# Modeling Spatial and Temporal Textures

by

Fang Liu

B.S., Beijing University, China (1984)
M.S., Northeastern University (1989)

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1997

© Massachusetts Institute of Technology 1997. All rights reserved.

Author ..........................................................................................
Program in Media Arts and Sciences
August 8, 1997

Certified by ..........................................................................................
Rosalind W. Picard
NEC Development Professor of Computers and Communications
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by ..........................................................................................
Stephen A. Benton
Chairman, Departmental Committee on Graduate Students
Program in Media Arts and Sciences

**Modeling Spatial and Temporal Textures**
by
Fang Liu

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on August 8, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy


## Abstract

Bridging the gap between the Wold random process decomposition theory and practical texture modeling, this research establishes Wold-based texture modeling as an important method for a wide range of applications that benefit from efficient and effective characterization of textural information.

- A robust and efficient algorithm is developed for spectral 2-D Wold decomposition of homogeneous or near homogeneous random fields.
- A psychophysical study is conducted to show that the Wold component energy of a texture pattern is a good computational measure for the most salient human texture perception dimension of repetitiveness vs. randomness.
- A shift, rotation, and scale invariant Wold-based texture model is presented. This model provides efficient and perceptually sensible features that are robust to many natural texture inhomogeneities.
- For model perspective invariance, a linear system characterization and a decomposition of image perspective transformations are proposed to form a basis for future algorithms to infer image perspective parameters from a single sample of harmonic texture data.
- Based on the Wold texture model, an algorithm is developed for textured image database retrieval. Compared to the state-of-the-art texture models, the new model appears to offer perceptually more satisfying retrieval results while matching or surpassing the best recognition performance of the others.
- A K-means-based image segmentation method is presented to demonstrate the use of Wold-based modeling in characterizing textured regions in natural scene images.
- Applying the principle of Wold decomposition to temporal texture modeling, a robust and efficient algorithm is developed for detecting and segmenting periodic motion. The use of periodicity templates is proposed for characterizing periodicity in space and time.


Thesis Supervisor: Rosalind W. Picard
Title: NEC Development Professor of Computers and Communications
Program in Media Arts and Sciences

# Doctoral Committee

Thesis Advisor . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

Rosalind W. Picard
NEC Development Professor of Computers and Communications
Program in Media Arts and Sciences

Thesis Reader . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

Michele M. Covell
Ph.D., Member of Research Staff
Interval Research Corp.

Thesis Reader . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

Alex P. Pentland
Toshiba Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Thesis Reader . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

Tomaso A. Poggio
Uncas and Helen Whitaker Professor of Vision Sciences and Biophysics
Dept. of Brain and Cognitive Sciences

# Acknowledgments

Grateful and sincere thanks to my advisor, Prof. Rosalind Picard, for her guidance and constant support. Her patience and the freedom she gave me in my research are also very much appreciated. Among other things, I learned a great deal from her about conducting research and pursuing a professional career.

Many thanks to the other members of my doctoral committee: Dr. Michele Covell, Prof. Alex Pentland, and Prof. Tomaso Poggio. Their comments and insightful suggestions have helped to strengthen the work and improve the clarity of the final document. Several long and informative conversations with Prof. Pentland have broadened my view on various aspects of this research.

Thanks also go to Prof. Edward Adelson, who led me into the area of visual perception, and Prof. Aaron Bobick, who has always been willing to help and offered many constructive critiques on my work. Thanks to Joseph Francos, for many early discussions on the 2-D Wold decomposition theory; to Thomas Minka, for benchmarking several texture models; and to Jocelyn Riseberg, for getting me started on human experiments.

The Vision and Modeling Group has been a stimulating and enjoyable place to work. Thanks to all: Ali Azarbayejani, Bea Bailey, Sumit Basu, Dave Becker, Judy Bornstein, Matthew Brand, Lee Campbell, Laureen Chapman, Trevor Darrell, Jim Davis, Irfan Essa, Raul Fernandez, Martin Friedmann, Bill Freeman, Monika Gorkani, Jennifer Healey, Bradley Horowitz, Stephen Intille, Yuri Ivanov, Tony Jebara, Jonathan Klein, Steve Mann, Baback Moghaddam, Kate Mongiat, Chahab Nastar, Nassir Navab, Karen Navarro, Nuria Oliver, Claudio Pinhanez, Kris Popat, Deb Roy, Stan Sclaroff, Alex Sherstinsky, Eero Simoncelli, Pawan Sinha, Flavia Sparacino, Thad Starner, Martin Szummer, Hong Z. Tan, Erik Trimble, Matthew Turk, Joshua Wachman, John Wang, Laurie Ward, Andy Wilson, Chris Wren, and many others.

Special thanks to my family, a perpetual source of support and encouragement, and specially to Steven Rosenthal, for his love, understanding, and patience.

*This research was conducted at the MIT Media Laboratory and was sponsored in part by BT, Interval Research Corp., IBM, and NEC.*

# Chapter 1

# Introduction

## 1.1 Texture

The most obvious property of texture is perhaps its ubiquity. While two-dimensional visual and three-dimensional haptic textures are the most intuitive, "texture" can also be used to characterize audio signals and spatiotemporal events, such as music and motion.

Ubiquitous as it is, a formal definition of texture remains elusive. In the literature, various textual properties are often used to serve as the definition of texture or, more precisely, to constrain the domain of problems. Commonly seen descriptions of textures include coarseness, contrast, directionality, regularity, uniformity, roughness, busyness, *etc.*. Categorically, texture has also been considered as "stuff", as opposed to "things"; the latter of which is usually associated with identifiable objects.

In this dissertation, the term texture refers to signals that exhibit statistically certain degrees of homogeneity and can be regarded as stationary or near stationary one-dimensional (1-D) or two-dimensional (2-D) random processes. Other types of homogeneity exist. An example is self-similar texture. The modeling of such texture has been addressed by others using fractal models [54][73], and will not be considered here.

## 1.2 Texture Modeling

### 1.2.1 Texture Models

Texture models provide computational features that facilitate tasks such as image understanding, representation, and synthesis. Historically, texture models are categorized as statistical, structural, or statistical-structural (hybrid) [36][37][88][93].

#### Statistical Models

The statistical approach focuses on the statistical properties of textures. A texture pattern is characterized either by statistics of image pixel gray scale values or by a stochastic model. Early methods include autocorrelation [49], run length [33], and co-occurrence [38]. In the 1980's, a large volume of literature appeared in the area of texture modeling using Markov-type random field models. The work of Whittle [94], Woods [97], Besag [12], and Kashyap [52] made fundamental contributions to this development. Connections between image pixel value co-occurrence matrices and Markov/Gibbs random field modeling were established in [25]. An excellent review of work in Markov random field modeling can be found in [22]. Markov-type random field models typically are effective for the analysis and synthesis of micro and random looking textures, but not much so for that of larger scale and more structured patterns [21].

### Structural Models

The structural methods represent a texture pattern by its textural primitives and their spatial placement rules. Examples for both texture analysis and synthesis using structural models can be found in [68][89][90]. The main deficiency of the structural methods is that they are incapable of capturing or generating the randomness that natural textures almost always possess.

### Statistical-structural Models

Analogies have been made between the structural/statistical characterization of texture models and the attentive/pre-attentive dichotomy in low-level human vision [95]. However, this division of texture models is based on the functional properties of the models, and natural textures usually contain both structural and statistical components. Texture models capable of representing both structure and randomness have been studied. Garand and Weinman proposed a hybrid model which uses structured low frequency Fourier information as the initial state for a Gibbs random field model to synthesize cloud textures [34]. Picard discussed the parameter selection of this hybrid model and explored the use of an external field to introduce structure to a Gibbs random field [74]. More recently, Francos *et al.* proposed a unified texture model based on the 2-D Wold decomposition of homogeneous random fields [27]. Wold-based texture modeling (Wold-based modeling or Wold model for short) is the focus of this dissertation.

## 1.3 Criteria for Texture Models

Different criteria can be applied when evaluating a texture model, often biased by the particular application in hand. Two of the most common considerations are:

1. The ability to faithfully reconstruct the data. The quality of the reconstruction can be measured in two ways. One is by the pixel-level mean-squared error criteria. The other is by the perceptual resemblance. For example, two pictures of a grass lawn may have large mean-squared pixel difference but look very similar.

2. The model efficiency. Two types of efficiencies are involved. One is in data representation, i.e., the ratio between the model feature size and data size. The other is the level of complexity in model implementation and feature computation. A model can be easy to compute but produce large feature set or vice versa.

Emerging applications such as image and video database retrieval pose new challenges to texture modeling. In those applications, the computer system serves the purpose of saving human users the time and effort of browsing the entire database. It is often expected that the retrieved images resemble the visual properties of a given prototype. For such a system, it is important that the computational image features used for pattern comparison are meaningful to human perception. This motivates the following additional criterion:

3. The model perceptual properties, *i.e.,* the perceptual interpretation of model computational features.

## 1.4 Focus: Wold-based Texture Modeling

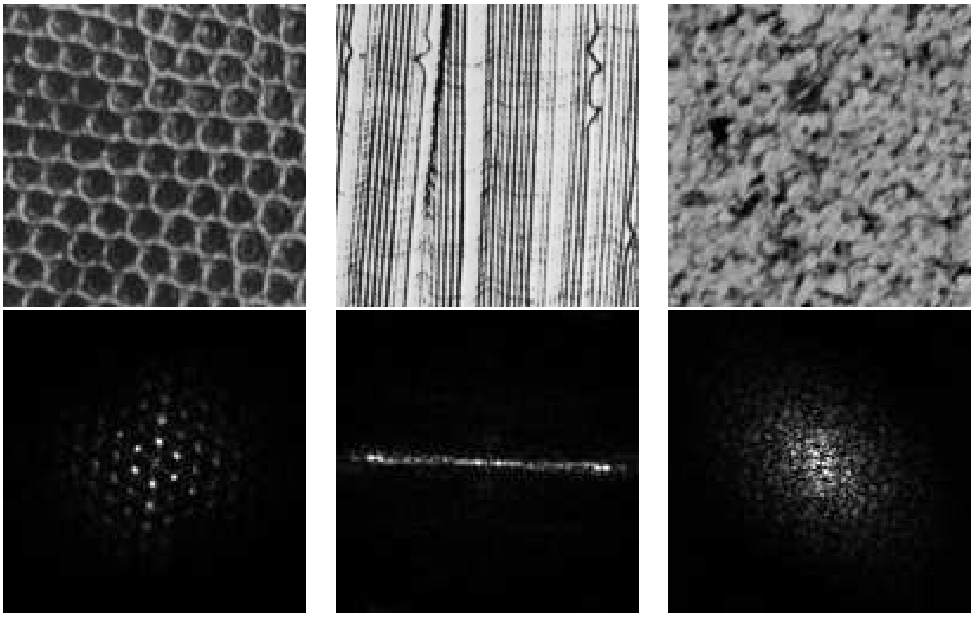
The focus of this dissertation research is on Wold-based Texture Modeling. The mathematical foundation of Wold-based texture modeling is the 2-D Wold decomposition of homogeneous random fields. The 2-D Wold theory allows a textured image to be decomposed into three mutually orthogonal components: harmonic, evanescent, and indeterministic (random). These component images can be characterized separately.

