METHODS FOR STOCHASTIC SPECTRAL SYNTHESIS

j.p. lewis

Computer Graphics Laboratory
New York Institute of Technology

ABSTRACT

A number of spectral modeling approaches in the engineering and estimation literature are potentially applicable to stochastic synthesis in computer graphics. Two specific approaches are developed. The orthogonality principle of estimation theory is used to derive a stochastic subdivision construction with specified autocorrelation and spectrum properties; this approach also provides an alternative theoretical basis for the popular fractal subdivision algorithms. A shaped Poisson point process is a second approach which conveniently separates the spectral and graphic modeling problems. Synthetic textures and terrains are presented as a means of visually evaluating the constructed noises.

KEYWORDS: stochastic models, texture synthesis, fractals, terrain modeling.

RESUME

Les méthodes de modèlisation spectrales empruntées aux sciences de l'ingénieur, ou dérivées de la theorie de l'estimation peuvent être appliquees à la synthèse stochastique dans le champ de l'informatique graphique. Deux points de vues sont présentés dans cette communication. A partir du principe d'orthogonalité de la theorie de l'estimation on peut dériver une méthode de subdivision stochastique possèdant certaines spécifications d'autocorrélation et propriétés spectrales; cette approche fournit aussi une base théorique nouvelle pour la construction d'algorithmes de subdivision fractale. Un processus utilisant un filtrage de l'impulsion de Poisson fournit une deuxieme approche, qui permet de déterminer une séparation claire des problemes de nature spectrale de ceux liés a la modèlisation graphique. Les textures synthétiques et les modèles de terrains présentés permettent d'évaluer visuellement les bruits ainsi générés.

MOTS CLEFS: Modèles stochastiques, Textures synthétiques, fractal, modèlisation de terrain.

1. Introduction

Stochastic techniques have assumed a prominent role in the synthesis of complex and naturalistic imagery, for example, [1][2][3][4][5][6][7]. This role has been termed *amplification* [5]: the image modeler specifies a pseudo-random procedure and its parameters; the procedure can then automatically generate the vast amount of detail necessary to create a realistically complex scene. The success of stochastic modeling depends both on its economy and on our ability to construct stochastic models to approximately emulate a variety of phenomena. The full power of stochastic modeling has not been achieved in existing techniques. For example, the widely-used stochastic fractal techniques model only spectra of the form f^{-d}, and thus cannot describe phenomena with scale-dependent detail or directional or oscillatory characteristics.

The problem of modeling a random process ("noise") with an arbitrary spectrum is well understood. Basically, the procedure is to filter an uncorrelated noise (as obtained from a random number generator) to obtain the desired spectrum. The spectrum of the filtered noise is simply the squared magnitude of the transfer function of the filter. Using this synthesis procedure, many of the filtering and spectral analysis approaches described in the literature are potentially applicable to the problem of stochastic modeling in computer graphics. This paper adopts two approaches, optimal mean-square estimation and a shaped point process model, to produce stochastic synthesis algorithms which are computationally suitable for computer graphics applications.

2. Generalized Stochastic Subdivision

The stochastic subdivision construction described by Fournier et. al. [1] may be generalized to synthesize a noise with an arbitrary prescribed spectrum (the generalized subdivision technique is described in more detail in [8]). The basis of the Fournier et. al. construction is a midpoint estimation problem: given two samples considered to be on the noise, a new sample midway between the two is estimated as the mean of the two samples, plus a random deviation whose variance is the single noise parameter. The construction is based on two properties of fractional Gaussian noise:

1) When the values of the noise at two points are known, the expected value of the noise midway between two known points is the average of the two values.

2) The increments of fractional Gaussian noise are Gaussian, with variance which depends on the lag and on the noise parameter.

Since only the immediately neighboring points are considered in making the midpoint estimation, the noise autocorrelation information is not used, and the constructed noise is Markovian. This is not a limitation as long as the construction is used as an approximate (stationary) model for Brownian motion. However, the construction has been applied to the non-Markovian fractional noises f^{-d}, $d \neq 2$ [9]; in these cases, disregarding the autocorrelation produces "creases".

