MODELING AND ESTIMATION FOR A CLASS OF MULTIRESOLUTION RANDOM FIELDS

Peyman Milanfar
SRI International, MS 40944
333 Ravenswood Ave., Menlo Park, CA 94025

Robert R. Tenney,
Robert B. Washburn,
and Alan S. Willsky
Alphatech, Inc.
50 Mall Rd, Burlington, MA 01803

ABSTRACT

In this paper we discuss a class of multiresolution models of random fields based on a generalization of the midpoint deflection construction of the 1-D Brownian motion. We then present Least Squares (LS) algorithms for the estimation of parameters which define these models and hence provide a framework for synthesizing and analyzing images with fractal-like properties such as those found in statistical representation of natural terrain and other geophysical phenomena. We also briefly discuss possible applications of this modeling framework to target detection in images.

1. INTRODUCTION

Random field models based on differential and spectral structures have a great deal of difficulty synthesizing realistic images of common physical entities such as terrain. The emerging field of wavelets and multiresolution modeling [2, 5, 4] can be viewed as providing an alternative for more accurate representation of non-linear and non-stationary effects in images. As with most multiscale models proposed to date such as [4, 1], our proposed models are based on the notion of causality in scale. That is, finer scale features of the images depend upon the presence or absence of coarser-scale features. Indeed, the finer scale features of the images are *derived* from the coarser scales by way of dynamical models evolving in scale. What distinguishes our models from those proposed in [4, 1] and elsewhere is the algebraic structure over which these dynamical models evolve in addition to the fact that, as a result of the underlying structure, our models can be described by linear or *non-linear* dynamical evolution equations in scale. Hence, the general modeling paradigm presented here is exceedingly rich and therefore capable of synthesizing realistic correlation structures such as those which characterize natural terrain.

An important feature of our proposed modeling framework is that in our setting the problem of multiscale system identification becomes quite easy to solve. This problem has, in fact, received little attention despite much recent activity in the area of multiscale stochastic modeling of signals and images [3]. Models such as [4, 1], while leading to efficient optimal estimation algorithms, inherently lead to quite complex multiscale system identification problems. In contrast, while our framework is not particularly well-suited to efficient optimal estimation algorithms, it yields a simple, powerful, and efficient framework for multiscale parameter estimation (system identification). We exploit this important property to advantage and develop Least Squares (LS) algorithms which estimate the multiscale model parameters recursively within each scale, and evolve from fine to coarse scale. We will show that once these parameters have been estimated from a given image, they can be used to generate detail (interpolation error) statistics at every scale. These can, in turn, be used to derive various useful quantities such as an indicator of fractal dimension for the given image. In addition, for a large class of natural images, these detail signals have been shown to be significantly spatially decorrelated within each scale. When these detail signals have ensemble variances that decay sufficiently fast as a function of increasing scale, the above properties may allow us to build efficient multiscale image *compression* algorithms.

2. MODELING

To see how the proposed multiscale image models work, let us first consider the *synthesis* of images. Without loss of generality, assume four pixel values to be given in a 2 pixel array. See Figure 1. A midpoint construction process then generates values along the boundaries of the image and at its center. A similar operation at

This work was supported by AFOSR under grant F49620-92-C-0068, and performed while the first author was at Alphatech, Inc.

0-8186-6950-0/94 $4.00 © 1994 IEEE

each scale produces the successive pixel values at the next scale as shown in Figure 1. In its most general formulation, this midpoint construction process need not be linear. However, linear midpoint construction processes are sufficiently rich that we choose to concentrate the remainder of this paper only on these models.