Wold-based texture modeling is capable of satisfying all three criteria discussed above. Francos *et al.* applied Wold-based models to image coding and reconstruction [26][27][85]. It was shown in their work that a handful of model parameters could reconstruct natural textures that are visually indistinguishable from the originals. The most important advantage of Wold-based texture modeling lies with the third criterion. An independent psychophysical study has identified the top three perceptual dimensions of human texture perception as "repetitiveness", "directionality", and "granularity and complexity" [79]. As shown by the Brodatz texture [15] examples in Figure 1-1, the perceptual properties of the Wold components can be described as "periodicity", "directionality", and "randomness", agreeing closely with the findings of the human study.

## 1.5 Goal and Contributions

The goal of this research is to establish Wold-based texture modeling as an important method for a wide range of applications that benefit from efficient and effective characterization of textural information. This goal is achieved by bridging the gap between the Wold theory and practical texture modeling. The main contributions of this work are as follows:

- A robust and efficient spectral 2-D Wold decomposition algorithm is developed for homogeneous or near homogeneous random fields.
- A computational measure for the most salient human texture perception dimension of pattern repetitiveness vs. randomness is proposed and validated by a psychophysical study.
- A shift, rotation, and scale invariant Wold-based texture model is presented. This model provides efficient and perceptually sensible features that are robust to many natural texture inhomogeneities. The superior performance of the model is demonstrated in comparison to state-of-the-art texture models in a textured image database retrieval experiment.
- For model perspective invariance, a linear system characterization of image perspective transformation and its decomposition into affine and chirp transformations are presented. The relation between geometric and spectral descriptions of perspective transformation is formulated to form a basis for future algorithms to infer image perspective parameters from a single sample of harmonic texture data.



(a) (b) (c)

Figure 1-1: Examples of Brodatz textures with different prominent Wold components. Top row: originals. Bottom row: corresponding Fourier magnitude images. (a) D3: Reptile skin, having a prominent harmonic component (spectral peaks supported by point-like regions). (b) D105: Cheesecloth, having a strong evanescent component (spectral peaks supported by line-like regions). (c) D29: Beach sand, having mostly an indeterministic component (relatively smooth spectrum).

---

- Based on the new texture model, an image retrieval algorithm is developed for textured image databases. Compared to other well-known models, the Wold model appears to offer perceptually more satisfying results in the image retrieval experiments while matching or surpassing the best recognition performance of state-of-the-art texture models.
- Applying the principle of Wold decomposition to temporal texture modeling, a robust and efficient algorithm is developed for detecting and segmenting periodic motion. The use of periodicity templates is also proposed for characterizing periodicity in space and time.

## 1.6 Organization

The rest of the dissertation is organized as follows.

**Chapter 2** A concise but comprehensive review of the 2-D Wold decomposition theory for random fields is presented. This mathematical framework is the theoretical foundation of this dissertation. Certain approximations to the theory are also discussed for practical applications.

**Chapter 3** A spectral 2-D Wold decomposition algorithm for homogeneous or near homogeneous random fields is presented. This algorithm detects the Fourier spectral harmonic and evanescent frequencies of a textured image and decomposes the image by extracting these frequency components from the image spectrum.

**Chapter 4** The perceptual property of Wold-based texture modeling is investigated by conducting a human texture ranking experiment. Human subjects and a computer order a set of texture samples along the perceptually most salient dimension of human texture perception. The ranking scores are analyzed to examine the concordance of the human rankings and the correlation between the computer behavior and that of the humans.

**Chapter 5** A Wold-based texture model is constructed. The model emphasizes the perceptually most salient harmonic structures in a texture pattern and is designed for use in large collections of natural textures. Based on this new model, an image retrieval algorithm is developed for textured image databases. The model invariance study leads to the decomposition of perspective transformation. A K-means-based image segmentation method is also presented to demonstrate how the Wold model can be used to characterize textured regions in natural scenes.

**Chapter 6** Based on the principle of 1-D Wold decomposition, an algorithm is developed to model temporal textures for image sequence analysis. This robust and computationally efficient method allows the detection, segmentation, and characterization of periodic motion to be accomplished simultaneously. The use of periodicity templates is also proposed for characterizing periodicity in space and time.

**Chapter 7** Conclusions.

**Chapter 8** Future research directions related to this work are suggested.

# Chapter 2

# Theoretical Background

## 2.1 Introduction

This chapter reviews the mathematical foundation for this dissertation work. The intent is to make a concise but comprehensive presentation of a theoretical framework to which many people have contributed over the years. References are given throughout the chapter for details and proofs.

The original decomposition theory due to Wold applies to the analysis of one-dimensional stationary random processes [96]. It provides a general representation of such processes, as well as an interpretation of the representation in terms of linear prediction. To analyze the structure of a 2-D discrete random field, a two-dimensional linear prediction problem can be similarly formulated.

In the following discussion, it is assumed that the 2-D random field $\{y(m,n)\}$, $(m,n) \in \mathcal{Z}^2$, is real and zero mean. In addition, the second-order moments of $y(m,n)$ are assumed to be finite,

$$\sup_{(m,n)\in\mathcal{Z}^2} E\left[y^2(m,n)\right] < \infty, \tag{2.1}$$

and $E\left[y^2(m,n)\right] > 0$ for at least one $(m,n) \in \mathcal{Z}^2$. Symbol $E\left[\cdot\right]$ denotes the expected value. The objective is to find the minimum-norm linear predictor of $y(m,n)$ as the projection of $y(m,n)$ on the Hilbert space spanned by all the field samples that are in the "past" relative to the "present".

Since there is no natural definition of "past" and "future" in a 2-D plane, different order definitions can lead to different orthogonal decompositions of a 2-D random field. Various choices of the "past", such as symmetric half-plane, non-symmetric half-plane (NSHP), quarter-plane, "vertical" and "horizontal" half-planes, etc., have been used and resulted in two-fold, three-fold and four-fold Wold-like decomposition of 2-D homogeneous random fields [18][19][42][43] [50][51][57]. It is shown by Francos *et al.* [28] that, by considering a countably infinite set of "total-order" and rational non-symmetric half-plane (RNSHP) support (see Section 2.2.1), a corresponding countably-infinite-fold Wold-like decomposition of 2-D homogeneous random field can be obtained and the two-fold, three-fold, and four-fold Wold-like decompositions are special cases of this countably-infinite-fold decomposition. Francos *et al.* also generalized the NSHP-based decomposition to the case of 2-D non-homogeneous random fields [31].

In the following, first the 2-D linear prediction problem is formulated. Then, the Wold-like decomposition of 2-D non-homogeneous random fields is presented, followed by the Wold-like decomposition of homogeneous random fields and the corresponding spectral decomposition. To apply

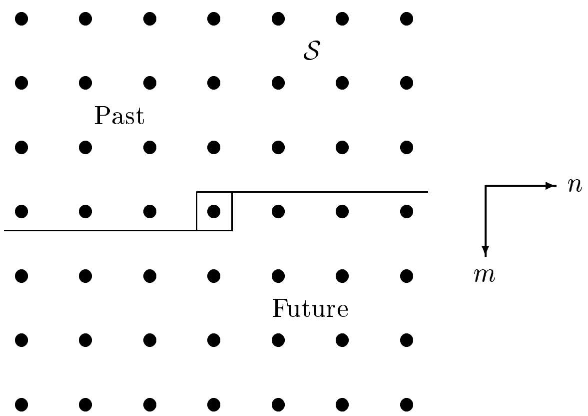



Figure 2-1: Totally ordered non-symmetric half-plane (NSHP) support $\mathcal{S}$ of a 2-D plane. The center location is "present".

---

the 2-D Wold decomposition to practical problems, certain approximations of the theory are made. Finally, examples are given for illustration.

## 2.2 Linear Prediction Formulation

In this section, the problem of 2-D linear prediction of a random field is formulated based on both infinite and finite supports.

### 2.2.1 Total-order and Support in 2-D Plane

**Definition 1** *In the 2-D plane, a* **total-order** *can be defined for the samples of a random field* $\{y(m,n)\}$, $(m,n) \in \mathcal{Z}^2$, *in a raster-scan manner: row after row, from left to right and top to bottom. The order* $\prec$ *is*

$$(i,j) \prec (s,t) \textbf{ iff } (i,j) \in \{(k,l) \mid k = s, l < t\} \cup \{(k,l) \mid k < s, -\infty < l < \infty\} \tag{2.2}$$

*and the order* $\preceq$ *is*

$$(i,j) \preceq (s,t) \textbf{ iff } (i,j) \prec (s,t) \text{ or } (i,j) = (s,t). \tag{2.3}$$

*Based on the total-order definition, a* **totally ordered, non-symmetric half-plane (NSHP) support** $\mathcal{S}$ *can be defined as follows. Given the* $(m,n)$*-th sample as "present", all* $(i,j) \prec (m,n)$ *are in the "past", and the rest are in the "future".*

This NSHP support is illustrated in Figure 2-1, where the "present" is at the center.

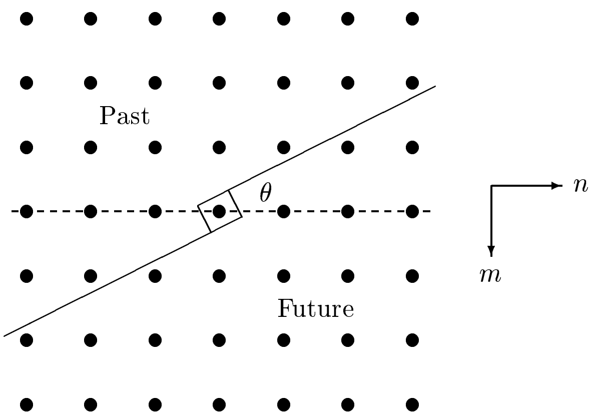



Figure 2-2: Example of a totally ordered rational non-symmetric half-plane (RNSHP) support, rotated from $\mathcal{S}$ by an angle $\theta = tan^{-1}(1/2)$. The center location is "present".

---

Obviously, the total-order and the NSHP support of Definition 1 is not unique on the 2-D lattice. Keeping the structure of the 2-D discrete sampling grid the same, multiple definitions of total-order and NSHP support can be made.

**Definition 2** *Let $\alpha$ and $\beta$ be co-prime integers and $\alpha \neq 0$. A new total-order and NSHP support can be defined on the original 2-D grid by rotating the NSHP total ordering $\mathcal{S}$ of Definition 1 counterclockwise by an angle*

$$\theta = tan^{-1}\left(\frac{\beta}{\alpha}\right)$$

*about the origin of its coordinate system. This new support is called the* **rational non-symmetric half-plane (RNSHP) support** *since its boundary line is of rational slope. Denote the set of all possible total order and RNSHP support defined in this manner by $\mathcal{O}$,*

$$\mathcal{O} = \{o \,|\, o = (\alpha, \beta);\ \alpha, \beta \ are\ co\text{–}prime\ integers\}\,.$$

Note that $\mathcal{O}$ is a countably infinite set.