The general problem of estimating the value of a stochastic process given knowledge of the process at other points is the subject of estimation theory and of the Wiener and Kalman filtering techniques [10]. The *orthogonality principle* indicates that the mean-square error of a stationary linear estimator will be minimum when the error is orthogonal in expectation to the known values on the process. It is also known that when the estimated process is Gaussian (as in the case of fractional noises), the linear estimate is optimal in the sense of being identical to the best nonlinear estimate given the same number of observations [11][12]. Stochastic subdivision is specifically similar to the application of digital Wiener filtering in the linear-predictive coding (LPC) of speech [13], since in both of these applications points on a stochastic process are estimated, and then perturbed and re-used as "observations".

In our case, the midpoint $\hat{x}$ at each stage in the construction will be estimated as a weighted sum of the noise values x known from the previous stages of the construction, in some practical neighborhood of size $2S$:

$$\hat{x}_{t+0.5} = \sum_{k=1-S}^{S} a_k x_{t+k}$$

(with t indexing the points known at the previous construction stage). The estimated value $\hat{x}_{t+0.5}$ will form a new noise point with the addition of a random number of known variance; the new points will in turn form some of the data in subsequent construction stages.

The orthogonality principle then takes the form

$$E\left\{ x_{t+m} \left(x_{t+0.5} - \sum_{k=1-S}^{S} a_k x_{t+k} \right) \right\} = 0$$

or

$$E\left\{ x_{t+m} x_{t+0.5} - \sum_{k=1-S}^{S} a_k x_{t+m} x_{t+k} \right\} = 0$$

for $1-S \leq m \leq S$. Recalling that the expectation of $x_{t+i} x_{t+j}$ is the value $R(i-j)$ of the noise autocorrelation function R (for a stationary noise), we obtain the equation

$$R(m-0.5) = \sum_{k=1-S}^{S} a_k R(m-k)$$

which can be solved for the coefficients a_k given R. The matrix $R(m-k)$ is Toeplitz, permitting the use of efficient algorithms available for the inversion of these matrices, such as the Levinson recursion [14]. The mean-square estimation error

$$\mathbf{E}\{(x - \hat{x})^2\} = R(0) - \sum_{k=1-S}^{S} a_k R(0.5-k)$$

is used to select the noise variance (and optionally the neighborhood size) at each construction stage [8]. Fig. 1 illustrates successive stages in generalized subdivision to an oscillatory noise with an autocorrelation $R(\tau) = \cos(\omega\tau)\exp(-\tau^2)$.

2.1. Subdivision in two dimensions

The significant difference from the one-dimensional solution is that there are now several classes of points to be estimated, categorized by their spatial relationship to the points computed at previous subdivision levels (this depends on the selected interpolation mesh). For the planar quadrilateral mesh shown in Fig. 2 the mid-face vertex 'x' will require different coefficients than the mid-edge vertices 'o'. For example, (using our coordinate system with the midpoints "indexed" by 1/2) the midpoint coefficients are obtained by solving

$$R(j-0.5, i-0.5) = \sum_{r=1-S}^{S} \sum_{c=1-S}^{S} a_{r,c} R(j-r, i-c).$$

for $1-S \leq j, i \leq S$. This equation can be considered as a system $A\mathbf{x} = \mathbf{b}$ by rewriting $R(y,x)$ and $a_{r,c}$ as vectors by a consistent ordering of the subscripts; the dimension of the matrix A is now the square of the neighborhood size $2S$.

2.2. Evaluation

The generalized subdivision technique produces high-quality noises with specified spectra and eliminates the creases associated with stochastic subdivision to non-Markovian noises. It also shares the attractive properties of the stochastic subdivision construction [1], i.e., the consistency properties described in [1] including the ability to model a noise at different resolutions, and the ability to model regions of a noise in any order (a "non-causal" property which is not available in Fourier synthesis and other spectral synthesis approaches). When a separable Markovian autocorrelation function $R(x,y) = exp(-|x|)exp(-|y|)$ is specified, the generalized subdivision reduces to a form of fractal subdivision, in the sense that only the coefficients for the nearest neighbors of an estimated midpoint are non-zero. Subdivision to non-Markovian spectra is computationally more expensive due to the larger neighborhood sizes required. Fig. 3 shows several textures produced with the generalized subdivision technique and Figs. 4, 5 illustrate two height fields produced using this technique, displayed as synthetic terrains.

