ON THE VISUAL DISCRIMINATION OF SELF-SIMILAR RANDOM TEXTURES

by

RONALD ANDY RENSINK

B.Sc.(Physics), The University of Waterloo, 1979
M.Sc.(Physics), The University of British Columbia, 1982

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
DEPARTMENT OF COMPUTER SCIENCE

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
September 1986

Department of Computer Science

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date 9 September 1986

Abstract

This work investigates the ability of the human visual system to discriminate self-similar Gaussian random textures. The power spectra of such textures are similar to themselves when rescaled by some factor $h > 1$. As such, these textures provide a natural domain for testing the hypothesis that texture perception is based on a set of spatial-frequency channels characterized by filters of similar shape.

Some general properties of self-similar random textures are developed. In particular, the relations between their covariance functions and power spectra are established, and are used to show that many self-similar random textures are stochastic fractals. These relations also lead to a simple texture-generation algorithm that allows independent and orthogonal variation of several properties of interest.

Several sets of psychophysical experiments are carried out to determine the statistical properties governing the discrimination of self-similar line textures. Results show that both the similarity parameter H and the scaling ratio h influence discriminability. These two quantities, however, are insufficient to completely characterize perceived texture.

The ability of the visual system to discriminate between various classes of self-similar random texture is analyzed using a simple multichannel model of texture perception. The empirical results are found to be compatible with the hypothesis that texture perception is mediated by the set of spatial-frequency channels putatively involved in form vision.

Table of Contents

List of Tables

List of Figures

List of Symbols

The following list contains a brief description of the symbols most commonly used in this work. As far as possible, compatibility has been maintained with the notational conventions used in other areas of study. As such, this occasionally leads to nonunique denotation. Where amibiguity arises, context should make clear the intended meaning of the symbol.

$a, \vec{a}$	deterministic fractal
$\mathbf{a}, \vec{\mathbf{a}}$	stochastic fractal
$a(t), \mathbf{a}(t)$	fractal with intrinsic parametrization
$a(x), \mathbf{a}(x)$	fractal with extrinsic parametrization
$\vec{\mathbf{b}}_0$	initial position of Brownian motion path
$\bar{c}(x)$	sample function for covariance of image
$\bar{\mathbf{c}}(x)$	estimator for covariance of image
$f(x, y), f_{xy}$	image, random texture
$\mathbf{f}(x, y), \mathbf{f}_{xy}$	random field
g	grey level
$h_d(\rho)$	generalized ball
k	spatial frequency in x–direction
$\vec{k}$	spatial frequency vector. $\vec{k} = (k, l)$
l	spatial frequency in y–direction
$m(k, l), m_{kl}$	modulation function
n	number of dimensions, number of generator segments
p	probability of error
$p(w)$	probability density function
r	geometric ratio
$t, \vec{t}$	intrinsic parameter, intrinsic parameter vector
u	displacement in x–direction

$\vec{u}$	displacement vector. $\vec{u} = (u, v)$
v	displacement in y–direction
$v(t)$	time series
$\mathbf{v}(t)$	stochastic process
$w(\vec{x})$	windowing function
x	spatial position, displacement in x–direction
$\vec{x}$	spatial position vector, displacement vector. $\vec{x} = (x, y)$
y	spatial position, displacement in y–direction
$\mathbf{z}_{kl}$	zero-mean, unit-variance Gaussian random variable

A	arbitrary scaling factor
$\vec{\mathbf{B}}(t)$	Brownian motion
$\vec{\mathbf{B}}_H(t)$	fractional Brownian motion
C	contrast of image
$C(\vec{x})$	covariance function
D	Hausdorff-Besicovitch dimension
D_k	spectral damping filter
E	dimension of embedding space
H	similarity parameter
$H_i(k)$	filter for channel i
L	luminance of display
$L(\lambda)$	length of line at resolution λ
M_i	measure on channel i
$P(k)$	spectral pattern function
$R(\vec{x})$	correlation function
$R(\theta)$	rotation operator
S_i	similitude for fractal subset i
$S(\vec{k})$	power spectrum
T	topological dimension

T_i	size of image in dimension i
V_i	relative contrast of channel i
Z_i	zero-crossing density of channel i

δ	parametric distance between endpoints of increment
θ_i	rotation angle for generator section i
κ_i	cut-off frequency of spectral damping filter
λ	lower cut-off scale for spatial self-similarity
μ	mean value of random field
σ	standard variation of random field
$\phi(k,l)$	phase of wave vector $\vec{k} = (k,l)$
$\psi(H, \Delta H)$	psychometric function
ω	lower cut-off frequency for spectral self-similarity

$\Gamma(x)$	Gamma function
Δ_i	sampling distance in dimension i
Λ	upper cut-off scale for spatial self-similarity
$\Pi(\Theta_i)$	orthonormal transformation for generator section i
Ω	upper cut-off frequency for spectral self-similarity

Acknowledgements

First of all, I would like to thank my supervisor, Bob Woodham, for all the guidance he has given me over the past few years. In the course of writing this thesis, I have learned a great deal from him about the formulation and investigation of scientific problems. Its shortcomings aside, I would like to think that this work attempts to meet his high standards.

I would also like to thank Anne Treisman of UBC Psychology for her comments on several aspects of the psychophysical experiments. From the all–too–few discussions we had, I learned a great deal about the design of experiments. Also thanks to Alan Mackworth of UBC Computer Science for his comments on an earlier draft of this thesis.

Several graduate students have helped with various aspects of this thesis. Marcia Grabowecky of UBC Psychology provided useful feedback about psychophysical testing. Debbie Aks, also of UBC Psychology, helped calibrate the display monitor. Jordan Brooks, UBC Computer Science, read and commented on a few chapters of an earlier draft of this thesis. Marc Majka, UBC Computer Science, provided assistance on several technical matters.

Last but certainly not least, I would like to acknowledge my great debt to Jennifer Brereton for her participation in the psychophysical experiments. She showed incredible perserverance in viewing thousands of texture displays over a period of several months. I thank her for all she has done.

This work has been supported by a Research Assistantship from the Department of Computer Science, University of British Columbia, and a University of British Columbia Graduate Fellowship.

Chapter 1

Introduction

One of the fundamental tasks of vision is the detection and recognition of objects in the surrounding environment. The surfaces of these objects often have characteristic textures distinguishing them from their surroundings. The effectiveness of a visual system is consequently increased if it can detect such structure. Indeed, many animals appear to make some use of texture — surface markings often promote high visibility or provide camouflage in natural habitats [BrGr85].

Even though some form of texture perception is used by many simple organisms, texture perception in general has proven difficult to analyze. Attempts to place it on a firm scientific basis have had only limited success. Various characterizations of texture exist, but none appears capable of capturing all aspects of structural and statistical regularity.

These difficulties arise in part because of the interdependence of texture perception and form perception. It is difficult to determine when the spatial structure of a surface is an intrinsic surface property, describable as texture, and when it is a collection of objects discriminable in their own right. For example, a distant field of wheat is seen as a single textured surface; at closer range, the same field is distinguishable as a collection of individual plants. The transition from one description to the other has no well-defined boundary. Texture perception and form perception may therefore share a set of common mechanisms.

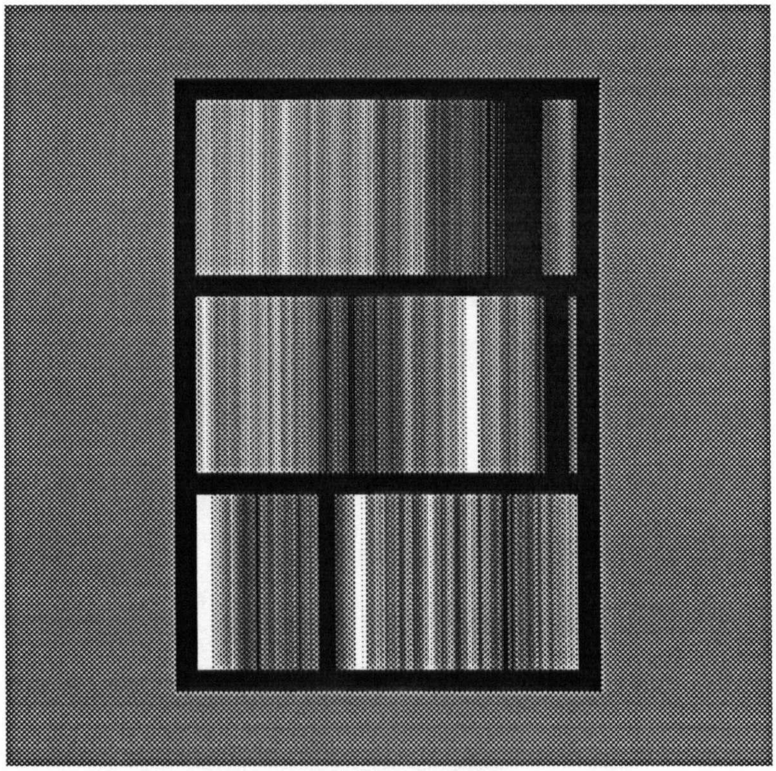
These matters must be resolved before a computational theory of texture perception can be established. To this end, an interesting class of textures for investigation is the self-similar random textures. For these textures, any characteristic present at a small scale is also present at a larger scale. Consequently, their spatial structure has no well-defined partition separating object boundary and intrinsic surface structure.

1.1 Overview of the Thesis

1.1.1 The Issues

This work examines the ability of the human visual system to discriminate among self-similar random textures (figure 1.1). The research hypothesis is that the performance of the human visual system in this domain can show whether common mechanisms underly both form perception and texture perception. In particular, evidence is sought that texture perception is based on measurements made in parallel on the set of spatially-filtered images constituting the basis of form vision.

Self-similar random textures have their origin in the work of Mandelbrot on stochastic fractals [MaNe68][Mand82]. Formally, fractals are the class of mathematical objects that have a non-integral Hausdorff-Besicovitch dimension D (see appendix B). These objects may be either deterministic or stochastic.



Upper texture: $H = 0.5$, $h \to 1$

Middle texture: $H = 0.5$, $h \to 1$

Lower texture: $H = 0.3$, $h \to 1$

Discriminability = percentage correct pairing = 75%

Figure 1.1: example of texture display

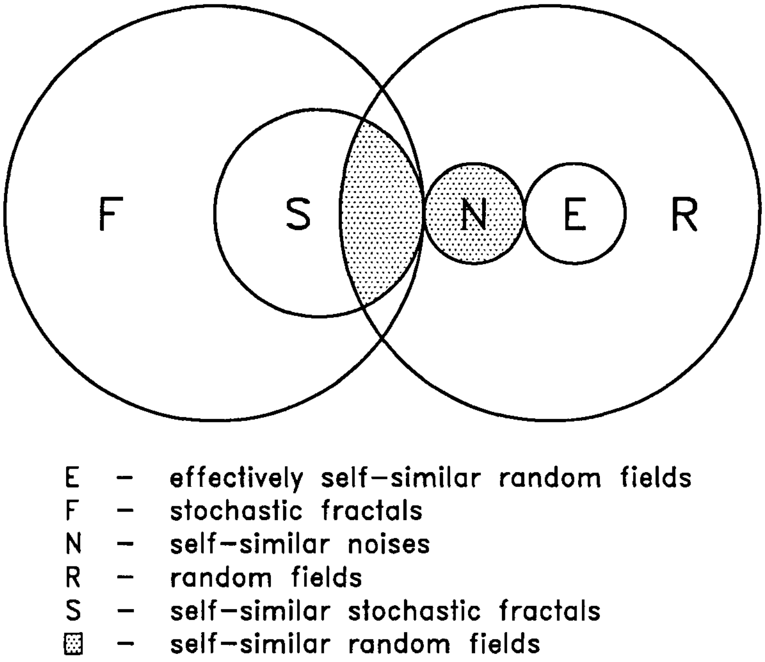


Figure 1.2: relation between fractals and random fields

For surfaces in three-dimensional space, the value of D ranges between 2 and 3. When $D \to 2$, the surface is smooth and almost planar. When $D \to 3$, it appears extremely rough and jagged. The fractal dimension is therefore a measure of the roughness of a surface. For reasons of mathematical convenience, D is often expressed in terms of the similarity parameter H. For the stochastic fractals considered here, $H = 3 - D$, so that $0 < H < 1$ (see appendix B).

Many fractals are self-similar, matching themselves completely when rescaled by a scaling ratio $h > 1$. Self-similar stochastic fractals are widely used in computer graphics to generate highly realistic images of clouds, landforms, and plants (e.g., [Mand75][FoFu82][Mand82]).

The self-similar random textures considered in this work are instances of self-similar random fields. The general class composed of such fields includes several self-similar stochastic fractals and self-similar noises (figure 1.2). Also considered here are a class of effectively self-similar random fields, for which self-similarity holds only over a limited range of scales. Taken together, the self-similar and effectively self-similar random fields form a useful domain for determining the ability of the human visual system to detect self-similarity. In particular, they allow measurement of its sensitivity to quantities such as the similarity parameter H and scaling ratio h.

1.1.2 Organization of the Work

The work is divided into three distinct sections:

1. Description of the properties of self-similar random fields, both in the spatial and the frequency domains.
2. Empirical investigation of the ability of the human visual system to discriminate among self-similar random textures.
3. Interpretation of the empirical results in light of current theories of texture and form perception.

a) Description of self-similar random fields

In this work, attention is restricted to random fields that are stationary. By definition, the statistical properties of such fields remain invariant under translation. A stationary random field is often represented by its covariance function $C(\vec{x})$, which describes the statistical correlation between the values of points separated by a displacement $\vec{x}$. Another measure is the power spectrum $S(\vec{k})$, which describes the contribution to the random field of the harmonic at spatial frequency $\vec{k}$ (see appendix A).

This work develops the relations between the covariance functions and power spectra of self-similar random fields. These relations are used to show that the class of self-similar random fields contains several stochastic fractals and self-similar noises. It is also shown that H and h are insufficient to completely specify a self-similar covariance function and power spectrum. This implies that other quantities must also enter the description of a self-similar random field.

The reformulation of stochastic fractals and self-similar noises provides the basis of a texture-generating algorithm. By taking the Fourier transform of a field of Gaussian random variables, it is possible to create a random field having a specific power spectrum, so that self-similar textures can be readily generated. This algorithm allows the independent and orthogonal variation of several properties of interest, including the similarity parameter H and scaling ratio h.

b) Psychophysical experiments

The texture-generating algorithm outlined above can produce a wide variety of self-similar random textures. Psychophysical experiments based on these textures are carried out to determine the discriminability of various statistical properties. Experiments are limited here to the class of monochromatic self-similar Gaussian line textures. These are formed by sweeping a horizontal instance of a one-dimensional self-similar Gaussian stochastic process down through a finite vertical distance (figure 1.1).

Although simpler than fully two-dimensional textures, line textures are not trivial, having been used before in psychophysical research (e.g., [StJu72][RiPo74][Rich79]). Line textures have the advantage of allowing the set of possible texture elements to be reduced to a bare minimum (viz., straight-line segments together with their endpoints). More importantly here, an analytical treatment of many of their statistical properties is possible. Results obtained using these textures can form a basis for the treatment of the more general case.

The texture-discrimination experiments involve a display composed of three line textures (figure 1.1). Two adjacent textures are from the same random field, while the third texture is an instance of a second field. Discriminability between the two fields is given by the percentage of correct pairings made over a series of presentations.

Results show that no abrupt change in discriminability occurs between self-similar fractals, self-similar noises, and effectively self-similar textures. They also show that H and h are insufficient to completely characterize the perception of all self-similar random textures.

c) Analysis of empirical results

The results of the texture-discrimination experiments are analyzed using a simple multiresolution model of texture perception. This model assumes that texture perception is based on measurements made in parallel on a set of filtered images of various spatial resolutions. The empirical results are consistent with the hypothesis that texture discrimination is based on measurements such as the relative contrast or the density of zero-crossings in each of these images.

The zero-crossings present at each level of resolution are the basic elements of visual perception in many theories of form vision (e.g., [Marr82]). As such, the results of the texture-discrimination experiments are consistent with the conjecture that texture perception and form perception share a set of common mechanisms.

1.1.3 Arrangement of the Thesis

A general framework for discussing the basic issues discussing texture perception is presented in chapter 2. It introduces basic concepts and definitions, briefly surveys the more popular methods of texture analysis, and examines several current models of texture perception. Chapter 3 develops the relation between the covariance functions and power spectra of self-similar random fields.

The texture-discrimination experiments are presented in chapter 4. Chapter 5 discusses the results using a multiresolution model of texture discrimination, and examines their significance for a general computational theory of texture perception. Chapter 6 summarizes the general conclusions reached, and suggests some possible directions for future work.

Appendix A is a short review of the basic concepts used in the analysis of time series and random fields. Appendix B introduces several of the main ideas of fractal geometry, emphasizing those aspects relevant to this work. Appendix C examines the effects a discrete spatial image and power spectrum have on perceived texture, describes the generation of the textures, and briefly describes the calibration of the monitor used to display the textures. Appendix D is a table containing the relative contrasts and zero-crossing densities of the reference textures used in the analysis of the texture-discrimination experiments.

Chapter 2

Approaches to Texture

Over the past few decades, rigorous bases have been established for several modules of low-level vision (e.g., shape-from-shading [Wood81], stereopsis [Grim81], and surface-boundary-from-velocity [Hild84]). In contrast, there has been little apparent progress on other modules such as colour and texture perception. For texture perception, principles and techniques have remained largely *ad hoc* [Hara79][Jule84][GoDe85]. This may be due to the inherent complexity of the processes involved. Indeed, it has been argued that the underlying mechanisms may be so complex that no concise theoretical treatment of texture perception can ever be given [Marr77].

Nevertheless, some progress has been made. Although a complete theoretical treatment is not yet possible, previous results can be described within a common framework. This is based on a general characterization of texture perception.

2.1 A General Characterization of Texture Perception

This section discusses the general nature of texture perception, emphasizing its contribution to early vision. Attention is restricted to monochromatic broadband images. Colour perception is considered to be a separate concern, and is not discussed here.

2.1.1 The Role of Texture Perception

The functions of the early visual system include determining the location and spatial extent of objects in the surrounding environment, and providing higher-level systems with enough information to identify the objects [Marr82]. Many sources of information are available to help with these tasks, including binocular disparity, accommodation, and motion.

The surface structure of the objects themselves can also be exploited for these purposes. Surface structure is the intrinsic spatial organization of a surface, together with its reflectance characteristics. It is largely determined by the basic physical and chemical composition of the object. Since many objects have a composition different from their surroundings, it follows that their surface structures should differ as well. These differences can help determine their location and spatial extent in an image. An important task of texture perception is therefore the segmentation of an image into distinct regions.

Texture can also assist in recovering three-dimensional shape. If a surface has an isotropic structure, its orientation can be determined from texture gradients [Kend79] or from foreshortening effects [Breu80].

Information about surface structure can be put to further use. Since many objects have a distinctive surface property, it would be advantageous for the low-level vision system to transmit a description of the surface to assist in higher-level identification or classification of the object. Another task of texture perception is therefore the extraction of information about intrinsic surface structure.

Texture segmentation and shape-from-texture are not examined here. Discussion is limited to uniform textures on flat, pre-segmented regions. Issues such as projection and foreshortening are bypassed, and attention focussed on the final task mentioned: the characterization of perceived texture.

2.1.2 Perceived Texture

The projection of a three-dimensional surface onto a two-dimensional image depends on the location, orientation, and illumination of the surface, as well as its intrinsic surface structure. In general, the effects of all these factors are confounded, so that surface structure cannot be completely recovered from an image.

Nevertheless, a perceptual system can recover some of the surface structure. It is limited in this task by several factors, including its ability to determine three-dimensional structure from the image, and its ability to represent spatial information. Those aspects of surface structure determined from an image are referred to here as the *perceived texture.* The term resists an exact definition — it is used loosely here to refer to the intrinsic surface structure of a perceived region not containing any perceived objects. The prohibition against perceived objects is essential if texture perception is to be studied apart from the perception of objects.

Perceived texture can be characterized in several different ways, depending on the complexity of the image and the conditions under which it is viewed. When an image is so disorganized that objects cannot be perceived in it without a considerable effort of will, it is commonly termed a *random texture*. Texture can be perceived in such images under all conditions. Some images contain spatial features that can be combined into simple objects when attended to consciously [Trei85]. To avoid the effects of conscious scrutiny, texture perception must be limited to non-attentive viewing [Marr76]. When an image contains only a few items, these are often perceived as objects in their own right. To study texture using such images, it becomes necessary to consider texture perception as a pre-attentive process, taking place within the first few hundred milliseconds of image presentation [Jule75].

2.2 Methods of Texture Analysis

One of the central problems in texture perception is to determine the particular aspects of surface structure that are most useful for identification and classification. Although work has been done on coding principles applicable to these tasks (e.g., [Cael84]), a complete theoretical understanding of these issues has not yet been achieved. Empirical evidence is consequently of great value. In this regard, results obtained from the machine analysis of texture are of interest.

Historically, several different approaches to texture analysis have been taken. This has led to a great variety of representations. Each emphasizes some particular aspect of an image, such as its periodicity, structural hierarchy, or intrinsic spatial features present. All approaches, however, describe a texture by its microstructure and macrostructure [Hara79][Breu80]. The *microstructure* is the set of basic elements, or texels, that form the texture. The *macrostructure* is the set of spatial relations that exist between the microstructure elements. Different approaches to texture analysis are characterized by the the microstructure and the macrostructure that they use.

2.2.1 Spatial Approaches

Spatial approaches treat texture as a collection of simple elements spread densely throughout a region. These elements form a continuum, parametrized by some co-ordinate system. Depending on the the continuum and the elements used, a spatial approach can be placed into one of three groups: point statistics, global transforms, or local transforms.

In the first group, the continuum is taken to be a two-dimensional geometric space, and the elements are the individual points in the image. These methods describe texture by the statistics of the intensity values at these points. The second group involves global transformation of the original image (e.g., using the Fourier transform). The continuum is given by the transform space of the new representation. Each point in this space represents a specific pattern of intensity values in the original image. The third group of methods is based on local transformation. The continuum is a two-dimensional geometric space, while the texture element at each point describes the structure in the local neighbourhood.

In all these approaches, the elements are parametrized by an underlying continuous space. The term 'spatial' is used here in this more general sense.

a) Point statistics

This form of texture analysis is based on the statistics of the intensity values of individual points in an image. Images are generally assumed to be instances of ergodic random fields (see appendix A), whose spatial averages reflect ensemble properties. The various methods used are distinguished by the statistical properties represented.

The simplest representations involve first-order statistics, which are based on the histogram of the intensity values present. To reduce the effects of unequal lighting or poor instrument calibration, the averages and standard deviations (see appendix A) of the images analyzed are often set to common values. This destroys much of the first-order information. On occasion, however, information from unequalized images is used. The earliest first-order representations (e.g., [Rose62][PrMe66][DaJo68]) made use of several properties, such as mode and skewness. However, first-order statistics generally contain little information apart from that contained the average and standard deviation [AhDa77]. In current practice, these are often the only first-order quantities measured [Hara79].

First-order statistics cannot completely describe a texture, since they have no reference to the spatial arrangement of the elements. To capture this structure, higher-order statistics must be used. Second-order statistics are based on the frequency of the joint intensities of pairs of pixels separated by various displacements. Julesz [Jule62] made the conjecture that the discriminability of random-dot textures is completely determined by their second-order statistics. Over the years, this conjecture has inspired many analytic methods based on second-order statistics.

Measures based on such statistics were among the earliest used for texture analysis: Kaizer, in 1955, used the *autocorrelation function* $C(\vec{u})$ as the basis of texture description [Hara79]. This function is the second-order moment of the joint probability density. For an image $f(x,y)$ of dimensions $T_x \times T_y$, it has the form

$$R(u,v) = \frac{1}{T_x T_y} \int_{T_x} \int_{T_y} f(x+u, y+v) f(x,y) \; dxdy,$$

where u is the horizontal displacement between the pair of pixels, and v the vertical displacement.

A related function is the *covariance function* $C(\vec{u})$, defined by

$$C(u,v) = \frac{1}{T_x T_y} \int_{T_x} \int_{T_y} (f(x+u, y+v) - \mu)(f(x,y) - \mu) \; dxdy,$$

where μ is the average intensity of the image (see appendix A). The two measures are related by

$$C(\vec{x}_1, \vec{x}_2) = R(\vec{x}_1, \vec{x}_2) - \mu^2,$$

showing that the second-order information they contain the same.

Various properties of the covariance function are used for texture classification, including spatial moments, autoregression parameters, and concavity/convexity of form [Laws80][ChKa81]. Measures based on the covariance function do not result in highly accurate texture analysis; other second-order quantities must be used as well [Laws80]. This agrees with the observation that for human perception, the mean, variance, and covariance function are insufficient to determine the perceived texture [PrFa78].

A more general system of second-order statistical features was proposed by Rosenfeld and Troy in 1970 [Hara79], and later developed by Haralick *et al* [HaSh73]. This approach is based on the grey level dependence matrix, which describes the frequency of joint intensities of pixel pairs as a function of their spatial separation. This approach is a development of the Markov models first used by Julesz [Jule62], who analyzed texture using the transition probabilities between the values of neighbouring pixels.

The grey level dependence matrix corresponds to the second-order joint density function of a stationary random field. As such, its description can be large: for an image of size $n \times n$ pixels with m grey levels per pixel, the complete matrix would have a size of order n^2m^2. Furthermore, few pixels are separated by displacements comparable to the size of the image. The determination of joint intensity distributions for such pixel pairs is therefore susceptible to statistical fluctuation.

To overcome these drawbacks, a small set of features based on the grey level dependence matrix is used. Haralick *et al* [HaSh73] proposed a set of 14 measures, one of which was the covariance. To further reduce the size of the description, only a few orientations and separation distances are chosen. The resultant descriptions prove to be useful for texture identification, leading to over 90% classification accuracy in certain texture domains [WeDy76][CoHa80a].

Less arbitrary methods of reducing the description size have been developed. An optimal set of pixel displacements can be determined by statistical tests on the matrices [ZuTe80]. These can lead to similar classification accuracy with fewer features. Absolute pixel values can also be discarded, keeping only the relative differences in pixel intensities [WeDy76]. Although generally not as powerful as the approach based on grey level dependency, the use of relative intensities leads to nearly similar classification accuracies in many texture domains [WeDy76][CoHa80a].

Second-order information is also contained in the fractal dimension of the image [Pent83][MeYa84]. This quantity is determined by the rate at which the increment $f(\vec{x}+\vec{\Delta})-f(\vec{x})$ increases as a function of displacement $\vec{\Delta}$ (see appendix B). Using only the fractal dimensions measured in the x and y directions, classification accuracies of up to 85% can be achieved for several classes of natural textures [Pent83].

b) Global transforms

Perhaps the simplest way to represent an image is to assign an intensity to each point. When searching for specific spatial patterns, however, it is often useful to determine a global transform. This describes an image in terms of a basis set of spatial functions. For example, the finite Fourier transform describes an image as a (possibly infinite) sum of sine and cosine functions. This transform makes clear the degree to which the image is periodic.

Global transforms contain no explicit reference to spatial position — the image is described only in terms of the basis functions. If the basis set is complete, however, the transform contains all the information present in the original image [RoKa82].

Various transforms have been used for texture analysis, including Hadamard transforms, slant transforms, and Fourier transforms. Although their effectiveness for texture discrimination appears to be similar [Kirv76], only the Fourier transform is widely used.