Referring to Figure 1, we define the *linear* mid-point construction process as follows. First, we define the following vectors of image pixel values:

$$y(s) = [\ y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5\ ]^T \tag{1}$$

$$x(s) = [\ x_1 \quad x_2 \quad x_3 \quad x_4\ ]^T \tag{2}$$

Given these vectors, the midpoint construction processe is defined by

$$y(s) = A(s)x(s) + B(s)w(s), \tag{3}$$

where $A(s)$ is a 5-by-4 linear interpolation kernel at scale s given by

$$A(s) = \begin{bmatrix} l(s) & 1-l(s) & 0 & 0 \\ 0 & u(s) & 1-u(s) & 0 \\ 0 & 0 & 1-l(s) & l(s) \\ u(s) & 0 & 0 & 1-u(s) \\ m_1(s) & m_2(s) & m_3(s) & 1-\sum_{i=1}^{3} m_i(s) \end{bmatrix}; \tag{4}$$

$w(s)$ is a zero-mean process assumed white both spatially and in scale; and finally, $B(s)$ is a noise shaping matrix given by

$$B(s)B(s)^T = \mathbf{diag}[\ q_l(s) \quad q_u(s) \quad q_l(s) \quad q_u(s) \quad q_m(s)\ ]. \tag{5}$$

Note that in the definition of $A(s)$, each row sum is equal to one so that each new pixel value at the next scale is the weighted average of a set of two or four pixels at the previous scale. Furthermore, the apparent symmetry in the definition of the elements of the first and third (and second and fourth) rows of $A(s)$ ensures that the resulting image has a self-simliar structure and that no ambiguity arises in assigning a value to a pixel that is shared by two 3×3 pixel fields (An example is pixel z in Figure 1). For these same reasons, the matrix $B(s)$ is chosen as above.

To produce the pixel values at the next scale, the midpoint construction process can then be applied to the following four sets of 2-by-2 pixel sets:

$$x^{(1)}(s+1) = [x_1\ y_1\ y_5\ y_4]^T \tag{6}$$

$$x^{(2)}(s+1) = [y_1\ x_2\ y_2\ y_5]^T \tag{7}$$

$$x^{(3)}(s+1) = [y_5\ y_2\ x_3\ y_3]^T \tag{8}$$

$$x^{(4)}(s+1) = [y_4\ y_5\ y_3\ x_4]^T \tag{9}$$

Finally, the scale-to-scale evolution of the vectors $x(s)$ can be written as

$$x^{(i)}(s+1) = T^{(i)}(s)x^{(j)}(s) + U^{(j)}(s)w^{(j)}(s) \tag{10}$$

for $i = 1, \cdots, 4$ and $j = 1, 2, \cdots, 4^s$. where $T_s(j)$ and $U_s(j)$ depend linearly on $A(s)$ and $B(s)$ respectively. (Note that a nonlinear midpoint construction process would have resulted in a nonliear evolution equation for the vectors $x(s)$.)

The dynamic model in (10) is composed of two terms. The term $T^{(i)}(s)x^{(j)}(s)$ denotes interpolation down to the next scale, while the term $U^{(j)}(s)w^{(j)}(s)$ represents new information (detail) added as the process evolves from one scale to the next. This process, and more generally most multiscale models proposed to date, are motivated by similar dynamical equations describing the scale-to-scale evolution of wavelet scaling coefficients. These coefficients can be generated as outputs of a bank of filters which satisfy certain orthogonality conditions (Quadrature Mirror Filters). Hence it is natural to expect that the 2-D process described by (10) may also be realized by a successive filtering process in the image domain. This is indeed the case as graphically depicted in Figure 2. In this Figure, a $(2^n+1) \times (2^n+1)$ pixel field I_s is assumed given. An *upsampling* process interleaves each row and column of this image with zeros, yielding I^0_{s+1}. This resulting image is then *convolved* with a $k \times k$ kernel $H(s)$ to yield I^*_{s+1}. This convolution process introduces $k-2$ extraneous pixels which are deleted next. Finally, detail (with statistics given by $B(s)$) is added to the appropriate "new" pixels where zeros were inserted. This last step yields the image I_{s+1} at scale $s+1$. It is easy to check that the choice of the following 3×3 convolution kernel yields *exactly* the same result as the midpoint deflection process depicted in Figure 1.