By Definition 2, the total-order and NSHP support $\mathcal{S}$ of Definition 1 can be denoted as $o = (1, 0)$. Figure 2-2 shows an example of RNSHP total ordering, with $\alpha = 2$ and $\beta = 1$.

In the following, certain definitions and theorems are stated with respect to (w.r.t.) a particular total-order and NSHP support definition $o \in \mathcal{O}$. In places, this total-order and NSHP dependency is noted by the superscript or the subscript $o$.

### 2.2.2 Linear Predictor Based on Infinite Support

Let $\mathcal{H}$ denote the **Hilbert space** formed by the random variables $y(m,n), (m,n) \in \mathcal{Z}^2$, with the inner product of any two random variables $y(m,n)$ and $y(s,t)$ defined as $E\left[y(m,n)y(s,t)\right]$. Then the closed linear manifold spanned by the set $\{y(s,t)\}, (s,t) \preceq (m,n)$ by the total-order $o \in \mathcal{O}$, is a subspace of $\mathcal{H}$:

$$\overset{o}{\mathcal{H}}{}^{y}{}_{(m,n)} = \overline{Sp}\{y(s,t) | (s,t) \preceq (m,n)\} \subset \mathcal{H}. \tag{2.4}$$

Note that this definition implies the **nesting property**:

$$\overset{o}{\mathcal{H}}{}^{y}{}_{(s,t)} \subset \overset{o}{\mathcal{H}}{}^{y}{}_{(m,n)}, \qquad \textit{if } (s,t) \preceq (m,n).$$

**Definition 3** *A predictor of $y(m,n)$ is* **causal** *and of* **continuous support** *w.r.t. the order defined in Definition 1 if it depends on all and only the preceding samples.*

**Definition 4** *The minimum-norm, causal, continuous support, linear predictor of $y(m,n)$ is the projection of $y(m,n)$ on the Hilbert space $\overset{o}{\mathcal{H}}{}^{y}{}_{(m,n-1)}$. Denote the predictor as $\hat{y}(m,n)$, then*

$$\hat{y}(m,n) = \sum_{(0,0)\prec(k,l)} b_{(m,n)}(k,l)\, y(m-k, n-l). \tag{2.5}$$

### 2.2.3 Linear Predictor Based on Finite Support

In practice, only a finite number of samples are available. Therefore, it is necessary to consider a finite support.

**Definition 5** *Define the 2-D* **discontinuous** *and* **finite half-plane support** *as*

$$S_{M,N} = \{(k,l) \mid k = 0,\, 1 \le l \le N\} \cup \{(k,l) \mid 1 \le k \le M, -N \le l \le N\}, \tag{2.6}$$

*where $M$ and $N$ are positive integers.*

Correspondingly, let

$$\overset{o}{\mathcal{H}}{}^{y}{}_{(m,n);S_{M,N}} = \overline{Sp}\{y(m-k, n-l) \,|\, (k,l) \in \{S_{M,N} \cup \{(0,0)\}\}\}.$$

**Definition 6** *The minimum-norm, causal, finite support, linear predictor of $y(m,n)$ is the projection of $y(m,n)$ on the Hilbert space $\overset{o}{\mathcal{H}}{}^{y}{}_{(m,n-1);S_{M,N}}$. Denote the predictor as $\hat{y}_{S_{M,N}}(m,n)$, then*

$$\hat{y}_{S_{M,N}}(m,n) = \sum_{(k,l)\in S_{M,N}} b'_{(m,n)}(k,l)\, y(m-k, n-l). \tag{2.7}$$

It is shown in [31] that the prediction of $y(m,n)$ based on the continuous infinite half-plane support can be approximated by the prediction based on the discontinuous finite half-plane support.

**Theorem 1**

$$\lim_{M\to\infty}\lim_{N\to\infty} E\left[\hat{y}(m,n)-\hat{y}_{S_{M,N}}(m,n)\right]^2 = 0. \tag{2.8}$$

Therefore, although the 2-D Wold-like decomposition theory presented in the following sections is based on infinite 2-D support, it can be applied to random fields defined on finite 2-D discrete grids.

## 2.3 Decomposition of Non-homogeneous Random Fields

The definitions and theorems in this section are stated with respect to the total-order and NSHP support $\mathcal{S}$ $(o=(1,0))$.

**Definition 7** *Let $\hat{y}(m,n)$ be the minimum-norm, causal, continuous support linear predictor of $y(m,n)$. Then the random field $\{u(m,n)=y(m,n)-\hat{y}(m,n)\}$ is called the* **innovation** *of the random field $\{y(m,n)\}$.*

**Definition 8** *A random field $\{y(m,n)\}$ is* **regular** *if $E\left[y(m,n)-\hat{y}(m,n)\right]^2>0$ for at least one $(m,n)\in\mathcal{Z}^2$, i.e., its innovation field $\{u(m,n)\}$ does not vanish.*

**Definition 9** *A random field $\{y(m,n)\}$ is* **deterministic** *if $E\left[(y(m,n)-\hat{y}(m,n))^2\right]=0$ for all $(m,n)\in\mathcal{Z}^2$, i.e., its innovation field $\{u(m,n)\}$ vanishes.*

Note that the deterministic field is a random field. It is deterministic only in the mean square sense.

**Definition 10** *A regular field $\{y(m,n)\}$ is* **purely indeterministic** *if $\overset{o}{\mathcal{H}}{}^{y}{}_{(m,n)}=\overset{o}{\mathcal{H}}{}^{u}{}_{(m,n)}$ for all $(m,n)\in\mathcal{Z}^2$, i.e., $\{u(m,n)\}$ spans the same Hilbert space spanned by $\{y(m,n)\}$.*

The following theorem is the basic theorem of the 2-D Wold-like decomposition of regular random fields. It is a generalization of Cramér's [20] 1-D Wold decomposition of non-stationary random processes by Francos *et al.* [31].

**Theorem 2** *If $\{y(m,n)\}$ is a 2-D regular random field, then it can be represented uniquely by the following orthogonal decomposition:*

$$y(m,n)=v(m,n)+w(m,n), \tag{2.9}$$

*where*

$$w(m,n) = \sum_{(0,0)\preceq(k,l)} a_{(m,n)}(k,l)u(m-k,n-l) \tag{2.10}$$

*and* $E[v(m,n)] = E[u(m,n)] = 0$. *Field* $\{v(m,n)\}$ *is deterministic and field* $\{w(m,n)\}$ *is regular and purely indeterministic. The innovation field* $\{u(m,n)\}$ *is white, i.e.,* $E[u(m,n)u(s,t)] = 0$, *for all* $(m,n) \neq (s,t)$. *Fields* $\{v(m,n)\}$ *and* $\{u(s,t)\}$ *are orthogonal, i.e.,* $E[v(m,n)u(s,t)] = 0$, *for all* $(m,n)$ *and* $(s,t) \in \mathcal{Z}^2$. *Thus fields* $\{v(m,n)\}$ *and* $\{w(s,t)\}$ *are also orthogonal. When* $E[u^2(m-k,n-l)] > 0$, *the coefficients* $a_{(m,n)}(k,l)$ *are given as*

$$a_{(m,n)}(k,l) = \frac{E[y(m,n)u(m-k,n-l)]}{E[u^2(m-k,n-l)]}. \tag{2.11}$$

*When* $E[u^2(m-k,n-l)] = 0$, $a_{(m,n)}(k,l)$ *are arbitrarily set to zero to accomplish the uniqueness of the sequence* $\{a_{(m,n)}(k,l)\}$.

It can be shown that field $\{u(m,n)\}$ is also the innovation field of $\{w(m,n)\}$ [31]. Since field $\{w(m,n)\}$ is purely indeterministic, $\overset{o}{\mathcal{H}^w}_{(m,n)} = \overset{o}{\mathcal{H}^u}_{(m,n)}$. Therefore, the purely indeterministic component $\{w(m,n)\}$ must exist for any regular field. If a random field $\{y(m,n)\}$ is regular and purely indeterministic, it can be represented completely by the white innovation driven moving average (MA) system

$$y(m,n) = \sum_{(0,0)\preceq(k,l)} a_{(m,n)}(k,l)u(m-k,n-l). \tag{2.12}$$

By Theorem 2, the subspace $\overset{o}{\mathcal{H}^y}_{(m,n)}$ has a direct sum representation

$$\overset{o}{\mathcal{H}^y}_{(m,n)} = \overset{o}{\mathcal{H}^v}_{(m,n)} \oplus \overset{o}{\mathcal{H}^u}_{(m,n)}. \tag{2.13}$$

Theorem 2 gives the basic decomposition of a regular random field into its deterministic and purely indeterministic components. Shown next are some properties of the deterministic field $\{v(m,n)\}$.

**Definition 11** *The* **remote past space** $\overset{o}{\mathcal{H}^y}_{(-\infty,-\infty)}$ *w.r.t. a specific total-order and NSHP definition is the intersection of all the Hilbert subspaces spanned by the samples of the regular field* $\{y(m,n)\}$, *i.e.,*

$$\overset{o}{\mathcal{H}^y}_{(-\infty,-\infty)} = \bigcap_{(m,n)\in\mathcal{Z}^2} \overset{o}{\mathcal{H}^y}_{(m,n)} \tag{2.14}$$

Let $\overset{o}{\mathcal{H}^v}_{(m,-\infty)} = \bigcap_{n=-\infty}^{\infty} \overset{o}{\mathcal{H}^v}_{(m,n)}$. Using Theorem 2, it can be shown that the deterministic field $\{v(m,n)\}$ has the property that $\overset{o}{\mathcal{H}^v}_{(m,n)} = \overset{o}{\mathcal{H}^v}_{(m,-\infty)}$ for all $m$ [31]. Define $\overset{o}{\mathcal{H}^v_m} = \overline{Sp}\{v \mid v \in \overset{o}{\mathcal{H}^v}_{(m,-\infty)},\, v \perp \overset{o}{\mathcal{H}^v}_{(m-1,-\infty)}\}$. Then,

$$\overset{o}{\mathcal{H}^v}_{(m,-\infty)} = \overset{o}{\mathcal{H}^v}_{(m-1,-\infty)} \oplus \overset{o}{\mathcal{H}^v_m}.$$

By induction, the Hilbert space spanned by the deterministic field $\{v(m,n)\}$ can be written as the direct sum of the remote past space and the **row-to-row innovations** of the deterministic field w.r.t. the specific total-order and NSHP definition:

**Theorem 3**

$$\overset{o}{\mathcal{H}}{}^{v}{}_{(m,n)} = \overset{o}{\mathcal{H}}{}^{v}{}_{(m,-\infty)} = \overset{o}{\mathcal{H}}{}^{y}{}_{(-\infty,-\infty)} \oplus \bigoplus_{k=-\infty}^{m} \overset{o}{\mathcal{H}}{}^{v}_{k}. \tag{2.15}$$

Theorem 3 implies that the remote past space and the row-to-row innovations of the deterministic field $\{v(m,n)\}$ are mutually orthogonal.