The Fourier transform $\tilde{f}(k,l)$ of a continuous image $f(x,y)$ is given by

$$\tilde{f}(k,l)=\int\int f(x,y)e^{-i2\pi(kx+ly)}dxdy,$$

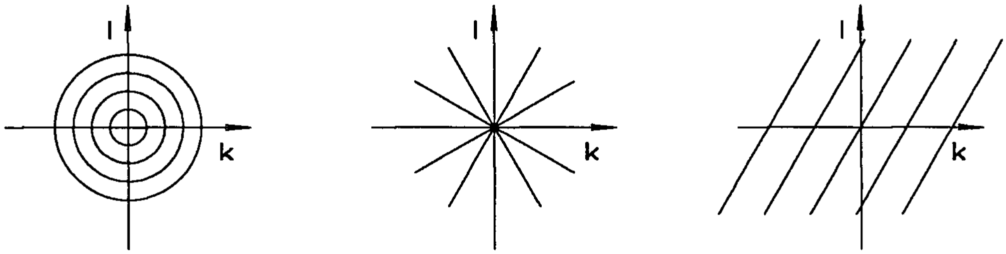
where k and l are spatial frequencies in the x and y directions respectively. This function is often written as a product, viz.,

$$\tilde{f}(k,l) = m(k,l)e^{i\phi(k,l)},$$

where $m(k,l) = |\tilde{f}(k,l)|$ is the *amplitude* of the waveform, and $\phi(k,l)$ is its *phase*. This class of representations emphasizes spatial periodicity. As such, it is most useful for the analysis of periodic patterns. However, it is also useful for random patterns as well.

Many ways exist to form equivalence classes of images based on their Fourier features, but only a few have been seriously investigated. Although the phase is important for images with global structure [JuBe83], the information it contains is generally of little use for classifying textures [Eklu79]. Some texture models are based on the sum of a few narrow-band sources of noise [Scha80], but only a few of these have been explored. Most approaches follow the lead of Lendaris and Stanley [LeSt70], who used the power spectrum of the image as the basis of texture analysis.

These approaches are based on the partitioning of frequency space into bins of varying shape. Description of a texture is given by the summed contributions of the power spectrum in each of the partitions. Three distinct types of partitioning are commonly used [CoHa80a]: annular rings, angular wedges, and parallel slits (figure 2.1). Annular rings provide a representation based on spatial frequency alone. Each ring corresponds to waveforms of arbitrary orientation, with frequency within some bounded range. Angular wedges allow a description of the directionality of the texture: each wedge corresponds to those waveforms oriented between two specified angles. The parallel-slit geometry is formed by a series of narrow, parallel rectangular regions. These are useful for



(a) annular rings (b) angular wedges (c) parallel slits

Figure 2.1: spectral partitions

detecting one-dimensional structure at a given orientation.

Fourier-based representations have been used for accurate classification of many natural textures [LeSt70][Bajc73][WeDy76], but in general are less useful than the statistical representations [CoHa80a].

c) Local transforms

This type of texture analysis is based on the local structure present at each point in the image. This is done using local transforms, which extract information from the neighbourhood surrounding each point. The form of these transforms depends on the the local structure considered relevant.

The results of several different transforms can be incorporated as feature planes into a composite description [Laws80]. Each element of the microstructure is then described by a vector quantity. These feature vectors can be given new bases in a generalized feature space, and be condensed down into a space of fewer dimensions. The various representations can be characterized by the local transform and the feature space used.

The earliest local transforms were spatial transforms, obtained by convolving a spatial filter over the image. Rosenfeld [Rose62] used the one-dimensional first-order derivative of an image as a basis for texture analysis. Classification was done via the first-order statistics of these derivatives. Linear filters emphasizing such shapes as lines, wedges, and spots have also been used [Hawk70].

More recent approaches use sets of general spatial filters. Laws [Laws80] uses a complete basis set of 3 × 3 and 5 × 5 masks, that describe averaging, first-differencing, and second-differencing operations. These filters are sums and differences of Gaussian functions [PiRo83]. By using texture energy measures based on the first-order statistics of the resulting elements, a classification accuracy of over 90% has been obtained for many classes of natural texture. Such accuracy generally depends on an appropriate choice of resolution size for the masks [Dumo85]. Methods have been developed [Ade83] to automatically select the best filter masks.

Another class of local transforms are the textural transforms introduced by Haralick [Hara75][Hara79]. The value of each element in the transformed image is a function of the grey level dependence matrix for the neighbourhood that surrounds it. Analysis is based on the first-order statistics of the elements of the transformed image. Accuracy is generally not as good as when the statistics of spatial transforms are used [Laws80].

d) General performance

Methods based on spatial approach have several common strengths and weaknesses. To begin with, they are all highly sensitive to the values of the intensities in an image — small changes to these values can lead to large changes of description. Such variations are almost always present between different images of the same texture, owing to uneven lighting, lack of camera calibration, etc. Some robustness can be instilled by equalizing the image histograms, so that all values are equally distributed [Hara79].

Methods based on local transforms must specify in advance the size of the neighbourhoods used. This renders texture description dependent on scale. More recent approaches, such as the fractal-based descriptions of Peleg *et al* [PeNa84], use measurements made at multiple scales of resolution to achieve a degree of scale-independence.

Another drawback is the inability of spatial approaches to capture the structure of higher-order groups of texture elements. This results largely from the homogeneous treatment of texture elements they employ.

In spite of these problems, spatial approaches are widely used. To begin with, they are indifferent to the pattern contained in the image — the computational resources required depend only on the size and the number of grey levels of the image. Furthermore, the descriptions are easily formed, so that many different measures can be created. These can then be combined into composite measures that lead to some of the highest classification accuracies yet achieved: over 80% for general classes of texture, and over 90% for more restricted domains [Hara79][Laws80].

2.2.2 Structural Approaches

Many images are highly regular in their spatial structure. Structural approaches take advantage of this regularity by restricting the ways in which basis functions can be combined. The resulting constraints enable compact descriptions to be made. Such representations generally involve only a few microstructure elements, arranged in patterns generated by a set of placement rules. Equivalent images are exactly those that can be described by the same microstructure and placement rules.

Although the possibility of such approaches has been discussed for many years (see, e.g., [Hawk70]), their development is a recent occurrence. Two groups of structural methods have been developed. In the first group, texture is considered to be composed of identical elements arranged in regular fashion throughout the plane. The second set makes use of syntactical techniques, representing texture as a parsing of the image.

a) Regular placement

These methods analyze texture by partitioning space into contiguous regions of identical spatial structure. The patterns in these regions form the microstructure of the texture. They may have a complex form, often being hierarchically composed of sub-elements [MaSa82]. The macrostructure is a two-dimensional periodic lattice, whose nodes describe the locations of the texture elements. Any image with a periodic structure can be partitioned this way [CoHa80b].

Two strategies are commonly used to create such descriptions: bottom-up and top-down. In the bottom-up approach, grouping processes are used to form the basic elements, while a clustering operation uses the locations of their centers to determine the macrostructure [MaSa82]. Such techniques can correct for missing of erroneous elements in the image, but generally remain extremely sensitive to noise, blur, and geometric distortion [MaMi83].

Most of these limitations can be bypassed by using top-down techniques.

The periodicity of the elements is first determined. This can be done via the grey level dependence matrices [CoHa80b][ZuTe80], or by Fourier analysis [MaMi83]. Elements can then be determined via region growing from the nodes of the macrostructure lattice.

Although not suitable for general use, these techniques can provide structural descriptions of many periodic and nearly periodic patterns. [CoHa80b][MaMi83].

b) Syntactic approaches

This approach is an outgrowth of picture grammars [Rose71]. A set of terminal symbols, made up of spatially-connected pixels, specifies the microstructure elements. A set of non-terminal symbols specifies the placement rules. Each non-terminal symbol corresponds to a fixed, two-dimensional template specifying the relative locations of several other terminal and non-terminal symbols. Such an approach is reminiscent of the grammars used for the syntactic analysis of languages. Indeed, the structural approach has been consciously developed along such lines — analysis is based on a parsing of the visual texture.

The various grammars differ in their specification of the terminal and non-terminal symbols. Shape grammars [BaBr82,ch6] use complex geometrical shapes for their terminal symbols. The placement rules are local in nature; they are represented by non-terminal markers that allow several adjacent symbols to be combined.

A more complete separation of microstructure and macrostructure is achieved in tree grammars [FuLu78]. The placement rules have the form of two-dimensional trees; these are combined to form the macrostructure of the texture (figure 2.2). Texture elements are then inserted into the resulting arrangement of terminal markers. To make this approach feasible, an image is first segmented into an array of rectangular windows, each of which is then analyzed. This is done to avoid the effect of large-scale warps of an ideal texture. Small-scale perturbations are handled by using a stochastic grammar, in which the placement rules can be selected nondeterministically. Combining these with an error-correcting bottom-up parser allows reasonably good discrimination among several classes of natural textures with large-scale structure [Fu82,ch12].

More complex approaches analyze the structure of the elements themselves. This is done by using levels of different grammars, the terminal symbols of one level being the starting symbols of the next one down [Jaya79]. The resulting descriptions are hierarchical in form, the placement rules at each level describing corresponding groupings in the image. Placement rules can also be recursively applied to scaled-down versions of themselves, resulting in descriptions with an infinitely many structural levels. Such grammars can be used to describe the self-similar deterministic fractals (see appendix B), that have similar spatial features at all levels of detail.

c) General performance

A purely structural approach is unsuitable for domains where few constraints exist on spatial structure, for spatial regularity is lost, and the descriptions become much larger. Furthermore, such descriptions are sensitive to noise in the image, a small perturbations in the image often leading to a large change in its description. To avoid some of these drawbacks, Zucker [Zuck76] proposed

$$
\begin{aligned}
G &= (V_N, V_T, P, S)\\
V_N &= \{S, A_1, A_2, A_3, A_4, N_0, N_1, 1, 0\}\\
V_T &= \{\square, \blacksquare\}
\end{aligned}
$$

$P:$

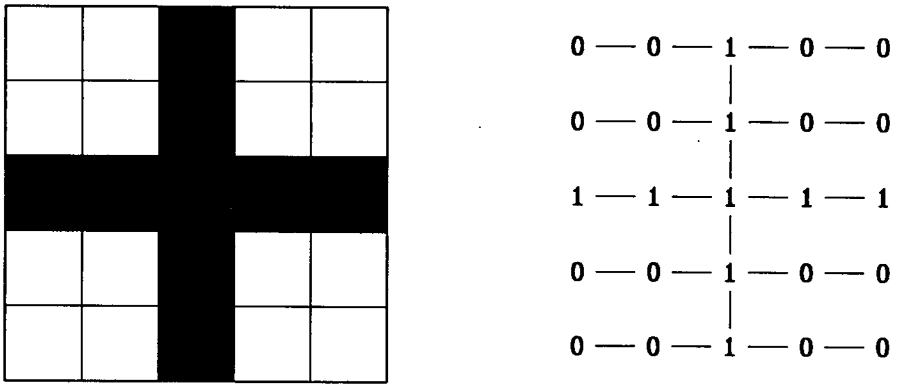
$$S \rightarrow A_1$$

$$
A_1 \rightarrow \begin{array}{ccccc} N_0 & - & 1 & - & N_0 \\ & & | & & \\ & & A_2 & & \end{array}
\qquad
A_4 \rightarrow \begin{array}{ccccc} N_0 & - & 1 & - & N_0 \\ & & | & & \\ & & A_4 & & \end{array}
\qquad
0 \rightarrow \square
$$

$$
A_2 \rightarrow \begin{array}{ccccc} N_0 & - & 1 & - & N_0 \\ & & | & & \\ & & A_3 & & \end{array}
\qquad
N_0 \rightarrow \begin{array}{ccc} 0 & , & 0 \\ | & & \\ N_0 & & \end{array}
\qquad
1 \rightarrow \blacksquare
$$

$$
A_3 \rightarrow \begin{array}{ccccc} N_1 & - & 1 & - & N_1 \\ & & | & & \\ & & A_4 & & \end{array}
\qquad
N_1 \rightarrow \begin{array}{ccc} 1 & , & 1 \\ | & & \\ N_1 & & \end{array}
$$

(a) tree grammar G



(b) pattern

(c) tree representation

Figure 2.2: example of tree grammar analysis

that there are two aspects of any natural texture: an ideal regular texture that forms its deep structure, and a spatial mapping that distorts it into the surface structure appearing in the image. This model of texture has led to several syntactic methods of texture analysis (e.g., [Fu82,ch12]).

Even for highly regular structures, a large amount of syntactic ambiguity is inevitable — many possible grammars exist for any given spatial pattern [Zuck76]. Before a structural description of a texture can be given, several *a priori* decisions must be made about its structure. When this is feasible (e.g., for classification of biological tissues), structural approaches prove useful for texture analysis.

2.2.3 Structural-Spatial Approaches

These approaches are hybrid, attempting to combine the best aspects of structural and spatial methods. As for a structural approach, microstructure elements are considered to be sparsely distributed throughout the image. The relations between them, however, are analyzed using spatial techniques. Local spatial structure can therefore be concisely described without imposing large constraints on the overall global structure of the texture.

Structural-spatial approaches characterize texture as a sparse set of spatially-ordered, structured elements. Each element is generally represented by a feature vector, whose values are obtained from the surrounding neighbourhood. These neighbourhoods are generally non-overlapping regions of finite extent, which may or may not form a partition of the plane.

Structural-spatial approaches can be divided into two groups, depending on how they characterize the neighbourhoods. The first group considers neighbourhoods as being contiguous areas of uniform grey level. The second group defines neighbourhoods using local extrema. All approaches can be characterized by the information extracted from each neighbourhood, and by the statistical properties of the resulting texture elements.

a) Uniform areas

One of the simplest ways of specifying neighbourhoods is to partition the image into a set of unidirectional grey level runs. These runs are defined as maximal collinear strings of constant grey level, oriented in some given direction [Gall75]; they are described by their run length, direction, and grey level. Description of a texture is based on the joint occurrence of grey level and run length in each direction. To reduce the size of a description, a set of five features is computed for each direction. These are similar in many ways to some of the measures developed by Haralick *et al* [HaSh73]. In general, run-length measures are not as useful as second-order statistics [WeDy76][CoHa80a].

Extensions of this approach to two-dimensional regions of constant grey level have been used for texture analysis [MaBr77][ToSH82]. Properties used for classification include the area, elongation, and grey level of the regions. Classification accuracies of over 80% can be achieved for several classes of natural texture [ToSh82].

b) Local extrema

Local extrema in an image form the basis for several structural-spatial methods of texture analysis. For one-dimensional extrema, the only features are height and width; these are measured using the neighbouring extrema. For two-dimensional extrema, neighbourhood boundaries can be formed in a variety of ways. One possibility is to associate with each extremum a *reachability set*, a set of points that can be reached from it along a monotonically increasing or decreasing path. Various properties such as its size, mean, and variance could be used [Hara79]. Such an approach, however, has not yet been thoroughly investigated.

One-dimensional local extrema form the basis for an extremely efficient technique for texture analysis — the max-min method [MiMy77]. The image is first smoothed to eliminate small fluctuations, and the logarithm taken to render the description independent of absolute intensity. Local extrema are then determined, being thresholded by a value of T above/below the neighbouring pixels. Description is given by the density of extrema for various values of T. Such a method is extremely fast, and has a classification accuracy comparable to those based on grey level dependence matrices [MiMy77].

Texture analysis has also been based on the local maxima of images filtered with the Laws masks [PiRo83]. Using only the first-order statistics of these maxima, a classification accuracy can be achieved that equals that of the original texture energy measures. This shows that the local maxima alone may contain all the essential information in texture [PiRo83].

Generalized co-occurrence matrices [DaJo79] describe texture using the relations between the local extrema present in an image. These matrices have a form much like the grey level dependence matrices, but their features are much more general: the joint occurrence of any property of neighbouring extrema can be used. When properties of local maxima of smoothed images are used in these matrices, classification accuracies can be achieved that are higher than those obtained using grey level dependency measures [DaJo79].

c) General performance

Structural-spatial analyses combine some of the best aspects of spatial and structural methods. Descriptions are readily determined, and are generally robust under small geometric distortions of the image. Furthermore, the analyses are also robust under monotonic changes of grey level.

Methods based on areas of uniform grey level, however, are sensitive to noise in the image. Smoothing the image would help somewhat, but the application of a smoothing filter would tend to alter the distribution of grey levels in many parts of the image, especially in areas near a boundary. This type of analysis is therefore inherently sensitive to noise.

Descriptions based on local structure are more suitable for texture analysis. Since the locations of extrema are invariant under monotonic transformations of grey level, descriptions tend to be robust, even under local filtering of the image. In addition, the classification accuracies of these methods are among the highest yet achieved [GoDe85]. This shows that local extrema can form the basis for robust and accurate analysis of texture.

2.3 Models of Texture Perception

Texture perception has been investigated using several psychophysical techniques. These generally involve restricted domains of synthetic textures, which are designed to isolate the spatial structure relevant to perceived texture. Although limited in scope, these methods have yielded valuable information about the ability of the human visual system to perceive texture.

Three different approaches to studying texture perception have been developed. Each is based on a somewhat different model of the process, and has its own distinctive character. Although some parts of the various models conflict, the three approaches are largely complementary, each modelling somewhat different aspects of a highly complex process.

2.3.1 Spatial-Feature Models

This approach concentrates on determining the spatial features that influence perceived texture. Texture perception is considered to be a pre-attentive process, occurring within the first few hundred milliseconds of presentation. Emphasis is placed on determining the necessary and sufficient conditions for two adjacent texture fields to be pre-attentively discriminable. This may equivalently be viewed as establishing the conditions under which two textures are perceptually identical, or *metameric*. This approach tends to be somewhat phenomenological in nature — little emphasis is placed on determining the underlying mechanisms involved.

The spatial-feature approach has its origins in the work of Julesz [Jule62] on the discrimination of random Markov textures. Based on these results, the conjecture was made that third- or higher-order statistics are irrelevant for texture discrimination. For textures with elements not locally distinguishable, the Julesz conjecture still appears to hold [PrFa78][Gaga81]. Second-order measures sufficient for discrimination are not known in general. The mean and variance, together with the covariance function, are not sufficient to describe texture completely [PrFa78].

All counterexamples to the original Julesz conjecture involve texture elements that are distinguishable locally [JuGi73][CaJu78]. This observation has led to the hypothesis that perceived texture depends only on the first-order densities of a specific set of localized spatial features [Jule75][Jule81][Beck82]. These *textons* are localized geometric shapes with simple properties; they include endpoints, elongated blobs, lines of various widths and lengths, and line-crossings. Texton properties include colour, binocular disparity, and orientation [Jule81]. Since only first-order densities are involved, the relative positions of textons to each other should not affect pre-attentively perceived texture. This prediction agrees with experiment [JuBe83][Jule84].

Textons have much in common with the set of pre-attentively distinguishable features found by Treisman [Trei85]. However, they are not to be identified with the elements of form vision, since they are considered to be part of a separate pre-attentive visual system [Jule84].

The texton theory, as developed by Julesz, has been largely based on the perception of simple texture elements scattered sparsely throughout an image — no algorithm need be given of how the descriptions are calculated. If the perception of more natural textures is to be understood, however, the determination of this process is essential. Caelli [Cael84] has taken a few steps toward this goal, showing that textons are members of a more general class of coding units.

2.3.2 Symbolic-Structure Models

A model of texture perception more concerned with underlying algorithm and mechanism is that proposed by Marr [Marr76][Marr82]. Texture perception is considered to be a non-attentive process, employing the same grouping operations and symbolic structures as used in form vision. As such, there is no separation between segmentation and classification. Texture discrimination is only one aspect of texture perception that can be treated using this approach — texture flow and grouping can be modelled as well.

As described by Marr, the basic elements of texture are exactly the basic elements of the primal sketch: blobs, endpoints, and lines. Each is represented by a token describing its size, location, contrast, orientation, etc. Various aggregation processes use local properties such as common orientation to create higher-level symbolic structures. This grouping can be done recursively, building up highly complex elements.

Texture discrimination is assumed to be based on the first-order density of the symbolic structures present locally. Such an approach can account for many of Julesz's results [Marr76]. In addition, several classes of metameric textures with different second-order statistics can be identified via the first-order statistics of virtual lines [Scha78]. These lines are purely symbolic structures, connecting pairs of dots in the primal sketch. More generally, they can connect arbitrary elements of the primal sketch [Marr82]. Virtual lines can also be used to show how local processes can cause the Moiré effect or texture flow seen in Glass patterns [Stev78].

The symbolic-structure approach has proven difficult to develop, largely owing to its inherent complexity. Even the tokens are difficult to ascertain [Rile81]. The operation of the grouping processes must also be reconciled with the indifference of perceived texture to the relative positions of the texture elements. More recent approaches (e.g., [Zuck84]) have tended to explain many of these processes by simple spatial operations applied to simple spatial elements.

The symbolic-structure and spatial-feature approaches are similar in several ways. Both make use of a basic set of simple elements that, apart from the line-crossings included by Julesz, are of much the same form. The properties that these elements have are also similar. For simple textures, then, these two models generally make similar predictions about which textures are discriminable.

Ontologically, however, these basic elements are distinct: primal-sketch elements are the basis of (attentive) form vision, while textons are part of a completely separate pre-attentive system. This is reflected in the distinction drawn between pre-attentive and non-attentive perception. The elements of the symbolic-structure approach, being part of a more powerful form vision system, can be grouped into higher-level features that may enter the description of a texture. The spatial-feature approach, on the other hand, explicitly rejects constructive processes as having a role in texture perception [Jule84]. In this view, texture perception involves only detection processes based on a set of simple spatial features.

2.3.3 Spatial-Frequency Models

This approach models the attentive perception of random textures using a set of parallel spatial-frequency channels. Each channel describes the convolution of the original image with a specific filter. By studying the apparent similarity of various random textures, some insight can be gained into the structure of these filters, since similar textures should be exactly those that have similar properties in each channel.

Spatial-frequency channels have their origin in the work of Campbell and Robson [CaRo68]. From studies on threshold spatial vision, Wilson *et al*

[WiBe79][WiGe84] determined the shape of the channel filters as being the sums and differences of several Gaussian functions. Since the Fourier transform of a Gaussian function is another Gaussian function, the general form of these filters are similar in both the spatial and the frequency domains. A set of $4-6$ filters is postulated, the individual filters being nearly identical in shape, and differing in size from each other by a factor of approximately two.

Richards and Polit [RiPo74] were the first to explain perceived texture using spatial-frequency channels. They established that only four different combinations of spatial frequencies are needed to serve as the basis functions of a perceptual space for line textures. Any line texture can therefore be perceptually matched by an appropriate linear combination of these functions; this suggests that there exist four physiological spatial filters mediating texture perception.

Interestingly, the shapes of these filters correspond closely to those later determined by Wilson *et al* [WiBe79] from work on threshold vision. These filters have also been shown to form a possible basis for the grouping of texture into classes of apparent similarity [HaGe78][HaGe81].

The spatial-frequency approach has several inherent advantages and disadvantages for texture representation. On the positive side, the descriptions are reliable [MaMo84]: they are easy to compute, are invariant under translation and rotation, vary continuously with change in the image, and capture information at several levels of detail. In addition, a metric can be established to determine the distance separating two dissimilar textures. On the other hand, this approach leaves unspecified the characteristic features (if any) being measured in each channel. This makes it difficult to generalize from line textures to the fully two-dimensional case. Furthermore, many shapes are possible for the filters characterizing the channels, so that assumptions must be made about their form. Unless firm links can be established between texture perception and other aspects of vision, any spatial-frequency model must contain a large degree of arbitrariness.

This work examines one possible link, investigating whether texture perception is based on measures such as the relative contrasts and zero-crossing densities in each of the filtered images. In this regard, the near-identity of the postulated filters under rescaling suggests that it is interesting to examine the discrimination of self-similar random textures. Assuming spatial-frequency filters of the form proposed by Wilson and Gelb [WiGe84], analysis shows that the texture-discrimination results are consistent with the multiple-channel model.

The use of spatial-frequency channels to model texture perception is compatible with the assumptions of the other two approaches. If the frequency bandwidth of a filter is sufficiently large, the corresponding convolution mask can have arbitrarily fine resolution in the spatial domain. As such, the features present in these multiresolution images may be precursors for the basic elements of the spatial-feature and symbolic-structure approaches.

Multiresolution representations have been successfully used for modelling several other aspects of vision [Grim81][Terz82][Burt84]. Whether such an approach also provides a good model for texture perception remains to be seen.

Chapter 3

Self-Similar Random Textures

Increased attention has recently been given to the modelling of random textures by self-similar stochastic fractals (e.g., [Pent83][MeYa84]). Such objects, introduced by Mandelbrot [MaNe68][Mand82], have a self-similar structure — any characteristic present at a small scale is also present at a larger scale (see appendix B). Their spatial structure is therefore complex, with no well-defined partition existing between object boundary and intrinsic surface structure. Many random textures can be accurately described as fractals, and calculation of their fractal dimension (see appendix B) has led to classification accuracies as high as 85% [Pent83][PeNa84]. Furthermore, the fractal dimension of a surface appears to correlate closely with its perceived roughness [Pent84].

In order to investigate the ability of the human visual system to discriminate among random textures with different fractal properties, it is useful to relate these properties to more conventional descriptions of texture. This chapter shows how this can be done. It also shows how textures with fractal properties can be viewed as special cases of a more general class of self-similar random textures. These are made up of instances of n-dimensional random fields with power spectra $S(\vec{k})$ such that for some $h, H, \gamma \in \Re$

$$[S(h\vec{k}) - \gamma\delta(0)] = h^{-n-2H}[S(\vec{k}) - \gamma\delta(0)],$$

where $\vec{k}$ is the spatial frequency, and $\delta(0)$ is the Dirac delta function.

This work examines the factors affecting the discriminability of self-similar random textures. A straightforward algorithm is developed to generate such textures for use in the psychophysical experiments described in chapter 4.

3.1 General Properties

All statistical properties of a stationary Gaussian random field are completely governed by its mean μ and covariance function $C(\vec{x})$, or equivalently, by its mean μ and power spectrum $S(\vec{k})$. Intuitively, any self-similar structure in such a field must be reflected in some form of self-similarity in its covariance function and power spectrum.

This section examines the form of the covariance function and power spectrum for self-similar random fields. Since self-similarity is characterized here by a two-point measure, third- and higher-order statistics are not relevant. The restriction that the field be Gaussian can be therefore be relaxed. In what follows, the field $\mathbf{f}(\vec{x})$ is taken to be any n-dimensional stationary random field.

3.1.1 Self-Similar Stochastic Fractals

A self-similar stochastic fractal $\vec{\mathrm{a}}(\vec{x})$ is characterized by the equation

$$\vec{\mathrm{a}}(\vec{x}_1 + h(\vec{x}_2 - \vec{x}_1)) - \vec{\mathrm{a}}(\vec{x}_1) = h^H[\vec{\mathrm{a}}(\vec{x}_2) - \vec{\mathrm{a}}(\vec{x}_1)],$$

for some $h, H \in \Re$, $h > 1$ (see appendix B). When $\vec{\mathrm{a}}(\vec{x})$ is a random field $\mathbf{f}(\vec{x})$, its description can be recast into more conventional form. This reformulation allows description of various properties in both the spatial and the frequency domains. Such a treatment shows that such fractals are special cases of self-similar random fields.

a) Self-similar covariance functions

Theorem 1: A stationary random field has a covariance function $C(\vec{x})$ such that within some range $\lambda < |\vec{x}| < \Lambda$

$$C(h\vec{x}) - C(\vec{0}) = h^{2H}[C(\vec{x}) - C(\vec{0})]; \quad h, H \in \Re, h > 1$$

iff within that range the field behaves as a stationary stochastic fractal, with scaling ratio h, and similarity parameter H.