$$H(s) = \begin{bmatrix} m_1(s) & u(s) & m_2(s) \\ l(s) & 1 & 1-l(s) \\ 1-\sum_i m_i(s) & 1-u(s) & m_3(s) \end{bmatrix} \tag{11}$$

3. PARAMETER ESTIMATION

Having established the linear midpoint construction image model, we can construct a model of this type for a given image $f(i,j)$ with $1 \le i,j \le 2^n+1$, by estimating the parameters $A(s)$ and $B(s)$ at every scale[1]. Given the image pixel values at two consecutive scales, we have

$$y^{(j)}(s) = A(s)x^{(j)}(s) + B(s)w^{(j)}(s) \tag{12}$$

for $j = 1, \cdots, 4^s$. We wish to estimate the parameters $A(s)$ and $B(s)$ from the given data. We proceed by first computing LS estimates of the parameters $l(s)$, $u(s)$, and $m_1(s)$, $m_2(s)$, and $m_3(s)$ which uniquely determine

[1] Note that for this f, the scale parameter s takes integer values between 1 and n

$A(s)$. Next, we estimate the parameters $q_l(s)$, $q_u(s)$, and $q_m(s)$ which determine $B(s)$.

By simply rewriting (12), we derive a *recursive* LS (RLS) estimate of the parameter vector $p(s)$ defined as follows

$$p(s) = \left[\begin{array}{cccccc} 1 & l(s) & u(s) & m_1(s) & m_2(s) & m_3(s) \end{array}\right]^T. \tag{13}$$

In particular, we have

$$y^{(j)}(s) = X^{(j)}(s)p(s) + B(s)w^{(j)}(s) \tag{14}$$

where

$X^{(j)}(s) =$

$$\begin{bmatrix} x_2^{(j)} & \Delta x_1 & 0 & 0 & 0 & 0 \\ x_3^{(j)} & 0 & \Delta x_2 & 0 & 0 & 0 \\ x_3^{(j)} & -\Delta x_3 & 0 & 0 & 0 & 0 \\ x_4^{(j)} & 0 & -\Delta x_4 & 0 & 0 & 0 \\ x_4^{(j)} & 0 & 0 & -\Delta x_4 & \Delta x_2 + \Delta x_3 & \Delta x_3 \end{bmatrix} \tag{15}$$

and

$$\Delta x_1 = x_1^{(j)} - x_2^{(j)} \tag{16}$$

$$\Delta x_2 = x_2^{(j)} - x_3^{(j)} \tag{17}$$

$$\Delta x_3 = x_3^{(j)} - x_4^{(j)} \tag{18}$$

$$\Delta x_4 = x_4^{(j)} - x_1^{(j)} \tag{19}$$

The LS estimate of the desired parameters is obtained by computing the LS estimate of the vector $p(s)$ subject to the following linear constraint

$$v^T p(s) = 1, \tag{20}$$

where

$$v = \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{21}$$

That is to say, we solve the following *constrained* least squares problem

$$\min \sum_{j=1}^{4^s} \|y^{(j)}(s) - X^{(j)}(s)p(s)\|^2, \quad \text{s. t.} \quad v^T p(s) = 1 \tag{22}$$

To find the unique solution $\hat{p}(s)$ to this problem, we first compute the *unconstrained* solution $\hat{p}^*(s)$ recursively, and then derive the constrained solution from this [6]. To this end, the unconstrained solution is given by the following RLS procedure.

