Fractional Brownian Motion: A Maximum Likelihood Estimator and Its Application to Image Texture

TORBJÖRN LUNDAHL, WILLIAM J. OHLEY, MEMBER, IEEE,
STEVEN M. KAY, MEMBER, IEEE, AND ROBERT SIFFERT

Abstract—**Fractals have been shown to be useful in characterizing texture in a variety of contexts. Use of this methodology normally involves measurement of a parameter H, which is directly related to fractal dimension. In this work the basic theory of fractional Brownian motion is extended to the discrete case. It is shown that the power spectral density of such a discrete process is only approximately proportional to $|f|^a$ instead of in direct proportion as in the continuous case. An asymptotic Cramer–Rao bound is derived for the variance of an estimate of H. Subsequently, a maximum likelihood estimator (MLE) is developed to estimate H. It is shown that the variance of this estimator nearly achieves the minimum bound. A generation algorithm for discrete fractional motion is presented and used to demonstrate the capabilities of the MLE when the discrete fractional Brownian process is contaminated with additive Gaussian noise. The results show that even at signal-to-noise ratios of 30 dB, significant errors in estimation of H can result when noise is present. The MLE is then applied to X-ray images of the human calcaneus to demonstrate how the line-to-line formulation can be applied to the two-dimensional case. These results indicate that it has strong potential for quantifying texture.**

INTRODUCTION

TEXTURE analysis can be applied to many different types of images. One particular textural model which has received recent interest, especially for its ability to model "natural" shapes and forms is that of fractals.

Fractals are a class of mathematical functions which have been used to characterize the geometrical properties of sets. They have been used to generate images [1], [4], [5], classify texture [9], [10], interpolate images [11], to describe the surface area of chemical reactants [12], and the length of coastlines [8].

The definition of a fractal is a set for which the Hausdorff–Besicovich dimension is strictly greater than the topological dimension [9]. Thus, a defining characteristic is that a fractal has fractional dimension. The fractal model is used to relate a metric property such as length of a line or area of a surface to the elemental length or area used as a basis for the calculation. For example, the length of a coastline may be determined by placing a 1 km ruler end-to-end along the shore line. If a 0.5 km ruler is used for the same coast, then the measured length will be longer. If the increase in length follows a consistent rule over a range of elemental rulers, then it may be called a measure of the coastline's geometrical properties. This property is found not only in coastlines but also in other naturally occurring functions [4]. The functional relationship between ruler size and length is

$$L = \lambda \epsilon^{1-D} \tag{1}$$

where

L = total length
ϵ = elemental ruler length
D = fractal dimension
λ = scaling constant.

In practice, D has been shown to correlate with the function's intuitive roughness. For $D = 1.0$, the curve is a smooth line, while for $D = 1.99$, the line is extremely rough.

While the above equation has been directly applied to lines and surfaces, in many applications it is appropriate to consider the function as a stochastic process. In order to accomplish this, fractional Brownian motion (FBM) has been proposed. This is a generalization of ordinary Brownian motion.

In FBM there is only one parameter of interest: H, which describes the roughness of the paths. H has been shown to be related directly to D, the fractal dimension, by

$$D = 2 - H \tag{2}$$

for a line-to-line function. The realization of FBM therefore describes a fractal set.

The problem in many applications of fractals is to estimate the parameter H from a given dataset. Many different estimators have been proposed [9], [10]. However, the estimators have been generally applied without any real verification of their statistical properties. Thus, a basic question is: what is the standard error of a computed fractal dimension?

In this study we develop an answer to this question by beginning with some basic definitions. This leads to the consideration of fractional Gaussian noise (FGN). The power spectral density (PSD) of FGN is computed and used to obtain an asymptotic Cramer–Rao bound for the variance of the estimator for H. Following this, a maximum likelihood estimator (MLE) is derived for H. To test

Manuscript received March 6, 1986.
T. Lundahl, W. J. Ohley, and S. M. Kay are with the Department of Electrical Engineering, University of Rhode Island, Kingston, RI 02881.
R. Siffert is with the Department of Orthopaedics, Mount Sinai School of Medicine, New York, NY 10029.
IEEE Log Number 8610145.

the estimator, a generative process for FGN is also developed using the probability density function (PDF) of the process. The MLE is then compared to another estimator for both noise-free and noise-corrupted data. Finally, the estimator is shown as a means for quantifying the texture in planar X-ray images of bone.

Fractional Brownian Motion

The definition of FBM was proposed by Mandelbrot and Van Ness [7]:

Let $B(t)$ be an ordinary Brownian motion with strength q such that for nonoverlapping intervals (t_0, t_1), (t_2, t_3)

$$B(0) = 0$$

$$B(t) - B(0) \quad \text{is contained within the set} \quad N(0, qt)$$

$$E[(B(t_1) - B(t_0))(B(t_3) - B(t_2))] = 0 \tag{3}$$

then FBM is defined as

$$B(0; H) = 0$$

$$B(t; H) - B(0; H) = \frac{q}{\Gamma(H + 0.5)} \left\{ \int_{-\infty}^{0} [(t - s)^{H-0.5} - (-s)^{H-0.5}]\, dB(s) + \int_{0}^{t} (t - s)^{H-0.5}\, dB(s) \right\} \tag{4}$$

where the integrals above are stochastic integrals.

It can be seen from the definition that $B(t; H)$ is a moving average of the increments of $B(t)$. The property of independent increments, which has made Brownian motion so useful, is not valid for FBM. This means also that the Markovian property has disappeared. For $H = 0.5$, FBM is an ordinary Brownian motion. The above definition (4) can be written in a more symmetric form:

$$B(t_2; H) - B(t_1; H) = \frac{q}{\Gamma(H + 0.5)} \left\{ \int_{-\infty}^{t_2} (t_2 - s)^{H-0.5} \cdot dB(s) - \int_{-\infty}^{t_1} (t_1 - s)^{H-0.5}\, dB(s) \right\}. \tag{5}$$

This formulation is a convergent difference of two divergent integrals.