Proof: If the random field $\mathbf{f}(\vec{x})$ is stationary, the behaviour of its increments can be described by

$$\mathbf{f}(\vec{x}_1 + h(\vec{x}_2 - \vec{x}_1)) - \mathbf{f}(\vec{x}_1) = w(h, H, \vec{x})[\mathbf{f}(\vec{x}_2) - \mathbf{f}(\vec{x}_1)]$$

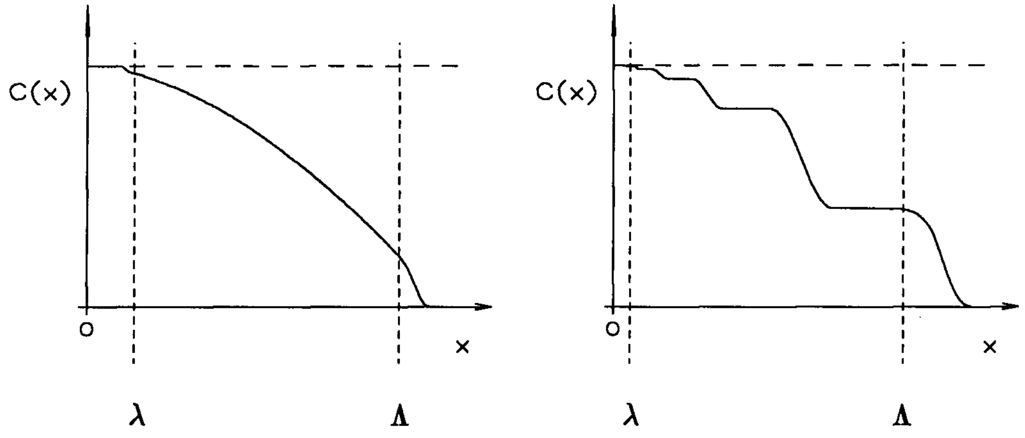
where $\vec{x}_1$ and $\vec{x}_2$ are arbitrary points, $\vec{x} = \vec{x}_2 - \vec{x}_1$, and $w(h, H, \vec{x})$ is a function as yet undetermined. Taking the variance of both sides and using the symmetry of the covariance function yields

$$[C(h\vec{x}) - C(\vec{0})] = w(h, H, \vec{x})^2[C(\vec{x}) - C(\vec{0})].$$

When $\lambda < |\vec{x}| < \Lambda$, $w(h, H, \vec{x})$ can be identified as h^H; the random field therefore exhibits fractal behaviour in this range.

Conversely, if the field exhibits fractal behaviour, its increments are such that

$$\mathbf{f}(\vec{x}_1 + h(\vec{x}_2 - \vec{x}_1)) - \mathbf{f}(\vec{x}_1) = h^H[\mathbf{f}(\vec{x}_2) - \mathbf{f}(\vec{x}_1)]; \quad \lambda < |\vec{x}| < \Lambda.$$



(a) $H = 0.85$, $h \to 1$ (b) $H = 0.85$, $h = 2$

Figure 3.1: examples of self-similar covariance functions

Taking the variance of both sides leads to a covariance function of the appropriate form. ■

Thus, a stationary random field has self-similar fractal behaviour iff its covariance function is of the form given in theorem 1. Figure 3.1 shows a few possible shapes for $C(x)$ in the one-dimensional case.

In higher dimensions, the random field is not necessarily isotropic, for $C(\vec{x})$ need not be rotationally symmetric. Indeed, the values of h and H may vary as a function of the direction of the displacement $\vec{x}$. The general case, however, is not developed here. Instead, the random fields are assumed isotropic.

The upper and lower cutoff scales are denoted by Λ and λ respectively. Taking the limits $\lambda \to 0$ and $\Lambda \to \infty$, the field becomes a true stochastic fractal, self-similar over all spatial scales. Given some initial displacement $\vec{x}_0$, it follows from self-similarity that the difference $C(\vec{x}) - C(0)$ is proportional to

h^{2Hj} for displacements of the form $\vec{x} = h^j \vec{x}_0$, $j \in Z$. As $j \to \infty$, this difference increases without bound, forcing $C(\vec{x})$ to become increasingly negative (figure 3.1). Since a covariance function is subject to the constraint that the variance $C(0) \geq |C(\vec{x})|$ [Papo84,Ch10], it follows that a random field with true self-similar behaviour must have an infinite variance.

To avoid such divergences in any physical realization, the fractal behaviour of a texture must be limited to some finite spatial range. This can be achieved by multiplying a true self-similar covariance function by a window function $w_p(\vec{x})$, where $p > 0$ is a measure of the window size. The function $w_p(\vec{x})$ may take such forms as $e^{-\alpha|\vec{x}|^2}$ or $\mathrm{sinc}(\beta|\vec{x}|)$, where $\alpha, \beta \in \Re > 0$. By appropriate choice of window parameters, self-similar behaviour of arbitrary accuracy can be achieved in any finite spatial range $|\vec{x}| < \Lambda$. At scales below this range, where theorem 1 still holds, the random field may be be considered a true fractal.

b) Self-similar power spectra

Theorem 2: Let $\mathbf{f}(\vec{x})$ be a stationary n-dimensional random field with a power spectrum $S(\vec{k})$ bounded above by $A|\vec{k}|^{-n-2H} + \eta\delta(0)$ for some $A, H, \eta \in \Re$, $0 < H < 1$. If $S(\vec{k})$ approaches a form such that

$$[S(h\vec{k}) - \gamma\delta(0)] = h^{-n-2H}[S(\vec{k}) - \gamma\delta(0)]; \quad h > 1$$

for some $\gamma \in \Re$, then the behaviour of $\mathbf{f}(\vec{x})$ approaches that of a true stochastic fractal, with scaling ratio h, and similarity parameter H.

Proof: Consider the function

$$\tilde{S}_a(\vec{x}) = \int_a^\infty S(\vec{k}) \exp\{i2\pi(\vec{x} \cdot \vec{k})\} d\vec{k}.$$

The central area $|\vec{k}| < a$ of $S(k)$ has been deleted, and its Fourier transform obtained. This transform can be rewritten as a function of radial distance $k = |\vec{k}|$ and $n-1$ angular parameters. Since $S(\vec{k}) < A|\vec{k}|^{-n-2H}$, its integral over the $n-1$ angular parameters is bounded by Bk^{-2H-1}, where $B \in \mathfrak{R}$ is some finite number. When $a > 0$ and $H > 0$, the integral over radial distance k is also finite. Since $\tilde{S}_a(0)$ is finite, $\tilde{S}_a(\vec{x})$ must exist for all $\vec{x}$.

Subtracting the term $\tilde{S}_a(0)$ and rescaling yields

$$\tilde{S}_a(h\vec{x}) - \tilde{S}_a(0) = h^{-n}\int_{ha}^{\infty} S(\vec{k}/h)[\exp\{i2\pi(\vec{x}\cdot\vec{k})\} - 1]d\vec{k}.$$

Due to the term $[\exp\{i2\pi(\vec{x}\cdot\vec{k})\} - 1]$, the contribution of $\gamma\delta(0)$ to any integral is zero. Since $\gamma\delta(0)$ does not influence later developments, the value of γ may conveniently be set to zero. Substituting the term $h^{n+2H}\tilde{S}_a(\vec{k}) = \tilde{S}_a(\vec{k}/h)$ into the above integral then leads to

$$\tilde{S}_a(h\vec{x}) - \tilde{S}_a(0) = h^{2H}[\tilde{S}_a(\vec{x}) - \tilde{S}_a(0)] - h^{2H}\int_{a}^{ah} S(\vec{k})[\exp\{i2\pi(\vec{x}\cdot\vec{k})\} - 1]d\vec{k}.$$

This last term describes the error from true self-similarity. Owing to the symmetry of $S(\vec{k})$, the sine component of $\tilde{S}_a(\vec{x})$ is zero. The exponential can therefore be replaced by a cosine. For $a|\vec{x}| << 1$, the magnitude of the error term obeys the inequality

$$\begin{aligned}
|h^{2H}\textstyle\int_a^{ah} S(\vec{k})[\cos(2\pi\vec{x}\cdot\vec{k}) - 1]d\vec{k}| &< |h^{2H}\textstyle\int_a^{ah} S(\vec{k})(2\pi\vec{x}\cdot\vec{k})^2 d\vec{k}| \\
&< |8\pi^{n+1}h^{2H}\textstyle\int_a^{ah} S(k)(|\vec{x}|k)^2 k^{n-1}dk| \\
&< |8\pi^{n+1}h^{2H}A\textstyle\int_a^{ah} k^{-n-2H}(|\vec{x}|k)^2 k^{n-1}dk| \\
&= 8\pi^{n+1}A|h^2 - h^{2H}|(2-2H)^{-1}a^{2-2H}|\vec{x}|^2.
\end{aligned}$$

This last result shows that the deviation in $\tilde{S}_a(\vec{x})$ from true self-similarity has an upper bound that goes as the square of the distance from the origin. For any given amount of error, then, a spatial range $|\vec{x}| < \Lambda$ can be found within which $\tilde{S}_a(\vec{x})$ has asymptotically self-similar behaviour.

Decreasing the value of a reduces the size of the error term, since $H < 1$. The range of self-similarity behaviour shown by $\tilde{S}_a(\vec{x})$ correspondingly increases. Since

$$\lim_{a \to 0} \tilde{S}_a(\vec{x}) = R(\vec{x}) = C(\vec{x}) + \mu^2,$$

it follows that the behaviour of $C(\vec{x})$ approaches

$$[C(h\vec{x}) - C(0)] = h^{2H}[C(\vec{x}) - C(0)].$$

Theorem 1 may then be invoked to show that the field exhibits self-similar fractal behaviour within a range that increases without bound as $a \to 0$.

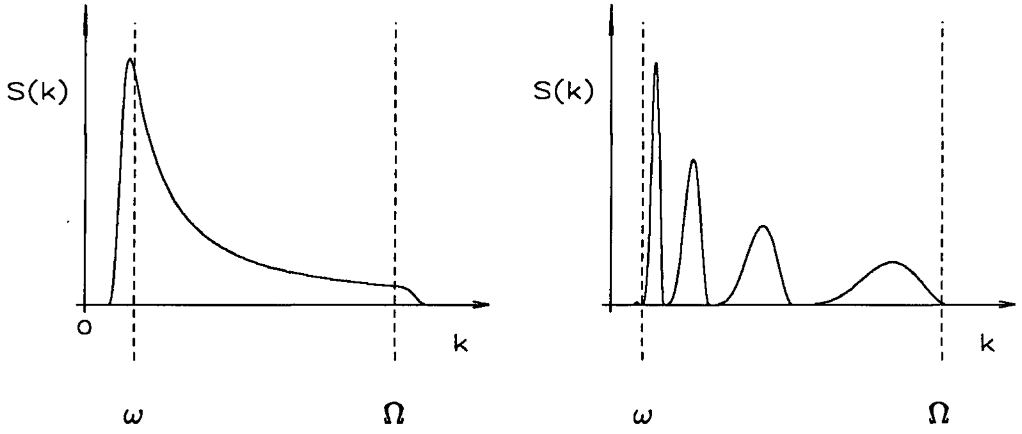
As noted above, the value of γ has no effect upon self-similarity. Since the mean μ of a random field contributes only a term $\mu^2\delta(0)$ to the power spectrum (see appendix A), this implies that the self-similar behaviour of the field is indifferent to the value of its mean. ■

For a one-dimensional power spectrum such that

$$[S(hk) - \gamma\delta(0)] = h^{-1-2H}[S(k) - \gamma\delta(0)],$$

it follows from theorem 2 that the corresponding random field is a self-similar fractal with similarity parameter H. When $h \to 1$, the fractal becomes self-similar under all scaling ratios. Setting $\gamma = 0$, this reduces to Mandelbrot's result, which states that $S(k) \propto k^{-1-2H}$ (see appendix B).

Any self-similar power spectrum of the form given in theorem 2 will correspond to a stationary stochastic fractal. Examples of such generalized one-dimensional spectra are shown in figure 3.2.



(a) $H = 0.20$, $h \to 1$ (b) $H = 0.00$, $h = 2$

Figure 3.2: examples of self-similar power spectra

Note that the scaling ratio h can vary independently of the similarity parameter H, and that their values can be common to a wide variety of spectral shapes.

3.1.2 Self-similar Noises

The restriction $0 < H < 1$ in theorem 2 stems from the requirement that the corresponding random field exhibit fractal behaviour. If this requirement is dropped, the only constraint governing H is that the covariance function $C(\vec{x})$ exists. Since the power spectrum remains self-similar, it follows that some form of self-similarity must also exist in the random field.

Theorem 3: Let $\mathbf{f}(\vec{x})$ be a stationary n-dimensional random field with a power spectrum $S(\vec{k})$ bounded above by $A|\vec{k}|^{-n-2H} + \eta\delta(0)$ for some $A, H, \eta \in \Re$, with $-n/2 < H < 0$. If $S(\vec{k})$ is the linear combination of a finite number of monotonic functions, and is such that when $\mu = 0$

$$[S(h\vec{k}) - \gamma\delta(0)] = h^{-n-2H}[S(\vec{k}) - \gamma\delta(0)]; \quad h > 1,$$

for some $\gamma \in \Re$, then $\mathbf{f}(\vec{x})$ has a covariance function $C(\vec{x})$ such that

$$[C(h\vec{x}) - \gamma] = h^{2H}[C(\vec{x}) - \gamma]; \quad \vec{x} \neq 0$$

for any value of the mean μ of the field.

Proof: To show that $C(\vec{x}) = \tilde{S}(\vec{x})$ exists, consider the case where $S(\vec{k})$ is a monotonic self-similar function of $|\vec{k}|$. Partition $\vec{k}$-space into the regions:

$$\begin{aligned} &i) \quad |\vec{k}| \leq a \\ &ii) \quad |\vec{k}| > a, \quad k_i \leq a, \quad i = 1, 2, ..., n \\ &iii) \quad k_i > a, \quad i = 1, 2, ..., n \end{aligned}$$

where the k_i are the components of $\vec{k}$, and $a \in \Re$ is some positive value. Since $H < 0$, the integral of $S(\vec{k})$ over region (i)

$$\int_0^a S(k)k^{n-1}dk < \int_0^a Ak^{-2H-1}dk \quad + \quad \eta$$

is finite. This implies that

$$\int S(\vec{k}) \exp\{i2\pi(\vec{x} \cdot \vec{k})\}d\vec{k} \leq 2\pi^{n-1} \int_0^a S(k)k^{n-1}dk$$

must also be finite.

The value of $S(k)$ over region (ii) is bounded from above by $A|\vec{k}|^{-n-2H}$. Since $S(k)$ is integrated over a finite range of $\vec{k}$-space, and $|\vec{k}| > a$, the contribution from region (ii) is finite as well.

The contribution from the third region can be expressed as

$$\int_a^\infty \ldots \int_a^\infty S(k) \exp\{i2\pi(\vec{x} \cdot \vec{k})\} \; dk_1 \ldots dk_n.$$

Let the n components of the displacement vector $\vec{x}$ be represented by $x_1, x_2, \ldots, x_n$. Owing to the rotational symmetry of $S(k)$, there is no loss of generality by assuming that all components of $\vec{x}$ are non-zero. Using the symmetry of the cosine function, the contribution of region (iii) can then be written

$$\int_a^\infty \ldots \int_a^\infty S(k) \cos(x_1 k_1) \ldots \cos(x_n k_n) dk_1 \ldots dk_n.$$

When $-n/2 < H$, $S(k)$ becomes a bounded, monotonically decreasing function in this region. Since $k^2 = \sum_i k_i^2$, monotonicity also holds for each single component k_i. The integral along any dimension i,

$$\int_a^\infty S(k) \cos(x_i k_i) dk_i$$

is therefore finite, owing to the monotonicity of $S(k)$ and the periodic symmetry of $\cos(xk)$ about zero when $x \neq 0$. The entire integral over all n dimensions must therefore also be finite.

Since the contributions of all three regions are finite when $\vec{x} \neq 0$, $\tilde{S}(\vec{x})$ must exist for all non-zero $\vec{x}$. By the linearity of Fourier transformation, the Fourier transform $\tilde{S}(\vec{x})$ must also exist if $S(k)$ results from the linear combination of several monotonically increasing or decreasing functions.

To obtain the formal relation between the self-similar behaviour of $S(\vec{k})$ and $C(\vec{x})$, consider the case $\mu = 0$. The formal relation between $S(\vec{k})$ and $C(\vec{x})$ (see appendix A) yields

$$C_0(\vec{x}) - \gamma = \int [S_0(\vec{k}) - \gamma\delta(0)] \exp\{i2\pi(\vec{x} \cdot \vec{k})\} d\vec{k},$$

where the subscript denotes that this holds only for the case $\mu = 0$. Rescaling by some $h > 1$ leads to

$$C_0(h\vec{x}) - \gamma = h^{-n} \int [S_0 - \gamma\delta(0)](\vec{k}/h) \exp\{i2\pi(\vec{x} \cdot \vec{k})\} d\vec{k}.$$

Direct substitution of $[S(\vec{k}) - \gamma\delta(0)] = h^{-n-2H}[S(\vec{k}/h) - \gamma\delta(0)]$ into this expression yields the result for $\mu = 0$.

The effect of setting the mean μ of a random field to a non-zero value is to add a term $\mu^2\delta(0)$ to the power spectrum $S_0(\vec{k})$ (see appendix A). The covariance function of the field is (see appendix A)

$$C(\vec{x}) = \int S(\vec{k}) \exp\{i2\pi(\vec{x} \cdot \vec{k})\} d\vec{k} \quad - \quad \mu^2,$$

so that

$$\begin{aligned} C(\vec{x}) - \gamma &= \int [S(\vec{k}) - (\gamma + \mu^2)\delta(0)] \exp\{i2\pi(\vec{x} \cdot \vec{k})\} d\vec{k} \\ &= \int [S_0(\vec{k}) - \gamma\delta(0)] \exp\{i2\pi(\vec{x} \cdot \vec{k})\} d\vec{k} \\ &= C_0(\vec{x}) - \gamma. \end{aligned}$$

This shows that the mean μ of the random field has no effect on the self-similar behaviour of the covariance function. ■

Thus, when the similarity parameter H has the values $-n/2 < H < 0$, a different type of self-similar random field results. These fields do not exhibit true fractal self-similarity: their covariance functions have the same type of self-similarity as their power spectra. Such random fields were first brought to attention by Mandelbrot [MaNe68], under the name of fractional Gaussian noises. The fields developed here are a generalization of these. They will be referred to as *self-similar noises.*

Fractional Gaussian noise is a one-dimensional random field, or stochastic process. Its power spectrum has the form $S(\vec{k}) = C_H|\vec{k}|^{-1-2H}$, where C_H is an arbitrary constant, and $-1/2 < H < 0$ (see appendix B). The scaling ratio h approaches unity, so that the power spectrum is self-similar over all scales. The

Fourier transform of $C_H|\vec{k}|^{-1-2H}$ can be evaluated via [GrRy65:3.762]

$$\int x^{-1+\beta}\cos(ax)dx = a^{-\beta}\Gamma(\beta)\cos(\beta\pi/2); \quad 0 < \beta < 1$$

to yield a covariance function of the form

$$C(u) = \frac{C_H}{\Gamma(2H+1)\sin(-H\pi)}|u|^{2H}; \quad -\frac{1}{2} < H < 0.$$

The self-similar behaviour of fractional Gaussian noise therefore agrees with the general results of theorem 3.

3.1.3 Effectively Self-Similar Textures

In general, any physical structure exhibits self-similar behaviour only within a certain range of spatial scales. For example, a coastline cannot have a definite structure at scales less than the size of a grain of fine sand, and is limited at the other extreme by the size of the Earth. The concept of true self-similarity must therefore be replaced by the notion of *effective self-similarity* : the similarity of measurements made over a limited range of scales. If this range falls well within the limits set by the upper and lower cut-off scales Λ and λ, there will be little difference in measurements made on true and effectively self-similar structures.

This notion of effective self-similarity can be used to develop a general class of random fields. The power spectra characterized by

$$[S(h\vec{k}) - \gamma\delta(0)] = h^{-n-2H}[S(\vec{k}) - \gamma\delta(0)]$$

correspond to well-defined random fields when $-n/2 < H < 0$ and $0 < H < 1$. Outside this range, their Fourier transforms do not necessarily exist. If the power spectrum is required to be self-similar only between the limits ω and Ω, however, this consideration does not apply — if $S(k)$ approaches zero quickly enough as $k \to 0$ and $k \to \infty$, the similarity parameter H may take on any real value, positive or negative. Such spectra describe a general class of effectively self-similar random fields.

The effectiveness of this self-similarity has an obvious dependence on the cut-off scales of the spatial structure being measured and the resolution of the measuring function used. For the human visual system, the greatest sensitivity to spatial frequency lies within the range 0.1 cyc/deg to 30 cyc/deg [CaRo68]. To be effectively self-similar, then, the random field need only have a power spectrum self-similar over this range of frequencies. Workers in computer graphics have discovered that such effectively self-similar fields are perceived as having much the same qualitative structure as fractals and self-similar noises [HaBa84].

3.2 Texture Generation

While suitable for the generation of many self-similar random textures, the fractal-generating algorithms described in the literature (e.g., [Mand75][FoFu82][HaBa84]) are not flexible enough to allow independent variation of all stochastic parameters of interest. More conventional techniques for producing random textures (e.g., [PrFa78][Scha80][Gaga81]), on the other hand, allow virtually complete control of statistical properties, but their specifications have usually lacked a direct connection to fractal properties. Such a connection, however, has been established in section 3.1, where various characteristics of self-similar random fields have been cast into terms involving covariance functions and power spectra. This provides a basis for generating self-similar random textures via the Fourier transformation of random variables. The Fourier approach has been used previously to generate conventional textures (e.g.,[Scha80]) as well as the Fourier-Brown-Wiener fractals [Mand82], but the framework established here allows production of a much larger class of self-similar textures.

3.2.1 Basis of the Algorithm

The generation of a one-dimensional Gaussian random field $\mathbf{f}(x)$ can be based on Fourier transformation, viz.,

$$\mathbf{f}(x) = \int_{-\infty}^{\infty} m(k)\mathbf{z}(k) \exp\{i2\pi kx\}dk,$$

where the $\mathbf{z}(k)$ are a one-dimensional field of identical, independent, zero-mean, delta-variance Gaussian random variables, and $m(k)$ is a modulating function. The power spectrum of $\mathbf{f}(x)$ is $S(k) = |m(k)|^2$; its covariance function $C(x)$ is the Fourier transform of $S(k)$. The field $\mathbf{f}(x)$, being a linear combination of independent Gaussian variables, is a stationary, zero-mean Gaussian random field, completely specified by $C(x)$ [Papo84]. This entails that $\mathbf{f}(x)$ is completely specified by $m(k)$, an easily-controllable quantity.

The generation of random self-similar textures, as developed here, is based on this result. Various one-dimensional random fields (i.e., stochastic processes) can be generated by specifying different forms for $m(k)$. The resulting images then need only be swept down through a finite distance to produce the line textures.

In what follows, only continuous functions are discussed. Any physical realization of an image, however, must be both bounded and discrete; its spectral representation must have a similar constraint. In Appendix C, it is shown that self-similarity can be effectively captured by discrete images. Textures can therefore be generated by the Fourier transformation of discrete unit-variance Gaussian random variables $\mathbf{z}_k$ modulated by a discrete function m_k.

3.2.2 Specification of Statistical Properties

A wide variety of random fields have a power spectrum $S(k)$ such that

$$[S(hk) - \gamma\delta(0)] = h^{-1-2H}[S(k) - \gamma\delta(0)].$$

For the case $h \to 1$, $S(k)$ must be proportional to k^{-1-2H}. For other values of h, its form is underconstrained, and various statistical properties can be specified independently of H and h.

The shape of $S(k)$ may be specified by using a fixed template function $P(k)$ to describe its values over the range $[1, h)$, where $h > 1$ is to be the scaling ratio. This pattern is then repeated for all intervals $[h^j, h^{j+1}), j \in Z$, with $P(k)$ being geometrically scaled up or down by $h^{(-1-2H)j}$, and its argument adjusted accordingly. This leads to a spectrum of the form

$$S(k) = cP(|k|/h^j)h^{(-1-2H)j}; \quad h^j \leq |k| < h^{j+1},$$

where j is the integer denoting the particular interval, and c is an arbitrary positive constant. The template function $P(k)$ is an arbitrary bounded function, constrained to be positive. The resultant spectrum is obviously self-similar; an example is given in figure 3.3. From theorems 2 and 3, such a power spectrum describes a fractal when $0 < H < 1$, and a self-similar noise when $-1/2 < H < 0$.

The square root $m(k)$ of $S(k)$ is the modulating function required to generate the random field $\mathtt{f}(x)$. By suitable design of $S(k)$, therefore, several statistical properties can be independently specified. Among these are:

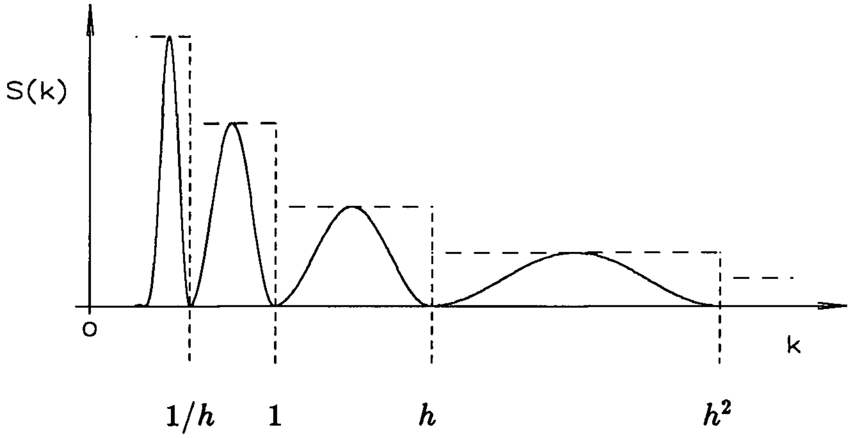


Figure 3.3: example of template construction

Similarity parameter

The similarity parameter H, closely related to the fractal dimension D (see appendix B), can have any real value. The choice of H determines whether $\mathbf{f}(x)$ is a true fractal, a self-similar noise, or an effectively self-similar random field.

Scaling ratio

The scaling ratio h can be given any value greater than unity. In the limit $h \to 1$, it leads to a random field self-similar for any scaling ratio, i.e., $S(k) \propto k^{-1-2H}$. Otherwise, it may be specified independently of H and $P(k)$.

Variance

All self-similar textures ideally have an infinite variance. In practice, however, any random field is only effectively self-similar, having a finite variance. This quantity can be varied by altering the value of c.

Moments

Different moments of $S(k)$ — or equivalently, of $C(x)$ — can be obtained by changing the form of the template function $P(k)$. Again, this can be done without altering the values of the similarity parameter and scaling ratio of the random field.

Chapter 4

Texture-Discrimination Experiments

This chapter describes the psychophysical experiments carried out to determine aspects of self-similar random textures relevant to human visual perception. The experiments used a two-alternative forced-choice (2AFC) method to measure the discriminability of line textures taken from different parent ensembles.