**Definition 12** *A 2-D deterministic random field $\{e(m,n)\}$ is* **evanescent** *w.r.t. a* specific *total-order and NSHP definition if it spans a Hilbert space that is identical to the one spanned by the row-to-row innovations of the deterministic random field at each coordinate $(m,n)$.*

By Theorem 3, under each total-order and NSHP support $o \in \mathcal{O}$, at most one evanescent field can be resolved. This is the one that generates the row-to-row innovations aligned to the row orientation of $o$. The subspaces spanned by all other evanescent components $e_{o'}$, where $o' \in \mathcal{O}$ and $o' \neq o$, are contained in the remote past space $\overset{o}{\mathcal{H}}{}^{y}{}_{(-\infty,-\infty)}$. Therefore, to resolve all the evanescent components of the deterministic field, it is necessary to check the random field against every possible total-order definition in $\mathcal{O}$.

So far, the orthogonal decomposition of a 2-D regular random field has been studied without any assumption on the homogeneity of the field. In general, the regularity and determinism of a 2-D random field are total-order dependent. Therefore, the use of multiple total-order and RNSHP supports can result in a family of orthogonal decompositions of a 2-D random field. It will be shown in the next section that the decomposition of a regular and homogeneous random field is NSHP support invariant. The homogeneity of the random field also makes its spectral decomposition possible.

**Summary:**

*By Theorem 2 and Theorem 3, a 2-D regular random field can be decomposed orthogonally into remote past, evanescent, and purely indeterministic random fields.*

## 2.4 Decomposition of Homogeneous Random Fields

### 2.4.1 Homogeneous Random Fields

**Definition 13** *A random field $\{y(m,n)\}$, $(m,n) \in \mathcal{Z}^2$, is* **homogeneous** *if*

$$E\left[y^2(m,n)\right] < \infty \tag{2.16}$$

*and*

$$r(k,l) = E\left[y(m+k,n+l)y(m,n)\right], \quad (k,l) \in \mathcal{Z}^2 \tag{2.17}$$

*is independent of $m$ and $n$.*

If field $\{y(m,n)\}$ is homogeneous, its innovation field $\{u(m,n)\}$, as well as its deterministic field $\{v(m,n)\}$, its evanescent field $\{e(m,n)\}$, and its purely indeterministic field $\{w(m,n)\}$, are also homogeneous. Furthermore, the variance of field $\{u(m,n)\}$ is a constant for all $(m,n) \in \mathcal{Z}^2$. Denote this constant by $\sigma^2$. If field $\{y(m,n)\}$ is also regular, then $\sigma^2$ is strictly positive and the 2-D Wold decomposition of homogeneous regular random fields is unique by Theorem 2.

### 2.4.2 Spectral Decomposition

For a homogeneous random field, its spectral representation exists in the form of a Fourier-Stieltjes integral. In the following, all spectral functions are defined on the rectangular region

$$\mathcal{K} = \left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right].$$

Let $F_y(\xi,\eta)$ be the **spectral distribution function** of the homogeneous field $\{y(m,n)\}$. Then the corresponding **spectral density function** $f_y(\xi,\eta)$ is the 2-D Lebesgue derivative of $F_y(\xi,\eta)$:

$$f_y(\xi,\eta) = \frac{\partial^2 F_y(\xi,\eta)}{\partial\xi\,\partial\eta}. \tag{2.18}$$

The spectral representation of the field $\{y(m,n)\}$ is

$$y(m,n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi j(m\xi+n\eta)} dZ(\xi,\eta), \tag{2.19}$$

where $j = \sqrt{-1}$, $Z(\xi,\eta)$ is a doubly orthogonal increment process,

$$E\left[dZ(\xi,\eta)dZ^*(\xi',\eta')\right] = 0, \qquad \xi \neq \xi', \eta \neq \eta', \tag{2.20}$$

and is related to $F_y(\xi,\eta)$ by

$$dF_y(\xi,\eta) = E\left[dZ(\xi,\eta)dZ^*(\xi,\eta)\right]. \tag{2.21}$$

The covariance function of $\{y(m,n)\}$ is

$$r_y(k,l) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi j(k\xi+l\eta)} dF_y(\xi,\eta). \tag{2.22}$$

Helson and Lowdenslager proved that a 2-D stationary (homogeneous) process can be orthogonally decomposed into three sub-processes: remote past, evanescent, and innovation [43]. The remote past and evanescent processes are deterministic and the innovation process is purely indeterministic. These are similar results as in Theorem 2, but for the homogeneous case. In the same paper, the decomposition of the spectral distribution function of a homogeneous regular random field is also given:

**Theorem 4** *The spectral distribution function $F_y(\xi,\eta)$ of a homogeneous regular random field $\{y(m,n)\}$ can be uniquely represented as*

$$F_y(\xi,\eta) = F_v(\xi,\eta) + F_w(\xi,\eta), \tag{2.23}$$

*where $F_v(\xi,\eta)$ and $F_w(\xi,\eta)$ are the spectral distribution functions of the deterministic and the purely indeterministic components of $\{y(m,n)\}$, respectively. Function $F_v(\xi,\eta) = F_y^s(\xi,\eta)$ is the singular part of $F_y(\xi,\eta)$ and function $F_w(\xi,\eta)$ is absolutely continuous. Thus, the spectral measure induced by $F_v(\xi,\eta)$ is singular w.r.t. the Lebesgue measure and is concentrated on a Borel set $\mathcal{L}$ with Lebesgue measure zero in $\mathcal{K}$. The derivative of $F_v(\xi,\eta)$ is zero except on the set $\mathcal{L}$. The spectral representations of the deterministic and the purely indeterministic component fields have the form*

$$v(m,n) = \int_{\mathcal{L}} e^{2\pi j(m\xi+n\eta)} dZ(\xi,\eta). \tag{2.24}$$

*and*

$$w(m,n) = \int_{\mathcal{K}\backslash\mathcal{L}} e^{2\pi j(m\xi+n\eta)} dZ(\xi,\eta) \tag{2.25}$$

Therefore, the 2-D Wold-like decomposition of a homogeneous regular random field into deterministic and purely indeterministic components corresponds to: (1) in terms of spectral measures, the decomposition of the spectral measure of the random field into the sum of two mutually singular spectral measures (2.24) and (2.25) that are concentrated on the sets $\mathcal{L}$ and $\mathcal{K}\backslash\mathcal{L}$, respectively; (2) in terms of spectral distributions, the representation of $F_y(\xi,\eta)$ as the sum of its singular and absolutely continuous components (2.23). Clearly, the orthogonal decomposition of a homogeneous random field into deterministic and purely indeterministic components can be achieved by performing a spectral **Lebesgue decomposition** [82], which separates the singular and the absolutely continuous components of the spectral distribution of the random field.

### 2.4.3 Invariability

As mentioned at the end of last section, the set of multiple total-order and RNSHP support $\mathcal{O}$ gives rise to a corresponding family of orthogonal decompositions of a random field since the regularity and determinism of a non-homogeneous random fields are total-order dependent. In [42], Helson and Lowdenslager showed the following:

**Theorem 5** *A 2-D homogeneous random field $\{y(m,n)\}$ is regular* **iff** *$f_y(\xi,\eta) > 0$ almost everywhere in $\mathcal{K}$ (in Lebesgue measure) and*

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \log f_y(\xi,\eta) d\xi d\eta > -\infty. \tag{2.26}$$

*In the regular case, the variance of the innovation field $\{u(m,n)\}$ is given by*

$$\sigma^2 = \exp \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \log f(\xi,\eta) d\xi d\eta. \tag{2.27}$$

The following theorem is based on Theorem 4, Theorem 5, and the fact that the results stated in these theorems are independent of the total-order and NSHP definition.

**Theorem 6** *The regularity property of a homogeneous random field and the decomposition of a regular and homogeneous random field into deterministic and purely indeterministic components are NSHP support invariant. The resulting component fields from the decomposition are unique.*

Therefore, if a homogeneous random field is regular w.r.t. one total-order and NSHP definition, it is regular w.r.t. any other total-order and NSHP definition. The deterministic (purely indeterministic) component obtained w.r.t. one total-order and NSHP definition is identical to the deterministic (purely indeterministic) component obtained w.r.t. any other total-order and NSHP definition.

### 2.4.4 Deterministic Field

Further analysis of the deterministic field is pursued in this subsection. As shown previously, under the definition of the set of total-order and RNSHP support $\mathcal{O}$, multiple evanescent fields can be resolved from the deterministic random field.

**Definition 14** *A random field $\{g(m,n)\}$ is called* **generalized evanescent** *if it can be represented as a linear (possibly infinite) combination of evanescent fields. Each of these evanescent fields generates row-to-row innovations w.r.t. a different total-order and RNSHP support.*

**Definition 15** *A 2-D random field $\{p(m,n)\}$ is* **half-plane deterministic** *if it spans the Hilbert space $\mathcal{H}^y_{(-\infty,-\infty)} = \bigcap_{o\in\mathcal{O}} \overset{o}{\mathcal{H}^y}_{(-\infty,-\infty)}$, where $\overset{o}{\mathcal{H}^y}_{(-\infty,-\infty)}$ is the remote past space of the random field w.r.t. the total-order and NSHP definition $o$.*

Since the half-plane deterministic field spans the intersection of all Hilbert spaces spanned by the random field samples $y(m,n)$, $(m,n) \in \mathcal{Z}^2$, w.r.t. all total-orders and RNSHP definitions, it contains no innovations w.r.t. any total-order and RNSHP definitions.

From Theorem 3,

$$\overset{o}{\mathcal{H}^v}_{(\infty,\infty)} = \overset{o}{\mathcal{H}^y}_{(-\infty,-\infty)} \oplus \bigoplus_{k=-\infty}^{\infty} \overset{o}{\mathcal{H}^v_k}. \tag{2.28}$$

By the uniqueness and NSHP support invariance of the deterministic field and using (2.28),

$$\begin{aligned}
\mathcal{H}^v_{(\infty,\infty)} &= \overset{o}{\mathcal{H}^v}_{(\infty,\infty)} = \bigcap_{o\in\mathcal{O}} \overset{o}{\mathcal{H}^v}_{(\infty,\infty)} \\
&= \bigcap_{o\in\mathcal{O}} \left( \overset{o}{\mathcal{H}^y}_{(-\infty,-\infty)} \oplus \bigoplus_{k=-\infty}^{\infty} \overset{o}{\mathcal{H}^v_k} \right) \\
&= \mathcal{H}^y_{(-\infty,-\infty)} \oplus \bigoplus_{o\in\mathcal{O}} \bigoplus_{k=-\infty}^{\infty} \overset{o}{\mathcal{H}^v_k}.
\end{aligned} \tag{2.29}$$

This leads to the following theorem [28]:

**Theorem 7** *The deterministic component $\{v(m,n)\}$ of a 2-D regular and homogeneous random field $\{y(m,n)\}$ can be decomposed uniquely as*

$$v(m,n) = p(m,n) + g(m,n), \tag{2.30}$$

*where field $\{p(m,n)\}$ is half-plane deterministic and field $\{g(m,n)\}$ is generalized evanescent. Fields $\{p(m,n)\}$ and $\{g(m,n)\}$ are mutually orthogonal, i.e., $p(m,n) \perp g(s,t)$ for all $(m,n),(s,t) \in \mathcal{Z}^2$. Field $\{g(m,n)\}$ consists of a countable number of mutually orthogonal evanescent fields:*

$$g(m,n) = \sum_{o \in \mathcal{O}} e_o(m,n), \tag{2.31}$$

*where $e_o(m,n)$ is the evanescent field of $\{y(m,n)\}$ w.r.t. the total-order and RNSHP definition $o$.*

The corresponding spectral decomposition of the deterministic random field is as follows [26].