FBM is a time-varying process which makes analysis difficult. However, the increments of FBM are a strict sense stationary process which has been termed fractional Gaussian noise (FGN). This is done by way of analogy with the notation for Brownian motion. Many of the properties of FGN have already been stated in [7]. Briefly, they are as follows.

1) The increments of FBM have zero mean.
2) The increments of FBM are strict sense stationary.
3) $\text{var}\,[B(t_2; H) - B(t_1; H)] = V_H(t_2 - t_1)^{2H}$

where var is the variance and V_H is a constant given by

$$V_H = \frac{q}{\Gamma(H + 0.5)^2} \left\{ \int_{-\infty}^{0} [(1 - s)^{H-0.5} - (-s)^{H-0.5}]^2\, ds + \frac{1}{2H} \right\}.$$

4) $$\Pr\left\{ \frac{B(t_2; H) - B(t_1; H)}{(t_2 - t_1)^H} < y \right\} = F(y)$$

where $F(y)$ is a Gaussian cumulative distribution function with zero mean.

5) $B(t + k; H) - B(t; H)$ is statistical self similar meaning that

$$B(t + k; H) - B(t; H) \triangleq h^{-H}[B(t + k \cdot h; H) - B(t; H)]$$

where $\triangleq$ means identical in distribution and h is a constant.

A. Discrete FBM (DFBM)

Since most of the real data encountered at present is discrete and not continuous, we define discrete Brownian motion

$$B[k; H] = B(k \cdot T_s; H) \tag{6}$$

where T_s is the sampling distance.

To differ between discrete and continuous data, squared brackets are used to denote the former.

Next we define

$$X[k; H] = B[k; H] - B[k - 1; H] \tag{7}$$

where $X[\,]$ denotes discrete fractional Gaussian noise (DFGN).

DFGN is a discrete strict sense stationary process with Gaussian distributed samples. Therefore, the mean and autocorrelation function comprise a sufficient description of the process.

From the definition of FBM it follows that:

$$E\{X[k, H]\} = 0. \tag{8}$$

Thus, $r[k]$, the autocorrelation function, can be written

$$r[k] = \frac{\sigma^2}{2} [|k + 1|^{2H} - 2 \cdot |k|^{2H} + |k - 1|^{2H}] \tag{9}$$

where σ^2 = the variance of $X[k]$.

This follows by considering the quantity

$$r[k] = E\{X[n + k; H]\, X[n; H]\}$$

and substituting by use of (7)

$$r[k] = E\{[B((n + k + 1)T_s; H) - B((n + k)T_s; H)]\, [B((n + 1)T_s; H) - B(nT_s; H)]\}$$

expanding

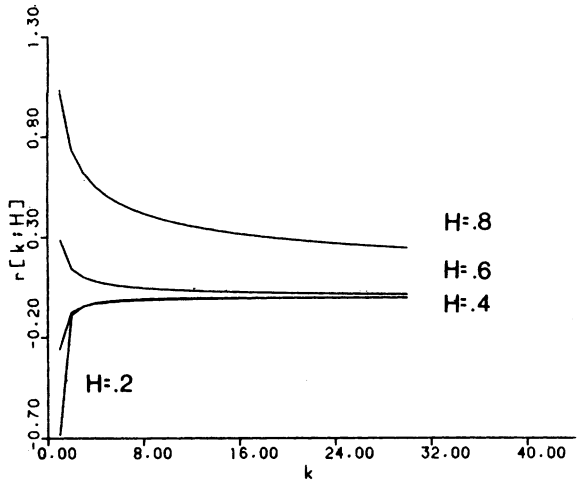

Fig. 1. The autocorrelation function for discrete fractional Gaussian noise (DFGN) $r[k; H]$ for $k = 1$ to 30 and $H = 0.2, 0.4, 0.6, 0.8$. The value at $k = 0$ was not plotted since it is the same for all values of H, this also enables the shape of the function be seen.

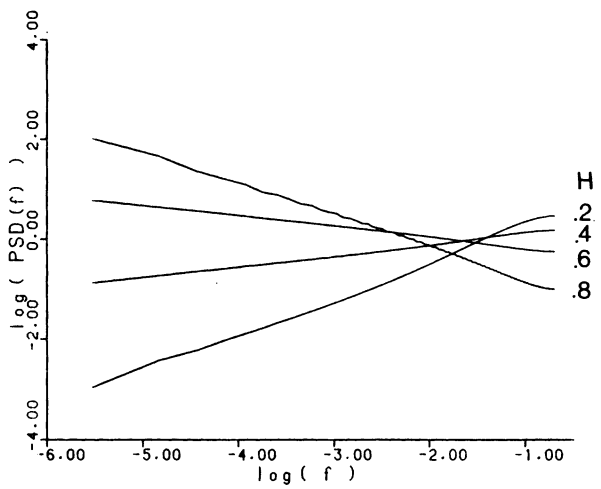

Fig. 2. The numerically computed power spectral density (PSD), equation (10), for discrete fractional Gaussian noise (DFGN). The PSD is shown for $H = 0.2, 0.4, 0.6, 0.8$.

$$r[k] = \tfrac{1}{2}\{E[(B((n+k+1)T_s; H) - B((n+k)T_s; H))^2] + E[(B((n+k)T_s; H) - B(nT_s; H))^2] - E[(B((n+k+1)T_s; H) - B((n+1)T_s; H))^2] - E[(B((n+k)T_s; H) - B(nT_s; H))^2]\}$$

and (9) follows directly.

Fig. 1 shows $r[k]$ for several values of H. The higher correlation for larger values of H is in keeping with the notion that the realizations for large H should be smoother than for lower H.

Power Spectral Density (PSD)

It has been shown in [7] that the PSD of the continuous FGN is proportional to $|f|^{(1-2H)}$. Thus, the slope in a log-log scale is $(1 - 2H)$. When FGN is sampled to obtain the DFGN, there is an obvious problem with aliasing.

To obtain the PSD of DFGN one has to take the Fourier transform of the autocorrelation function

$$p(f) = r[0] + 2\sum_{k=1}^{\infty} r[k]\cos(2\pi fk) \qquad (10)$$

where $r[k]$ is given in (9).