4.1 General Format

4.1.1 Subjects

Two volunteers participated in the experiments. Subject A, the author, had vision corrected to normal. Subject B had uncorrected normal vision, and was unaware of the purpose of the experiments.

4.1.2 Stimuli and Apparatus

All stimuli were composed of three rectangular line textures, placed one above the other (figure 4.1). The dimension of each rectangle was 128 × 256 pixels, or 7.5 cm × 15 cm. A dark border of width 16 pixels surrounded each texture,

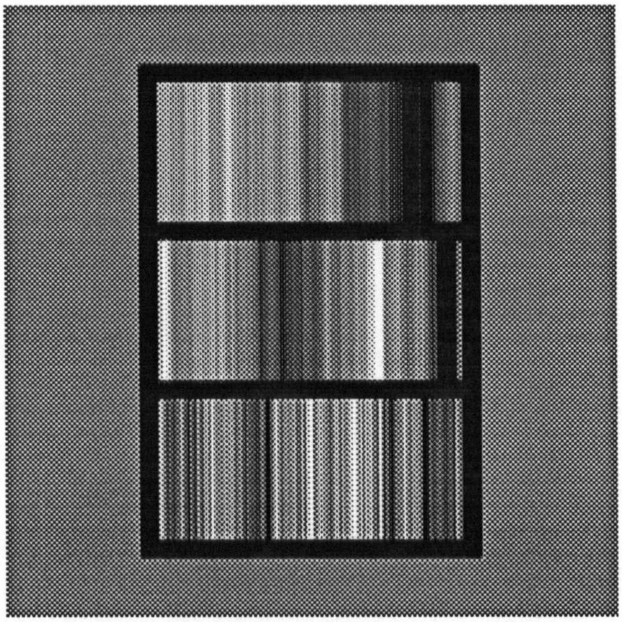


Figure 4.1: display format

separating it from its neighbours. This array was surrounded by a uniform field with a luminance equal to the average value of the textures.

Each display contained line textures from the two classes being investigated. Two of the three textures came from the same class (i.e., they were generated using the same spectral parameters), while the third was an instance of the other random field. A set of 20 different instances was generated for each class. Selection and positioning of the instances in a given display were done randomly, subject only to the constraint that textures from the same class be adjacent. This meant that the task of the subject was to pair the middle texture with either the top or the bottom texture.

Stimulus patterns were displayed on a Hitachi HM-2719B-C-11 monitor. Each texture had an average luminance of 30.0 cd/m^2 and contrast of 0.7 (see appendix C). The distance from the subject to the screen was set to approximately 200 cm, so that each texture subtended an angle of $2° \times 4°$. The textures consequently had half-power bandlimits of 0.12 cyc/deg and 31.9 cyc/deg (see appendix C).

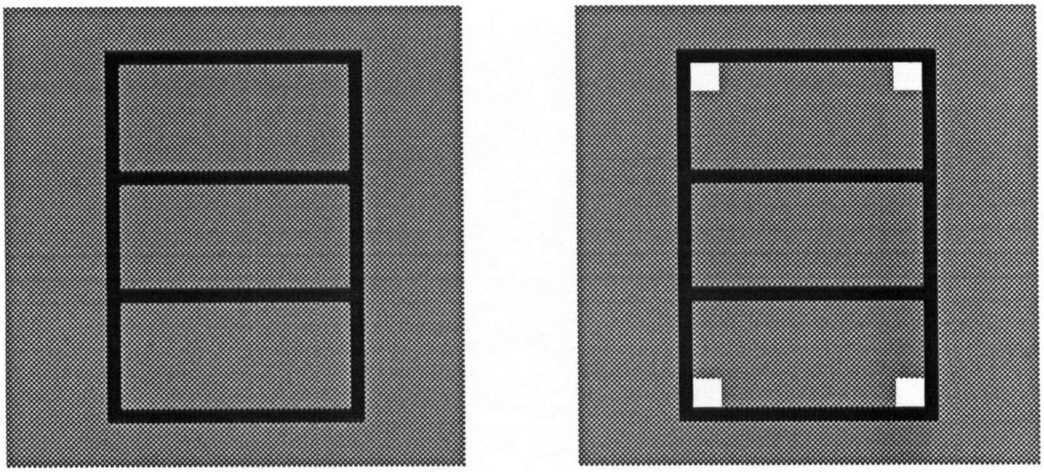
4.1.3 Presentation

At the start of an experiment, subjects were presented with a display similar to the stimulus pattern, but with rectangles of uniform intensity in place of the textures (figure 4.2(a)). The luminance of these rectangles was set equal to the average luminance of the textures. This minimized any effects of sudden luminosity changes in the display when line textures replaced the rectangles in a presentation.

Each presentation of a stimulus pattern was preceded by an acoustic warning signal. This was followed by a visual warning signal: a one-second flash of four white squares on the display (figure 4.2(b)).

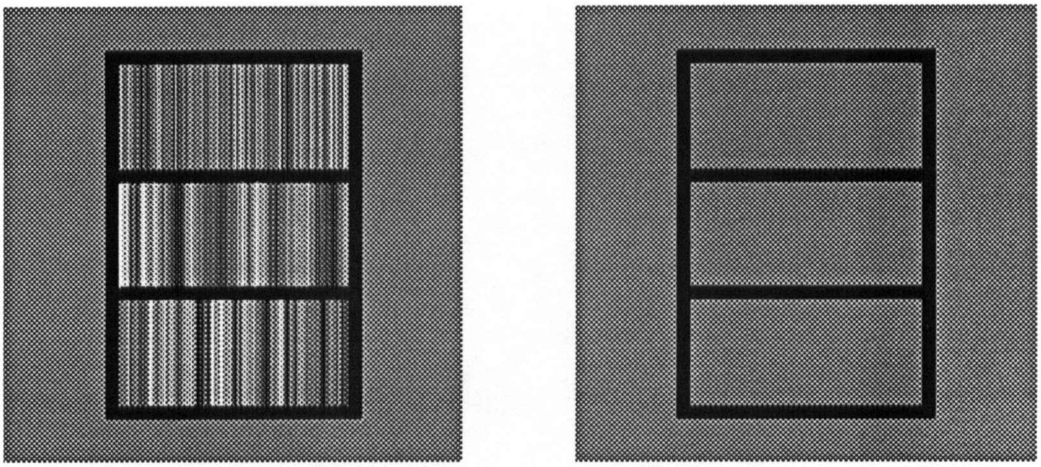
One second later, the stimulus pattern was displayed (figure 4.2(c)). The presentation lasted for five seconds, after which the line textures were replaced by the original uniform rectangles. Subjects were then asked whether the upper or lower pair of textures appeared more similar. Following standard psychophysical practice [GrSw66], they were subsequently informed of the correctness of their response by the experimenter. To avoid biasing the responses, the experimenter did not know the correct answer until after the response of the subject had been recorded.

The duration of stimulus presentation was found not to affect the performance of the subjects. Several presentation intervals, ranging from 3 seconds to 9 seconds, were tested. Performance did not vary significantly.



(a) uniform rectangles

(b) visual warning signal



(c) line textures

(d) uniform rectangles

Figure 4.2: presentation sequence

To measure the discriminability between two classes of texture, subjects were given a series of consecutive presentations, each involving randomly-selected elements of the two classes. To accustom subjects to this format, a set of learning trials was first performed. Subjects were then given sets of 50 consecutive presentations, each set done in a single sitting. For most textures, two sets of trials were used, each set tested on a different day. Discriminability of the two random fields was measured by the fraction of correct responses in the combined set.

4.2 Similarity Parameter

The first set of experiments was designed to measure the discriminability threshold ΔH_θ for textures of different similarity parameter H. When $0 < H < 1$, this quantity is closely related to the fractal dimension D (see appendix B) of the texture, viz.,

$$D = 3 - H.$$

These experiments determined the discrimination thresholds ΔH_θ of several self-similar textures with $h \rightarrow 1$. In accord with common psychophysical practice [GrSw66], this was taken to be the difference in similarity parameters separating textures distinguishable 75% of the time.

4.2.1 Procedure

A set of self-similar line textures was generated for each of four reference classes. These textures had $H \in \{-1/2, 0, 1/2, 1\}$ and $h \rightarrow 1$. Each reference class was tested against eight classes of comparison textures with similar h but different

H. The similarity parameters of the comparison sets were greater than those of the reference classes, and were separated in steps of 0.025 from each other.

The discriminability between each reference and comparison class was determined in the fashion outlined in section 4.1. Fifty presentations were given to each subject for every pair of classes tested. Examples of two different comparison classes used against reference class $H = 0$ are shown in figure 4.3.

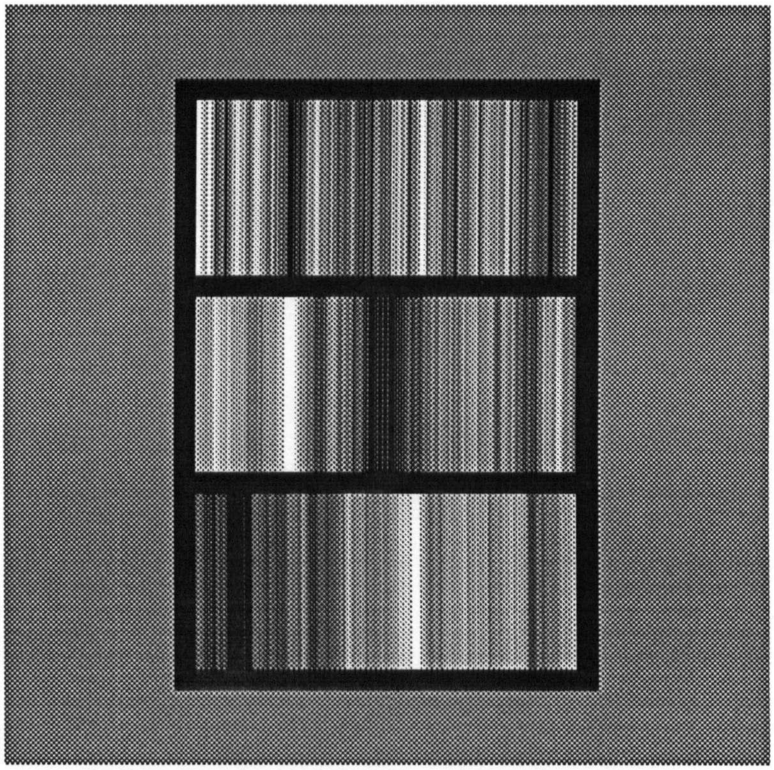
To test for symmetry of discriminability about the reference values, a second set of observations was collected from subject A. This set was similar to the first, except that the similarity parameters of the comparison classes were below the reference values.

4.2.2 Results and Discussion

The eight points obtained for each reference class were used as the basis for a psychometric function $\psi(H, \Delta H)$, describing the percentage correct identification as a function of the reference value H, and the difference ΔH. Probit analysis [Finn71][McKl85] was used to determine the threshold values (75% correct identification) for each of these curves. By convention, ΔH is taken to be positive for comparison textures with values of H higher than those of their reference classes.

Results are shown in table 4.1. Performance was consistent for both subjects. The chi-square values were calculated, and used to determine the quantity p, the probability of error in the fit of the curve. The values of p show that the $\psi(H, \Delta H)$ calculated from the data have tolerably good fits to lognormal form.

The threshold value ΔH_θ was found to generally decrease with increasing H. Results were similar for both positive and negative thresholds, except that



Upper texture: $H = 0.00$, $h \to 1$

Middle texture: $H = 0.30$, $h \to 1$

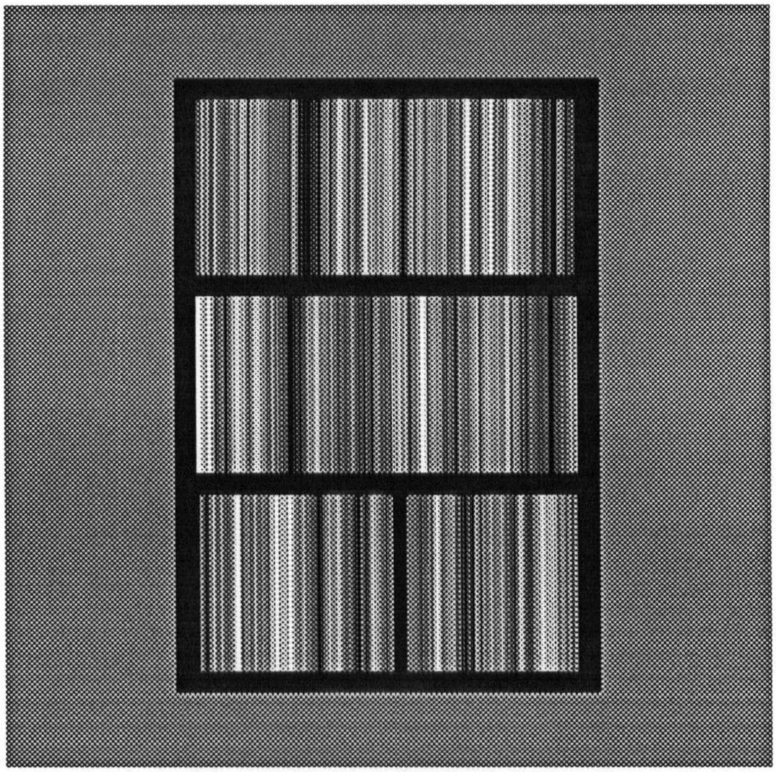
Lower texture: $H = 0.30$, $h \to 1$

$\Delta H = 0.30 > \Delta H_\theta = 0.17$

Discriminability = 88%

(a) textures above threshold difference ΔH_θ

Figure 4.3: line textures above and below discrimination threshold



Upper texture: $H = 0.00$, $h \to 1$

Middle texture: $H = 0.00$, $h \to 1$

Lower texture: $H = 0.10$, $h \to 1$

$\Delta H = 0.10 \ < \ \Delta H_\theta = 0.17$

Discriminability = 56%

(b) textures below threshold difference ΔH_θ

Figure 4.3 (continued)

the negative thresholds tended to have somewhat lower magnitudes.

These results have two major implications. First, there appears to be no appreciable perceptual distinction between fractals and self-similar noises. The positive and negative thresholds for $H = 0$ involve fractals and self-similar noises respectively; no large difference was found in their values. Furthermore, the threshold values for the self-similar noise ($H = -1/2$) showed no great deviation from the general pattern common to all other self-similar textures.

The results also show that line textures that have an effective spectral self-similarity behave in much the same way as do true self-similar textures. The two discrimination thresholds for $H = -1/2$ have nearly similar values, even though the negative threshold was measured using textures with $H < -1/2$. Similarly, the positive threshold for $H = 1$ has a magnitude virtually equal to that of its negative counterpart, in spite of being based on comparison textures with $H > 1$.

In summary, then, the results of this experiment show that self-similar line textures with $h \rightarrow 1$ give rise to a continuum of perceived textures. No large changes in discriminability occur when stimuli change from fractals to self-similar noises to effectively self-similar textures.

4.3 Scaling Ratio

The scaling ratio h is the minimum factor by which a self-similar texture can be compressed or expanded to match itself statistically (see appendix B). Consequently, such a texture will also match itself under rescaling by factors of h^j, $j \in Z$. A second set of experiments was carried out to measure the sensitivity

	discrimination thresholds — subject A			
H	positive ΔH_θ[a]	p[b]	negative ΔH_θ	p
−0.5	0.19 ±0.03	0.01	0.22 ±0.04	0.33
0.0	0.17 ±0.02	0.07	0.22 ±0.04	0.25
0.5	0.17 ±0.02	0.27	0.20 ±0.03	0.19
1.0	0.10 ±0.02	0.07	0.07 ±0.01	0.33

	discrimination thresholds — subject B			
H	positive ΔH_θ	p	negative ΔH_θ	p
-0.5	0.17 ±0.02	0.38	–[c]	–
0.0	0.15 ±0.02	0.38	–	–
0.5	0.17 ±0.02	0.22	–	–
1.0	0.10 ±0.02	0.60	–	–

[a] (tolerances are for 5% error)

[b] (probability of deviation from lognormal form)

[c] (dash indicates experiment not performed)

Table 4.1: discriminability of similarity parameter H

of the human visual system to this quantity when H was kept fixed.

As was shown in chapter 3, the form of $S(k)$ for a self-similar random field is not uniquely specified by H and h when $h > 1$. To examine the discriminability of h, arbitrary choices must be made for the shape of the pattern function $P(k)$. Several different shapes – hopefully representative – were investigated.

4.3.1 Procedure

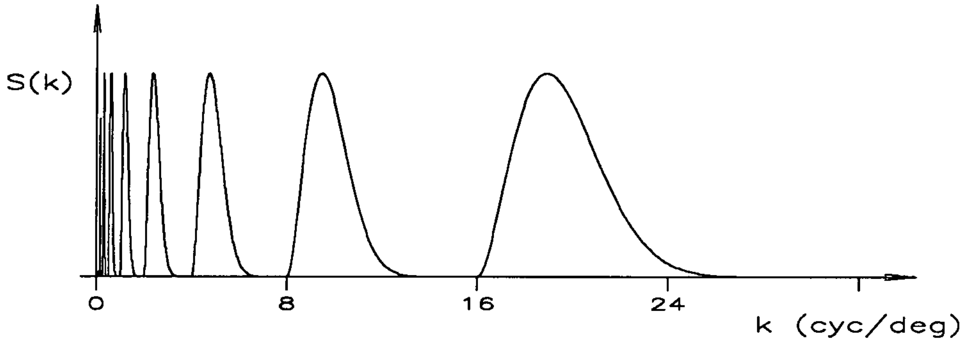
A set of reference textures $\{U_H\}$ with $h \to 1$ was generated for the values $H \in \{-1/2, 1/2\}$. These textures have power spectra $S_{U_H}(k) \propto k^{-1-2H}$, and are therefore spectrally self-similar over all scaling ratios. Each set served as a reference for comparison against other textures of similar H. Discriminability was measured using one hundred presentations for each pair of texture classes compared.

A set of comparison textures $A_{H,h}$ was generated for each of the similarity parameters $H \in \{-1/2, 1/2\}$, and scaling ratios $h \in \{1.73, 2.0, 3.0, 4.0\}$. The spectrum $S_{A_{H,h}}(k)$ of these textures was based on a template function of the form

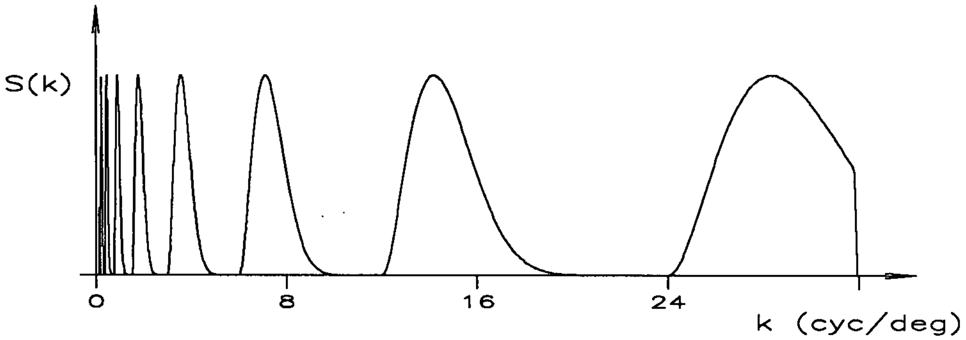
$$p_{A_{H,h}}(k) = C_{A_{H,h}}(k - hk_A^{base})^5[1 - \cos(\frac{2\pi(k - k_A^{base})}{(h-1)k_A^{base}})]; \quad k_A^{base} \leq k < hk_A^{base},$$

where the base frequency k_A^{base} was set to 1/128 cyc/deg. Zeroes of $S_{A_{H,h}}(k)$ consequently occurred at the frequencies $h^j k_A^{base}$, $j \in Z$. Such a power spectrum is a series of relatively narrow peaks that increase in width and spacing as $k \to \infty$ (figure 4.4(a)).

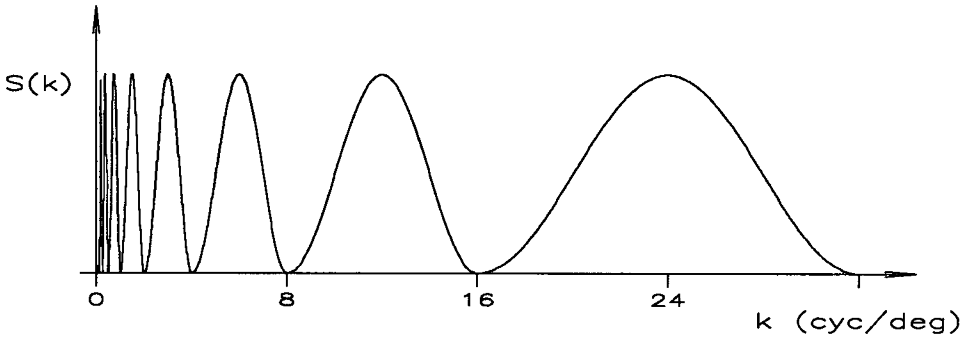
Both subjects were also tested on a second family of comparison textures $\{B_{H,h}\}$. The power spectrum $S_{B_{H,h}}(k)$ of these textures was similar to $S_{A_{H,h}}(k)$,



(a) power spectrum $S_{A_{H,h}}(k)$ for $H = -1/2$, $h = 2.0$



(b) power spectrum $S_{B_{H,h}}(k)$ for $H = -1/2$, $h = 2.0$



(c) power spectrum $S_{C_{H,h}}(k)$ for $H = -1/2$, $h = 2.0$

Figure 4.4: power spectra for $\{A_{H,h}\}$, $\{B_{H,h}\}$, and $\{C_{H,h}\}$

but with $k_B^{base} = \frac{h+1}{2h}k_A^{base}$, so that the peaks of $S_{B_{H,h}}(k)$ fell midway between the peaks of $S_{B_{H,h}}(k)$. Figure 4.4(b) shows the spectrum $S_{B_{H,h}}(k)$ for $H = 1/2$, $h = 4.0$.

The discriminability of a third family of textures $\{C_{H,h}\}$ was also tested using subject A. These textures were similar to the $\{A_{H,h}\}$, but with a template function of the form

$$p_{C_{H,h}}(k) = C_{C_{H,h}}[1 - \cos(\frac{2\pi(k - k_C^{base})}{(h-1)k_C^{base}})]; \quad k_C^{base} \leq k < hk_C^{base},$$

where $k_C^{base} = k_A^{base}$. The power spectrum $S_{C_{H,h}}(k)$ for $H = 1/2$, $h = 4.0$ is shown in figure 4.4(c). Since $k_C^{base} = k_A^{base}$, the zeroes of these two functions coincide.

4.3.2 Results and Discussion

Results are shown in Table 4.2. Again, performance was consistent for both subjects. In all cases tested, textures became more discriminable as the difference in h increased. The exact shape of the template function did not greatly affect the results. Textures with $H = -1/2$, however, were more discriminable for a given difference Δh than were those with $H = 1/2$.

These results show that the similarity parameter H — or equivalently, the fractal dimension D — is not the only second-order quantity relevant to discriminability. Although having little effect when $H = 1/2$, the value of h does influence discriminability when $H = -1/2$.

Interestingly, the scaling ratio does not have a large effect on discriminability when its value is two or less. White noise ($H = -1/2$, $h \to 1$) is virtually indistinguishable from textures with power spectra composed of pulses spaced

$H = -1/2$	percentage correct [a] — subject A		
h	$\{U_H\}$ vs $\{A_{H,h}\}$ [b]	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	64 ±4.8	55 ±5.0	56 ±5.0
2.00	70 ±4.6	61 ±4.9	55 ±5.0
3.00	80 ±4.0	69 ±4.6	64 ±4.8
4.00	93 ±2.5	91 ±2.9	72 ±4.5

$H = -1/2$	percentage correct — subject B		
h	$\{U_H\}$ vs $\{A_{H,h}\}$	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	61 ±4.9	50±5.0	– [c]
2.00	68 ±4.7	61 ±4.9	–
3.00	78 ±4.1	76 ±4.3	–
4.00	92 ±2.7	93 ±2.6	–

[a] (tolerances represent ±1 standard error)
[b] (reference $\{U_H\}$ has $h \to 1$)
[c] (dash indicates experiment not performed)

Table 4.2: discriminability of scaling ratio h

$H = 1/2$	percentage correct [a] — subject A		
h	$\{U_H\}$ vs $\{A_{H,h}\}$ [b]	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	52 ±5.0	49±5.0	47 ±5.0
2.00	61 ±4.9	62 ±4.9	59 ±4.9
3.00	69 ±4.6	62 ±4.9	67 ±4.7
4.00	77 ±4.2	64 ±4.8	69 ±4.6

$H = 1/2$	percentage correct — subject B		
h	$\{U_H\}$ vs $\{A_{H,h}\}$	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	53 ±5.0	58±4.9	–[c]
2.00	56 ±5.0	52 ±5.0	–
3.00	61 ±4.9	64 ±4.8	–
4.00	61 ±4.9	59 ±4.9	–

[a] (tolerances represent ±1 standard error)

[b] (reference $\{U_H\}$ has $h \to 1$)

[c] (dash indicates experiment not performed)

Table 4.2 (continued)

apart by a factor of two; when the pulses are spaced apart by a factor of four, however, discriminability is increased dramatically (figure 4.5).

4.4 Discriminability of Other Properties

To examine whether the second-order quantities H and h are sufficient to characterize the perception of self-similar random textures, a final set of experiments measured the discriminability of textures with identical H and h, but with different template functions $P(k)$ for their power spectra.

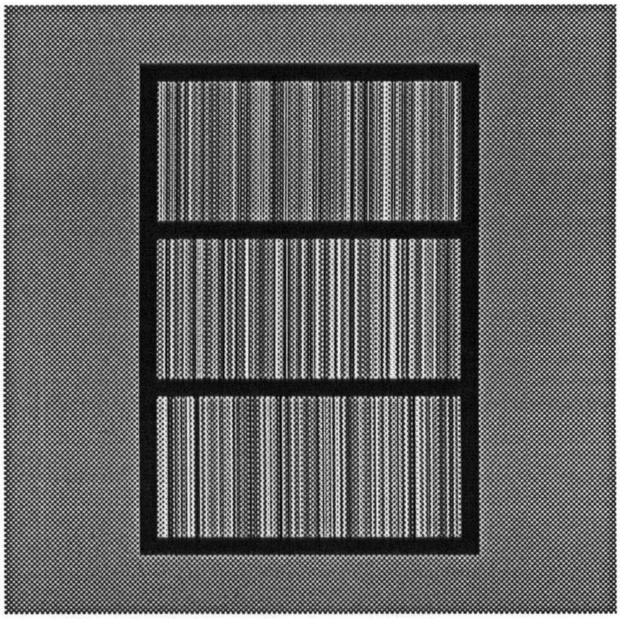
4.4.1 Procedure

The textures in the families $\{A_{H,h}\}$, $\{B_{H,h}\}$, and $\{C_{H,h}\}$ were tested against each other for all combinations involving similar H and h. Subject A made the full range of observations $H \in \{-1/2, 1/2\}$, $h \in \{1.73, 2.0, 3.0, 4.0\}$. Subject B was tested on the range $H \in \{-1/2, 1/2\}$, $h \in \{2.0, 4.0\}$. Again, each pair of classes was tested using one hundred presentations.