**Theorem 8** *Let $F_p(\xi,\eta)$ be the spectral distribution function of the half-plane deterministic component of a homogeneous regular random field $\{y(m,n)\}$ and $F_g(\xi,\eta)$ be the spectral distribution function of the generalized evanescent component of $\{y(m,n)\}$. The spectral distribution function $F_v(\xi,\eta)$ of the deterministic component of $\{y(m,n)\}$ can be uniquely represented as*

$$F_v(\xi,\eta) = F_p(\xi,\eta) + F_g(\xi,\eta) = F_p(\xi,\eta) + \sum_{o \in \mathcal{O}} F_{e_o}(\xi,\eta), \tag{2.32}$$

*where $F_{e_o}(\xi,\eta)$ is the spectral distribution function of the evanescent field w.r.t. the total-order and RNSHP definition $o$. The spectral measures induced by the distribution functions in (2.32) are mutually singular.*

From the definition of the evanescent field and Theorem 8, the spectral measure of the evanescent field w.r.t. the total-order and RNSHP definition $o$ is a linear combination of spectral measures of the form

$$dF_{e_o}(\xi^o,\eta^o) = k(\xi^o)\,d\xi^o\,dF^s(\eta^o), \tag{2.33}$$

where $F^s(\eta^o)$ is a one-dimensional singular spectral distribution function and $k(\xi^o)$ is a one-dimensional spectral density function. Thus, the spectral distribution function of each evanescent field is absolutely continuous in one dimension and singular in the orthogonal dimension.

**Summary:**

*By Theorem 2 and Theorem 7, a homogeneous regular random field $\{y(m,n)\}$ can be represented uniquely by the following orthogonal decomposition:*

$$\begin{aligned} y(m,n) &= p(m,n) + g(m,n) + w(m,n) \\ &= p(m,n) + \sum_{o \in \mathcal{O}} e_o(m,n) + w(m,n). \end{aligned} \tag{2.34}$$

*Correspondingly, the Hilbert space spanned by the homogeneous regular random variables* $\{y(m,n)\}$, $(m,n) \in \mathcal{Z}^2$, *can be decomposed into three mutually orthogonal subspaces: the subspace containing no innovations and spanned by the half-plane deterministic random field* $\{p(m,n)\}$, *the subspace representing the row-to-row innovations and spanned by the generalized evanescent random field* $\{g(m,n)\}$, *and the subspace containing the innovations of the 2-D field* $\{y(m,n)\}$ *and spanned by the purely indeterministic random field* $\{w(m,n)\}$, *which can be described by the white innovation driven MA term in (2.10).*

*By Theorem 4 and Theorem 8, the spectral distribution function of the homogeneous regular random field* $\{y(m,n)\}$ *can be uniquely represented as*

$$\begin{aligned} F_y(\xi,\eta) &= F_p(\xi,\eta) + F_g(\xi,\eta) + F_w(\xi,\eta) \\ &= F_p(\xi,\eta) + \sum_{o \in \mathcal{O}} F_{e_o}(\xi,\eta) + F_w(\xi,\eta). \end{aligned} \tag{2.35}$$

*The orthogonal decomposition of a homogeneous random field into deterministic and purely indeterministic components can be achieved by the spectral Lebesgue decomposition, which separates the singular and the absolutely continuous components of the spectral distribution of the field.*

## 2.5 Approximations

Certain approximations to the 2-D Wold-like decomposition theory can be made for practical applications. (See [26] for more details.)

Since the spectral measure of the deterministic component of a homogeneous regular field is concentrated on a set with Lebesgue measure zero (Theorem 4), the derivative of the spectral distribution function of the deterministic component is zero almost everywhere in $\mathcal{K}$. In practice, the spectral density function of the deterministic component can be considered as being non-zero only on a countable set of points and curves in $\mathcal{K}$.

One frequently observed component of the half-plane deterministic field is the **harmonic random field** $\{h(m,n)\}$. In the "spectral density function", the harmonic field generates 2-D $\delta$-functions supported by discrete points in $\mathcal{K}$. Hence, field $\{h(m,n)\}$ has the form of a countable sum:

$$h(m,n) = \sum_{p=1}^{P} \left\{ A_p \cos 2\pi(m\xi_p + n\eta_p) + B_p \sin 2\pi(m\xi_p + n\eta_p) \right\}, \tag{2.36}$$

where $(\xi_p, \eta_p)$ are the spatial frequencies of the $p$-th harmonic and the $A_p$'s and $B_p$'s are mutually orthogonal random variables with $E[A_p^2] = E[B_p^2] = \sigma_p^2$. The autocorrelation function of $\{h(m,n)\}$ is

$$r_h(m,n) = \sum_{p=1}^{P} \sigma_p^2 \cos 2\pi(m\xi_p + n\eta_p). \tag{2.37}$$

Similarly, the "spectral density function" of each evanescent component $f_{e_{(\alpha,\beta)}}(\xi,\eta)$ can be considered as containing 1-D $\delta$-functions that are supported by lines of angle $\theta = tan^{-1}(\beta/\alpha)$ to the $\xi$ axis in $\mathcal{K}$. Therefore, these "spectral density functions" are continuous along their support lines and singular in the orthogonal dimensions, *i.e.*,

$$f_{e_o}(\xi^o, \eta^o) = k(\xi^o) \sum_{i=1}^{I^o} \gamma_i^{o2} \left[\delta(\eta^o - \eta_i^o) + \delta(\eta^o + \eta_i^o)\right]. \tag{2.38}$$

Hence, the evanescent field $e_o(m^o, n^o)$ has the form

$$e_o(m^o, n^o) = s(m^o) \sum_{i=1}^{I^o} \left(C_i^o \cos 2\pi n^o \eta_i^o + D_i^o \sin 2\pi n^o \eta_i^o\right), \tag{2.39}$$

where $\{s(m^o)\}$ is a purely indeterministic 1-D process with spectral density $2k(\xi^o)$ and the $C_i^o$'s and $D_i^o$'s are mutually orthogonal random variables with $E[C_i^{o2}] = E[D_i^{o2}] = \gamma_i^{o2}$. The generalized evanescent field and its spectral density function have the form of a countable sum of representations in the form of (2.39) and (2.38) respectively.

As shown in (2.10), the purely indeterministic component of a regular random field has a white noise driven MA representation. In practice, one may want to explore the possibility of using other models. One such model is the 2-D autoregressive (AR) model:

$$w(m,n) = -\sum_{(0,0)\prec(k,l)} b(k,l)w(m-k, n-l) + u(m,n). \tag{2.40}$$

The validity of a MA to AR inversion can be determined by either examining the invertibility of the MA representation itself or testing whether the spectral density function $f_w(\xi, \eta)$ possesses certain properties. One sufficient condition for the existence of an AR representation is that $f_w(\xi, \eta)$ is strictly positive on and analytic in some neighborhood of the unit bicircle [24][94]. In practice, this condition is usually satisfied and an AR representation of the purely indeterministic field can be found [26].

In the following, the harmonic, the evanescent, and the purely indeterministic components of a random field are referred to as the **Wold components**. Shown in the next section, the spatial patterns of these Wold components appear to be visually repetitive, directional, and random.

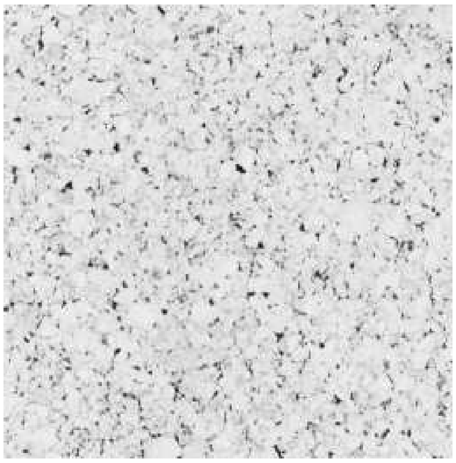
**Summary:**

*A 2-D homogeneous regular random field can be represented as the sum of a harmonic component, a countable number of evanescent components, and a purely indeterministic component. These Wold components are mutually orthogonal.*

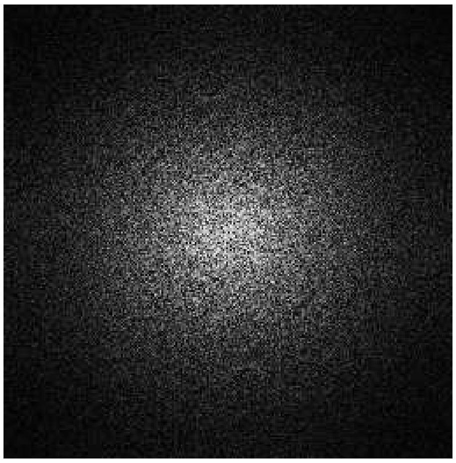
*The "spectral density function" of the harmonic (evanescent) component has the form of 2-D (1-D) $\delta$-functions supported by points (lines) in the 2-D spectral domain. The spectral distribution function of the purely indeterministic component is absolutely continuous. In practice, an AR representation of the purely indeterministic component usually exists.*

## 2.6 Examples of Wold Components

The Wold components of three natural textured images are shown here. Figure 2-3 shows the Brodatz texture Pressed Cork (D32). Since this pattern is mainly indeterministic, its Fourier magnitude image has no large peak. The sweater texture in Figure 2-4 has a prominent harmonic



(a) Original



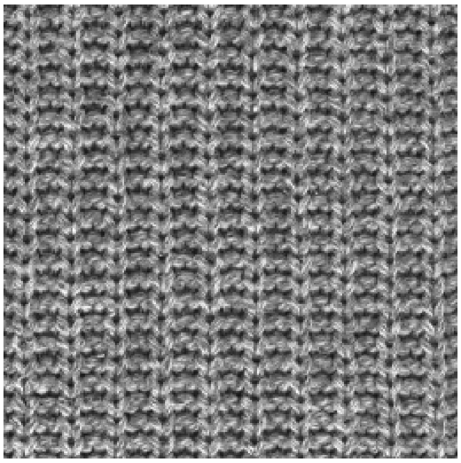
(b) Fourier Magnitudes

Figure 2-3: Wold components of Brodatz texture D32: Pressed Cork. (a) Original. (b) Fourier magnitude image of (a). Since this pattern is mainly indeterministic, its Fourier magnitude image has no large peak.