Several problems make the above expression difficult to evaluate analytically. First, $r[k]$ consists of a convergent difference of divergent terms. Second, $p(f)$ tends to be unbounded at $f = 0$ and $f = 0.5$. Although no explicit expression was found for the PSD, the following statements can be made.

The value of $p(0)$ is given by

$$p(0) = r[0] + \lim_{n\to\infty} 2\sum_{k=1}^{n} r[k] = r[0] + \lim_{n\to\infty} 2 \cdot s(n) \qquad (11)$$

where

$$s(n) = \sum_{k=1}^{n} r[k]$$

$$s(n) = \frac{\sigma^2}{2}[(n+1)^{2H} - 1 - n^{2H}]$$

$$s(n)/n^{2H} = \frac{\sigma^2}{2}\left[\left(1 + \frac{1}{n}\right)^{2H} - \frac{1}{n^{2H}} - 1\right] \approx \frac{\sigma^2}{2}\left[\frac{2H}{n} - \frac{1}{n^{2H}}\right]$$

when n is large

$$s(n) \approx \frac{\sigma^2}{2}[2Hn^{2H-1} - 1]$$

one gets

$$p(0) = \begin{cases} 0 & 0 < H < 0.5 \\ r[0] & H = 0.5 \\ \infty & 0.5 < H < 1. \end{cases} \qquad (12)$$

This indicates that the process is very different for the three different cases above. The first case is a kind of high-frequency process, the second is discrete white noise, and the last is a low-frequency process.

The expression for the PSD, (10) was evaluated numerically. The sum was truncated when $|r[k]| < 0.005$ or after 10 000 terms. However, at least 100 terms were always included in the sum. The results of this calculation are shown in Fig. 2.

In general, the PSD is very much a straight line with the slope a function of the parameter H. The ripple seen for $H = 0.8$ is most likely due to a truncation artifact (Gibbs-type phenomena). However, the slope of the line does not correspond to exactly $(1 - 2H)$ as in the continuous case. Linear regression was used to compute the

TABLE I
RESULTS OF LINEAR REGRESSION ON log (PSD) VERSUS log (f)

True H	H by PSD Slope	r
0.1	−0.05	0.9943
0.2	0.12	0.9948
0.3	0.26	0.9955
0.4	0.38	0.9958
0.5	0.5	—
0.6	0.61	0.9965
0.7	0.72	0.9971
0.8	0.83	0.9970
0.9	0.93	0.9011

slope of the family of PSD's. The result is presented in Table I.

B. Asymptotic Cramer–Rao Bound

In general, the exact Cramer–Rao bound is very difficult to evaluate. Thus, a conventional approach is to make use of asymptotic statistics. The problem can be stated as follows. Given a discrete PSD proportional to $|f|^a$ for $f < 0.5$, what is the bound of the variance of the estimate of a? Using the result given in [13] this bound can be expressed as

$$\text{var}\,(\hat{a}) \geq \frac{2}{N} \cdot \frac{1}{\displaystyle\int_{-0.5}^{0.5} \left[\frac{d \ln p(f)}{da}\right]^2 df} \tag{13}$$

where N is the number of data points

$\hat{a}$ is the estimate of a.

Evaluating (13) gives

$$\text{var}\,(\hat{a}) \geq 1/2N \tag{14}$$

since it is known that a is almost equal to $1 - 2H$, the variance of the estimate of H is

$$\text{var}\,(\hat{H}) \geq 1/8N. \tag{15}$$

How large N has to be for this result to be valid is impossible to say in an analytical sense since it depends on the PDF of the data. However, simulations as presented in the subsequent section show that at $N = 100$, the variance of the maximum likelihood estimator is approximately equal to this asymptotic bound. It is also interesting to note that the bound is independent of the true value of H. This implies that the relative estimation error (percent error) will be a function of H.

THE PROBABILITY DENSITY FUNCTION AND ITS APPLICATIONS

In this section, the probability density function of DFGN is introduced. Using this function, first an algorithm for generation of sample paths is developed; second a maximum likelihood estimator (MLE) is developed. The MLE is then applied to generated data with known H. The result shows that the MLE nearly obtains the asymptotic Cramer–Rao bound. Since the asymptotic statistics of a maximum likelihood estimate are known [14], a confidence interval for the estimates is then stated.

A. The Probability Density Function

The probability density function can be expressed straightforwardly, since all samples are jointly Gaussian distributed and the covariance matrix is known.

$$p(\boldsymbol{x};\, H) = \frac{1}{(2\pi)^{N/2}|R|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{x}^T R^{-1}\boldsymbol{x}\right\} \tag{16}$$

where $\boldsymbol{x} = \{x_0 x_1 \cdots x_{(N-1)}\}^T$ is the dataset and R is the covariance matrix.

$$R = E[\boldsymbol{x}\boldsymbol{x}^T] \quad \text{or} \quad [R]_{ij} = r[|i - j|] \tag{17}$$

where $r[k]$ is the autocorrelation function given in (9). Note that the dependence on H is incorporated into R. The likelihood function was evaluated on a dataset using values of $N = 30, 40, 50, 60$ and for $H = 0.4$ and 0.6. The dataset was generated according to the algorithm given in the following section.

The results of the evaluation are given in Fig. 3. As can be seen, the curve becomes more peaked as N increases. However, most importantly the function is unimodal. Thus, when a maximum is found, it is unique.

B. Generative Algorithm

Different methods to generate FBM have been proposed [1], [6]. The problem is that one cannot generate the samples independently. In addition, because of the strong correlation, $r[k]$ cannot be truncated. It is preferable to uncorrelate the samples by a transform, realize the samples, and transform back again to the desired covariance matrix. Since the correlation matrix is positive definite, it is a true covariance matrix for a wide sense stationary process. It can be decomposed as

$$R = LL^T \tag{18}$$

where L is a lower triangular matrix.

Since R is positive definite, L is positive definite as well, and specifically L is invertible. L was determined by use of Cholesky decomposition [14].