Several limitations of the generalized subdivision technique are:

One must know or invent the noise autocorrelation function. Since the autocorrelation function is the Fourier transform of the power spectrum (Wiener-Khinchine relation), and the latter must be non-negative, the autocorrelation function must be *non-negative definite*. Unless this constraint is well understood, it may be easier to design the power spectrum and obtain the autocorrelation by transformation, or to restrict one's choice to paradigmatic or empirically estimated autocorrelation functions.

A second restriction of the generalized subdivision technique derives from the variable-resolution property of subdivision constructions. The identification of different stages in the construction with different resolutions is strictly incorrect. This can be seen from one point of view by considering the problem of obtaining a half-resolution version of a given noise record. A half-resolution noise which preserves the spectral content of the original up to the new, lower Nyquist rate is achieved by low-pass filtering, followed by dropping every other sample ("decimation"). The half-resolution noise resulting from reducing the recursion level in a stochastic subdivision construction is achieved by decimating without filtering. A half-resolution noise does not in general coincide with every other sample of the original noise unless the latter has no detail at frequencies above half its Nyquist rate. Thus, any spectral energy above half the original Nyquist rate is aliased in changing the resolution through the subdivision construction depth.

Significantly, an aliased noise does not form coherent artifacts such as Moire patterns; rather, the noise at the lower resolution appears as a somewhat different noise than the original, so the subject may appear to "bubble" during a zoom. The aliasing is limited for noises with monotonically decreasing spectra such as fractal noises, since the majority of the spectral energy remains unaliased in any resolution change. However, serious anomalies may occur if the resolution of a noise whose spectrum is flat or increasing at some frequencies (as may be achieved with the generalized subdivision technique) is varied by changing the subdivision recursion depth.

3. Shaped Point Process

A second stochastic synthesis algorithm is suitable when samples of the desired noise are available. An analysis-synthesis approach would analyze the noise to determine parameters of a stochastic model, and then apply the model to generate a synthetic noise. If the only goal is to synthesize the noise, however, a more direct approach is feasible: the noise x is produced by a (discrete) convolution

$$x_t = \sum_{k=-S}^{S} h_k u_{t-k}$$

of a uncorrelated noise u with the (windowed) noise sample h of size $2S+1$, with h playing the role of a filter kernel. The autocorrelation of x is easily derived:

$$R(\tau) = \mathrm{E}\{x_t x_{t+\tau}\}$$

$$= \mathrm{E} \sum_k \sum_m h_k h_m u_{t-k} u_{t+\tau-m}$$

The noise u is stationary and uncorrelated so the expectation of the factors $u_{t-k} u_{t+\tau-m}$ is $\mathrm{E}\{u^2\}\, \delta(\tau+k-m)$, so

$$R(\tau) = \sum_k h_k h_{k+\tau} \tag{1}$$

The power spectrum of x is the Fourier transform of R. Since (1) is a convolution $h_t * h_{-t}$, its transform is (by the convolution theorem [15])

$$S(\omega) = H(e^{j\omega}) H(e^{-j\omega}) = |H(e^{j\omega})|^2$$

so the spectrum of x is that of h (as expected). The spectrum of the noise sample h will in turn resemble that of the prototype noise if it is large enough to include any low-frequency components characteristic of the prototype and if it is windowed to reduce the effects of discontinuities at the sample boundary.

Convolution with a large noise sample is inefficient and the convolution would usually be implemented in the frequency domain by FFT. Computational economy can also be achieved by replacing the noise u with a 'sparse noise' or particle system (sampled Poisson point process) $\hat{u}$ which is non-zero at a limited number of points under the sample h. The reduced convolution takes the form

$$x_t = \sum_k \hat{u}_k h(t - t_k) \tag{2}$$

where t_k is the location of the kth non-zero point of the process, and the summation is now over these points rather than over h (a similar technique was described as one of the methods in [16] but its use as a general spectral modeling approach was not fully developed there). The autocorrelation and spectrum are unchanged provided the values of $\hat{u}$ are independent. This "shaped point process" resembles both shot noise (in which the noise $\hat{u}$ is defined to be a constant-amplitude Poisson impulse process), and a generalized form of pulse amplitude modulation reconstruction, which requires t_k to be evenly spaced.