To determine the effect of a different compression of the power spectrum, a fourth family of textures $\{D_{H,4}\}$ was generated for $H \in \{-1/2, 1/2\}$, $h = 4$. These had the same form of template function as the $\{A_{H,h}\}$, but with k_D^{base} set to $\frac{3}{8}k_A^{base}$. These sets were tested for discriminability against the classes $A_{H,4}$ and $C_{H,4}$.

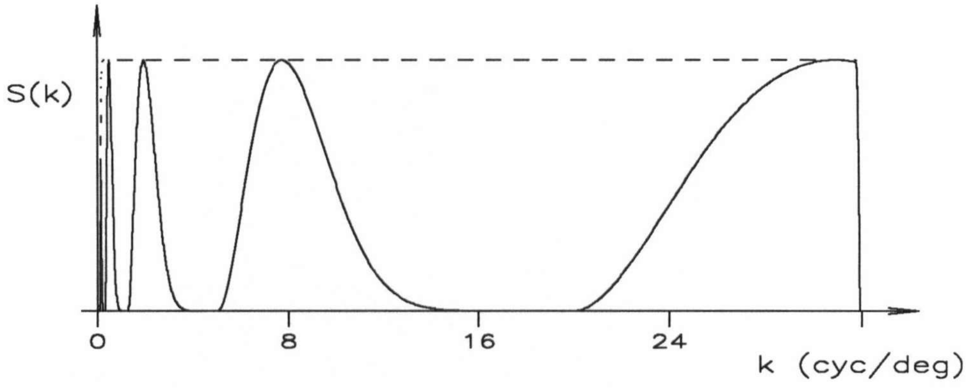
4.4.2 Results and Discussion

Results are given in Table 4.3. Performance of both subjects again remained consistent. Textures with similar H and h were found to be discriminable. This



Upper texture: $B_{H,h}$, $H = -0.50$, $h = 4$
Middle texture: U_H, $H = -0.50$, $h \to 1$ (white noise)
Lower texture: U_H, $H = -0.50$, $h \to 1$ (white noise)

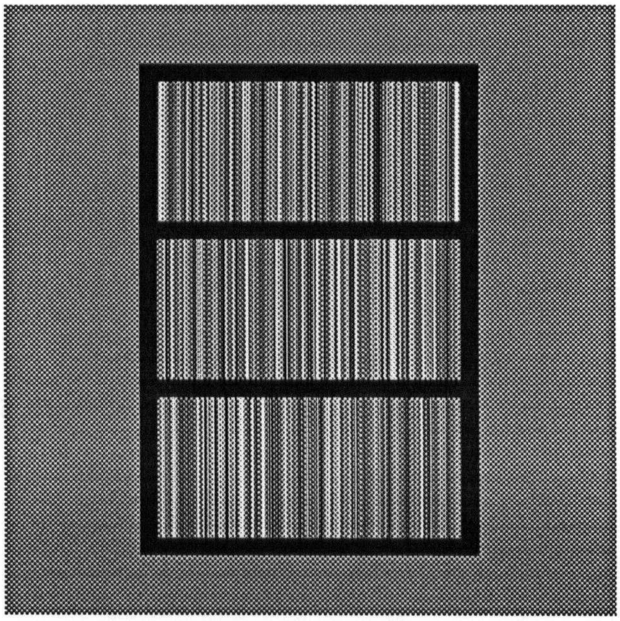
Discriminability = 91%



- - - - white noise ——— $\{B_{H,h}\}$

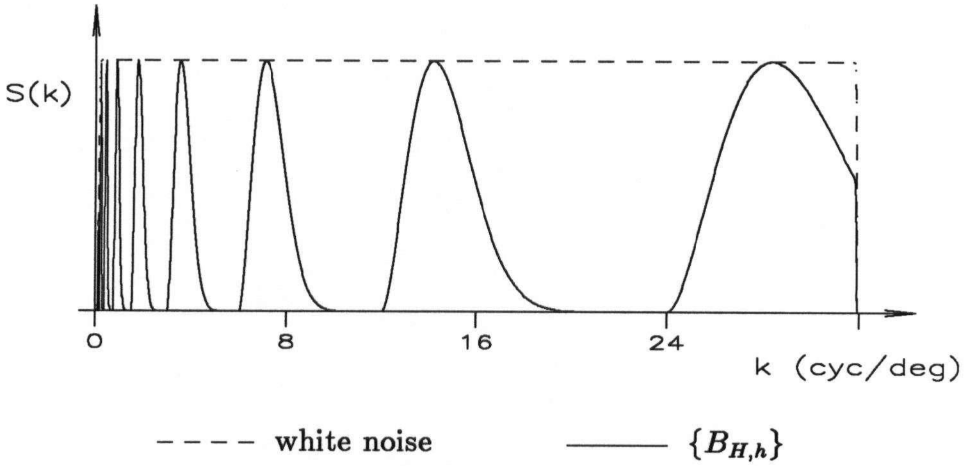
(a) comparison of white noise and $\{B_{H,h}\}$ for $H = -0.5$, $h = 4$

Figure 4.5: white noise vs $\{B_{H,h}\}$



Upper texture: U_H, $H = -0.50$, $h \to 1$ (white noise)
Middle texture: U_H, $H = -0.50$, $h \to 1$ (white noise)
Lower texture: $B_{H,h}$, $H = -0.50$, $h = 2$

Discriminability = 61%



(b) comparison of white noise and $\{B_{H,h}\}$ for $H = -0.5$, $h = 2$

Figure 4.5 (continued)

discriminability generally increased as H decreased or h increased. As in the previous experiment, discriminability was always low when h was two or less.

These results show that H and h are not the only second-order quantities relevant to the perception of self-similar random textures. Other quantities, dependent on $P(k)$, must also be involved. This is consistent with the observation that when $h \to \infty$ (i.e., when the texture is no longer self-similar), the function $P(k)$ completely describes the power spectrum. From the low discriminability between textures with $h \leq 2$, however, it follows that a difference in the template functions does not generally suffice for high discriminability of the corresponding textures.

$H = -1/2$	percentage correct [a] — subject A		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
1.73	51 ±5.0	53 ±5.0	48 ±5.0
2.00	50 ±5.0	65 ±4.8	75 ±4.3
3.00	85 ±3.6	87 ±3.4	76 ±4.3
4.00	97 ±1.7	86 ±3.5	99 ±1.0

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	– [b]
4.00	98 ±1.4	57 ±5.0	–

$H = -1/2$	percentage correct — subject B		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
2.00	70 ±4.6	56 ±5.0	67 ±4.7
4.00	100 ±0.0	88 ±3.2	97 ± 1.7

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	–
4.00	100 ±0.0	60 ±4.9	–

[a] (tolerances represent ±1 standard error)

[b] (dash indicates experiment not performed)

Table 4.3: discriminability of template function $P(k)$

$H = 1/2$	percentage correct [a] — subject A		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
1.73	47 ±5.0	46 ±5.0	50 ±5.0
2.00	54 ±5.0	65 ±4.8	54 ±5.0
3.00	69 ±4.6	58 ±4.9	76 ±4.3
4.00	78 ±4.1	76 ±4.3	73 ±4.4

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	– [b]
4.00	77 ±4.2	53 ±5.0	–

$H = 1/2$	percentage correct — subject B		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
2.00	47 ± 5.0	55 ±4.9	51 ±5.0
4.00	58 ±4.9	81 ±3.9	76 ±4.3

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	–
4.00	73 ±4.4	59 ±4.9	–

[a] (tolerances represent ±1 standard error)

[b] (dash indicates experiment not performed)

Table 4.3 (continued)

Chapter 5

Discussion

The central goal of this work is to determine aspects of self-similar random textures perceived by the human visual system. Results from the preceding chapter show that both the similarity parameter and the scaling ratio have an influence on perceived texture, but that they are insufficient to characterize it completely. Some implications of these results for a general model of texture perception are now examined.

Since neither the texton theory nor the symbolic-structure approach are sufficiently developed to allow quantitative predictions to be made about the discriminability of random line textures, discussion is focussed on the relevance of the results for spatial-frequency models. The quantitative nature of these models allows the results to be checked for compatibility with the hypothesis that texture perception is mediated by measurements made on each of several parallel spatial-frequency channels.

channel	peak frequency (cyc/deg)	β_2	β_3	σ_1(deg)	σ_2(deg)	σ_3(deg)
A	0.8	0.267	0.000	0.198	0.593	0.000
B	1.7	0.333	0.000	0.098	0.294	0.000
C	2.8	0.894	0.333	0.084	0.189	0.253
D	4.0	0.894	0.333	0.059	0.132	0.177
E	8.0	1.266	0.500	0.038	0.060	0.076
F	16.0	1.266	0.500	0.019	0.030	0.038

Table 5.1: values of constants for spatial-frequency channels

5.1 Spatial-frequency Channels

For concreteness, the channels proposed by Wilson and Gelb [WiGe84] are used in the analysis of the results. In this formulation, six parallel linear channels are postulated, with each channel i based on an isotropic filter $H_i(k)$ of the form

$$H_i(k) = \alpha\pi^{1/2}[\sigma_1 \exp\{-(\pi\sigma_1 k)^2\} - \beta_2\sigma_2 \exp\{-(\pi\sigma_2 k)^2\} + \beta_3\sigma_3 \exp\{-(\pi\sigma_3 k)^2\}],$$

where k denotes the magnitude of the spatial frequency. The values of the β_j and σ_j, taken from [WiGe84], are given in Table 5.1. Variation of these parameters with eccentricity from the center of the fovea has been ignored. The value of α, which describes the absolute sensitivity of each channel, is not of importance here — only the relative responses of a channel to different stimuli are of concern for the present analysis. The value of α is therefore arbitrarily set to unity.

Any multichannel model of texture perception must specify the measure M_i

used for each channel i. Two sets of possible measures are considered here; they are not intended to be exhaustive. The first is the *relative contrast*

$$V_i = \left(\frac{\int H_i^2(k) S(k) dk}{\int S(k) dk} \right)^{1/2},$$

where $S(k)$ denotes the power spectrum of the unfiltered image. This quantity describes the standard deviation of each filtered image, given that the unfiltered image has unit variance. Since contrast is proportional to the amplitude of the constituent waveforms of the image, V_i is proportional to the contrast of the image filtered by $H_i(k)$.

The second measure is the zero-crossing density Z_i in each channel. For a Gaussian stochastic process [Papo84,11-4],

$$Z_i = 2 \left(\frac{\int k^2 H_i^2(k) S(k) dk}{\int H_i^2(k) S(k) dk} \right)^{1/2}.$$

Zero-crossing densities resulting from a set of $\nabla^2 G$ filters at differing spatial scales were briefly considered by Riley [Rile81], but a thorough investigation of their suitability for representing texture has never been carried out.

Various metrics for the perceptual distance between two texture classes can be based on $\vec{V}$ and $\vec{Z}$. In what follows, the only constraint placed on the postulated metric is that it be *weakly monotonic*: if the ensemble values of measures M_i made on texture classes T_0, T_1, and T_2 are such that

$$M_i(T_2) \geq M_i(T_1) \geq M_i(T_0)$$

for each channel i, then the discriminability between T_1 and T_0 must be less than or equal to that between T_2 and T_0. If the measurements made do not obey such an ordering, then no prediction can be made.

The use of such a weak constraint entails that predictions can only be made about the relative discriminability between some of the texture classes. The advantage of this approach, however, is that no detailed mechanisms of texture perception need be specified. This provides a way to test the compatibility of the results with the general class of models based on spatial-frequency channels.

5.2 Analysis

The values of $\vec{V}$ and $\vec{Z}$ for each of the reference classes $H \in \{-1/2, 0, 1/2, 1\}$, $h \rightarrow 1$ are given in Appendix D. Also given are the values $\vec{V}$ and $\vec{Z}$ for the comparison classes at the upper and lower discrimination thresholds. These values have been calculated only for subject A, since the performance of subject B was similar over virtually all textures tested.

The results concerning the discriminability of H show that texture classes on the discrimination threshold generally have $\Delta V_i/V_i$ in the range $10-50\%$ for all channels. The values of ΔV_i are roughly symmetric about the reference values for the upper and lower thresholds. This symmetry also holds for the ΔZ_i. The relative quantities $\Delta Z_i/Z_i$ almost always fall in the range of $2-10\%$, with most values falling in the range of $2-6\%$. This last figure is interesting, since the relative difference in frequency for sine-wave gratings at the discrimination threshold is $2-5\%$ [WiGe84]. The discriminability of the similarity parameter H would therefore appear to be compatible with a multichannel mechanism for texture perception.

To determine whether the discriminability of the scaling ratio h is also compatible with such a model, the values of $\vec{V}$ and $\vec{Z}$ for all comparison textures used in section 4.2 were compared against those of the threshold texture classes used in the previous analysis. Assumption of a monotonic discrimination mechanism allows two types of prediction to be made:

1. If all measures M_i of a comparison class T_1 fall between the corresponding values for one of the threshold textures T_θ and the values for the reference class T_0, then the discriminability between T_1 and T_0 must be less than or equal to the threshold level of 75%.
2. If all measures M_i for one of the threshold classes T_θ fall between the corresponding values for the comparison class T_1 and the values for the reference class T_0, then the discriminability between T_1 and T_0 must be greater than or equal to the threshold level of 75%.

Table 5.2 shows the predicted discriminability of the texture classes, using $\vec{V}$ and $\vec{Z}$ as two independent sets of measures. Comparison with the experimental results (Table 4.2) shows agreement with almost all predictions made. The sole exception occurs for $H = 1/2$, $h = 4$. Prediction based on $\vec{V}$ states that the discriminability between $\{U_H\}$ and $\{A_{H,h}\}$ should be below 75%. Comparison with table 4.2 shows that the observed discriminability is 77%. The standard error, however, could push this value down below the discrimination threshold, so that this exception is not statistically significant.

The predictions — in particular, those based on $\vec{V}$ — correctly describe the diminished discriminability between textures with $H = 1/2$. Especially interesting is the result that discriminability should generally be poor between textures with $h \leq 2.0$. These predictions are not sufficiently detailed to determine whether $\vec{V}$, $\vec{Z}$, or some combination of the two should be favoured as the set of measurements involved in texture perception. However, they do show that a multichannel model remains consistent with the discriminability of the scaling ratio.

This approach can also be used for texture classes that have different template functions $P(k)$. The differences ΔV_i and ΔZ_i can be used as before. If the values V_i and Z_i for the textures differ only slightly from those of the $(h \rightarrow 1)$ references, the same discrimination thresholds as in the previous analyses can be used as the basis of a consistency check.

Table 5.3 shows the predicted discriminabilities of the texture classes. Comparison with experimental results (Table 4.3) shows that the low discriminability of the $H = 1/2$ textures is correctly predicted, as is the generally low discriminability between textures with $h = 1.73$. Again, the V_i appear to be more sensitive measures of discriminability than the Z_i. However, predictions are still too weak to allow one set to be preferred over the other.

The low discriminability between many texture pairs cannot be predicted using this model. For example, no prediction can be made for the discriminability between $C_{H,4}$ and $D_{H,4}$ when $H = -1/2$. However, a comparison of $\vec{V}$ and $\vec{Z}$ for the two classes of texture (Table 5.4) shows that ΔV_i and ΔZ_i are generally quite small. This situation is typical of many cases where prediction cannot be made.

Taken together, these results imply that the ability of the human visual system to perceive self-similar random textures is compatible with models of texture perception based on spatial-frequency channels.

$H = -1/2$	percentage correct — measure = $\vec{V}$		
h	$\{U_H\}$ vs $\{A_{H,h}\}$ [a]	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	≤ 75 [b]	≤ 75	≤ 75
2.00	— [c]	≤ 75	≤ 75
3.00	—	—	—
4.00	—	—	—

$H = -1/2$	percentage correct — measure = $\vec{Z}$		
h	$\{U_H\}$ vs $\{A_{H,h}\}$	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	≤ 75	≤ 75	≤ 75
2.00	≤ 75	—	≤ 75
3.00	—	—	—
4.00	—	—	—

[a] (reference $\{U_H\}$ has $h \to 1$)

[b] ('$\leq$ 75' indicates discriminability is below discrimination threshold)

[c] (dash indicates no prediction can be made)

Table 5.2: predicted discriminability of scaling ratio h

$H = 1/2$	percentage correct — measure = $\vec{V}$		
h	$\{U_H\}$ vs $\{A_{H,h}\}$ [a]	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	≤ 75 [b]	≤ 75	≤ 75
2.00	≤ 75	≤ 75	≤ 75
3.00	≤ 75	≤ 75	≤ 75
4.00	≤ 75	≤ 75	≤ 75

$H = 1/2$	percentage correct — measure = $\vec{Z}$		
h	$\{U_H\}$ vs $\{A_{H,h}\}$	$\{U_H\}$ vs $\{B_{H,h}\}$	$\{U_H\}$ vs $\{C_{H,h}\}$
1.73	≤ 75	≤ 75	≤ 75
2.00	— [c]	≤ 75	≤ 75
3.00	—	—	—
4.00	—	—	—

[a] (reference $\{U_H\}$ has $h \to 1$)

[b] ('≤ 75' indicates discriminability is below discrimination threshold)

[c] (dash indicates no prediction can be made)

Table 5.2 (continued)

$H = -1/2$	percentage correct — measure = $\vec{V}$		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
1.73	— [b]	—	≤ 75 [a]
2.00	—	—	—
3.00	—	—	—
4.00	—	—	—

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	—
4.00	—	—	—

$H = -1/2$	percentage correct — measure = $\vec{Z}$		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
1.73	—	≤ 75	≤ 75
2.00	—	—	—
3.00	—	—	—
4.00	—	—	—

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	—
4.00	—	—	—

[a] ('≤ 75' indicates discriminability is below discrimination threshold)
[b] (dash indicates no prediction can be made)

Table 5.3: predicted discriminability of template function $P(k)$

$H = 1/2$	percentage correct — measure = $\vec{V}$		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
1.73	≤ 75 [a]	≤ 75	≤ 75
2.00	≤ 75	≤ 75	≤ 75
3.00	≤ 75	≤ 75	≤ 75
4.00	— [b]	≤ 75	≤ 75

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	—
4.00	≤ 75	≤ 75	—

$H = 1/2$	percentage correct — measure = $\vec{Z}$		
h	$\{A_{H,h}\}$ vs $\{B_{H,h}\}$	$\{B_{H,h}\}$ vs $\{C_{H,h}\}$	$\{C_{H,h}\}$ vs $\{A_{H,h}\}$
1.73	≤ 75	≤ 75	≤ 75
2.00	—	≤ 75	—
3.00	—	—	—
4.00	—	—	—

h	$\{A_{H,h}\}$ vs $\{D_{H,h}\}$	$\{D_{H,h}\}$ vs $\{C_{H,h}\}$	—
4.00	—	≤ 75	—

[a] ('$\leq$ 75' indicates discriminability is below discrimination threshold)

[b] (dash indicates no prediction can be made)

Table 5.3 (continued)

$H = -1/2$	measure = $\vec{V}$		measure = $\vec{Z}$	
channel	$\{C_{H,4}\}$	$\{D_{H,4}\}$	$\{C_{H,4}\}$	$\{D_{H,4}\}$
A	5.40×10^{-2}	5.41×10^{-2}	2.22	2.25
B	3.13×10^{-2}	3.02×10^{-2}	4.43	4.35
C	1.93×10^{-2}	1.82×10^{-2}	7.76	8.24
D	2.08×10^{-2}	2.14×10^{-2}	9.47	9.49
E	5.33×10^{-3}	4.79×10^{-3}	15.87	14.03
F	6.21×10^{-3}	6.45×10^{-3}	37.05	37.27

Table 5.4: comparison of ensemble values for $C_{H,4}$ and $D_{H,4}$

Chapter 6

Conclusions

This work investigates the ability of the human visual system to discriminate between self-similar random textures. The properties of such textures are determined in both the spatial and the frequency domain, and their relation to the class of self-similar stochastic fractals is established. It is shown using psychophysical experiments that the similarity parameter H and the scaling ratio h influence the discrimination of self-similar line textures, but that they are insufficient to completely characterize perceived texture. Analysis shows that the results of the experiments are compatible with a multiscale model of texture perception.

These results are relevant to three areas of study. First, they suggest new methods of texture analysis. Previous techniques for analyzing textures by their fractal properties (e.g., [Pent83][PeNa84]) have considered only the similarity parameter H, often using one-dimensional measures for its determination. This work shows that H is insufficient to characterize a random texture: other properties, such as the scaling ratio h, must also be taken into account. The treatment of self-similar random textures given in chapter 3 provides a basis for determining these quantities, and to do so using two-dimensional spatial and spectral measures.

More generally, the multichannel model of texture perception suggested in chapter 5 can be implemented on a machine. Using only the relative measurements in each channel, algorithms can be developed that are translationally and rotationally invariant, and are robust under changes of grey level and scale. Indeed, initial work on one such algorithm shows promising results for the segmentation of natural images [Litt86].

The second area of relevance is computer graphics. Self-similar stochastic models are widely used to represent various objects and surfaces (e.g., [FoFu82][HaBa84]). The treatment of self-similarity given in chapter 3 forms a rudimentary basis for relating these self-similar constructs to true fractals. Further, the texture-discrimination experiments described in chapter 4, although based only on self-similar line textures, determine the sensitivity of the human visual system to several properties of interest. This enables an estimate to be made of whether an algorithm can generate objects that appear truly self-similar.

Finally, the results of this work are of relevance to the computational study of the human early vision system. The results of the psychophysical experiments are found to be compatible with the hypothesis that texture perception is mediated through measurements made in parallel on the spatial-frequency channels putatively involved in form vision. This suggests that texture perception and form perception could share common mechanisms.

Open Questions

Although sufficient for the purposes of this work, the treatment of self-similar random fields given in chapter 3 is incomplete in several respects. First, the relation between self-similar random fields and self-similar stochastic fractals established in in theorem 2 (section 3.1.1) applies only in one direction: if the n-dimensional power spectrum $S(\vec{k})$ is self-similar, with $0 < H < 1$, then the corresponding random field is a stationary stochastic self-similar fractal. The converse relation, however, is not established, and its existence is an open issue.

Another issue also involves theorem 2. For the proof of the theorem to work, the power spectrum must be limited from above by a function $A|\vec{k}|^{-n-2H}$, where A is some positive number. Although presenting no constraint for any practical application, this restriction places a theoretical limitation on the type of spectral self-similarity that can correspond to spatial self-similarity. In the interest of completeness, it would be useful to establish whether the relation given by theorem 2 holds for all self-similar power spectra.

Limitations on the form of $S(\vec{k})$ also apply to theorem 3 (section 3.1.2), which relates the self-similar form of the covariance function to the self-similar form of the power spectrum when $-n/2 < H < 0$. Again, the boundedness required of $S(\vec{k})$ is of no practical concern. The condition that $S(\vec{k})$ be composed of several monotonic functions is likewise of little practical consequence. From a theoretical viewpoint, however, it would be interesting to determine whether theorem 3 would still apply if the restrictions on the form of $S(\vec{k})$ were removed. If so, the converse of theorem 3 would follow as a natural result.

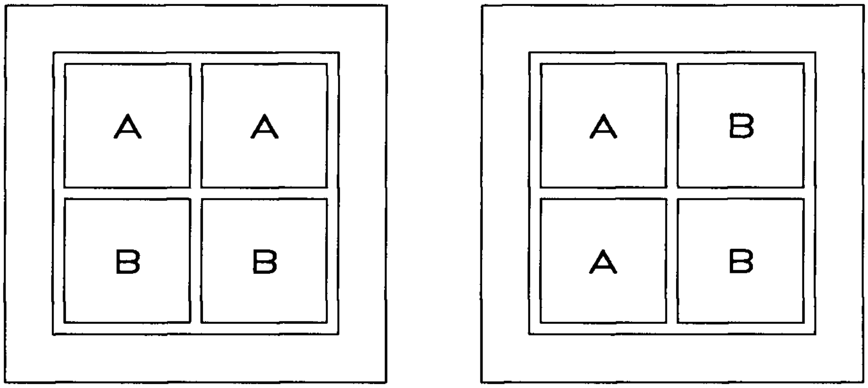
Directions for Future Work

The approach used in this work can be extended in several ways. First, a much wider range of discrimination experiments could be carried out, using the techniques described in chapters 3–5. This would not only allow a more precise estimate of the discrimination thresholds, but would also provide additional evidence either for or against various multichannel mechanisms. Experiments involving textures with different first-order statistics (i.e., mean and variance) would also contribute toward this end. Although time-consuming, these experiments would be straightforward to carry out.

The techniques described in chapters 3–5 could also be used to examine the multichannel hypothesis itself, without specific reference to self-similar textures. The requirement of self-similarity could be dropped, and the power spectra of texture classes designed expressly to distinguish between competing multichannel models. To make this approach feasible, a method is required for the design of the appropriate spectra.

A more elegant route to the same goal, however, would be to develop techniques for determining the form of the putative spatial-frequency channels directly from the observed discriminabilities. It is difficult to estimate the amount of effort required to develop such techniques. Once constructed, however, they would be of great value in determining the exact mechanisms of texture perception.

An extension of these techniques to self-similar non-Gaussian random textures would also be of interest. Such textures could be readily created, e.g., by using dithering techniques to binarize Gaussian textures. Although the theorems developed in chapter 3 would still hold, the analysis of the results would almost invariably be difficult, for the simplifications available for the Gaussian case are not generally applicable. However, special cases might be found for which the analysis would be tractable. These could provide useful checks on



(a) horizontal alignment (b) vertical alignment

Figure 6.1: cross display format

the results achieved for the Gaussian case.

Perhaps the most obvious extension of the approach developed in this work is to apply it to fully two-dimensional textures. The theorems on self-similarity developed in chapter 3 apply to the general n-dimensional case, so that generation of various two-dimensional self-similar textures would be straightforward. Discrimination experiments analogous to those of chapter 4 could be carried out via a cross display format: two pairs of textures would be aligned at random either in the horizontal or in the vertical direction (figure 6.1).

This technique could be used to determine various anisotropies of the visual system. Analysis of the results, however, would be difficult — there is at present no analytic formulation of the distribution of the angles, lengths, curvatures, etc., of the zero-crossings in a two-dimensional Gaussian random field.

A more reasonable approach to analyzing the two-dimensional nature of texture perception might be to investigate the discriminability of line textures at various orientations. Such a study would be relatively easy to carry out, since the analytical techniques have already been developed. It would be interesting to compare such results against those expected from fully two-dimensional spatial-frequency filters.

One last suggestion for future work is to investigate the discriminability of textures of various colours. Virtually all work in texture perception has involved monochromatic textures that were broadband, i.e., black and white. It would be straightforward to do analogous experiments and analyses on narrowband textures, composed of just a few spectral colours. Of particular interest would be the discriminability of textures whose chromatic components have different dimensions. The results could provide new insights into the relation between texture perception and colour perception.