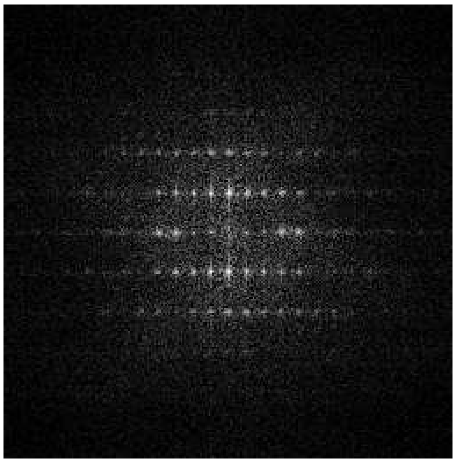
---

component, which appears in its Fourier magnitude image as large peaks with point-like support. Figure 2-5 shows another Brodatz texture, Oriental Straw Cloth (D78). This pattern has strong evanescent components, which appear in its Fourier magnitude image as large values supported by line-like regions. Both Sweater and Oriental Straw Cloth patterns have certain amount of indeterministic component, which appears as the low value "smooth" background in their Fourier magnitude images. The extracted harmonic, evanescent, and indeterministic components of these two images are also shown.

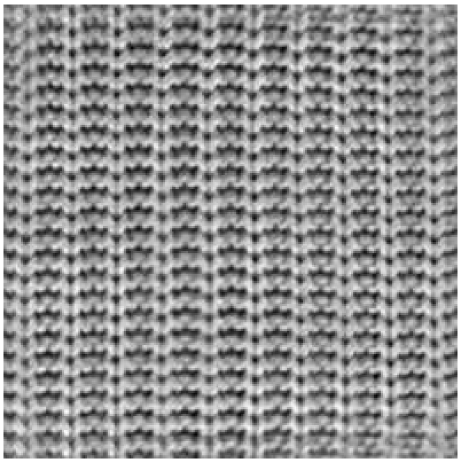
Visually, the spatial images of the indeterministic components in the three examples are random looking, while the harmonic component in the Sweater pattern is very regular in both dimensions and the evanescent components in the Oriental Straw Cloth image appear to be directional.



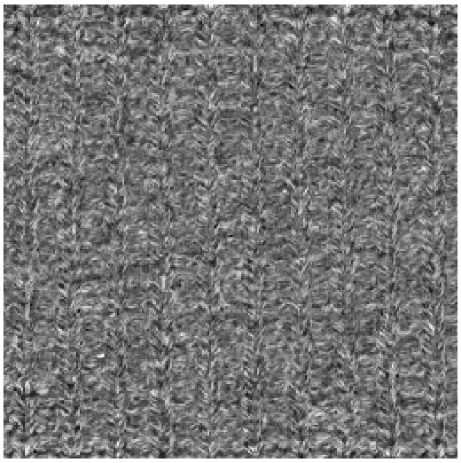
(a) Original



(b) Fourier Magnitudes

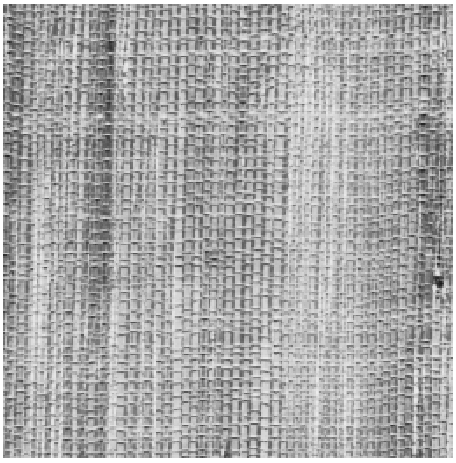


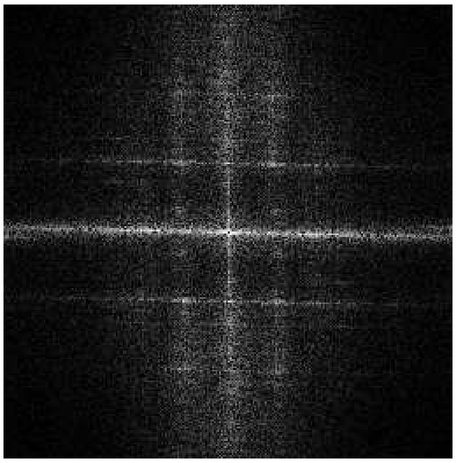
(c) Harmonic Component

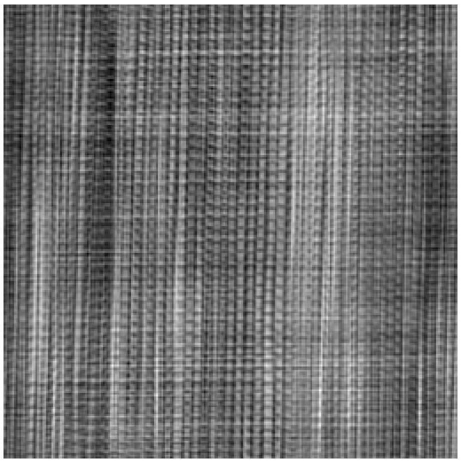


(d) Indeterministic Component

Figure 2-4: Wold components of texture Sweater. (a) Original. (b) Fourier magnitude image of (a). (c) Harmonic component. (d) Indeterministic component. This pattern has a prominent harmonic component, which appears in its Fourier magnitude image as large peaks with point-like support.


(a) Original


(b) Fourier Magnitudes


(c) Evanescent Component

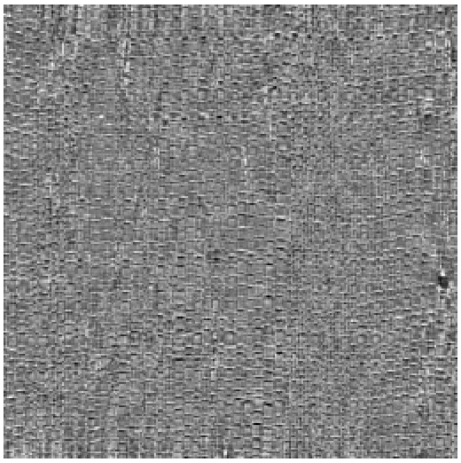

(d) Indeterministic Component

Figure 2-5: Wold components of Brodatz texture D78: Oriental Straw Cloth. (a) Original. (b) Fourier magnitude image of (a). (c) Evanescent component. (d) Indeterministic component. This pattern has a strong evanescent component, which appears in its Fourier magnitude image as large values supported by line-like regions.

# Chapter 3

# Spectral Decomposition

## 3.1 Introduction

The 2-D Wold-like decomposition theory presented in Chapter 2 provides the basis for two types of approach to the decomposition of a homogeneous random field. One approach is the direct parameter estimation from spatial data, and the other is the spectral decomposition based on the Lebesgue decomposition of the singular and continuous spectral components of the random field.

The effectiveness of the algorithms must be gauged in the context of their applications. For image coding, it is important that the estimated decomposition parameters provide enough information about the original image so that the reconstructed image resembles the original at least visually. For image similarity comparison, the discriminatory power of the extracted image features is more valuable than their reconstructive ability, and the form of the features should facilitate the distance computation in the feature space. In applications such as image database retrieval, automation and fast processing can be critical, especially when users introduce new images as query prototypes and the features need to be computed on the fly.

In the following, the existing 2-D Wold-based decomposition algorithms seen in the literature are first summarized and discussed. Then a new robust and computationally efficient spectral decomposition algorithm is presented.

## 3.2 Previous Work

Francos *et al.* proposed two 2-D Wold decomposition methods. One is a maximum-likelihood (ML) direct parameter estimation procedure and the other a periodogram thresholding scheme.

### 3.2.1 Direct Parameter Estimation

In [30][29], a conditional maximum-likelihood direct parameter estimation procedure was devised based on the assumption that the purely indeterministic component is a real-valued, Gaussian distributed, AR random field whose model is given by Equation (2.40) with $(k, l) \in \mathcal{S}_{M,N} \backslash \{(0,0)\}$.

The estimation problem is one of simultaneously estimating all the decomposition parameters from a finite number of samples taken from a single observation of the random field. Staring from the ML formulation, it was shown that the original ML problem can be transformed into a nonlinear problem of minimizing a new objective function over the deterministic component spectral support parameters, which include all the harmonic frequencies $\{\xi_p, \eta_p\}_{p=1}^{P}$, all the total orderings $(\alpha, \beta) \in \mathcal{O}$ that correspond to the evanescent components, and the frequencies $\{\eta_i^{(\alpha,\beta)}\}_{i=1}^{I(\alpha,\beta)}$ for each $(\alpha, \beta)$. The amplitudes of the harmonic components and the AR parameters of the indeterministic components are then computed by solving a linear least-squares problem. The parameters of the modulating 1-D purely indeterministic processes of the evanescent components are estimated by solving a 1-D two-channel ARMA problem.

Difficult to solve analytically, the nonlinear minimization problem was dealt with numerically. To avoid an exhaustive search in the parameter space, a two-stage procedure was used. In the first stage, a parametric Fourier spectral estimation is conducted by fitting a high-order linear prediction model to the observed data and then computing the magnitude of the predictor transfer function inverse. Isolated peaks, as well as peaks that form continuous lines in the magnitude function, are identified as the candidates of the deterministic component spectral support parameters. In the second stage, a conjugate gradient procedure is used to refine the candidates. This procedure does not guarantee convergence to the global minimum unless the initial estimates are sufficiently close to the true values. It is reported in [30] that this iterative procedure can be computationally expensive, especially when the energy in the spectral peaks are not very high comparing to that in the neighboring Fourier frequencies. Unfortunately, this situation often arises in nature textures.

The effectiveness of the ML algorithm was demonstrated in [29] and [30] on two synthesized and six natural textures. The main advantage of this method is that it provides parametric descriptions of all Wold components in a random field. The main disadvantage is its computational cost. The entire procedure involves a high-order prediction model fitting, ARMA fittings, and gradient-based search.