3.1. Spectral and graphic modeling

The primary advantage of this algorithm is not efficiency, however, but that it suggests manipulating the point process as an entity itself. For example, to produce a 'fluid texture' by animating the point process requires only updating the location of each point by a dynamic equation, whereas manipulating a uniformly sampled noise field to the same effect requires computations more analogous to those of a fluid flow problem on a uniform grid. Similarly, points may be restricted to an area of the plane with conceptually simple algorithms such as Monte Carlo or an ad hoc placement procedure, whereas restricting a noise field requires scan converting the boundary of the region or a global windowing operation.

The non-causal property of subdivision methods is achieved in a shaped point process using an appropriate (non-causal and consistent) construction of the point process. A simple construction is to divide the noise domain into numbered cells and approximate the Poisson point process by N points in each cell, with the random number generator seeded by the cell number. The value of the shaped noise at a particular point is obtained by (2) summed over only the points in those cells which are closer than a radius the size of the kernel.

The kernel h can also be manipulated independently of the point process. The spectral bandwidth of a shaped point process is entirely determined by the kernel. If the size of the kernel is small compared to the depth in a perspective view of a shaped point process noise, the noise can be accurately and efficiently anti-aliased by selecting appropriate precomputed bandlimited versions of the kernel as a function of depth. The kernel can be varied as a function of the position of each point to produce a *non-stationary* noise. For example, wind-blown clouds or terrain ridges where the directional tendency varies over the scene could be emulated by rotating the kernel as a function of position. This type of control is not directly available in most filtering techniques; e.g. it is achieved in a Fourier transform method only by breaking the noise into small overlapping stationary regions and interpolating the synthesis on these regions (overlap-add method for short-time Fourier transformation).

Thus, a shaped point process provides a convenient separation between the spectral modeling problem (obtaining the kernel) and the graphic modeling problem of shaping the noise to form a subject. (A similar separation occurs in 'waveform' speech synthesis: a kernel is used to model the formant (spectral) shape; it is convolved with a impulse sequence or noise representing the voice pitch and amplitude [17]).

3.2. Evaluation

The shaped point process is a simple means of approximately "resynthesizing" noises. The method also generalizes directly to several dimensions. Fig. 6 shows the shaped point process resynthesis of several texture samples from the Brodatz album [18]. Resynthesis is of course more intuitive than specifying the parameters of a texture model, and it allows the generation of homogeneous noises of arbitrary extent. Periodic noises can be produced by altering the addressing in (2) to wrap around specified boundaries; this is a useful property in applications such as texture mapping. The shaped point process can also be applied with an analytically defined kernel; the lower right plot in Fig. 6 is a perspective view of a wave-like texture created with a bandpass kernel of the form $R(x,y) = cos(\alpha x + \beta y)\exp(-\sqrt{x^2+y^2})$.

The textures in Fig. 6 also suggest the limitations of the shaped point process method, and of spectral methods in general. The phase spectrum in a spectral synthesis method is that of the driving noise, which is random. Thus spectral synthesis cannot produce a coherent-phase texture such as a brick wall pattern. In fact, given a step function for the kernel h, the shaped impulse process will result in a f^{-2} noise -- the spectrum of the kernel is reproduced but the visual appearance is quite different.

The grey levels in a texture photograph reflect the illumination of the texture and may not directly correspond to 'physical' properties of the texture such as color or relief depth. Thus, a texture synthesized from a photographic sample will reproduce the spectral character of the texture as illuminated rather than as we perceive it. One common effect is that sharp cast shadows produce discontinuities in the texture kernel and so introduce f^{-2} noise into the synthesized texture.

Unlike subdivision constructions, a shaped point process noise has definite inner and outer scales. The autocorrelation (1) is zero beyond the width of the kernel, so the noise is uncorrelated at scales larger than this width (this can be seen in Fig. 6 as the scale at which the textures become "blotchy"). The inner scale is of course the Nyquist rate determined by the (fixed) sample rate of the noise. The bandwidth available in a shaped point process is nevertheless considerably greater than that available in many artificial texturing methods (e.g. [19]) and is adequate for many purposes, since a stochastic model will rarely be applicable over a very broad range of scales in any case. Also, some phenomena such as waves, fire, and bark which might be modeled by stochastic methods are often fairly smooth above and below a range of scales.