Bibliography

Ade83 : Ade,F., 'Characterization of textures by "eigenfilters"', *Signal Processing*, **5**, pp 451–457 (1983)

AhDa77: Ahuja,N.,Davis,L.,Haralick,R., and Panda,D., *Image Segmentation Based on Local Gray Level Patterns*, TR-551, University of Maryland, June 1977

BaBr82: Ballard, D., and Brown, C., *Computer Vision*, Prentice-Hall, 1982

Bajc73: Bajcsy,R., *Computer description of textured surfaces*, IJCAI-73, pp 572–579 (1973)

Beck82: Beck,J., 'Textural segmentation and second-order statistics', TR-1181, Computer Vision Laboratory, Computer Science Department, University of Maryland, June 1982

Breu80: Breu, H., *Slant from Texture: Computational Methods for Recovering Surface Slant from Images of Textured Scenes*, MSc Thesis, Dept. of Computer Science, University of British Columbia, Apr 1980

BrGr85: Bruce, V., and Green, P., *Visual Perception: Physiology, Psychology and Ecology*, Lawrence Erlbaum Associates, 1985

Burt84: Burt,P., 'The pyramid as a structure for efficient computation'. In *Multiresolution Image Processing and Analysis'*, A. Rosenfeld, ed. Springer-Verlag, 1984

Cael84: Caelli, T., 'On the specification of coding principles of visual image processing'. In *Figural Synthesis*, P. Dodwell and T. Caelli, eds. Lawrence Erlbaum Associates, 1984

CaHü84: Caelli, T., and Hübner, M., 'On the number of intensity levels discriminated in texture', *Perception*, **13**, pp 21-31 (1984)

CaJu78: Caelli,T., Julesz,B., and Gilbert,E., 'On perceptual analyzers underlying visual texture discrimination: Part II', *Biol Cybernetics*, **29**, pp 201–214 (1978)

CaRo68: Campbell, F., and Robson, J, 'Application of Fourier Analysis to the Visibility of Gratings', *J Physiology*, **197**, pp 551–566 (1968)

ChKa81: Chellappa,R., and Kashyap,R., 'On the correlation structure of random field models of images and textures', *IEEE PRIP*, pp 574–576, 1981

CoHa80a: Conners,R., and Harlow,C., 'A theoretical comparison of texture algorithms', *IEEE Trans*, **PAMI-2**, pp 204–222 (1980)

CoHa80b: Conners,R., and Harlow,C., 'Toward a structural textural analyzer based on statistical methods', *Comp Graphics Image Processing*, **12**, pp 224–256 (1980)

DaJo79: Davis,L., Johns,S., and Aggarwal,J., 'Texture analysis using generalized co-occurrence matrices', *IEEE Trans* **PAMI-1**, pp 251–259 (1979)

Dumo85: Dumoulin, F., *Using texture energy measures for the segmentation of forest scenes*, MSc thesis, Department of Forestry / Remote Sensing, University of British Columbia, Dec 1985

Eklu79: Eklundh,J., 'On the use of fourier phase features for texture discrimination', *Comp Graphics Image Processing*, **9**, pp 199–201 (1979)

Finn71: Finney,J., *Probit Analysis*, 3rd ed., Cambridge University Press, 1971

FoFu82: Fournier,A., Fussell,D., and Carpenter,L., 'Computer rendering of stochastic models', *CACM*, **25**, 6, pp 371–384 (1982)

Fu82 : Fu, K., *Syntactic Pattern Recognition and Application*, Prentice-Hall, 1982

FuLu78: Fu,K., Lu,S., 'Computer generation of texture using a syntactic approach', *Computer Graphics*, **12**, pp 147–152 (1978)

Gaga81: Gagalowicz, A., 'A new method of texture field synthesis: Some applications to the study of human vision', *IEEE Trans* **PAMI-3**, pp 520–533 (1981)

GoDe85: van Gool,L., Dewaele,P., and Oosterlinck,A., 'Survey: texture analysis anno 1983', *Comp Vision Graphics Image Processing*, **29**, pp 336–357 (1985)

Grim81: Grimson,W.E., *From Images to Surfaces – a computational study of the human early visual system*, MIT Press, 1981

GrRy65: Gradshteyn,I.S., and Ryzhik,I.M., *Table of Integrals, Series, and Products* , fourth ed., Academic Press, New York, 1965

GrSw66: Green, D., and Swets, J., *Signal Detection Theory and Psychophysics*, John Wiley and Sons, 1966

HaBa84: Haruyama,S., and Barsky, B., 'Using stochastic modeling for texture generation', *IEEE CG&A*, Mar 1984, pp 7–19

HaGe78: Harvey,L., and Gervais, M., 'Visual texture perception and fourier analysis', *Perception and Psychophysics*, **24**, pp 534–542 (1978)

HaGe81: Harvey,L., and Gervais, M., 'Internal representation of visual texture as the basis for the judgement of similarity', *J Exp Psych: Human Perception and Performance*, **7**, pp 741–753 (1981)

Hara75: Haralick,R., 'A resolution preserving textural transform for images', *IEEE Proc Conf on Comp Graphics, Patt Recog, and Image Processing*, pp 51–61, 1975

Hara79: Haralick,R., 'Statistical and structural approaches to texture', *Proc IEEE* **67**, pp 786–804 (1979)

HaSh73: Haralick,R., Shanmugam,K., and Dinstein,I., 'Textural features for image classification', *IEEE Trans* **SMC-3**, pp 610–621 (1973)

Hawk70: Hawkins,J., 'Textural properties for pattern recognition'. In *Picture Processing and Psychopictorics*, B. Lipkin and A. Rosenfeld, eds. Academic Press, 1970

Hild84: Hildreth, E., *The Measurement of Visual Motion*, MIT Press, 1984

Hutc81: Hutchinson, J., 'Fractals and self-similarity', *Indiana Univ Math Journal*, **30**, pp 713–747 (1981)

Jaya79: Jayaramamurthy, S., 'Multilevel array grammars for generating texture scenes', *Proc PRIP*, pp 391-398 (1979)

JeWa68: Jenkins,G., and Watts,D., *Spectral Analysis and its Applications*, Holden-Day, San Francisco, 1968

JuBe83: Julesz,B., and Bergen,J., 'Textons, the fundamental elements in preattentive vision and perception of texture', *Bell Sys Tech J* , **62**, pp 1619–1645 (1983)

JuGi73: Julesz,B., Gilbert,E., Shepp,L., and Frisch,H., 'Inability of humans to discriminate between visual textures that agree in second- order statistics – revisited', *Perception*, **2**, pp 391–405 (1973)

Jule62: Julesz,B., 'Visual pattern discrimination', *IRE Trans*, **IT-8**, pp 84–92 (1962)

Jule75: Julesz,B., 'Experiments in the visual perception of texture', *Sci Am*, 232(4), pp 34–43 (1975)

Jule84: Julesz,B., 'Toward an axiomatic theory of preattentive vision'. In *Dynamic Aspects of Neocortical Function*, G.Edelman, W.Gall, and W.Cowan, eds. Neurosciences Research Foundation, 1984

Kend79: Kender, J., 'Shape from texture: an aggregation transform that maps a class of textures into surface orientation', *IJCAI-79*, pp 475–480 (1979)

Kirv76: Kirvida,L., 'Texture measurements for the automatic classification of imagery', *IEEE Trans*, **EMC-18**, pp 38–42 (1976)

Knut81: Knuth, D., *The Art of Computer Programming*, 2nd ed., Addison Wesley, 1981

Laws80: Laws, K., *Textured image segmentation.* Report 940, Image Processing Institute, University of Southern California, Jan 1980

LeSt70: Lendaris,G., and Stanley, G., 'Diffraction pattern sampling for automatic pattern recognition', *Proc IEEE*, **58**, pp 198–216 (1970)

Litt86: Little,J., AI Lab, MIT. Personal communication.

MaBr77: Maleson,J., Brown,C., and Feldman,J., 'Understanding natural texture', *Proc ARPA Image Understanding Workshop*, Palo Alto, CA, Oct 1977, pp 19–27

MaHi80: Marr, D., and Hildreth, E., 'Theory of edge detection', *Proc Roy Soc*, **B207**, pp 187–217 (1980)

MaMi83: Matsuyama,T., Miura,S.-I., and Makoto,N., 'Structural analysis of natural textures by fourier transform', *Comp Vision Graphics Image Processing*, **24**, pp 347–362 (1983)

Mand67: Mandelbrot,B., 'How long is the coast of Britain? Statistical self-similarity and fractional dimension', *Science*, **155**, pp 636–638 (1967)

Mand71: Mandelbrot,B., 'A fast fractional Gaussian noise generator', *Water Resources Res*, **7**, pp 543–553 (1971)

Mand75: Mandelbrot,B., 'Stochastic models for the Earth's relief, and the shape and the fractal dimension of the coastlines, and the number-area rule for islands', *Proc Nat Acad Sci USA*, **72**, pp 3825–3828 (1975)

Mand82: Mandelbrot,B., *The Fractal Geometry of Nature*, W.H. Freeman & Co., San Francisco, 1982

Mand84: Mandelbrot,B., 'Fractals in physics : Squig clusters, diffusions, fractal measures, and the unicity of fractal dimensionality', J Stat Phys, **34**, pp 895–930 (1984)

MaNe68: Mandelbrot,B., and van Ness,J., 'Fractional Brownian motions, fractional noises and applications', *SIAM Review*, **10**, pp 422–437 (1968)

Marr76: Marr,D., 'Early processing of visual information', *Phil Trans R Soc London*, **B275**, pp 483–519 (1976)

Marr77: Marr,D., 'Artificial intelligence — a personal view' *Artificial Intelligence*, **9**, pp 37–48 (1977)

Marr82: Marr,D., *Vision – a computational investigation in the human representation and processing of visual information*, W.H. Freeman & Co., San Francisco, 1982

MaSa83: Matsuyama,T., Saburi,K, and Makoto,N., 'A structural analyzer for regularly arranged textures', *Comp Graphics Image Processing*, **18**, pp 259–278 (1982)

McKl85: McKee,S., Klein,S., and Teller,D., 'Statistical properties of forced-choice psychometric functions: implications of probit analysis', *Perc & Psychophys*, **37**, pp 286–298 (1985)

MeYa84: Medioni,G., and Yasumoto,Y., 'A note on using the fractal dimension for segmentation', *IEEE Computer Vision Workshop*, Annapolis, MD, Apr 30 – May 3, 1984, pp 25–30

MiMy77: Mitchell,O., Myers,C., and Boyne,W., 'A max-min measure for image texture analysis', *IEEE Trans*, **C-26**, pp 408–414 (1977)

NaBa66: Naylor,T., Balintfy, J., Burdick,D., and Chu,K., *Computer Simulation Techniques*, John Wiley & Sons, 1966

Papo84: Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 2nd ed., McGraw-Hill, 1984

PeNa84: Peleg,S., Naor,J., Hartley,R., and Avnir,D., 'Multiple resolution texture analysis and classification', *IEEE Trans*, **PAMI-6**, pp 518–523 (1984)

Pent83: Pentland, A., 'Fractal-based description', IJCAI-83, pp 973–981 (1983)

Pent84: Pentland, A., 'Fractal-based description of natural scenes', *IEEE Trans*, **PAMI-6**, pp 661–674 (1984)

PiRo83: Pietkäinen,M., Rosenfeld, A., and Davis,L., 'Experiments with texture classification using averages of local pattern matches', *IEEE Trans*, **SMC-13**, pp 421–426 (1983)

PrFa78: Pratt, W.K., Faugeras, O.D., and Gagalowicz, A., 'Visual discrimination of stochastic texture fields', *IEEE Trans*, **SMC-8**, pp 796–804 (1978)

PrMe66: Prewitt, J., and Mendelsohn,M., 'The analysis of cell images', *Ann N.Y. Acad Sci*, **128**, pp 1035–1053 (1966)

Rega82: Regan,D., 1982. 'Visual information channeling in normal and disordered vision'. *Psychological Review*, **89**, pp 407-444 (1982)

Rich79: Richards, W., 'Quantifying sensory channels: Generalizing colorimetry to orientation and texture, touch, and tones', *Sensory Processes*, **3**, pp 207–229 (1979)

Rile81: Riley, M., *The representation of image texture*, AI-TR-649, Artificial Intelligence Laboratory, Massachusettes Institute of Technology, Sept 1981

RiPo74: Richards, W., and Polit, A., 'Texture Matching', *Kybernetik*, **16**, pp 155–162 (1974)

RoKa82: Rosenfeld, A., and Kak, A., *Digital Picture Processing*, 2nd ed., Academic Press, 1982

Rose62: Rosenfeld,A., 'Automatic recognition of basic terrain types from aerial photographs', *Photogrammetric Eng*, **28**, pp 115–132 (1962)

Rose71: Rosenfeld, A., 'Isotonic grammars, parallel grammars, and picture grammars'. In *Machine Intelligence 6*, B. Meltzer, and D. Mitchie, eds. American Elsevier, 1971

Scha78: Schatz,B., *The computation of immediate texture discrimination*, CMU-CS-78-512, Computer Science Department, Carnegie-Mellon University, Dec 1978

Scha80: Schachter, B., 'Long crested wave models', *Comp Graphics and Image Proc*, **12**, pp 187–201 (1980)

ScSh75: Schwartz, M., and Shaw, L., *Signal Processing: discrete spectral analysis, detection, and estimation*, McGraw-Hill, 1975

Stev78: Stevens, K. 'Computation of locally parallel structure', *Biol Cybernetics*, **29**, pp 19–28 (1978)

StJu72: Stromeyer, C., and Julesz, B., 'Spatial-frequency masking in vision: critical bands and spread of masking', *J Opt Soc Am*, **62**, pp 1221–1232 (1972)

Terz82: Terzopoulos,D.,*Multi-level reconstruction of visual surfaces*, MIT AI Lab, AI Memo No 671, 1982

ToSh82: Tomita,F., Shirai,Y., and Tsuji,S., 'Description of textures by a structural analysis', *IEEE Trans*, **PAMI-4**, pp 183–191 (1982)

Trei85: Treisman, A., 'Preattentive Processing in Vision', *Comp Vision Graphics Image Processing*, **31**, pp 156–177 (1985)

WeDy76: Weszka,J., Dyer,C., and Rosenfeld,A., 'A comparative study of texture measures for terrain classification', *IEEE Trans* **SMC-6**, pp 269–285 (1976)

WiBe79: Wilson, H., and Bergen, J., 'A four mechanism model for threshold spatial vision', *Vis Res*, **19**, pp 19–32 (1979)

WiGe84: Wilson, H., and Gelb, D., 'Modified line-element theory for spatial-frequency and width discrimination', *J Opt Soc Am A*, **1**, pp 124–131 (1984)

Wood81: Woodham, R., 'Analysing images of curved surfaces', *Artificial Intelligence*, **9**, pp 117–140 (1981)

Zuck76: Zucker,S., 'Toward a model of texture', *Comp Graphics Image Processing*, **5**, pp 190–202 (1976)

Zuck84: Zucker,S., 'Two constraints on early orientation selection in dot patterns'. In *Figural Synthesis*, P. Dodwell and T. Caelli, eds. Lawrence Erlbaum Associates, 1984

ZuTe80: Zucker,S., and Terzopoulos,D., 'Finding structure in co-occurrence matrices for texture analysis', *Comp Graphics Image Processing*, **12**, pp 286–308 (1980)

Appendix A

Random fields

This appendix describes some of the basic concepts and methods used for the analysis of two-dimensional random fields. Much of the material is based on Jenkins and Watts [JeWa68] and Papoulis [Papo84].

A.1 Introduction

The term *random image*, as used here, refers to an image containing no apparent regularities of any kind. In other words, there is no algorithm available to the observer that would allow compression of the size of the image description. Since only a limited amount of information can be used for a representation, the description of an entire random image is often impractical. Instead, representations must be used that allow a maximum of information to be captured by a minimum of description.

Among the representations commonly used are sets of average properties. These properties and their relations can be determined by the methods of random-field analysis. Such methods originated in the analysis of time series, one-dimensional random functions for which the value of a function at a given

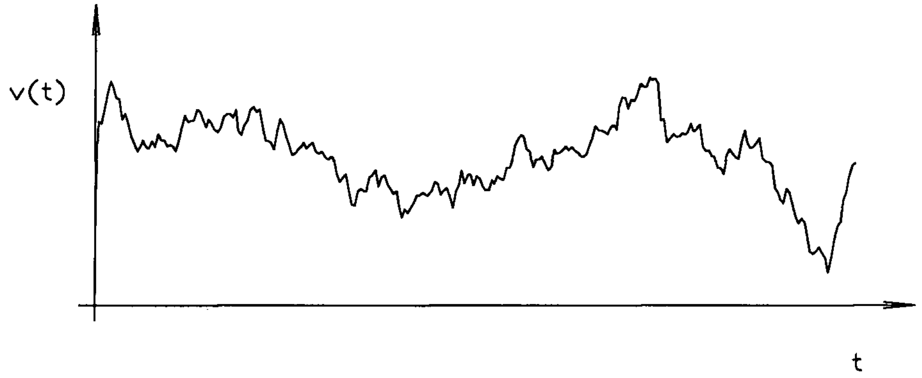


Figure A.1: example of time series

time cannot be predicted exactly from the knowledge of its values at previous times (figure A.1).

Although different sections of a time series $v(t)$ over similar intervals Δt have little similarity in a strict sense, their average properties are often nearly identical. This leads to the idea of modelling a time series by a *stochastic process*, an ordered set of random variables $\mathbf{v}(t)$ that describes the ensemble of functions that could possibly be realized. The function $v(t)$ is simply one of the infinitely many values that the process $\mathbf{v}(t)$ could have taken. Such a treatment makes it possible to relate the measured averages of $v(t)$ to the ensemble properties of $\mathbf{v}(t)$. This allows the relations between the average properties to be treated in an exact fashion.

To maintain the distinction between the ensemble and one of its instances, random variables are always denoted by bold-faced characters. Any particular instance is denoted by a character of standard font.

The concept of a stochastic process can be extended to obtain that of a *random field*, a two-dimensional space of (complex) random variables $\mathbf{f}(x, y)$. A random image $f(x, y)$ may then be interpreted as an instance of the ensemble of functions described by $\mathbf{f}(x,y)$. Time series and random fields can be considered as special cases of n-dimensional random fields $\mathbf{f}(\vec{x})$, where $\vec{x} = (x_1, x_2, \ldots, x_n)$ is an n-dimensional position vector. In what follows, only the two-dimensional case will be developed, since it is of most relevance to image analysis. Analogous developments, however, can be carried out for any finite number of dimensions.

A random field can be either continuous or discrete, depending on the parameter set $\{\vec{x}\}$. Continuous random fields are denoted here by standard functional notation (e.g., $\mathbf{f}(x,y), \tilde{\mathbf{g}}(k,l)$). Discrete fields, on the other hand, are denoted using subscript pairs (e.g., $\mathbf{f}_{xy}, \tilde{\mathbf{g}}_{mn}$). For convenience, only the continuous case is described in this appendix. The discrete case can be developed in a parallel fashion by replacing integration by summation, and the continuous Fourier transforms by their discrete counterparts.

A.2 Mean and Covariance

A random field $\mathbf{f}(x,y) = \mathbf{f}(\vec{x})$ is defined to be a set of random variables parametrized by a two-dimensional space, with each point $\vec{x}_0$ being the location of a (complex) random variable $\mathbf{f}(\vec{x}_0)$. A random field is often represented by a multivariate probability distribution, which completely describes the joint statistical properties of all its constituent random variables.

Each variable $\mathbf{f}(\vec{x}_0)$ has an associated probability density function $p_{\mathbf{f}(\vec{x}_0)}(w)$, where w is, in general, a complex quantity. The consideration given to complex random variables is not only for the sake of generality, but also because of the simplifications brought to the formal description of Fourier transformation.

Real-valued random fields are easily treated as a special case of this more general approach.

Two significant properties of a complex random variable are its *mean*

$$\mu(\vec{x}_0) = E\{\mathbf{f}(\vec{x}_0)\} = \int_{\Im} w \cdot p_{\mathbf{f}(\vec{x}_0)}(w)dw, \tag{A.1}$$

and its *variance*

$$\sigma^2(\vec{x}_0) = E\{|\mathbf{f}(\vec{x}_0) - \mu(\vec{x}_0)|^2\} = \int_{\Im} |w - \mu(\vec{x}_0)|^2 p_{\mathbf{f}(\vec{x}_0)}(w)dw, \tag{A.2}$$

where the region of integration is taken to be the complex plane $\Im$.

Pairwise correlations amongst the random variables are described by the *correlation function*

$$R(\vec{x}_1, \vec{x}_2) = E\{\mathbf{f}(\vec{x}_1)\mathbf{f}(\vec{x}_2)^*\}, \tag{A.3}$$

and the *covariance function*

$$\begin{aligned} C(\vec{x}_1, \vec{x}_2) &= E\{[\mathbf{f}(\vec{x}_1) - \mu(\vec{x}_1)][\mathbf{f}(\vec{x}_2) - \mu(\vec{x}_2)]^*\} \\ &= R(\vec{x}_1, \vec{x}_2) - \mu(\vec{x}_1)\mu(\vec{x}_2)^*, \end{aligned} \tag{A.4}$$

which describe the linear dependence of $\mathbf{f}(\vec{x}_1)$ and $\mathbf{f}(\vec{x}_2)$ on each other.

When $\mu(\vec{x}) = \mu$, the two functions differ only by the constant $|\mu|^2$, and it is common practice to use only one of them for description. In such a case, this work uses the covariance function $C(\vec{x}_1, \vec{x}_2)$ to describe the second-order moments of random fields. Note that $C(\vec{x}_1, \vec{x}_1) = \sigma^2(\vec{x}_1)$, i.e., the covariance of a single random variable is its own variance.

If the random field has a multivariate Gaussian distribution associated with it, the field is completely specified by $\mu(\vec{x})$ and $C(\vec{x}_1, \vec{x}_2)$ [Papo84:9-2]. For such a case, the condition $C(\vec{x}_1, \vec{x}_2) = 0$ implies that the corresponding random variables are independent.

A.3 Stationarity

The random field $\mathbf{f}(x, y)$ is said to be *stationary* if its statistical properties remain invariant under translation, viz,

$$\mathbf{f}(x, y) = \mathbf{f}(x + \Delta x, y + \Delta y).$$

Such fields represent processes that are independent of any particular location — they have an equilibrium distribution that has the same statistical properties everywhere. In the remainder of this report, attention is restricted to stationary random fields.

For a multivariate Gaussian distribution, it follows that the random field is stationary iff $\mu(\vec{x}) = \mu$ and $C(\vec{x}_1, \vec{x}_2) = C(\vec{x}_1 - \vec{x}_2)$.

A.4 Sample Functions

Given an instance $f(x, y)$ of $\mathbf{f}(x, y)$, various sample functions may be defined on it. These functions determine various average properties, which in turn can provide estimates of ensemble properties.

For any given instance $f(x, y)$ a *sample average* $\bar{f}$ may be defined as

$$\bar{f} = \frac{1}{T_x T_y} \int_{-T_x/2}^{T_x/2} \int_{-T_y/2}^{T_y/2} f(x, y) dx dy,$$

where the random field $\mathbf{f}(x, y)$ has been assumed to be zero outside the domain $[-T_x/2, T_x/2] \times [-T_y/2, T_y/2]$. The *sample correlation function* is similarly defined as

$$\bar{c}(\Delta x, \Delta y) = \frac{1}{T_x T_y} \int_{-T_x/2}^{T_x/2} \int_{-T_y/2}^{T_y/2} [f(x + \Delta x, y + \Delta y) - \bar{f}][f(x, y) - \bar{f}]^* dx dy.$$

These sample functions may be considered to be particular instances of ensemble estimators. Ensemble estimators are functionals of the random field, assigning a random variable to any $\mathbf{f}(x,y)$. The estimator for the mean, for example, is

$$\bar{\mathbf{f}} = \frac{1}{T_x T_y} \int_{-T_x/2}^{T_x/2} \int_{-T_y/2}^{T_y/2} \mathbf{f}(x,y)\,dxdy,$$

and the covariance estimator is

$$\bar{\mathbf{c}}(\Delta x, \Delta y) = \frac{1}{T_x T_y} \int_{-T_x/2}^{T_x/2} \int_{-T_y/2}^{T_y/2} [\mathbf{f}(x+\Delta x, y+\Delta y) - \bar{\mathbf{f}}][\mathbf{f}(x,y) - \bar{\mathbf{f}}]^* dxdy.$$

When $f(x,y)$ is stationary, it can be shown that the estimators $\bar{\mathbf{f}}$ and $\bar{\mathbf{c}}(\Delta x, \Delta y)$ asymptotically approach the constant ensemble values of μ and $C(\Delta x, \Delta y)$ respectively as $T_x, T_y \rightarrow \infty$ [JeWa68,ch5]. Both sample functions are consequently *ergodic*, the spatial averages reflecting the ensemble averages.

A.5 Fourier Analysis

The techniques of Fourier analysis can be usefully applied to the study of random fields. Attention is focussed here on the Fourier series and Fourier transforms of continuous functions. Analogous developments for the discrete case can be done using finite discrete Fourier series [JeWa68].

An instance $f(x,y)$ of a random field $\mathbf{f}(x,y)$ may be analyzed into its constituent harmonics in the same way as any other function. Let $f(x,y)$ denote an image that is non-zero only inside the domain $[-T_x/2, T_x/2] \times [-T_y/2, T_y/2]$. Such an image can be represented by the Fourier series

$$f(x,y) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \tilde{f}_{kl} \exp\{i2\pi(n/T_x + m/T_y)\},$$

where

$$\tilde{f}_{kl} = \frac{1}{T_x T_y} \int_{-T_x/2}^{T_x/2} \int_{-T_y/2}^{T_y/2} f(x,y) \exp\{-i2\pi(kx+ly)\} dxdy.$$

The term $\tilde{f}_{kl}$ is the frequency-space representation of $f(x,y)$. It is almost always random for any given instance $f(x,y)$, its value at a given point having no definite relation to the values at other points. Increasing the size of T_x and T_y does not cause $\tilde{f}_{kl}$ to settle down to some deterministic function. Apart from windowing effects arising from the finite sizes of T_x and T_y, the average properties of $\tilde{f}_{kl}$ (as determined from its sample functions) usually converge to definite values as $T_x, T_y \rightarrow \infty$. (See [JeWa68] for illustrations of the one-dimensional case.)

With this is mind, it is natural to regard $\tilde{f}_{kl}$ as an instance of an ensemble of possible functions $\mathbf{f}_{kl}$. When $T_x, T_y \rightarrow \infty$, this leads to the the random field

$$\tilde{\mathbf{f}}(k,l) = \tilde{\mathbf{f}}(\vec{k}) = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} \mathbf{f}(x,y) \exp\{-i2\pi(kx+ly)\} dxdy. \qquad \text{(A.5)}$$

Similarly, $\mathbf{f}(x,y)$ can be expressed as the inverse Fourier transform of $\tilde{\mathbf{f}}(k,l)$, viz.,

$$\mathbf{f}(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\mathbf{f}}(k,l) \exp\{i2\pi(kx+ly)\} dkdl. \qquad \text{(A.6)}$$

A.6 Power Spectra

When $\mathbf{f}(x,y)$ is stationary with mean μ and covariance $C(\vec{x}_1 - \vec{x}_2)$, it follows from eq (A.5) that

$$E\{\tilde{\mathbf{f}}(\vec{k})\} = \mu\delta(0),$$

and

$$
\begin{aligned}
&E\{\tilde{\mathbf{f}}(\vec{k}_1)\tilde{\mathbf{f}}^*(\vec{k}_2)\} \\
&\quad = \textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(\vec{x})\exp\{-i2\pi(\vec{k}_1\cdot\vec{x})\}\exp\{-i2\pi(\vec{x}_2\cdot(\vec{k}_1-\vec{k}_2))\}d\vec{x}_1 d\vec{x}_2,
\end{aligned}
$$

where $\vec{x} = \vec{x}_1 - \vec{x}_2$, and $\delta(\vec{k})$ is the Dirac delta function. This last term can be written

$$
\begin{aligned}
&E\{\tilde{\mathbf{f}}(\vec{k}_1)\tilde{\mathbf{f}}^*(\vec{k}_2)\} \\
&\quad = \textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(\vec{x})\exp\{-i2\pi(\vec{k}_1\cdot\vec{x})\}\exp\{-i2\pi(\vec{x}_2\cdot(\vec{k}_1-\vec{k}_2))\}d\vec{x}d\vec{x}_2 \\
&\quad = \textstyle\int_{-\infty}^{\infty} R(\vec{x})\exp\{-i2\pi(\vec{k}_1\cdot\vec{x})\}\delta(\vec{k}_1-\vec{k}_2)d\vec{x} \\
&\quad = S(\vec{k}_1)\delta(\vec{k}_1-\vec{k}_2),
\end{aligned}
$$

where

$$S(\vec{k}) = \int_{-\infty}^{\infty} R(\vec{x})\exp\{-i2\pi(\vec{k}\cdot\vec{x})\}d\vec{x}$$

is the *power spectrum* of the random field.