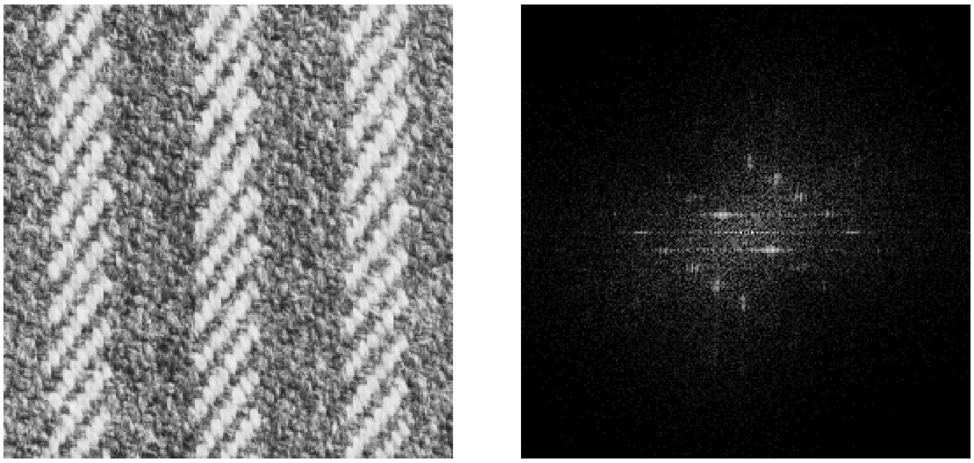
### 3.2.2 Periodogram Thresholding

The decomposition procedure proposed in [27] uses image periodogram thresholding to identify large Fourier spectral values as the harmonic and the evanescent components. The initial threshold is set to be the maximal value of the periodogram. Then the threshold is gradually lowered to qualify more frequency components as spectral peaks until "additional detected peaks are too wide to be considered as the contribution of harmonic components". The evanescent frequencies are determined by checking if the large spectral peaks are located in nearby frequencies along one dimension while fast decay of the periodogram values are observed in the orthogonal dimension.

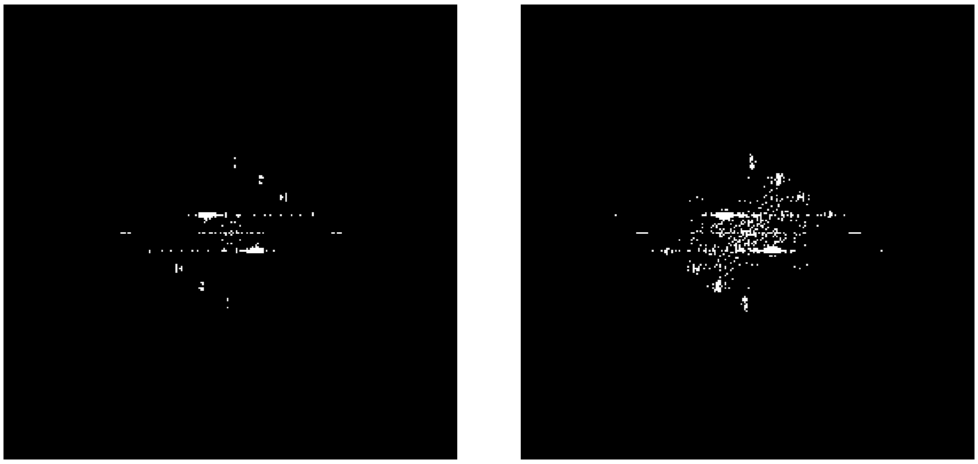
After the removal of the deterministic component, the remaining purely indeterministic component is modeled by an AR model using a 2-D Levinson-type algorithm.

In the periodograms of natural texture images, the support region of each harmonic peak is usually not a point, but a small spread from the central frequency. There are essentially two issues in spectral Wold decomposition. One is to detect the spectral peaks; the other is to determine the peak support regions. The procedure in [27] resolves these two issues by using global thresholding of the periodogram values. However, there are cases in which this method fails. An example is shown in Figure 3-1. The pattern is Brodatz texture D11, Homespun Woolen Cloth, which has high frequency spectral peaks that are only locally large in value. Global spectral thresholding gives either poor segmentation of the peak support regions as in Figure 3-1 (c) or inaccurate peak identifications as in Figure 3-1 (d). In natural textures, this type of spectra abounds.

The main advantage of the periodogram thresholding approach is its computational simplicity. However, as shown above, this algorithm has serious limitations.



(a) Original (b) Fourier Magnitudes



(c) Use High Threshold (d) Use Low Threshold

Figure 3-1: Harmonic peak identification on Brodatz texture D11, Homespun Woolen Cloth, using global thresholding. (a) Original. (b) Fourier magnitude image of (a). (c) A high threshold results in poor segmentation of the support regions. Note that this threshold value is already not high enough since some low frequency random peaks are picked up. (d) A low threshold for better peak support gives inaccurate peak identification.

## 3.3 A New Spectral Decomposition Algorithm

### 3.3.1 Overview

The objective here is to develop a robust practical algorithm to decompose the deterministic and the indeterministic components of a homogeneous regular random field. The new algorithm takes a spectral decomposition approach, which is based on the principle of Lebesgue decomposition. The focus is to detect and extract the spectral singularities, which appear in Fourier spectra as peaks supported by point-like and line-like regions. As mentioned before, two issues are essential: one is to locate the singularities, and the other is to determine their support regions.

The algorithm for extracting the deterministic components consists of three main parts: the harmonic peak detection, the evanescent line detection, and the peak support segmentation. After the spectral support of the deterministic component is determined, the component frequencies are separated from the rest, which comprises the indeterministic component. The spatial values of the deterministic and the indeterministic components can be obtained by taking the inverse Fourier transform of the corresponding frequency components. As illustrated by the example in Figure 3-1, both spectral peak detection and peak support determination should be local, as opposed to global, operations.

#### Notations

In the following, the spatial and frequency samples are indexed by $(m, n)$ and $(k, l)$ respectively, where $m$ and $k$ are the row indices and $n$ and $l$ the column indices. Unless specified otherwise, the samples are defined on the 2-D region

$$\mathcal{D} = \{(i,j) \mid 0 \leq i \leq N-1,\ 0 \leq j \leq N-1\}. \tag{3.1}$$

At times, vector notation $\mathbf{f}$ is used to denote frequency index $(k, l)$.

### 3.3.2 Spectral Estimation

The first step of a spectral approach is to compute the spectrum of a random field. There exist a large variety of spectral estimation methods. In general, the ones providing better spectral estimates in terms of frequency resolution and estimation bias and consistency are computationally more expensive. To facilitate applications such as image database retrieval, the signal periodogram is used in this algorithm for its computational efficiency.

The basic periodograms can be computed as the squared magnitudes of the signal discrete Fourier transform (DFT). Given an image $y(m,n)$, $(m,n) \in \mathcal{D}$, its DFT and inverse DFT are defined as

$$Y(k,l) = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} y(m,n)\, e^{-j\frac{2\pi}{N}km} e^{-j\frac{2\pi}{N}ln}, & (k,l) \in \mathcal{D} \\ 0, & otherwise \end{cases} \tag{3.2}$$

and

$$y(m,n) = \begin{cases} \dfrac{1}{N^2} \displaystyle\sum_{k=0}^{N-1} \sum_{l=0}^{N-1} Y(k,l)\, e^{j\frac{2\pi}{N}mk} e^{j\frac{2\pi}{N}nl}, & (m,n) \in \mathcal{D} \\ 0, & otherwise \end{cases} \tag{3.3}$$

The periodogram estimate of the image spectrum is

$$\hat{P}_y(k,l) = \frac{1}{N^2}|Y(k,l)|^2, \qquad (k,l) \in \mathcal{D}. \tag{3.4}$$

It is well known that basic periodograms are "noisy". The variance of $\hat{P}_y(k,l)$ is often on the order of the true spectra $P_y(k,l)$, independent of the image size [59]. Typically, some type of smoothing technique is used in the estimation, as in the Bartlett estimator and Welch estimator [45]. However, in the spectral Wold decomposition, an image is decomposed via the inverse Fourier transform of its decomposed spectra. Hence, the spectral decomposition has to be based on the actual DFT values of the original image, so smoothing techniques can not be liberally applied. The noisiness of the periodogram poses a serious challenge to the decomposition algorithm.

Another issue related to the periodogram estimation is the image boundaries. Proper handling of the boundaries is especially important when the image has irregular boundaries or is not quite homogeneous. Circular Gaussian tapering is used in this algorithm when necessary. The Gaussian tapering window is

$$g_t(m,n) = e^{-\frac{(m-N/2)^2+(n-N/2)^2}{2\sigma^2}}, \qquad (m,n) \in \mathcal{D}, \tag{3.5}$$

where the standard deviation $\sigma$ is 0.375, normalized by the image half-size $N/2$. Since a spatial multiplicative tapering corresponds to a spectral circular convolution, the periodogram of the tapered image is a low-pass filtered version of the one without tapering. The smoothing effect of Gaussian tapering is illustrated in Figure 3-2.

### 3.3.3 Spectral Harmonic Peak Detection

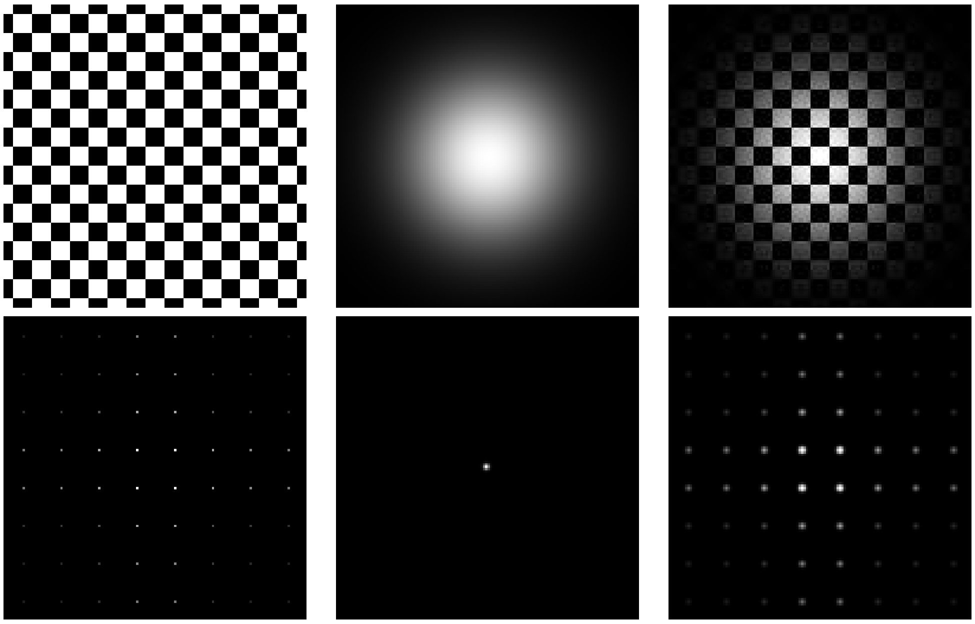
The harmonic peak detection is conducted on the 2-D Fourier magnitude image, which is the scaled square root of the periodogram. The basic idea is to first find the local maxima of the Fourier magnitudes. These local maxima provide the candidate locations of harmonic peaks. A local maximum qualifies to be a harmonic peak only when its frequency is either a fundamental or a harmonic.