Consider the transformation

$$\boldsymbol{y} = L^{-1}\boldsymbol{x} \Leftrightarrow \boldsymbol{x} = L\boldsymbol{y} \tag{19}$$

then

$$E[\boldsymbol{y}\boldsymbol{y}^T] = E[L^{-1}\boldsymbol{x}\boldsymbol{x}^T L^{-T}] = I \tag{20}$$

where I is the identity matrix.

This shows that the y components are uncorrelated and have unit variance. The y components can therefore easily be generated using a Gaussian random generator, and x is obtained by (19). The DFBM is then generated by adding

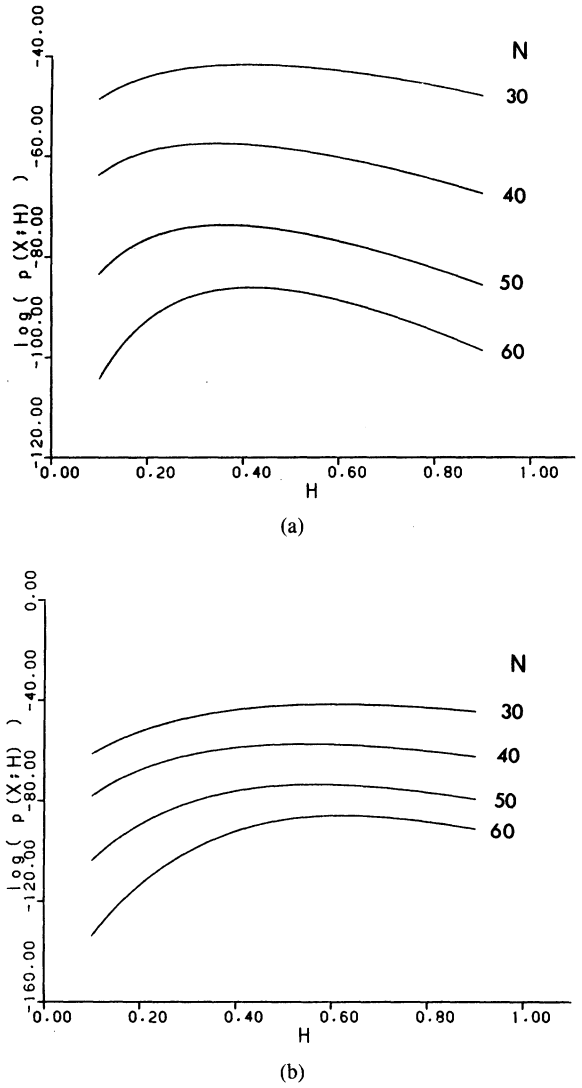

Fig. 3. The logarithm of the likelihood function with $N = 30, 40, 50, 60$. (a) $H = 0.4$, (b) $H = 0.6$.

up the increments

$$B[n; H] = \sum_{k=0}^{n} x[k; H]. \tag{21}$$

One hundred datasets were generated for each value of $H = 0.2, 0.4, 0.6, 0.8$. Each set contained one hundred data points. A sample for each different H value can be seen in Fig. 4. DFGN does not appear to be greatly different for different H values but DFBM is very different. As expected, the function is quite smooth for $H = 0.8$ but very rough for $H = 0.2$.

C. Maximum Likelihood Estimator (MLE)

The MLE is often considered to be the best obtainable estimator. In addition, the estimate is asymptotically unbiased, asymptotically efficient, i.e., it obtains the Cramer–Rao bound, and is asymptotically Gaussian distributed.

In most cases, as in this case, it is impossible to obtain an explicit form for the estimate as a function of the data. Instead, numerical methods are used to find the maximum of the likelihood function. As indicated above, $p(x; H)$

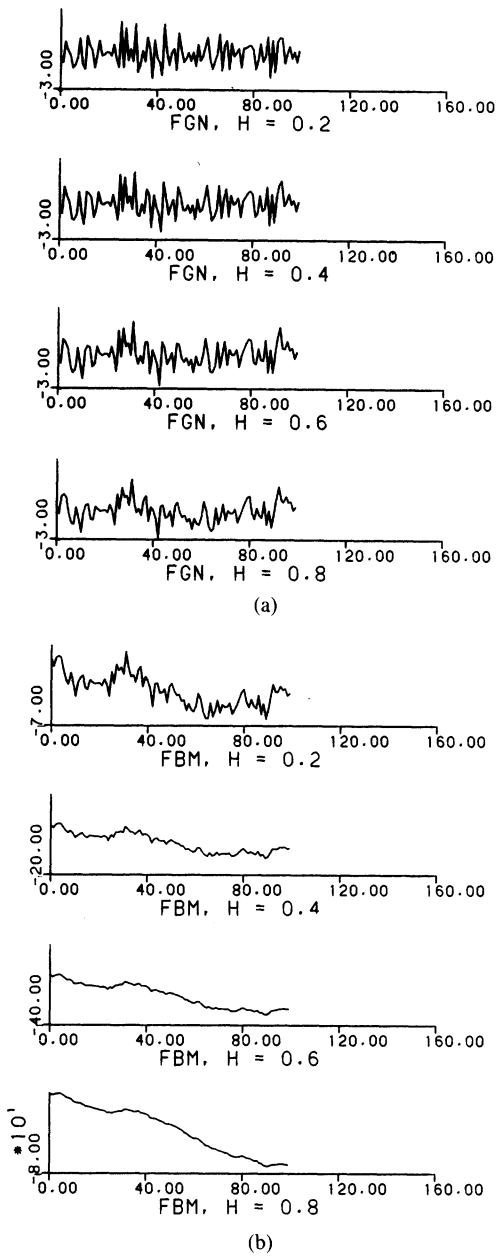

Fig. 4. Samples of the generated datasets for $H = 0.2, 0.4, 0.6, 0.8$. (a) Discrete fractional Gaussian noise (DFGN). (b) Discrete fractional Brownian motion (DFBM). Note that the y-axis scales are not identical for each value of H.