First define the following quantities:

$$\Gamma_k(s) = \sum_{j=1}^{k} X^{(j)^T}(s)X^{(j)}(s) \tag{23}$$

$$\Gamma_0(n) = \frac{1}{2^n} \sum_{i,j=1}^{2^n+1} (f(i,j) - \bar{f})^2 I \tag{24}$$

$$\Gamma_0(s-1) = \Gamma_{4^s}(s) \tag{25}$$

$$p_0^*(s) = \left[\begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]^T \tag{26}$$

where $\bar{f}$ denotes the average value of $f(i,j)$, and I denotes the 6×6 identity matrix. Note that in the above, the "sample covariance" matrix Γ at a given scale is initialized with the computed covariance at the next higher scale to reflect the relative a-priori uncertainty in the starting guess.

We carry out the following recursions [6] starting at the finest scale n to arrive at the solution.

$$\begin{aligned} p_k^*(s) &= p_{k-1}^*(s) + G_k(s)(y^{(k)}(s) - X^{(k)}(s)p_{k-1}^*(s)) \\ G_k(s) &= \Gamma_{k-1}(s)X^{(k)^T}(s)(I + X^{(k)}\Gamma_{k-1}(s)X^{(k)^T})^{-1} \\ \Gamma_k(s) &= \Gamma_{k-1}(s) - G_k(s)X^{(k)}(s)\Gamma_{k-1}(s) \end{aligned}$$

After 4^s iterations, the resulting unconstrained estimate $\hat{p}^*(s)$ is then used to compute the constrained estimate as follows:

$$\hat{p}(s) = \hat{p}^*(s) + \Gamma(s)v(1 - v^T\hat{p}^*(s))(v^T\Gamma(s)v)^{-1}, \tag{27}$$

where $\Gamma(s) = \Gamma_{4^s}(s)$.

Given the RLS estimate of $p(s)$ (and hence $A(s)$), we estimate the parameters of the noise shaping matrix $B(s)$ by defining the residual vectors

$$e^{(j)}(s) = y^{(j)}(s) - X^{(j)}(s)\hat{p}(s). \tag{28}$$

By referring to Figure 1, it becomes evident that the vectors $e^{(j)}(s)$ contain redundant elements at every scale. For instance, for $s = 2$, the second element in the vector $e^{(1)}(2)$ coincides with the fourth element in the vector $e^{(2)}(2)$. To estimate the parameters of $B(s)$, we remove these redundancies first [2]. Define the following *non-redundant* column vectors

$$\begin{aligned} E_l(s) &= \text{vector of } 1^{st} \text{ and } 3^{rd} \text{ elements of } e^{(j)}(s) \\ E_u(s) &= \text{vector of } 2^{nd} \text{ and } 4^{th} \text{ elements of } e^{(j)}(s) \\ E_m(s) &= \text{vector of } 5^{th} \text{ elements of } e^{(j)}(s) \end{aligned}$$

Given these vectors, the LS estimates of the desired parameters is given by:

$$\hat{q}_l(s) = \frac{\left(E_l(s) - \bar{E}_l(s)\right)^T \left(E_l(s) - \bar{E}_l(s)\right)}{4^s + 2^s - 1}, \tag{29}$$

$$\hat{q}_u(s) = \frac{\left(E_u(s) - \bar{E}_u(s)\right)^T \left(E_u(s) - \bar{E}_u(s)\right)}{4^s + 2^s - 1}, \tag{30}$$

$$\hat{q}_m(s) = \frac{\left(E_m(s) - \bar{E}_m(s)\right)^T \left(E_m(s) - \bar{E}_m(s)\right)}{4^s - 1} \tag{31}$$

where $\bar{\cdot}$ denotes the mean of the argument.

4. NUMERICAL RESULTS AND CONCLUSIONS

The results presented above have been applied to a variety of real images from a number of different sources,

[2] We note that we have ignored an analogous redundancy in the measurement equations (14) which lead to an estimate of $A(s)$.

including SAR, elevation maps, and satellite imagery. The results show that, overall, the above proposed modeling framework is very rich and is capable of modeling the spatial characteristics of a wide variety of random fields. In particular, those of images of natural scenes. Preliminary analysis has also shown that the detail images[3] resulting from this framework can in many cases be compressed efficiently due to resulting small spatial correlations and very Gaussian-like aggregate histogram statistics. In addition, the inherently efficient and simple process of calculating the detail images makes them extremely amenable to target detection problem where the multiscale model may be used to model the background, hence enhancing the target visibility within the resulting detail images. Due to limited space, below we present only one simple application of our modeling framework to the target detection problem.