4. Non-Gaussian Noises

By a loose version of the central limit theorem, the probability density of a noise produced with spectral synthesis will tend to be Gaussian regardless of the density of the driving noise, since the spectral shaping operation is effectively a linear filter or a weighted sum of the input noise values [15]. It is sometimes desirable to model non-Gaussian processes. For example, with respect to the uniform or normal distributions, a distribution such as $\exp(-|x|)$ has an increased number of 'events' far removed from the mean. Transforming a Gaussian noise to have a higher-variance non-Gaussian distribution tends to differentially exaggerate the most pronounced portions of the noise and so can produce the impression of a 'subject' against a background, or of a non-stationary noise. Some of the published fractal landscape pictures depict fractional noises passed through a square or cube nonlinearity which improves their appearance.

The probability density of a random process can be shaped by means of a memoryless nonlinear transformation $g(x)$. For this purpose it is sufficient to consider only monotonically increasing $g(x)$. Then, by "conservation of probability", the probability of an event $\mathbf{y} < y$ where $y = g(x)$ is identical to that of the event $\mathbf{x} < x$:

$$F_y(g(x)) \equiv \mathbf{P}\{\mathbf{y} < y\}$$

$$= F_x(x) \equiv \mathbf{P}\{\mathbf{x} < x\}$$

or

$$f_y(y) = f_x(g^{-1}(y))\frac{dg^{-1}(y)}{dy}$$

so

$$g(x) = F_y^{-1}\left[F_x(x)\right]$$

Two cases are particularly useful. When $\mathbf{x}$ is uniformly distributed in (0,1), $F_x(x) = x$ so the nonlinear function $g(x)$ which shapes a uniform noise $\mathbf{x}$ to have a desired distribution F_y is just $g = F_y^{-1}$. When the desired distribution F_y is uniform, $g = F_x$. Thus, the procedure to transform a noise to have a desired distribution is to first pass the noise through its own distribution function to make a uniform (0,1) noise, and then use the result to index the inverse of the desired distribution function. Both of these operations can be implemented by table lookup for reasonably smooth functions, so distribution shaping can be very efficient.

4.1. Effect on spectrum

The nonlinearity which shapes the distribution can also have a powerful effect on the spectrum of a correlated noise, however. This can be appreciated by considering the potential effect of a nonlinearity $g(x)$ on a single "frequency component" $\cos(\omega t)$. By choosing $g(x) = f(\cos^{-1}(x))$, an arbitrary periodic waveform $f(\omega t)$ is produced at the output of the nonlinearity given the single frequency as input. The envelope of the spectrum at the output of the nonlinearity also depends on the amplitude of the input signal. A signal passed through a nonlinearity does not obey either the superposition or homogeneity principles of linear systems, so the effect of a nonlinearity on a noise cannot be analyzed as the superposition of its frequency components.

A general expression for the autocorrelation function at the output of $g(x)$ is [20]

$$R(\tau) = \int\int g(x_1)g(x_2)f_x(x_1,x_2,\tau)dx_1dx_2$$

where f_x is the second-order joint probability density of the input. The spectrum of the output is the transform of this. However, this integral is difficult to evaluate and analytic solutions are known only for some special cases, including various cases where f_x is Gaussian. Beckmann [20] gives a expression for the distorted correlation function of a Gaussian noise as a series involving weighted powers of the input autocovariance. The output spectrum is the transform of this series, which by the modulation (or convolution) theorem is a weighted series of nth-order self convolutions of the input spectrum. This effect is illustrated in Fig. 7 . In theory it should be possible to design the spectrum of the undistorted noise so that a desired spectrum is achieved *after* distortion, but this approach has not been formulated to the author's knowledge.

We conclude that nonlinear distortion is a powerful means of generating correlated non-Gaussian noises. However, this approach should be used carefully if accurate control of the spectrum and probability density are required. For example, some of the "fractal Gaussian" terrains we have seen are probably neither Gaussian nor of the attributed spectral exponent or fractal dimension as a result of squaring or other nonlinear distortions (e.g. a squared Gaussian noise has a one-sided probability density

$$f_y(y) = \frac{1}{\sqrt{2\pi y}}e^{-y/2}, \quad y \geq 0$$

which is quite different from the Gaussian density).