appears to be unimodal. This property simplifies the search for a maximum. Since the logarithm is a monotonic function, the log of $p(x; H)$ can be maximized instead of $p(x; H)$. From (17)

$$\log p(\boldsymbol{x}; H) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\log |R| - \frac{1}{2}\boldsymbol{x}^T R^{-1}\boldsymbol{x} \tag{22}$$

where the elements of R are given by

$$[R]_{ij} = r[|i-j|] = \frac{\sigma^2}{2}\left[|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right] \quad \text{for } k = |i-j|. \tag{23}$$

There are really two parameters to be estimated: H and

σ^2. R can be decomposed as

$$R = \sigma^2 R'. \tag{24}$$

Using this, the likelihood function becomes

$$\log p(x; H) = -\frac{N}{2}\log 2\pi - \frac{N}{2}\log \sigma^2 - \frac{1}{2}\log |R'| - \frac{1}{2\sigma^2} x^T R'^{-1} x \tag{25}$$

(25) can then be minimized over σ^2 by taking the derivative with respect to σ^2, and letting the derivative go to zero.

$$\hat{\sigma}^2 = x^T R'^{-1} x/N. \tag{26}$$

Inserting this in (26) gives the final function to be maximized over H

$$\log p(x; H) = -\frac{N}{2}\log 2\pi - \frac{N}{2}\log (x^T R'^{-1} x/N) - \frac{N}{2} - \frac{1}{2}\log |R'|. \tag{27}$$

One good thing about the search is that the range of H is limited to the interval [0, 1]. Some notes how to evaluate (27) can be worth mentioning. The problem is the matrix R'. This matrix has to be inverted and the determinant has to be computed. Since R' is positive definite, symmetric, and Toeplitz, Levinson's algorithm [14] can be used to decompose R' as

$$R'^{-1} = L \cdot D^{-1} L^T \tag{28}$$

where L is a lower triangular matrix with one's on the diagonal and D is a diagonal matrix with positive elements on the diagonal.

The determinant of R' is then equal to the determinant of D. By applying the Levinson algorithm both the inverse of R' and the determinant are computed in one step. The maximum of (27) was then found by using the Golden search method [15].

The MLE estimator was applied to all of the generated datasets. This yielded 100 estimates of H for true H = 0.2, $\cdots$, 0.8. The results of the estimation can be seen in Table II where the true H, the mean of the estimates, and the standard deviation of the estimates are listed. Note that for N = 100, by (15), var $(\hat{H}) \geq 0.035$.

Assuming that the estimates are normally distributed with mean equal to the true H and standard deviation equal to $0.5/\sqrt{N}$, a 99 percent confidence interval is given by

$$\Pr(\hat{H} - 2.6\sigma < H < \hat{H} + 2.6\sigma) = 0.99 \tag{29}$$

where σ is the standard deviation of the estimate of H, $\hat{H}$.

The MLE results were directly compared to results using a previously described estimator [9]. This H estimator employs a property that the variance of the increments of the process increases with i as

$$\text{var}\,(B[k + i; H] - B[k; H]) = i^{2H} \text{var}\,(B[k + 1; H] - B[k; H]). \tag{30}$$

TABLE II
RESULTS FROM THE APPLICATION OF THE MAXIMUM LIKELIHOOD ESTIMATOR (MLE) TO THE GENERATED DATASETS

True H	Mean of Hest	STD of Estimates
0.2	0.200	0.05
0.4	0.399	0.06
0.6	0.599	0.06
0.8	0.796	0.06

TABLE III
RESULTS FROM THE VARIANCE ESTIMATOR APPLIED TO THE DATASETS

True H	Mean of Hest	STD of Estimates
0.2	0.196	0.06
0.4	0.392	0.08
0.6	0.591	0.09
0.8	0.778	0.10

First, the variance for different lags, i, is estimated. Then H is determined by regression from the slope of the variance plotted as a function of the lag in a log–log scale.

This estimator was applied to the same dataset as the MLE. The results are given in Table III using the same format as for the MLE in Table II.

NOISE CORRUPTED DATA

In the real world one never has a clean signal. If no other noise is present in digital data, there is always quantization noise. One might then ask how much of an effect does noise have on the estimates of H. In the approach taken here we assume that the noise is zero mean white Gaussian and added to the signal.

$$y[k] = B[k; H] + w[k]. \tag{31}$$

Because of the high degree of nonlinearity in the estimator there is no convenient analytical way of predicting the effect of noise. Although the Cramer–Rao bound could be found by use of (13), here we have taken the simulation approach. Noise was added to the signals generated previously for signal-to-noise ratio (SNR) = 10, 20, and 30 dB defined as

$$P_S = \frac{1}{N}\sum_{k=1}^{N} B[k; N]^2 \tag{32}$$

$$P_N = \sigma_w^2, \text{ i.e., the variance for the white noise} \tag{33}$$

$$\text{SNR} = 10 \log (Ps/Pn). \tag{34}$$

Samples of the realizations are shown in Fig. 5.

As in the previous section, both the MLE for the noise-free case and variance estimator method were applied to the datasets. Note that under these conditions the use of the term MLE is a misnomer, since the estimator has not considered the noise.

The results of these estimations are presented in Table IV. Both estimators seem to be unbiased, but the variance estimator has about 50 percent higher standard deviation than the MLE. Also interesting to see is that already at

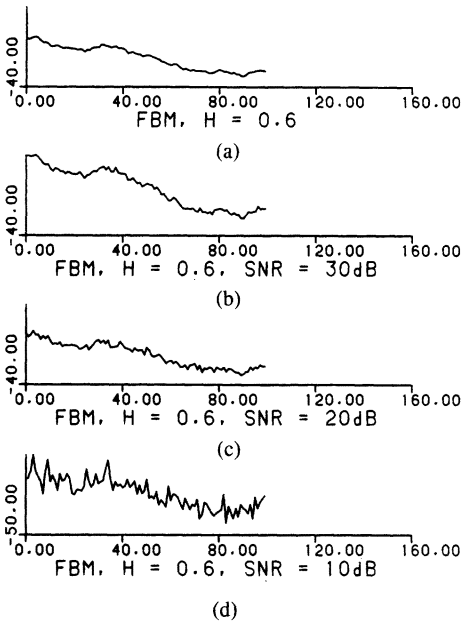

Fig. 5. Samples of discrete fractional Brownian motion (DFBM) with white Gaussian noise. (a) No noise. (b) SNR = 30 dB. (c) SNR = 20 dB. (d) SNR = 10 dB.