Figure 3 shows a 257 × 257 image of a helicopter against a natural background. The corresponding estimated multiscale parameters are shown in Figure 4. We can see that these parameters evolve in a non-trivial fashion as a function of scale. In particular, the evolution of the $B(s)$ parameters shows that the detail signals have monotonically smaller variances between scales 4 and 8. In fact, these estimated variances decay rather quickly as a function of increasing scale.

In the finest scale detail signal shown in Figure 5, it is apparent that the size of the detail is relatively large where the "target" (chopper) was present. This indicates that the multiscale modeling framework proposed here has captured the *background* reasonably well. Hence, the chopper, being different from the background stands out in the detail image (at the finest scale). To further illustrate the point, consider the cross correlation of the columns of the resulting detail image in Figure 5. This is plotted as an image itself in Figure 6. Large correlation is evident where the chopper was present in the original image.

[3]The detail image d_s at scale s is obtained by 1) downsampling the given image f to scale $s-1$, yielding f_{s-1}. 2) using $A(s)$, interpolating f_{s-1} to scale s yielding $\hat{f}_s$. The detail is then give by $d(s) = f_s - \hat{f}_s$.

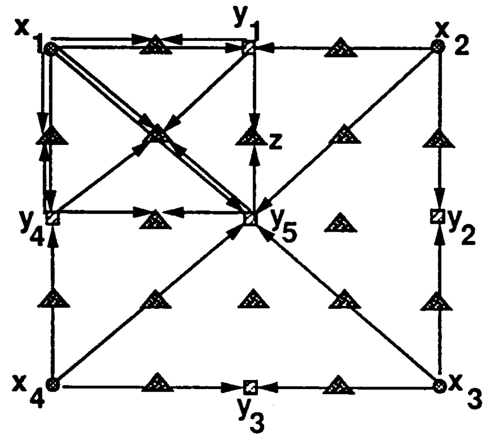

Figure 1: Midpoint Construction Process

5. REFERENCES

[1] C. Bouman. A multiscale image model for bayesian image segmentation. Technical report, Purdue University, School of Electrical Engineering, Dec 1991. TR-EE-91-53.

[2] K.C. Chou, A.S. Willsky, A. Benveniste, and M. Basseville. Recursive and iterative estimation algorithms for multiresolution stochastic processes. In *Proceedings of IEEE Conference on Decision and Control*, September 1989.

[3] V. V. Digalakis and K. C. Chou. Maximum likelihood identification of multiscale stochastic models using the wavelet transform and the EM algorithm. In *Proceedings of IEEE ICASSP.*, 1993.

[4] M.R. Luettgen, W.C. Karl, A.S. Willsky, and R.R. Tenney. Multiscale representation of markov random fields. *IEEE Transactions on Signal Processing*, 41(12):3377–3396, December 1993.

[5] Stephane Mallat. A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

[6] Louis L. Scharf. *Statistical Signal Processing.* Addison Wesley, 1991.

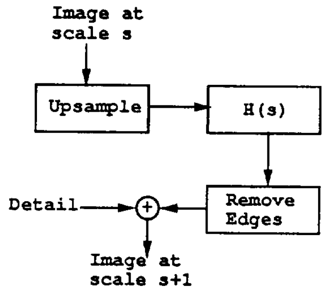

Figure 2: The filterbank interpretation.