The $\tilde{\mathbf{f}}(\vec{k})$ form a field of independent random variables. When $\vec{k} \neq 0$, they are zero-mean, and have a variance of $S(\vec{k})\delta(0)$. It is often convenient to factor the $\tilde{\mathbf{f}}(\vec{k})$ into

$$\tilde{\mathbf{f}}(k,l) = m(k,l)\mathbf{z}(k,l) + \mu\delta(0),$$

where the $\mathbf{z}(k,l)$ are zero-mean random variables with a delta variance (i.e., $E\{\mathbf{z}(\vec{k}_1)\mathbf{z}^*(\vec{k}_2)\} = \delta(\vec{k}_1 - \vec{k}_2)$), and where $m(k,l)$ is a real-valued non-negative function that modulates them. Using this factorization, eq (A.6) can be written

$$\mathbf{f}(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[m(k,l)\mathbf{z}(k,l) + \mu\delta(0)]\exp\{i2\pi(kx+ly)\}dkdl. \tag{A.7}$$

A similar result holds for the discrete case, where $E\{\mathbf{z}_{k_1}\mathbf{z}^*_{k_2}\} = \delta_{k_1k_2}$ [Papo84].

Using the definition

$$
\begin{aligned}
R(\vec{x}_1,\vec{x}_2) &= E\{\mathbf{f}(\vec{x}_1)\mathbf{f}^*(\vec{x}_2)\} \\
&= \textstyle\int_{-\infty}^{\infty}[m^2(\vec{k}) + |\mu|^2\delta(0)]\exp\{i2\pi(\vec{k}\cdot(\vec{x}_1-\vec{x}_2))\}d\vec{k},
\end{aligned}
$$

it follows that

$$S(\vec{k}) = m^2(k,l) + |\mu|^2\delta(0).$$

Since the $\mathbf{z}(k,l)$ have delta variance, it follows that the power spectum describes the contribution of the harmonic at (k,l) to the rms power of $\mathbf{f}(x,y)$.

An interesting relation exists between $m^2(k,l)$ and the covariance function $C(\vec{x}_1,\vec{x}_2)$ of a stationary field. Since

$$\begin{aligned} C(\vec{x}_1-\vec{x}_2) &= R(\vec{x}_1-\vec{x}_2) - |\mu|^2 \\ &= \int_{-\infty}^{\infty}[S(k,l) - |\mu|^2\delta(0)]\exp\{i2\pi(\vec{k}\cdot(\vec{x}_1-\vec{x}_2))\}d\vec{k} \\ &= \int_{-\infty}^{\infty} m^2(k,l)\exp\{i2\pi(\vec{k}\cdot(\vec{x}_1-\vec{x}_2))\}d\vec{k}, \end{aligned}$$

the function $C(\vec{x}_1-\vec{x}_2)$ is the Fourier transform of the quantity $m^2(k,l) = S(\vec{k}) - |\mu|^2\delta(0)$.

When the $\mathbf{z}(k,l)$ are delta-variance Gaussian random variables, the function $\mathbf{f}(x,y)$ is multivariate Gaussian, since it is a linear combination of Gaussian random variables. The random field is then completely determined by $\mu(\vec{x})$ and $C(\vec{x}_1,\vec{x}_2)$. In this work, the $\mathbf{z}(k,l)$ are always taken to be zero-mean, delta-variance Gaussian random variables. This allows the power spectrum and covariance function to be equivalent descriptions of the random field.

A.7 Real-Valued Random Fields

If the random field $\mathbf{f}(x,y)$ is real-valued, then $\mathbf{f}(x,y)$ must equal $\mathbf{f}^*(x,y)$. From eq (A.7), it then follows that $\mathbf{z}(\vec{k})$ must be conjugate-symmetric about the origin (i.e., $\mathbf{z}(\vec{k}) = \mathbf{z}^*(-\vec{k})$), and that μ be real-valued.

It is important to note that if $\mathbf{f}(x,y)$ is real-valued, then the $\mathbf{z}(\vec{k})$ must be complex quantities with random phases. If the $\mathbf{z}(\vec{k})$ were real-valued, with conjugate symmetry, then

$$E\{\mathbf{z}(\vec{k}_1)\mathbf{z}^*(\vec{k}_2)\} = \delta(\vec{k}_1 - \vec{k}_2) + \delta(\vec{k}_1 + \vec{k}_2).$$

This relation, together with eq (A.7), entails that

$$\begin{aligned}
C(\vec{x}_1 - \vec{x}_2) &= E\{[\mathbf{f}(\vec{x}_1) - \mu][\mathbf{f}(\vec{x}_2) - \mu]^*\} \\
&= \textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} m(k_1, l_1)m(k_2, l_2)E\{\mathbf{z}(k_1, l_1)\mathbf{z}^*(k_2, l_2)\}\cdot \\
&\qquad \exp\{i2\pi((k_1x_1 - k_2x_2) + (l_1y_1 - l_2y_2))\}d\vec{k}_1 d\vec{k}_2 \\
&= \textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} m^2(k, l)\exp\{i2\pi(k(x_1 - x_2) + l(y_1 - y_2))\}d\vec{k} \\
&\quad + \textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} m^2(k, l)\exp\{i2\pi(k(x_1 + x_2) + l(y_1 + y_2))\}d\vec{k}.
\end{aligned}$$

The covariance function therefore depends upon $\vec{x}_1 + \vec{x}_2$, showing that stationarity is lost when the $\mathbf{z}(\vec{k})$ are real-valued.

Appendix B

Fractals

This appendix provides a brief overview of some of the basic ideas of fractal geometry. It is an extension of the expositions given by Mandelbrot [Mand67][MaNe68][Mand75][Mand82][Mand84].

B.1 Introduction

The piecewise-differentiable curves and surfaces commonly used to describe shape do not adequately represent all aspects of the forms found in nature. For example, the length of a coastline is not a well-defined quantity — as the scale of measurement is made finer, small indentations and promontories previously unresolved become noticeable, thereby increasing the total length measured. As the resolution is increased, the length of the coastline tends toward infinity. In a similar fashion, the measured surface area of a rugged terrain depends on the scale of measurement, increasing without bound as the scale tends to zero. These are not isolated phenomena.

Richardson showed empirically [Mand82] that jagged objects such as coastlines could be characterized by the rate at which their length increased as a function of measurement resolution. When the basic measuring-scale λ used in some method is replaced by λ/a, the number of the scale-lengths measured increases as a^D, where D is some constant such that $1 \leq D \leq 2$. Thus the total measured length $L(\lambda)$ can be written as

$$L(\lambda) = F\lambda^{-D} \cdot \lambda = F\lambda^{1-D}, \tag{B.1}$$

with F a constant depending on the particular object measured, and the method of measurement used.

The parameter D holds constant over a wide range of spatial scales, reflecting the inherent jaggedness of a coastline. For $D \approx 1$, the line is smooth. As D increases, the line becomes much more jagged. For $D \rightarrow 2$, a coast becomes extremely rough and convoluted, with many islands and fiords. The average value of D for the world's coastlines is estimated to be 1.2 [Mand82].

This approach can be applied equally well to jagged surfaces and volumes. When the value of D is the same as the intuitive dimension of the object (e.g., $D = 1$ for coasts, $D = 2$ for surfaces), the object can be adequately described by standard Euclidean geometry. When D has a different value, the object is regarded as a *fractal.*

B.2 Definitions

Fractals are a class of mathematical objects largely developed by Mandelbrot [MaNe68][Mand82], who defines a fractal as a set whose Hausdorff-Besicovitch dimension is greater than its topological dimension.

The topological dimension T is that captured by the intuitive idea of dimension. All surfaces, for example, have $T = 2$. The Hausdorff-Besicovitch dimension is obtained via a test function or generalized ball

$$h_d(\rho) = \frac{[\Gamma(1/2)]^d}{\Gamma(1+(d/2))} \cdot \rho^d, \tag{B.2}$$

where $\Gamma(x)$ is the gamma function. The ball takes the following forms for $d \in \{1,2,3\}$:

$$\begin{aligned} d = 1: &\qquad h_1(\rho) = 2\rho \\ d = 2: &\qquad h_2(\rho) = \pi\rho^2 \\ d = 3: &\qquad h_3(\rho) = \tfrac{4}{3}\pi\rho^3 \end{aligned}$$

Let the quantity $M(d,\rho)$ be the smallest possible covering of a set S with h_d-balls of radius $\rho_m < \rho$. The measure $M(d)$ of S is then given as the limit of $M(d,\rho)$ as the radius of the balls approaches zero

$$M(d) = \lim_{\rho \to 0} \left\{ \inf_{\rho_m < \rho} \sum_S h_d(\rho_m) \right\} \tag{B.3}$$

Only one value of d results in $M(d)$ taking on a non-zero, finite number. This is the Hausdorff-Besicovitch dimension D of the set S. For example, if S is a square area of unit dimensions,

$$\begin{aligned} d = 1: &\qquad h_1(\rho) \propto \rho. \quad \text{Thus } M(1) \propto \rho/\rho^2 \to \infty \\ d = 2: &\qquad h_2(\rho) \propto \rho^2. \text{ Thus } M(2) \propto \rho^2/\rho^2 \to 1 \\ d = 3: &\qquad h_3(\rho) \propto \rho^3. \text{ Thus } M(3) \propto \rho^3/\rho^2 \to 0 \end{aligned}$$

The only non-vanishing, finite measure is for the ball h_2, so that $D = 2$.

This approach readily extends to arbitrary sets. For many of these, D is not an integer, and is greater than the topological dimension T. These sets are termed *fractals*, and D is called the *fractal dimension.*

An embedding set X is defined to be a Euclidean space that contains the fractal set S i.e., $S \subseteq X$. The dimension E of X provides an upper bound on D. Thus, D must always obey the double inequality $T \leq D \leq E$.

B.3 Deterministic fractals

The deterministic fractals are those that are constructed according to fixed, definite rules. The simplest such fractals are sets made up of the union of a finite number $n \in Z^+$ of compact sets [Hutc81] such that

$$a = S_1(a) \cup S_2(a) \cup \cdots \cup S_n(a),$$

where the S_i are ***similitudes***, i.e., mappings composed of a translation, an orthonormal transformation, and a homothety (a uniform scaling). The scaling factor $r > 1$ of the homothety describes the ratio of the set a to its subset $S_i(a)$. It is often referred to as the ***geometric ratio***. The form of the set a is constrained to produce self-similarity, but may otherwise be arbitrarily chosen.

The dimension of such a set can be easily ascertained when the intersection of the $S_i(a)$ has a smaller dimension than that of the set itself. (This condition is almost always the case [Mand82].) Given that a self-similar set a has a dimension D, it follows that each of the $S_i(a)$ must have the same dimension D, since this quantity is invariant under translation, scaling, and orthonormal transformation. Taking the measure of a to be the sum of the measures of the $S_i(a)$, this becomes

$$M(D) = \lim_{\frac{\rho}{r} \to 0} \left\{ \inf_{\frac{\rho_m}{r} < \frac{\rho}{r}} \sum_{\bigcup S_i(a)} \frac{[\Gamma(\frac{1}{2})]^D}{\Gamma(1 + \frac{D}{2})} (\frac{\rho_m}{r})^D \right\}$$

Equating this with the direct measure of a, viz.,

$$M(D) = \lim_{\rho \to 0} \left\{ \inf_{\rho_m < \rho} \sum_a \frac{[\Gamma(\frac{1}{2})]^D}{\Gamma(1 + \frac{D}{2})} \rho_m^D \right\}$$

leads to the condition

$$\frac{n}{r^D} = 1.$$

The dimension of the set may therefore be expressed as

$$D = \log(n) / \log(r). \tag{B.4}$$

It is important to distinguish clearly between fractal sets and self-similar sets, for neither is a strict subclass of the other. For example, a union of straight horizontal line segments connected at their endpoints forms a horizontal line, which is technically not a fractal, since $D = T = 1$. Fractals, on the other hand, are not necessarily characterized by self-similarity, for the similitudes S_i are not the only mappings that produce them [Hutc81]. In this work, attention is restricted to the class of objects that are both self-similar and fractal, as they provide a convenient domain containing all the properties of the general class of fractals.

Example: Koch curve

The Koch curve K is a simple self-similar curve, constructed in a recursive fashion. In the first stage of construction, the initial base is replaced by a generator made up of four smaller sections (see figure B.1), each having a geometric ratio $r = 3$ with respect to the base.

Each of these first-stage sections is then replaced in its turn by a scaled generator to obtain the second-stage figure. The Koch curve is defined to be the limit of the process as the number of stages approaches infinity.

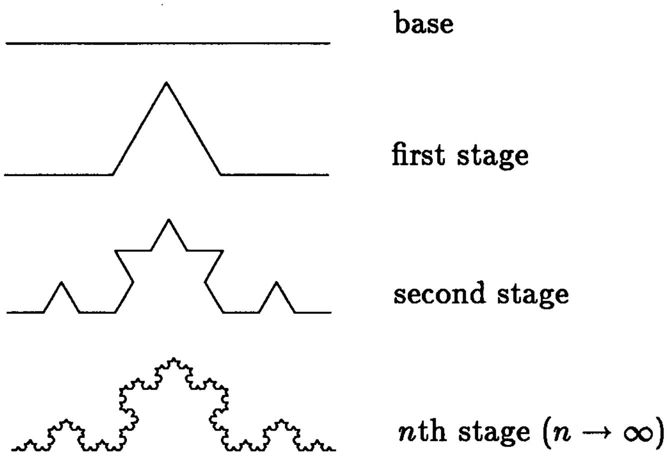


Figure B.1: Construction of Koch curve

The resulting curve is self-similar in its embedding space ($E = 2$), with $r = 3$ and $n = 4$. Its dimension is therefore

$$D = \log(n)/\log(r) = \log(4)/\log(3) \approx 1.26.$$

In many respects, the Koch curve roughly approximates a coastline. Indeed, its length L increases as

$$L(k) = C(n/r)^k = C(1/r)^{k(1-D)},$$

where k is the level of the stage generated, and C is a constant depending on the size of the base. Denoting the length of the sections at stage k by

$$\epsilon = (1/r)^k,$$

the length may be written

$$L(\epsilon) = C\epsilon^{1-D}, \tag{B.5}$$

The similarity between eqs (B.1) and (B.5) is readily apparent.

The example also shows that although the condition $D > T$ implies that the curve is non-differentiable, it does not imply that continuity is necessarily lost.

B.3.1 Parametric Representation

An alternate characterization of self-similarity is to regard a fractal as an E-dimensional vector quantity $\vec{a}$ parametrized by a T-dimensional vector $\vec{t}$. The parameterization process is illustrated for the case of a curve ($T = 1$) in a two-dimensional embedding space ($E = 2$). This process can be readily extended to the general case.

A general class of self-similar curves can be constructed by generalizing the process used to form the Koch curve. An initial straight-line base is replaced by a generator composed of n equal-length sections. Apart from the requirement of contiguity of the sections, the form of the generator is arbitrary (figure B.2). Each section is in turn replaced by a scaled-down generator, the process being recursively continued until a limiting form is reached.

Parametrization by a real-valued quantity t is done in a fashion which parallels that of the construction process. To begin with, the location of the left and right endpoints of the initial section are left undisturbed by later stages of construction. They may therefore be unambiguously assigned correspondences to the parametric values t_0 and t_n respectively, where $t_0 < t_n$.

Consider now the first stage of construction. The locations of the corners between the n line sections in the generator remain unaffected by later stages. They may therefore be assigned correspondences to the parametric values

$$t_i = t_0 + (i/n)[t_1 - t_0] \mid i \in \{1, 2, ..., (n-1)\},$$

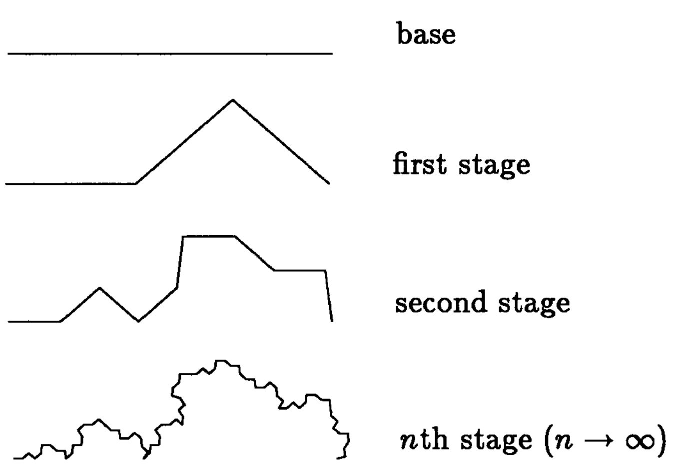


Figure B.2: Construction of generalized curve

the assignment being such that the values along the line form a strictly increasing sequence. This parameterization is continued in a similar way for all later stages of the construction. Each point on the curve therefore uniquely corresponds to a value of t.

For an arbitrary point on such a parametrized curve, its location $\vec{a}(t)$ may be specified in two different ways. The first is with respect to the origin $\vec{a}(t_0)$ of the base section; the second is with respect to the starting point $\vec{a}(t_i)$ of the generator section giving rise to that point (figure B.3). Comparing the two formulations, it follows that

$$\vec{a}(t) - \vec{a}(t_0) = \vec{a}(t_i) - \vec{a}(t_0) + (1/r)R(\theta_i)[\vec{a}(t_0 + n(t - t_i)) - \vec{a}(t_0)] \qquad \text{(B.6)}$$

for all $t_0 < t < t_n$. The $R(\theta_i)$ are rotation operators, that relate the orientation of the generator sections to that of the base. For the section running from t_i to

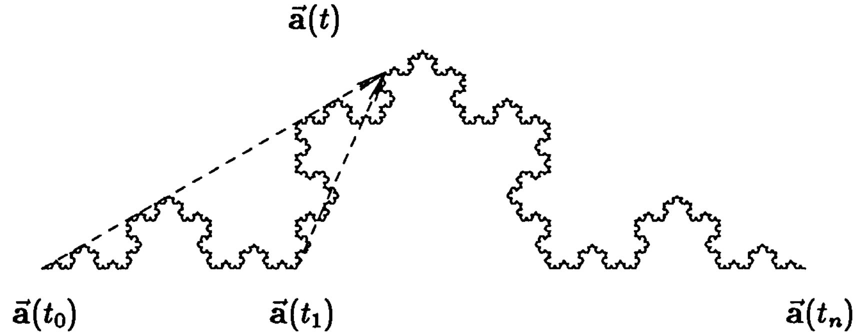


Figure B.3: Relation of descriptions of self-similar curve

t_{i+1}, the rotation angle θ_i is given by

$$\theta_i = \arctan(\frac{a_y(t_{i+1}) - a_y(t_i)}{a_x(t_{i+1}) - a_x(t_i)}) - \arctan(\frac{a_y(t_n) - a_y(t_0)}{a_x(t_n) - a_x(t_0)}). \qquad (B.7)$$

For these curves, then, the similitudes S_i are composed of a translation $\vec{a}(t_i) - \vec{a}(t_0)$, a rotation $R(\theta_i)$, and a uniform rescaling by a factor of r.

The quantity n describes the scaling of the parameter t. It is referred to as the *(parametric) scaling ratio.* In general, this quantity is not uniquely defined, since a self-similar set with $n = n_0$ is also self-similar with $n = n_0^j$; $j \in Z^+$. To provide a unique characterization of this aspect of a fractal's behaviour, only the smallest value of n greater than unity is taken as the value of the scaling ratio.

For convenience, the geometric ratio r is often expressed indirectly, using n and the *similarity parameter* $H = \log(r)/\log(n)$. The similarity parameter describes the relation between the geometric ratio of the embedding-space and the scaling ratio. Comparison with eq.(B.4) shows that $D = 1/H$ when $1/E \leq H \leq 1/T$. Although fully equivalent to the use of $1/D$ for deterministic fractals, the use of the similarity parameter H is more advantageous in the stochastic domain. Mandelbrot regards the quantity $1/H$ as the latent dimension of the fractal, capturing its self-similar behaviour better in most respects than the Hausdorff-Besicovitch dimension D [Mand84]. For many purposes, then, H is the most appropriate measure of self-similarity.

Generalizing from self-similar curves, the parametric equation for a self-similar set $\vec{a}$ may be written in the form

$$\vec{a}(\vec{t}) - \vec{a}(\vec{t}_0) = \vec{a}(\vec{t}_i) - \vec{a}(\vec{t}_0) + n^{-H}\Pi(\vec{\Theta}_i)[\vec{a}(\vec{t}_0 + n(\vec{t} - \vec{t}_i)) - \vec{a}(\vec{t}_0)] \tag{B.8}$$

where $\Pi(\vec{\Theta}_i)$ is an orthonormal transformation, and Θ denotes its parameters. Note that it is the relative changes of $\vec{a}(\vec{t})$ that are translated, scaled, and transformed. This allows self-similarity to be well-defined even when $\vec{a}(\vec{t}_0)$ is a divergent quantity.

B.3.2 Fractal Functions

A special class of fractals are the *fractal functions*, comprising those fractals that are parametrized by an extrinsic co-ordinate system in the embedding space itself. Two different characterizations may be used to describe such functions. The first is the standard one, using the intrinsic parameter $\vec{t}$. The second is the behaviour of the fractal in an embedding space containing $\vec{t}$ as one of the geometric dimensions. The intrinsic parameter $\vec{t}$ then becomes an extrinsic position vector $\vec{u}$. Using this latter characterization, the function is therefore a generalized graph.

For self-similar curves in two-dimensional space, $\vec{u}$ is a scalar, and the fractal functions become a linked pair, having the form

$$\begin{aligned} a_x(u) - a_x(u_0) = {} & a_x(u_i) - a_x(u_0) \\ & + n^{-H}\{[a_x(u_0 + n(u - u_i)) - a_x(u_0)]\cos(\theta_i) \\ & \quad - [a_y(u_0 + n(u - u_i)) - a_y(u_0)]\sin(\theta_i)\} \end{aligned} \tag{B.9}$$

$$\begin{aligned} a_y(u) - a_y(u_0) &= a_y(u_i) - a_y(u_0) \\ &\quad + n^{-H}\{[a_x(u_0 + n(u - u_i)) - a_x(u_0)]\sin(\theta_i) \\ &\qquad + [a_y(u_0 + n(u - u_i)) - a_y(u_0)]\cos(\theta_i)\} \end{aligned} \tag{B.10}$$

The general relation between the dimension D of these functions and H is not known. For a few scalar functions in two dimensions, however, it has been shown that $D = 2 - H$. It is conjectured that this relation holds in general, and that for functions embedded in three dimensions, $D = 3 - H$ [Mand82].

B.4 Stochastic fractals

Stochastic fractals are defined as sets of random variables that are self-similar with regard to their statistical properties, i.e.,

$$\mathbf{a} = S_1(\mathbf{a}) \cup S_2(\mathbf{a}) \cup \cdots \cup S_n(\mathbf{a}),$$

holds for their joint probability density functions. Using parametric representation, this becomes

$$\vec{a}(\vec{t}) - \vec{a}(\vec{t}_0) = \vec{a}(\vec{t}_i) - \vec{a}(\vec{t}_0) + n^{-H}\Pi(\vec{\Theta}_i)[\vec{a}(\vec{t}_0 + n(\vec{t} - \vec{t}_i)) - \vec{a}(\vec{t}_0)] \tag{B.11}$$

where the $\vec{t}_i$ are the parametric values of the origin of the appropriate generator section. The $\vec{\Theta}_i$ are random variables specifying the parameters of the orthogonal transformation Π in the similitude S_i. These fractals may be defined in a recursive fashion similar to that of the deterministic case, except that the generator is no longer a fixed pattern, but is an ensemble of patterns. The ensemble determines the (joint) density functions of the $\vec{\Theta}_i$.

For the case of random curves through a two-dimensional embedding space, the stochastic counterpart to equation (B.6) is

$$\vec{a}(t) - \vec{a}(t_0) = \vec{a}(t_i) - \vec{a}(t_0) + n^{-H}R(\vec{\Theta}_i)[\vec{a}(t_0 + n(\vec{t} - \vec{t}_i)) - \vec{a}(t_0)],$$

where

$$\vec{\Theta}_i = \arctan(\frac{\mathbf{a}_y(t_{i+1}) - \mathbf{a}_y(t_i)}{\mathbf{a}_x(t_{i+1}) - \mathbf{a}_x(t_i)}) - \arctan(\frac{\mathbf{a}_y(t_n) - \mathbf{a}_y(t_0)}{\mathbf{a}_x(t_n) - \mathbf{a}_x(t_0)}).$$

An instance of self-similar random sets can be constructed using a recursive procedure similar to that for the deterministic figure, except that each straight-line section present at a given stage is replaced by a (different) instance of the generator ensemble. Such a construction process is used in computer graphics to produce self-similar random curves; a corresponding process produces self-similar random surfaces (e.g., [FoFu82][HaBa84)].

B.4.1 Stationary Increments

When the increments $\vec{\mathbf{a}}(t_2) - \vec{\mathbf{a}}(t_1)$ of a stochastic fractal are stationary (see appendix A), its description can be simplified in two ways. First, the translational component of the similitude S_i can be eliminated, since the probability densities of all increments must be of the same form. This implies that the translation $\vec{\mathbf{a}}(\vec{t}_i) - \vec{\mathbf{a}}(\vec{t}_0)$ of any sections relative to the base must be zero. Second, the distribution of transformation parameters $\vec{\Theta}_i$ must be the same for any section $\vec{\mathbf{a}}(\vec{t}_{i+1}) - \vec{\mathbf{a}}(\vec{t}_i)$. If the generator sections have isotropic probability densities, any orthonormal transformation is compatible with the description. If the sections have anisotropic densities, however, $\Pi(\vec{\Theta}_i)$ may be taken as the identity operator, for otherwise the sum of adjoining increments would have a density function of a form different from that of the original set. In general, then, stochastic fractals with stationary increments may be considered as having neither a translational nor a transformational component to their similitudes.

Taken together, these two conditions imply that the stochastic properties of all generator sections are identical. This implies that only the first section need be used to describe the fractal, so that the parametric value t_i may be set to t_0. Abolishing the translational and rotational components of the similitude in eq (B.11) then leads to

$$\vec{a}(\vec{t}) - \vec{a}(\vec{t}_0) = n^{-H}[\vec{a}(\vec{t}_0 + n(\vec{t} - \vec{t}_0)) - \vec{a}(\vec{t}_0)].$$

This can be written in the more symmetrical form

$$\vec{a}(\vec{t}_0 + n(\vec{t} - \vec{t}_0)) - \vec{a}(\vec{t}_0) = n^{H}[\vec{a}(\vec{t}_0 + (\vec{t} - \vec{t}_0)) - \vec{a}(\vec{t}_0)]$$

which emphasizes that it is the behaviour of the increments that characterizes the fractal. In particular, note that stationarity of the increments does not imply stationarity of the fractal itself.