TABLE IV
RESULTS FROM TWO ESTIMATORS: THE MLE, AND THE VARIANCE AS APPLIED TO NOISE-CORRUPTED DATA

	MLE		Variance	
	Hest	STD	Hest	STD
SNR = 30 dB				
H = 0.2	0.204	0.05	0.194	0.06
H = 0.4	0.392	0.06	0.385	0.08
H = 0.6	0.532	0.06	0.550	0.08
H = 0.8	0.592	0.08	0.660	0.07
SNR = 20 dB				
H = 0.2	0.194	0.05	0.184	0.05
H = 0.4	0.330	0.06	0.331	0.07
H = 0.6	0.377	0.08	0.405	0.09
H = 0.8	0.366	0.09	0.406	0.12
SNR = 10 dB				
H = 0.2	0.136	0.05	0.125	0.05
H = 0.4	0.181	0.06	0.179	0.08
H = 0.6	0.176	0.07	0.167	0.10
H = 0.8	0.161	0.06	0.138	0.09

SNR = 30 dB the estimates for high H are affected by the noise. The effect seems to be about the same for both estimators and indicates that noise must be either filtered out or included in the estimation model.

A. Application to Image Texture

The MLE for fractional Brownian motion presented above is for line-to-line functions. It might be expected that a two-dimensional extension of the estimator is necessary to apply it directly to images. However, as other investigations have shown [9], useful image descriptors can be formed by using the line-to-line analysis. Accordingly, the MLE was applied as an example to X-ray images of bones.

X-ray images of bones generally exhibit a variety of textures which have been resistant to automated methods of classification. Four images were acquired as 256 × 256 × 8 bit pixels. The images [Fig. 6(a)–(d)] were taken from a normal human calcaneus (heel) during the course of immobilization for fracture and subsequent recovery. The regions are approximately a 2 × 2 cm section in the center of the heel. They are: a) before injury and casting; b) two months of immobilization; c) three months post injury (six weeks of partial weight bearing) d) nine months after injury. In a), very fine, coarse structural patterns are present. In b), there is a slight degradation in the coarseness of the structure. This is typical of the process of osteoporosis, in which calcium is dissolved from the bone structure following immobilization. It is continued in c) with further loss. In d) the calcaneus has undergone ap-

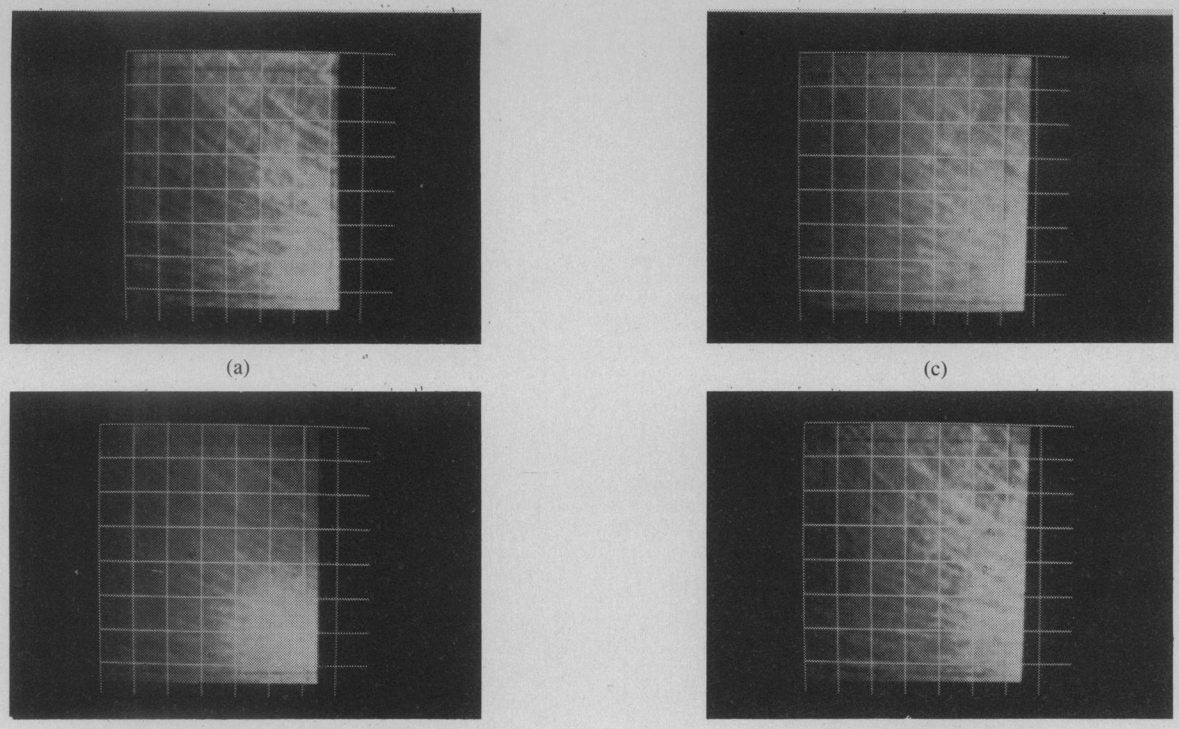

Fig. 6. X-ray images of human calcaneus sections. (a) Normal. (b) Same bone after two months immobilization in cast. (c) Following six weeks partial weight bearing. (d) Six months following cast removal (nine months post injury). The grid pattern is overlayed to provide a reference frame for regional computations. Because of edge effects in the image within the video frame, the outer two rows and columns were not used. The H calculations in Fig. 7 are for the inner 16 regions only.

proximately six months of mobility. Some of the structure returns but not exactly as before the incident. The films from which the images were taken were not acquired in any special manner, and exhibit typical variations in magnification and intensity. The dark straight lines at the boundary of the images were drawn by clinicians to indicate the same anatomical location.

