = 0.7$, $b_1 = 0.4$, $b_0 = 1.2$ and $b_1 = 1.0$. a) represents the process at the 10th scale and b) represents the process at the 9th scale.

5. Conclusion

We have discussed a scheme for generation of cyclostationary processes in the multiscale framework. The scheme can also be used to model and generate stationary stochastic processes. The estimation procedure for the model parameters for stationary processes is also given in this paper. This scheme is in contrast to the work by Basseville et. al. [1] who have dealt with isotropic

processes. In the proposed scheme coarser scale information is used to estimate the finer scale version. To this estimate are added innovations, which are modeled as the output of an AR with white noise as the input. The methods for estimation of parameters, for both the given data as well as the given statistics case, are presented.

6. References

- [1] M. Basseville, A Benveniste, and A.S. Willsky, "Multiscale Autoregressive Processes – PartI: Schur-Levinson Parameterization," *IEEE Trans. Signal Processing*, Vol. SP-40, pp. 1915-1934, Aug. 1992.
- [2] --, "Multiscale Autoregressive Processes - Part II: Lattice structures for whitening and modelling," *IEEE Trans. Signal Processing*, Vol. 40, pp. 1935-1954, Aug. 1992.
- [3] M.R. Luetgen, W.E. Karl, A.S. Willsky, and R.R. Tenney, "Multiscale Representation of Markov Random Fields," *IEEE Trans. Signal Processing*, Vol. 41, pp. 3377-3395, Dec.1995.
- [4] M.R.Luetgen, and A.S.Willsky, "Likelihood Calculation for a class of Multiscale Stochastic Models, with Application to Texture Discrimination," *IEEE Trans. Image Processing*, Vol. 4, pp. 94-207, Feh.1995.
- [5] Ke. Chou, A.S. Willsky, and A. Benveniste, "Multiscale Recursive Estimation, Data Fusion, and Regularization," *IEEE Trans. Automatic Control*, Vol. 39, pp. 464-478, March 1994.
- [6] Ke. Chou, A.S. Willsky, and R. Nikoukhah, "Multiscale systems, Kalman Filters, and Riccati Equations," *IEEE Trans. Automatic Control*, Vol. 39, pp. 479-492, March 1994.
- [7] B. Friedlander, and M. Morf, "Least Squares algorithms for Adaptive Linear-Phase Filtering," *IEEE Trans. Acoust. Speech and Signal Processing*, Vol. ASSP-30, pp. 381 390, June 1982.
- [8] I. Daubechies, "Ten lectures on wavelets," SIAM Appl. Math., 1991.
- [9] S. Mallat, "A theory for multi resolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, pp. 674-693, July 1989.
- [10] --, "Multifrequency channel decomposition of images and wavelet models." *IEEE Trans. Acoust. Speech and Signal Processing*, vol. 37, pp. 2091-2110, Dec. 1990.
- [11] B. Lall, S. D. Joshi, and R. K. P. Bhatt, "Second-Order Stastistical Characterization of the Filter Bank and its Elements." *IEEE Trans. Signal Processing*, vol. 47, pp. 1745-1749, June. 1999.