A more general class of stochastic fractals is obtained by removal of the constraint that the scaling ratio n be an integer [MaNe68]. This takes advantage of the fact that for stochastic fractals with stationary increments, only the first generator section is required for their specification. The requirement of an integral number of sections may therefore be relaxed — the integral quantity n may be replaced by the real-valued quantity $h > 1$. If the fractal is self-similar for all parametric scaling factors, the scaling ratio $h \to 1$.

By using the symmetry of fractals to expansion and contraction, the range of h may be extended to the positive real numbers, for the behaviour of a fractal with scaling ratio h is identical to one with scaling ratio $1/h$. This yields

$$\vec{a}(\vec{t}_0 + h(\vec{t} - \vec{t}_0)) - \vec{a}(\vec{t}_0) = h^{H}[\vec{a}(\vec{t}_0 + (\vec{t} - \vec{t}_0)) - \vec{a}(\vec{t}_0)] \ ; \quad h > 0$$

as the general description of fractals with stationary increments. Since fractals with scaling ratios of h and $1/h$ are similar, however, the convention is made that $h > 1$.

Fractal functions

Stationarity also leads to a simplification of the graphs of the fractal functions. For the case of the self-similar curve $\vec{\mathbf{a}}(t)$ in a two-dimensional embedding space, the component graphs $\mathbf{a}_x(u)$ and $\mathbf{a}_y(u)$ become a pair of independent equations (cf. eqs (B.9) and (B.10))

$$\mathbf{a}_x(u_0 + h(u - u_0) - \mathbf{a}_x(u_0) = h^H[\mathbf{a}_x(u_0 + (u - u_0)) - \mathbf{a}_x(u_0)] \tag{B.12}$$

$$\mathbf{a}_y(u_0 + h(u - u_0) - \mathbf{a}_y(u_0) = h^H[\mathbf{a}_y(u_0 + (u - u_0)) - \mathbf{a}_y(u_0)] \tag{B.13}$$

since the translational components have been removed, and $\Pi(\Theta_i) = 1$ implies $\theta_i = 0$. Although independent, these equations do not necessarily describe a stationary curve — for example, the relative scales of the two curves could differ by some non-zero finite ratio.

Example: Brownian motion

Brownian motion $\vec{\mathbf{B}}(t)$ is the motion undergone by a small particle as it is randomly bombarded by the atoms and molecules of the surrounding fluid. Given any set of time steps $\{t_i\}$, where

$$t_i = t_0 + i\Delta t \mid i \in Z^+,$$

the probability distribution of the increments $\vec{\mathbf{B}}(t_i) - \vec{\mathbf{B}}(t_{i+1})$ is a stationary, zero-mean Gaussian distribution, with variance of $A|t_i - t_{i+1}| = A|\Delta t|$, where A is some positive number. Rescaling the time steps by an arbitrary factor $h > 0$ leads to a similar distribution, with variance $A|h\Delta t|$. Thus,

$$\vec{\mathbf{B}}(t_i + h\Delta t) - \vec{\mathbf{B}}(t_i) = h^{1/2}[\vec{\mathbf{B}}(t_i + \Delta t) - \vec{\mathbf{B}}(t_i)],$$

for arbitrary t_i and h. Thus, Brownian motion is self-similar, with a similarity parameter $H = 1/2$, and a scaling ratio $h \rightarrow 1$. The relation $D = 1/H$ implies that $\vec{\mathbf{B}}(t)$ is two-dimensional. Indeed, the curve is capable of completely filling regions in the plane.

It is important to realize that the detail of structure existing at all scales in deterministic fractals also exists in stochastic ones. For example, straight lines cannot accurately interpolate $\vec{\mathbf{B}}(t)$ from its values at time steps $\{t_i\}$. Between any two points t_i and t_{i+1}, its stochastic behaviour is completely re-created at a smaller scale.

B.4.2 Fractional Brownian Motion

Brownian motion can be generalized to obtain a class of self-similar stochastic fractals – the fractional Brownian motions $\vec{\mathbf{B}}_H(t)$, defined by the conditions [MaNe68]:

1. $\vec{\mathbf{B}}_H(0) = \vec{b}_0$, where $\vec{b}_0$ is an arbitrary vector
2. $\vec{\mathbf{B}}_H(t) - \vec{\mathbf{B}}_H(0) = \frac{1}{\Gamma(H+\frac{1}{2})}\{\int_{-\infty}^{0}[(t-s)^{H-1/2} - (-s)^{H-1/2}]d\vec{\mathbf{B}}(s)$
   $+ \quad \int_0^t (t-s)^{H-1/2} d\vec{\mathbf{B}}(s)\}$,

where $t > 0$ and $0 < H < 1$. This is a moving average of $\vec{\mathbf{B}}(t)$, weighted by the factor $(t-s)^{H-1/2}$.

The increments $\Delta\vec{\mathbf{B}}_H(t) = \vec{\mathbf{B}}_H(t+\Delta t) - \vec{\mathbf{B}}_H(t)$ are stationary and are characterized by [MaNe68]:

1. $\vec{\mathbf{B}}_H(t+\Delta t) - \vec{\mathbf{B}}_H(t)$ has a Gaussian distribution, since it is the sum of Gaussian random variables

2. The mean increment $\left\langle \vec{\mathbf{B}}_H(t+\Delta t) - \vec{\mathbf{B}}_H(t) \right\rangle = 0$, since it is the sum of zero-mean random variables.

3. The variance $\left\langle |\vec{\mathbf{B}}_H(t+\Delta t) - \vec{\mathbf{B}}_H(t)|^2 \right\rangle \propto |\Delta t|^{2H}$, since the weighting factor has an exponent of $H - 1/2$, while $d\vec{\mathbf{B}}(t)$ has a similarity parameter of $1/2$.

The square root of the variance is proportional to $|\Delta t|^H$ for all Δt. Therefore, for any value h,

$$\vec{\mathbf{B}}_H(t + h\Delta t) - \vec{\mathbf{B}}_H(t) = h^H[\vec{\mathbf{B}}_H(t+\Delta t) - \vec{\mathbf{B}}_H(t)].$$

This shows that $\vec{\mathbf{B}}_H(t)$ is self-similar, with similarity parameter H, and scaling ratio $h \to 1$. Fractional Brownian motion is therefore an appropriate generalization of regular Brownian motion, since $\vec{\mathbf{B}}(t)$ is $\vec{\mathbf{B}}_H(t)$ when $H = 1/2$.

Since $\vec{\mathbf{B}}_H(t)$ is isotropic, all component functions $[\vec{\mathbf{B}}_H(t)]_i$ are of the same form, denoted here simply by $\mathbf{B}_H(x)$. The change in argument shows that these components are functions of an extrinsic parameter. The function $\mathbf{B}_H(x)$ is also a fractal, with dimension $D = 2 - H$ [Mand82].

Adjacent increments of $\mathbf{B}_H(x)$ have a correlation

$$\begin{aligned} &E\{[\vec{\mathbf{B}}_H(x) - \vec{\mathbf{B}}_H(x-\delta)][\vec{\mathbf{B}}_H(x+\delta) - \vec{\mathbf{B}}_H(x)]\} \\ &\qquad = \tfrac{1}{2}\, E\{[\vec{\mathbf{B}}_H(x+\delta) - \vec{\mathbf{B}}_H(-x-\delta)]^2\} - \tfrac{1}{2}\, E\{[\vec{\mathbf{B}}_H(\delta) - \vec{\mathbf{B}}_H(-\delta)]^2\} \\ &\qquad\qquad - E\{[\vec{\mathbf{B}}_H(x+\delta) - \vec{\mathbf{B}}_H(-\delta)]^2\} \\ &\qquad = c_H/2[|2x + 2\delta|^{2H} + (2\delta)^{2H} - 2|x + 2\delta|^2], \end{aligned}$$

where δ is the interval of the increments, and c_H is some positive factor.

The sign of the correlations depends only on the value of H [MaNe68]. For $H = 1/2$ (Brownian motion), the increments are uncorrelated — the motion is completely random, the past having no effect on the future. When $H > 1/2$, however, the correlations between increments are always positive, so that persistence occurs between the values of successive increments. As $H \to 1$, the function becomes smoothly varying. For $H < 1/2$, the correlations are negative, giving rise to antipersistence, with successive increments alternating between positive and negative values.

Although the function $\mathbf{B}_H(x)$ is nonstationary, its power spectrum can be calculated. It has the form [MaNe68]

$$S(k) = V_H|k|^{-1-2H},$$

where V_H is a constant, parametrized by H. The power spectrum therefore has the same self-similarity properties as its corresponding graph.

B.4.3 Fractional Gaussian Noise

Since the graph $\mathbf{B}_H(x)$ is non-differentiable, its derivative $\mathbf{B'}_H(x)$ does not exist in a strict mathematical sense. However, $\mathbf{B'}_H(x)$ can be represented as a random Schwarz distribution: the limit of the derivative of a smoothed $\mathbf{B}_H(x)$ as the amount of smoothing goes to zero. The resultant process is referred to as a fractional Gaussian noise [MaNe68], a stationary stochastic process of infinite variance.

Fractional Gaussian noise is zero-mean, with a correlation function

$$R(\tau) = r_H|\tau|^{2H-2},$$

and a power spectrum

$$S(k) = w_H|k|^{1-2H},$$

where r_H and w_H are positive quantities parametrized by H, and $1/2 < H < 1$ [MaNe68]. The lower bound on H prevents divergence of $S(k)$ as $k \to \infty$. The form of the power spectrum is consistent with that of the power spectrum for fractional Brownian motion, since differentiation in the spatial domain corresponds to multiplication by $+ik$ in the frequency domain.

When the parameter H is formally replaced by $H + 1$, it is seen that fractional Brownian motion and fractional Gaussian noise both belong to a generalized class of functions $\mathbf{B^+}_H(t)$, for which H is a non-zero quantity such that $-1/2 < H < 1$. The fractional Brownian motions are true fractals; the fractional Gaussian noises are not. However, the spectral behaviour of all functions in this class is identical — a self-similarity of form over all possible scales.

Appendix C

Technical Considerations

The following sections are concerned with various technical aspects of the psychophysical experiments described in chapter 4. In particular, they discuss issues involved with the creation and display of the textures used.

C.1 Discretization of Power Spectra

In practice, any physical realization of an image must be quantized, bounded, and discrete. The effect of quantization on perceived texture is small when the number of grey levels involved exceeds 16 [CaHü84]. For the textures used in the experiments of chapter 4, the standard deviation was set to 32 grey levels, and the textures were displayed using 256 grey levels. The effect of quantization is therefore considered negligible upon the perception of the textures displayed.

The issues of boundedness and discretization, however, are more complex, and must be treated in greater detail.

C.1.1 Discrete Fourier Transform

Consider the discrete image f_{xy} with spacing Δ_x and Δ_y in the $x-$ and the $y-$directions respectively, and with bounds $x = \pm T_x/2$, $y = \pm T_y/2$. For convenience, the number of points $N_i = T_i/\Delta_i$ in direction i is assumed to be even and equal to $2n_i$. The finite discrete Fourier transform $\tilde{f}_{kl}$ is then defined as [JeWa68]

$$\tilde{f}_{kl} = \frac{1}{N_x N_y} \sum_{r=-n_x}^{n_x-1} \sum_{s=-n_y}^{n_y-1} f_{r\Delta_x\ s\Delta_y} \exp[-i2\pi(rk/N_x + sl/N_y)].$$

This transformation considers $f_{xy} = f_{r\Delta_x\ s\Delta_y}$ to be a spatially periodic function with period T_i in direction i.

The transform $\tilde{f}_{kl}$ is itself discrete and periodic, with period N_x and N_y in the $k-$ and $l-$ directions respectively. The original image f_{xy} may be recovered by the inverse transformation

$$f_{xy} = \sum_{p=-n_x}^{n_x-1} \sum_{q=-n_y}^{n_y-1} \tilde{f}_{kl}\ \exp[i2\pi(px/N_x\Delta_x + qy/N_y\Delta_y)].$$

Discrete transforms are analogous to continuous transforms in several ways — in particular, the correlation function R_{xy} of a random field is the discrete Fourier transform of its power spectrum S_{kl} [ScSh75,ch4]. Note that R_{xy} described in this way is based on the assumption that the random field is periodic. When the displacement is much less than the size of the image, the error from the true value of the non-periodic R_{xy} is small.

The discrete Fourier transform can be used to produce a random field $\mathbf{f}_{xy}$, since

$$\mathbf{f}_{xy} = \sum_{p=-n_x}^{n_x-1} \sum_{q=-n_y}^{n_y-1} m_{kl}\ \mathbf{z}_{kl} \exp[i2\pi(px/N_x\Delta_x + qy/N_y\Delta_y)],$$

where the $\mathbf{z}_{kl}$ are an array of independent zero-mean, unit-variance complex Gaussian random variables, and m_{kl} is a modulation function (see appendix

A). The power spectrum S_{kl} of the generated field is given by $|m_{kl}|^2$, and its (periodic) correlation function is given by $\tilde{S}_{xy}$.

Owing to considerations of computational efficiency, the finite discrete Fourier transform (DFT) is often implemented as the fast Fourier transform (FFT). Apart from requiring the dimensions of the image to be integral powers of two, the performance of the FFT is no difference from that of the full DFT. Implementations of the FFT are robust, with little sensitivity to numerical error [Knut81,4.3.3].

C.1.2 Self-Similarity and Discrete Images

The theorems on self-similarity developed in chapter 3 are applicable only to continuous random fields. Discretization and boundedness destroy true self-similarity, both at spatial scales less than the spacing Δ_i and greater than the period T_i. These effects are related, since discretization of a continuous function $g(x)$ by a spacing Δ corresponds to the convolution of its transform $\tilde{g}(k)$ by the translated functions $\tilde{g}(k + n/\Delta)$, $n \in Z$ [JeWa68,2.2]. A converse relation also holds for discretization in the spatial-frequency domain. The use of finite discrete representations therefore destroys self-similarity in both the spatial and frequency domains.

However, if the value of $g(\vec{x})$ is always small beyond the range $T_i/2$, and if $\tilde{g}(\vec{k})$ has no significant values outside the Nyquist limit $N_i/2T_i$, the effect of discretization and boundedness on the shape of the functions is small. Therefore, when both $R(x, y)$ and $S(k, l)$ are effectively bandlimited, the random field generated by the finite discrete Fourier transform is effectively self-similar over the corresponding range of intermediate scales.

C.2 Generation of Textures

Instances of a one-dimensional self-similar random field were generated by the Fourier transformation of an array of complex-valued random variables. The following subsections describe some of the technical details involved.

C.2.1 Fourier Transformation

A damping filter D_k was designed to minimize the ranges of both the correlation function R_x and the power spectrum S_k, while simultaneously keeping the shape of S_k effectively self-similar inside the Nyquist limit. It has the form

$$\begin{aligned} D_k &= 0 & ;\quad & 0 \le k < \kappa_1 \\ &= (e^{-(k-\kappa_1)^2/\sigma^2} - 1)^2 (e^{-(k-\kappa_2)^2/\sigma^2} - 1)^2 & ;\quad & \kappa_1 < k < \kappa_2 \\ &= 0 & ;\quad & \kappa_2 < k \\ &= D_{-k} & ;\quad & k < 0. \end{aligned}$$

The parameters κ_1 and κ_2 correspond to the lower and the upper cut-off frequencies of the filter. The parameter σ governs the range of the edge of the filter. Between the bounds of $\kappa_1 + \sigma$ and $\kappa_2 - \sigma$, D_k has a value approaching unity.

Transforms were based on arrays of 8192 points. The value of σ was set to 6 pixels, and κ_1 and κ_2 were chosen so that the half-power points were 16 pixels and 4080 pixels. All transforms were found to exhibit smooth behaviour at the limits of their spatial range, and the values found there were generally several orders of magnitude smaller than those at the origin. The power spectra and correlation functions of the textures used in the experiments were therefore effectively bandlimited, so that R_x and S_k approximated their continuous counterparts $R(x)$ and $S(k)$.

The target textures were produced from an array of 8192 independent Gaussian random variables $\mathbf{z}_k$, each with a mean of zero and a unit variance. The random variables were complex-valued, with random phases produced by generating the real and imaginary components independently. Since the output image was real-valued, $\mathbf{z}_k = -\mathbf{z}^*_{-k}$, so that only half of the array needed to be generated directly. Each of the $\mathbf{z}_k$ was multiplied by a non-negative real-valued function m_k. Fourier transformation of this array then created a random field with a power spectrum $S_k = m_k^2$.

The textures displayed were 256-pixel sections of the output of the FFT. Since these textures subtended a longitudinal angle of 4°, the half-power points of D_k corresponded to 0.12 cyc/deg and 31.9 cyc/deg. These values nicely bracket the limits of human spatial vision [CaRo68], so that the resulting textures were effectively self-similar over all scales relevant to the human visual system.

C.2.2 Random Number Generation

The texture-generation algorithm outlined in the previous section involved the Fourier transformation of a field of independent Gaussian random variables. These quantities were generated via the polar method developed by Box, Muller, and Marsaglia [Knut81,3.4.1], which relies on two independent random variables that are both uniformly distributed between 0 and 1. The random variables that served to this procedure were obtained from the pseudonumber generator random available on the Unix operating system.

Since these random variables were produced by a finite-precision generator of pseudorandom numbers, it was necessary to use a series of statistical tests to check the quality of the numbers generated. The following tests, described in [NaBa66], were adapted slightly to fit the Gaussian case. They were applied to sequences of 8192 successively-generated numbers:

frequency test: this checks the distribution of the values generated. Each test involved 200 sets of 40 numbers each, using 10 levels of quantization.

serial test: this checks the degree of randomness between successive numbers. Each test involved 200 sets of 40 numbers each, using 10 levels of quantization.

run tests: these check the distribution of runs of values above and below the mean, as well as runs of steadily increasing and decreasing values. Each test involved 8192 numbers.

correlation test: this checks the distribution of the correlation product of numbers separated by a given lag. Each test involved 200 sets of 40 numbers each; lags examined ranged from 1 to 25.

Several hundred sets of numbers were tested against the hypothesis that they formed a set of true Gaussian random variables. The sets selected for use were those with the lowest χ^2 values, or equivalently, those with the lowest probability p of being non-random. The sets chosen typically had $p < 0.15$ for the frequency and the serial tests, $p < 0.10$ for each of the various run tests, and $p < 0.20$ for the correlation test at each of the lags examined.

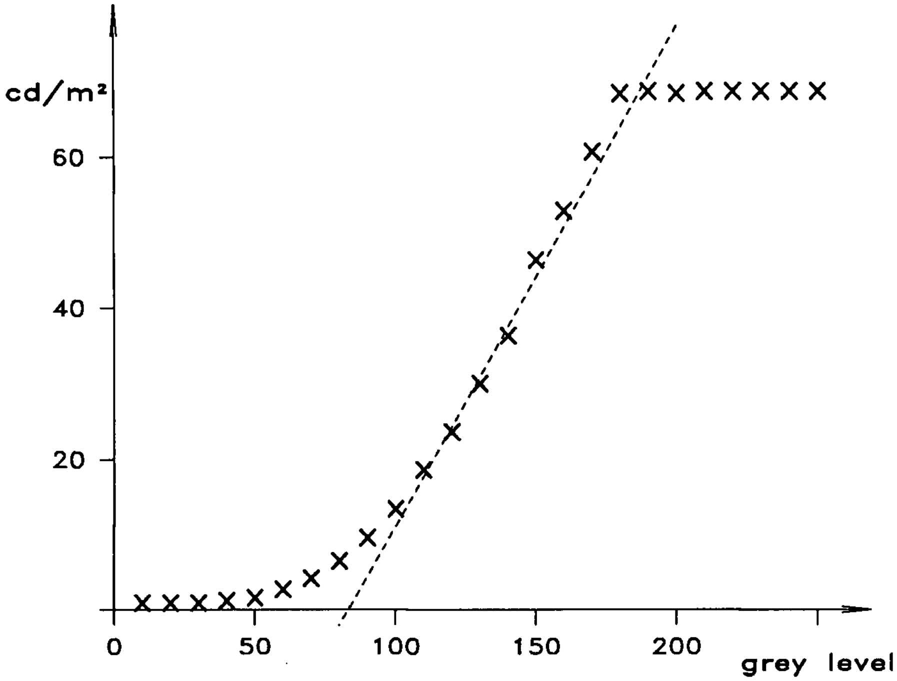


Figure C.1: calibration curve for monitor

C.3 Monitor Calibration

All textures used in the experiments of chapter 4 were displayed on a Hitachi HM-2719B-C-11 monitor. Calibration at the settings used was done using a spot photometer. Luminance of the screen was measured at 25 different grey levels ranging from 10 to 250 in steps of 10. The results are shown in figure C.1.

A least-squares fit was made of the data for grey levels in the range 100−160. This range corresponded to the values within one standard deviation of the mean $g = 128$ used for the displays. The calibration equation obtained was

$$L(g) = 0.66g - 54.9, \tag{C.1}$$

where L is the luminance in cd/m^2, and $g \in \{0, 1, \ldots, 255\}$ is the grey level.

The curve is plotted on the graph of figure C.1.

Since the mean grey level of the textures was set to 128, and the standard deviation to 32, it follows from eq (C.1) that the mean luminance of the display was

$$L_{mean} = 0.66 g_{mean} - 54.9 = 30.0 cd/m^2.$$

The contrast of the displayed textures was therefore

$$C = \frac{\sigma}{L_{mean}} = 0.7.$$

Appendix D

Values of $\vec{V}$ and $\vec{Z}$ for Threshold Textures

Let $H_i(k)$ be a filter that is applied to a set of one-dimensional textures with power spectrum $S(k)$. Among the possible measures on the set of filtered images are the *relative contrast*

$$V_i = \left(\frac{\int H_i^2(k) S(k) dk}{\int S(k) dk} \right)^{1/2}$$

and the *zero-crossing density*

$$Z_i = 2 \left(\frac{\int k^2 H_i^2(k) S(k) dk}{\int H_i^2(k) S(k) dk} \right)^{1/2} .$$

This last relation is taken from [Papo84,11-4]. When a series of m different filters is applied in parallel, the V_i and Z_i can be formed into the composite measures $\vec{V} = (V_1, V_2, \ldots, V_m)$ and $\vec{Z} = (Z_1, Z_2, \ldots, Z_m)$. These quantities may be used as the bases for a multiresolution representation of texture.

This appendix contains the values of $\vec{V}$ and $\vec{Z}$ calculated for the reference classes $H \in \{-1/2, 0, 1/2, 1\}$, $h \to 1$ used in the first set of experiments described in chapter 4. Also calculated are the corresponding values for the textures at the upper and lower discrimination thresholds determined for subject

A. These threshold values form the bases for the predictions made in section 5.2.

All values calculated are based on the formulation of $H_i(k)$ given by Wilson and Gelb [WiGe84], viz.,

$$H_i(k) = \alpha\pi^{1/2}[\sigma_1 \exp\{-(\pi\sigma_1 k)^2\} - \beta_2\sigma_2 \exp\{-(\pi\sigma_2 k)^2\} + \beta\sigma_3 \exp\{-(\pi\sigma_3 k)^2\}],$$

where $\alpha = 1$, and the values of the β_j and σ_j are given in Table 5.1.

$h \to 1$	measure $= \vec{V}$		
channel	$H = -0.500$	$H = -0.720$	$H = -0.310$
A	5.36×10^{-2}	2.94×10^{-2}	8.58×10^{-2}
B	3.57×10^{-2}	2.32×10^{-2}	4.90×10^{-2}
C	2.22×10^{-2}	1.58×10^{-2}	2.79×10^{-2}
D	1.85×10^{-2}	1.43×10^{-2}	2.17×10^{-2}
E	7.62×10^{-3}	6.81×10^{-3}	7.89×10^{-3}
F	5.38×10^{-3}	5.60×10^{-3}	4.89×10^{-3}

$h \to 1$	measure $= \vec{Z}$		
channel	$H = -0.500$	$H = -0.720$	$H = -0.310$
A	2.15	2.33	1.99
B	4.54	4.86	4.27
C	6.56	6.82	6.34
D	9.36	9.72	9.04
E	17.70	18.21	17.25
F	35.27	36.26	34.39

Table D.1: values of $\vec{V}$ and $\vec{Z}$ for $h \to 1$ textures

$h \to 1$	measure = $\vec{V}$		
channel	$H = 0.000$	$H = -0.225$	$H = 0.175$
A	1.48×10^{-1}	1.03×10^{-1}	1.69×10^{-1}
B	6.48×10^{-2}	5.48×10^{-2}	6.29×10^{-2}
C	3.14×10^{-2}	2.99×10^{-2}	2.76×10^{-2}
D	2.19×10^{-2}	2.26×10^{-2}	1.81×10^{-2}
E	6.47×10^{-3}	7.77×10^{-3}	4.74×10^{-3}
F	3.23×10^{-3}	4.54×10^{-3}	2.10×10^{-3}

$h \to 1$	measure = $\vec{Z}$		
channel	$H = 0.000$	$H = -0.225$	$H = 0.175$
A	1.72	1.92	1.57
B	3.83	4.15	3.58
C	5.99	6.24	5.79
D	8.54	8.90	8.26
E	16.50	17.04	16.07
F	32.94	34.00	32.10

Table D.1 (continued)

$h \to 1$	measure = $\vec{V}$		
channel	$H =$ 0.500	$H =$ 0.305	$H =$ 0.675
A	1.68×10^{-1}	1.73×10^{-1}	1.58×10^{-1}
B	4.51×10^{-2}	5.69×10^{-2}	3.51×10^{-2}
C	1.61×10^{-2}	2.31×10^{-2}	1.11×10^{-2}
D	9.41×10^{-3}	1.45×10^{-2}	6.05×10^{-3}
E	1.97×10^{-3}	3.46×10^{-3}	1.12×10^{-3}
F	6.95×10^{-4}	1.40×10^{-4}	3.52×10^{-4}

$h \to 1$	measure = $\vec{Z}$		
channel	$H =$ 0.500	$H =$ 0.305	$H =$ 0.675
A	1.27	1.45	1.12
B	3.12	3.40	2.87
C	5.42	5.64	5.21
D	7.74	8.05	7.46
E	15.26	15.75	14.79
F	30.46	31.46	29.42

Table D.1 (continued)

$h \to 1$	measure = $\vec{V}$		
channel	$H =$ 1.000	$H =$ 0.940	$H =$ 1.095
A	1.39×10^{-1}	1.43×10^{-1}	1.35×10^{-1}
B	2.13×10^{-2}	2.37×10^{-2}	1.84×10^{-2}
C	5.26×10^{-3}	6.17×10^{-3}	4.25×10^{-3}
D	2.49×10^{-3}	3.03×10^{-3}	1.90×10^{-3}
E	3.74×10^{-4}	4.74×10^{-4}	2.74×10^{-4}
F	1.01×10^{-4}	1.29×10^{-4}	7.46×10^{-5}

$h \to 1$	measure = $\vec{Z}$		
channel	$H =$ 1.000	$H =$ 0.940	$H =$ 1.095
A	0.85	0.90	0.78
B	2.41	2.51	2.28
C	4.72	4.84	4.52
D	6.95	7.06	6.79
E	13.62	13.94	13.04
F	25.26	26.71	22.37

Table D.1 (continued)