5. Conclusion

Two spectral methods for stochastic synthesis were described. Spectral approaches allow the synthesis of noises with arbitrary power spectra, and so can describe both narrowband deterministic-like noises such as [19] and broadband random noises such as fractals, as well as noises which exhibit a mixture of structure and randomness.

References

1. Fournier, A., Fussel, D., and Carpenter, L., Computer rendering of stochastic models, *Communications ACM* **25**, 6, June 1982, 371-384.
2. Kawaguchi, Y., A morphological study of the forms of nature, *Siggraph 82 Proceedings* July 1982.
3. Reeves, W., Particle Systems--A Technique for Modeling a Class of Fuzzy Objects, *ACM Trans. Graphics* **2**, 2, April 1983, 91-108.
4. Reeves, W. and Blau, R., Approximate and probabilistic algorithms for shading and rendering structured particle systems, *Siggraph 85 Proceedings* July 1985, 313-322.
5. Smith, A. R., Plants, Fractals, and Formal Languages, *Computer Graphics* **18**, 3, July 1984, 1-10.
6. Voss, R., Random fractal forgeries, *Siggraph conference tutorial notes* July 1985.
7. Perlin, K., An image synthesizer, *Siggraph 85 Proceedings* July, 1985, 287-296.
8. Lewis, J.P., Generalized stochastic subdivision, *submitted for publication.* Jan. 1986.
9. Fournier, A. and Milligan, T., Frame buffer algorithms for stochastic models, *Proceedings, Graphics Interface* May 1985, 9-16.
10. Deutsch, R., *Estimation Theory,* Prentice-Hall, New Jersey, 1965.
11. Yaglom, A., *An Introduction to the Theory of Stationary Random Functions,* Dover, New York, 1973.

12. Papoulis, A., *Signal Analysis*, McGraw Hill, New York, 1977.

13. Markel, J. and Gray, A., *Linear Prediction of Speech*, Springer-Verlag, New York, 1976.

14. Levinson, N., The Wiener RMS (root mean square) error criterion in filter design and prediction, *J. Math. Phys.* **25**, 1947, 261-278.

15. Bracewell, R., *The Fourier Transform and Its Applications*, McGraw-Hill, New York, 1965.

16. Lewis, J.P., Texture synthesis for digital painting, *Computer Graphics* **18**, 3, July 1984, 245-252.

17. Baumwolspiner, M., Speech generation through waveform synthesis, *IEEE Acoustics, Speech and Signal Processing Conference* 1978.

18. Brodatz, P., *Textures: A Photographic Album*, Dover, New York, 1966.

19. Schachter, B., Long Crested Wave Models, *Computer Graphics and Image Processing* **12**, 1980, 187-201.

20. Beckmann, P., *Probability in Communication Engineering*, Harcourt-Brace, New York, 1967.

21. Vanmarcke, E., *Random Fields*, MIT Press, Cambridge, 1983.

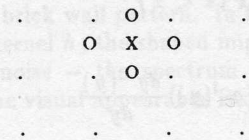

Fig. 2 : Planar quadrilateral subdivision mesh using a 4^2 neighborhood. The vertices 'o' and 'x' are estimated using the surrounding 'observation' points '.'.

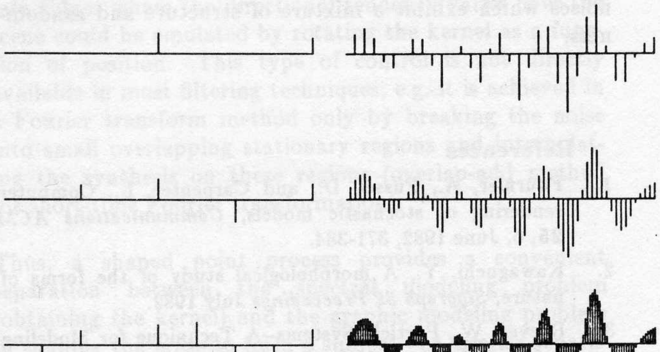

Fig. 1: (top to bottom) Stages in generalized subdivision to a non-fractal (oscillatory) noise.