Each image was broken into 64 regions of 32 × 32 pixels each. The ML estimator was applied to each region along a row, along a column, and along the 45 and 135 degree diagonals. This provided four estimates of H for each region, one for the horizontal, one for the vertical, and two for the diagonals.

The results of these calculations were then displayed as a composite set of plots consisting of four-lined "stars." Each star was produced from the four H estimates in each region. Each line in the "star" corresponds to the H value in that direction. The larger the star, the larger the H in that direction. Because of edge effects in the 256 × 256 image, only the inner 16 regions (4 × 4) are shown in Fig. 7. As evidenced by Fig. 7, there is a clear reduction in H from (a) to (b) and to (c) for most of the regions in the image. To further illustrate this change, a composite average H was formed for each image by averaging all the H values for each image. These results are presented in Table V. The composite index follows closely the clinically expected course of osteoporosis in that a lower H indicates a greater degree of osteoporosis.

Discussion

Fractional Brownian motion has been shown by many authors to be a useful description of many different types of natural phenomena. However, when employing a continuous mathematical framework to discrete data, it is important to provide optimal methods for estimating the correct parameters of the original process. In this work we have extended the notion of FBM into the discrete domain and then developed a method for determining basic parameters with a minimum variance. As can be seen in Table II, the MLE provides a variance which is nearly equal to the asymptotic Cramer–Rao bound. Furthermore, the variance is approximately one half of that obtained with previously reported estimators.

In previous work [9] it had been assumed that the frequency property of continuous FBM which relates the PSD to H by an inverse power law could be applied directly to sampled data. Fig. 2 and Table I clearly show that this will be in error since the discrete PSD does not exactly follow this rule. While it still may be possible to use a frequency based method for estimation, methods for this approach were not included in this investigation. The cause of the errors, while not explicitly elucidated here, are most likely due to aliasing and spectral truncation effects since the spectrum in the continuous case is not band-limited. Most of the results of the work presented here have relied upon the use of generated FBM. There have been other methods proposed to realize FBM. However,

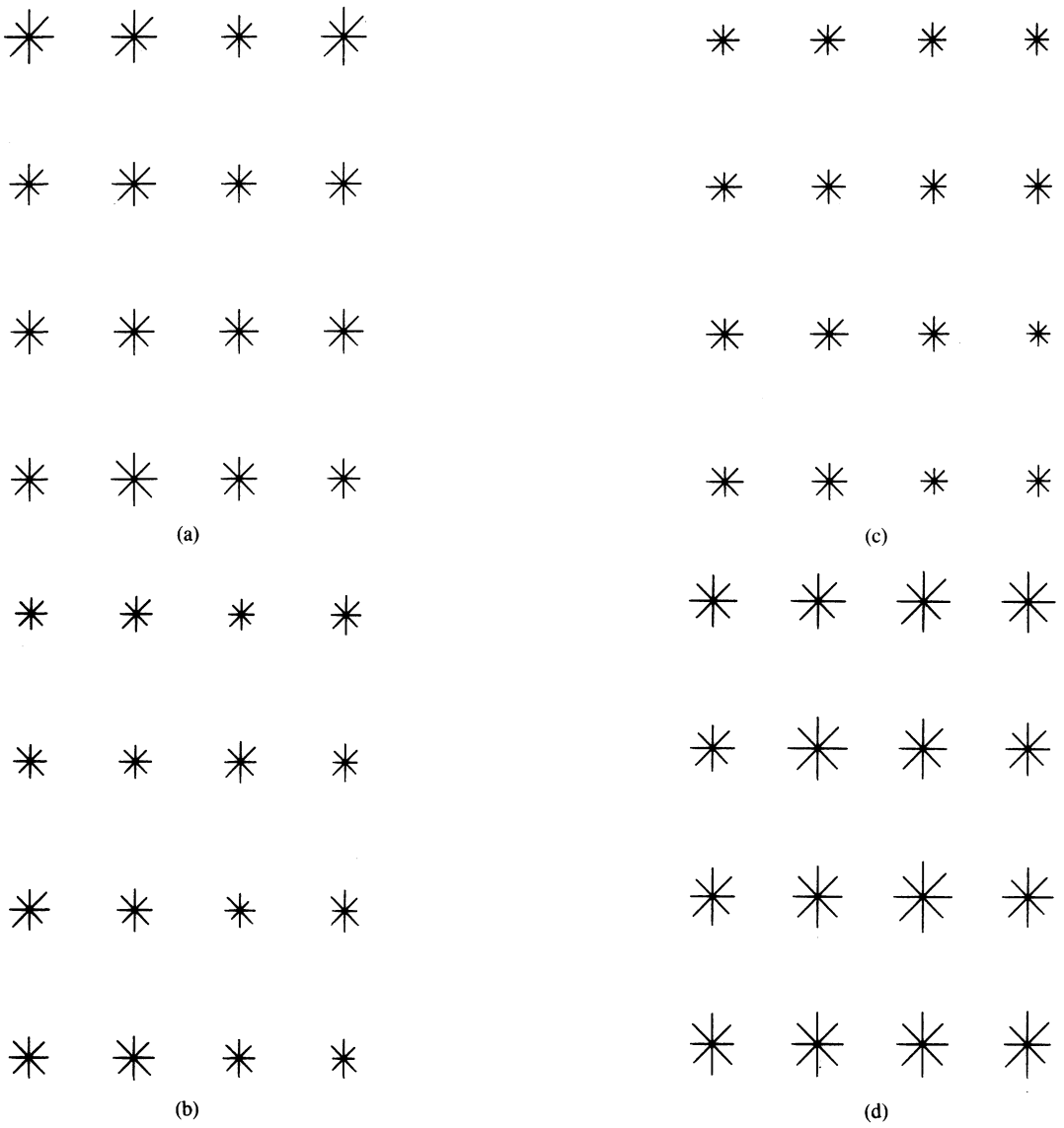

Fig. 7. Graphical presentation of the results seen in Table V, each star pattern represents the H parameter values in the given direction for the inner 16 regions of Fig. 6(a)-(d). The line segment length is proportional to H. Averages of these results were used to produce Table V.