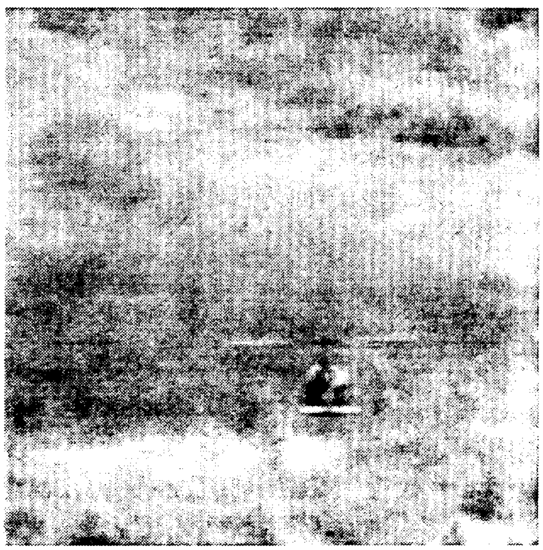

Figure 3: Chopper Image

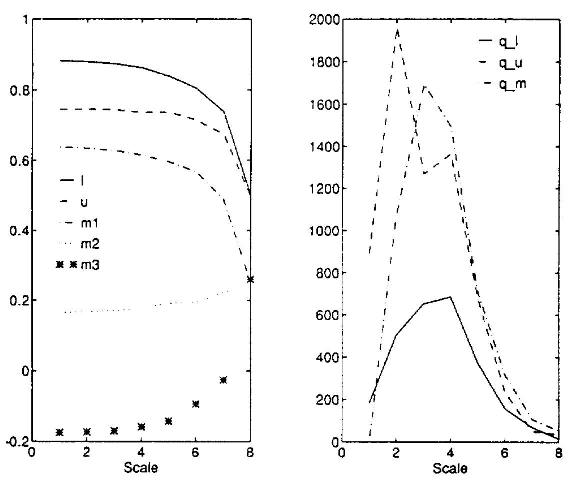

Figure 4: Estimated multiscale parameters

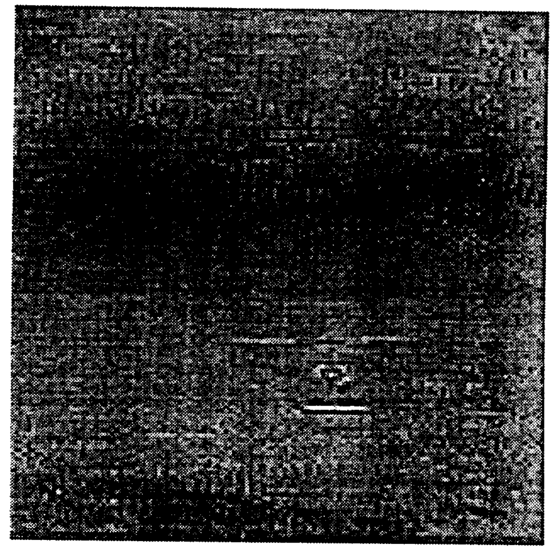

Figure 5: Detail image at the finest scale

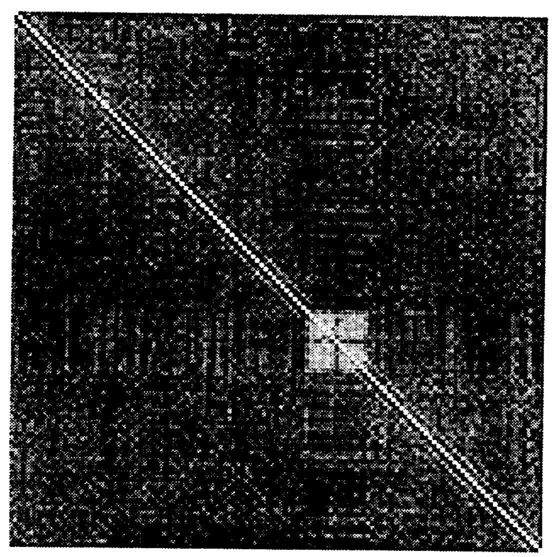

Figure 6: Horizontal correlation matrix