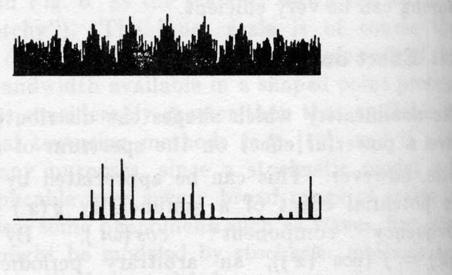

Fig. 7 : Bandpass noise and spectrum (lower figures) and noise and spectrum at the output of a pair of nonlinearities effecting a hyperbolic probability density. The self-convolution of the input spectrum produced by the nonlinearities results in an odd-harmonic spectrum structure.

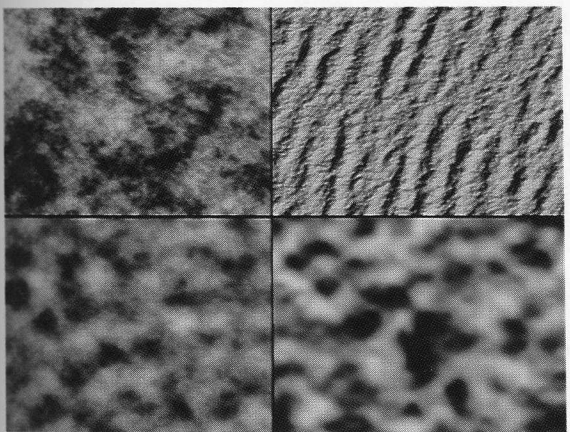

Fig. 3 : Several textures produced using the generalized subdivision technique. Clockwise from top left: Markovian, oscillatory (shaded as an obliquely illuminated height field), Gaussian, and highpass isotropically oscillatory textures.

Fig. 5 : Synthetic sky and terrain with directional trend produced with generalized subdivision.

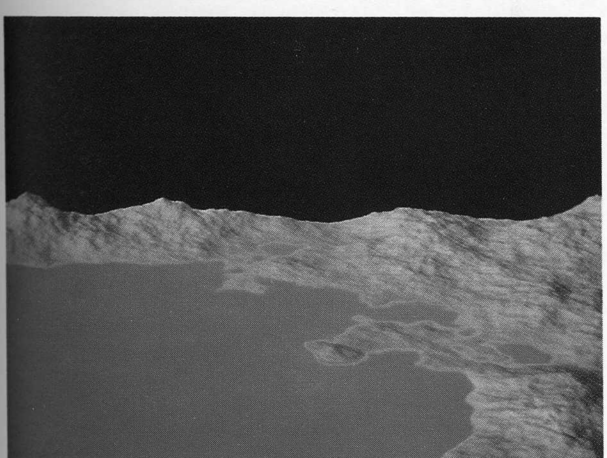

Fig. 4 : Generalized subdivision terrain with an isotropic autocorrelation $R(x,y) = \exp(-(x^2+y^2)^{0.7})$. This figure resembles a power-modified fractal terrain but it can be distinguished (in being smoother) in a comparison.

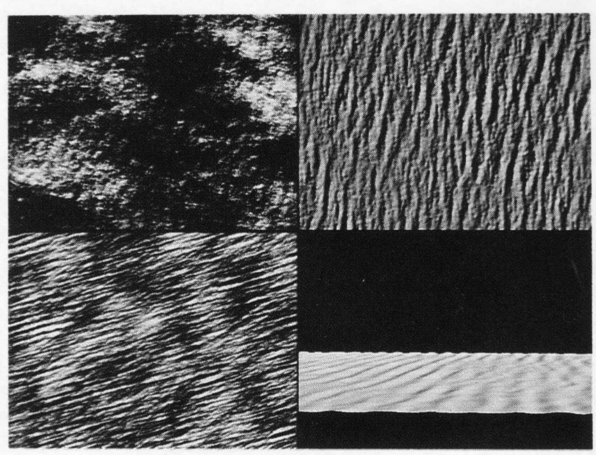

Fig. 6 : Several shaped impulse process textures. Counterclockwise from top right: rough waves (shaded as an obliquely illuminated height field), fieldstone, and straw synthesized from Brodatz [18]. The last figure is a perspective view of a wave-like texture.