TABLE V
THE COMPOSITE H ESTIMATES OBTAINED BY AVERAGING THE RESULTS FROM FIG. 7, THUS THIS REPRESENTS THE AVERAGE VALUE OF H FOR EACH IMAGE IN FIG. 6

Image	Condition	Composite H
(a)	Preinjury	0.422
(b)	Two months cast	0.358
(c)	Three months cast	0.325
(d)	Six months no cast	0.552

the formulation shown above is straightforward and simple to implement. Even though it uses an extensive amount of computer time, this is more than offset by the decrease in software development time.

An interesting outcome of the simulations performed with the generated data is the performance of H parameter estimators in the presence of additive Gaussian noise. It is difficult to imagine a set of experimental data that is not in part contaminated with noise. In the case of FBM, it is seen that noise can have a large effect on the data; furthermore, the effect is not uniform. Instead, the resulting errors in estimation are greater for higher values of H. Since a high value of H implies that the line is smooth, when noise is added, the line appears to be rougher. Thus, the estimates in these conditions give lower values of H than the uncontaminated values. Exactly why the low values are not as bad off is probably related to the scale compression for values of H. By definition they must lie in the interval $[0,1]$.

The performance of the estimator in noise is important especially if one is attempting to make structural inferences about an object from which the data were acquired. If noise is considered to be present, then it must be either removed from the data before processing, or incorporated into the estimation model. Either of these approaches should provide superior results to those shown here. Such formulations are the subject of future investigations.

In several portions of this work, both the MLE and variance estimator are compared. In all cases the results with the MLE are better. However, due to the greatly reduced computations for the latter case, the variance estimator may be the best choice in some applications. However, in the sample image analysis, the increased variance of this estimator could potentially contribute greatly to the scatter of H values even for adjacent and visually similar textures.

Even though the results presented here are for line-to-line functions, it has been shown how they can be applied directly to images by use of a directional H estimate. This not only provides a measure of the absolute value of H but tells of the image's uniformity as well. The results shown in Fig. 7 and in Table V are also in keeping with the physical process which was used to alter the bone structure. Following a period of immobilization, some of the fine trabecular structure has been dissolved away, while the coarse structure is left behind. The original bone, (a) exhibited large H values, indicating that there is a long-term spatial dependency between the coarse and fine structure. When the fine structure is removed, the spatial long-term dependency is reduced. Hence, a lower H is found in (b) while still lower value is seen in (c). Following the cast removal and remobilization, the composite H index returns to a larger value. This preliminary analysis indicates that such methodology may be extremely powerful in the quantification of osteoporosis.

Certainly, there are other texture analysis methods which would provide an osteoporotic index for the four images shown here. However, the estimation of H, theoretically, is immune to additional "noise" processes which may not be obvious. For example, the ordinary clinical X-ray image is not taken with exact scale or positional references, nor is film exposure always constant. The parameter, H, over limited scale ranges is independent of scale (self similarity). In addition, it has been shown that H is independent of changes in the mean value of the data (H is a second-order statistic) [9]. While we have not demonstrated these properties to be true in an experimental sense, there is a firm basis for continued investigation.

To further enhance the utility of this approach in examining texture patterns, the results must be expanded to two or more dimensions, however, this may not be such a trivial task, especially when we also seek its use as a practical solution to real problems.

REFERENCES

[1] A. Fournier, D. Fussell, and L. Carpenter, "Computer rendering of stochastic models," *Commun. ACM*, vol. 25, no. 6, pp. 371-384, 1982.

[2] B. Hughes, E. Montroll, and M. Schlesinger, "Fractal random walks," *J. Statist. Phys.*, vol. 28, no. 1, pp. 111-126, 1982.

[3] S. Liu, "Fractal model for the AC response of a rough interface," *Amer. Phys. Soc.*, vol. 55, no. 5, pp. 529-532, 1985.

[4] B. B. Mandelbrot, *The Fractal Geometry of Nature.* New York: Freeman, 1982.

[5] —, *Fractals: Form, Chance and Dimension.* New York: Freeman, 1977.

[6] —, "A fast fractional Gaussian noise generator," *Water Resources Res.*, vol. 7, no. 3, pp. 543-553, 1971.

[7] B. B. Mandelbrot and B. J. Van Ness, "Fractional Brownian motion, fractional noises and applications," *SIAM*, vol. 10, no. 4, pp. 422-438, 1968.

[8] B. B. Mandelbrot, "Stochastic models for the Earth's relief, the shape and the fractal dimension of the coastlines and the number-area rule for islands," *Proc. Nat. Acad. Sci.*, vol. 72, no. 10, pp. 3825-3828, Oct. 1975.

[9] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Trans. Pattern Anal. Machine Intell.*, no. 6, pp. 661-674, Nov. 1984.

[10] S. Peleg, J. Naor, R. Hartley, and D. Avnir, "Multiple resolution texture analysis and classification," *IEEE Trans. Pattern Anal. Machine Intell.*, no. 4, pp. 518-523, July 1984.

[11] T. Lundahl, W. J. Ohley, W. S. Kuklinski, D. O. Williams, H. Gewirtz, and A. S. Most, "Analysis and interpolation of angiographic images by use of fractals," *IEEE Comput. Cardiol.*, 1985.

[12] P. Pfeeifer, D. Avinir, and A. Farin, "Ideally irregular surfaces of dimension greater than two in theory and practice," *Surface Sci.*, vol. 126, pp. 569-572, 1983.

[13] P. Whittle, "On stationary process in the plane," *Biometrica*, vol. 41, pp. 434-449, 1954.

[14] S. Kay, *Modern Spectral Estimation.* Englewood Cliffs, NJ: Prentice-Hall, 1986, Ch. 6, to be published.

[15] R. P. Brent, *Algorithms for Minimization Without Derivatives.* Englewood Cliffs, NJ: Prentice-Hall, 1973.