#### Frequency Half-plane

Given an image of size $N \times N$, its DFT magnitude image has the same size, with the zero frequency (DC) at sample $(0,0)$. Translate the magnitude image by $(N/2, N/2)$ and wrap the image around at the edges so that the DC is at $\mathbf{f}_o = (N/2, N/2)$ of the frequency plane. (Note that the translation and wrapping are equivalent to swapping the quadrants.) When the image is real, the translated magnitude image is symmetric to $\mathbf{f}_o$. Define the **frequency half-plane** as

$$\mathcal{D}_h = \left\{(k,l) \mid 0 \le k \le \frac{N}{2} - 1,\ 0 \le l \le N-1\right\} \cup \left\{(k,l) \mid k = \frac{N}{2},\ 0 \le l \le \frac{N}{2}\right\}. \tag{3.6}$$

For a symmetric magnitude image, its frequency half-plane contains no symmetric frequency components. In the following discussion, all 2-D Fourier magnitudes and spectra are translated such that the DC components are at $\mathbf{f}_o$. Also, unless specified otherwise, only the frequency half-plane is under consideration.



(a) Original (b) Gaussian Window (c) Tapered

Figure 3-2: Smoothing effect of multiplicative Gaussian tapering. Top row: spatial images. Bottom row: corresponding Fourier magnitude images. (a) Original checkerboard pattern. (b) Gaussian tapering window of $\sigma = 0.375$ ($\sigma$ is normalized by the image half-size $N/2$). (c) Tapered checkerboard pattern: the top image is the multiplication of the other two top images and the bottom image is the circular convolution of the other two bottom images.

---

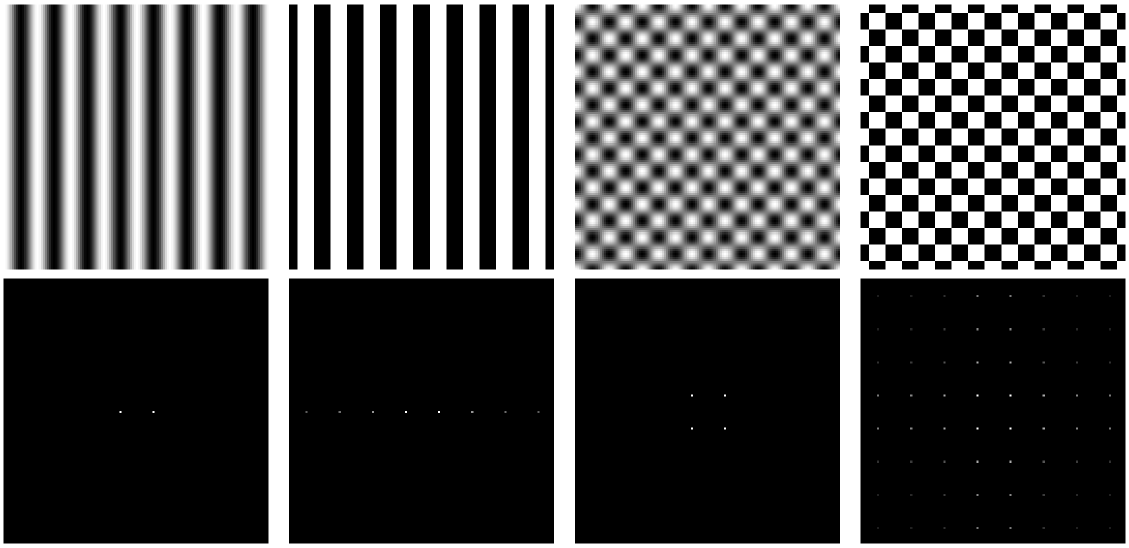
### Local Maxima

Although harmonic peaks usually correspond to large magnitude values, as shown previously, some of them may be large only locally. For this reason, local maxima of the Fourier magnitudes are first detected as candidates of harmonic peaks.

Since the spacing between the harmonic peaks can be small, the estimation window for local maxima detection should not be large. However, the indeterministic "background" of the Fourier magnitude image is usually "lumpy". Under these conditions, many detected local maxima do not correspond to any harmonic peaks and are located rather randomly. Some intrinsic properties of the harmonic random field should be used to discriminate the local maxima for the true harmonic peaks.

### Fundamental-harmonic Relationship

One important property of the harmonic random field is the fundamental-harmonic relationship among its spectral peaks. This is illustrated in Figure 3-3. The top row of the figure contains the



(a) (b) (c) (d)

Figure 3-3: Fundamental-harmonic relationship among harmonic peak frequencies. Top row: spatial images. Bottom row: corresponding Fourier magnitude images. (a) 1-D sine grating containing only one frequency component (considering half of the frequency plane). (b) Black and white stripes have one fundamental frequency, which is the same as the single frequency in (a), and three harmonic frequencies. (c) 2-D sine grating with two frequencies. (d) Checkerboard pattern, whose fundamentals are the same as the frequencies in (c), has a series of harmonics associated with its fundamentals.

---

spatial images and the bottom row the corresponding Fourier magnitudes. The 1-D sine grating pattern in Figure 3-3 (a) contains only one frequency component in $\mathcal{D}_h$. The black and white stripe pattern in (b) have one fundamental frequency, which is the same as the single frequency in (a), and three harmonic frequencies. Similar phenomenon can be observed in (c) and (d). While the 2-D sine grating pattern in (c) has only two frequency components, the checkerboard pattern, whose fundamentals are the same as the frequencies in (c), has a series of harmonics associated with its fundamentals.

Since the edges in a natural pattern usually do not have strictly sinusoidal profiles, one can expect to find harmonics associated with the fundamentals in the spectra of harmonic fields. Therefore, the fundamental-harmonic relationship can be used to identify the true harmonic peaks among the detected local maxima of the Fourier magnitude images.

For harmonic peak detection, the fundamental and the harmonic frequencies are defined as follows.

**Fundamental**

A Fundamental is a frequency that can be used to linearly express the frequencies of some other local maxima of the Fourier magnitudes.

**Harmonic**

A harmonic is a frequency that can be represented as a linear combination of some fundamentals.

Hence, a harmonic frequency $\mathbf{f}^h \in \mathcal{D}_h$ can be expressed as

$$\mathbf{f}^h = \mathbf{f}_o + \sum_{i=1}^{I_h} \alpha_i \left( \mathbf{f}_i^f - \mathbf{f}_o \right), \qquad \alpha_i \in \mathcal{Z}, \tag{3.7}$$

where $\mathbf{f}_i^f \in \mathcal{D}_h$ is the $i$-th contributing fundamental of the harmonic $\mathbf{f}^h$.

### Effect of Frequency Sampling

The discrete Fourier spectrum of an image can be regarded as the sampling of the image continuous spectrum. In general, the sampling points do not fall right on to the very peaks of the continuous spectrum. Therefore, in the DFT plane, the discrete harmonics are in general not the exact multiples of their discrete fundamentals. Due to this sampling effect, certain amount of tolerance should be considered when examining the fundamental-harmonic relations among the local maxima. Furthermore, the fundamental frequency values should be refined to subsample precision since the accuracy of the fundamental frequencies is critical, especially when the harmonics are at the high multiples of the fundamentals. In the algorithm, the fundamental frequencies are refined by using the frequencies of their detected harmonics.

### Implementation

To detect the harmonic peaks, the image is first zero-meaned and Gaussian tapered, and then its DFT magnitudes are computed. The local maxima of the magnitudes are found by searching a small neighborhood of each frequency sample, typically a $5 \times 5$ region. To save computation, frequencies whose magnitude values are below 5% of the entire magnitude range are not considered.

Next, the fundamental-harmonic relationship among the local maxima is examined. Starting from the lowest frequency to the highest, each local maxima is checked first for its harmonicity — if its frequency can be expressed as a linear combination of the existing fundamentals, and then for its fundamentality — if the multiples of its frequency, combined with the multiples of existing fundamentals, coincide with the frequency of another local maximum. To reduce the frequency sampling effect, a tolerance of two sample points in both row and column directions is used in the frequency matching.

When a new fundamental $\mathbf{f}^f$ is found, the algorithm detects the harmonics for which the new fundamental $\mathbf{f}^f$ is *solely* responsible and refines the fundamental frequency value after each harmonic is found. Denote the $j$th harmonic found as

$$\mathbf{f}_j^h = \beta_j \mathbf{f}_{j-1}^f + \Delta\mathbf{f}_j, \tag{3.8}$$

where $\mathbf{f}_{j-1}^f$ is the fundamental frequency value after the $(j-1)$-th refinement, $|\Delta\mathbf{f}_j| < |\mathbf{f}_{j-1}^f|$ and $\beta_j$

is a positive integer. Shown in Appendix A.1, the $j$-th fundamental refinement using $\mathbf{f}_j^h$ is

$$\mathbf{f}_j^f = \frac{\sum_{i=0}^{j} \mathbf{f}_i^h}{\sum_{i=0}^{j} \beta_i} = \mathbf{f}_{j-1}^f + \frac{\Delta \mathbf{f}_j}{\sum_{i=0}^{j} \beta_i}, \tag{3.9}$$

where $\beta_0 = 1$ and $\mathbf{f}_0^h = \mathbf{f}^f$.

All the local maxima whose frequencies are either fundamental or harmonic are qualified as the harmonic peaks.

#### Examples

Figure 3-4 shows two examples of harmonic peak detection. Shown in each example are the original image, the Fourier magnitude image, and the locations of detected harmonic peaks.

### 3.3.4 Spectral Evanescent Line Detection

The spectral evanescent line detection utilizes the Hough transformation [40] of the Fourier magnitude image. Consider a line in a conventional Cartesian $X$-$Y$ coordinate system. Denote the line normal angle by $\phi$ and the line normal length by $d$. Then the line equation is

$$d = x \cos \phi + y \sin \phi. \tag{3.10}$$

A 2-D histogram is built for the normal angle $\phi$ and normal length $d$. The histogram bin size is half of a degree for $\phi$ and unity for $d$. For each pixel $(x, y)$ of the magnitude image, the pixel value is added to every histogram cell whose index $(\phi, d)$ satisfies (3.10). Large local maxima of the histogram correspond to the prominent lines in the image while the bin indices of the maxima provide the line parameters. Note that for accurate evanescent line detection, the large Fourier magnitude values associated with the harmonic peak frequencies should be removed first since these frequencies leave strong signatures in the Hough transform histogram.

An example of the spectral line detection is shown in Figure 3-5. The original is the Brodatz texture D64: Handwoven Oriental Rattan. The Fourier magnitude image in Figure 3-5 (b) has six lines. Image (c) is the Hough transform of (b). Six large local maxima are detected in (c). Lines corresponding to the local maxima are shown in (d).

### 3.3.5 Determining Peak Support Regions

The support regions of the harmonic peaks and evanescent lines are determined by an iterative algorithm. At the beginning of each iteration, a 2-D Gaussian surface is fitted to the Fourier magnitude image, from which all the identified peak support frequencies are removed, to coarsely model the indeterministic component. Based on the local standard deviation (square root of the local variance) of the fitting residual, new support frequencies are identified. This process terminates automatically when the averaged local standard deviation at the near neighbor of the estimated support regions is comparable to that in the ambient areas.

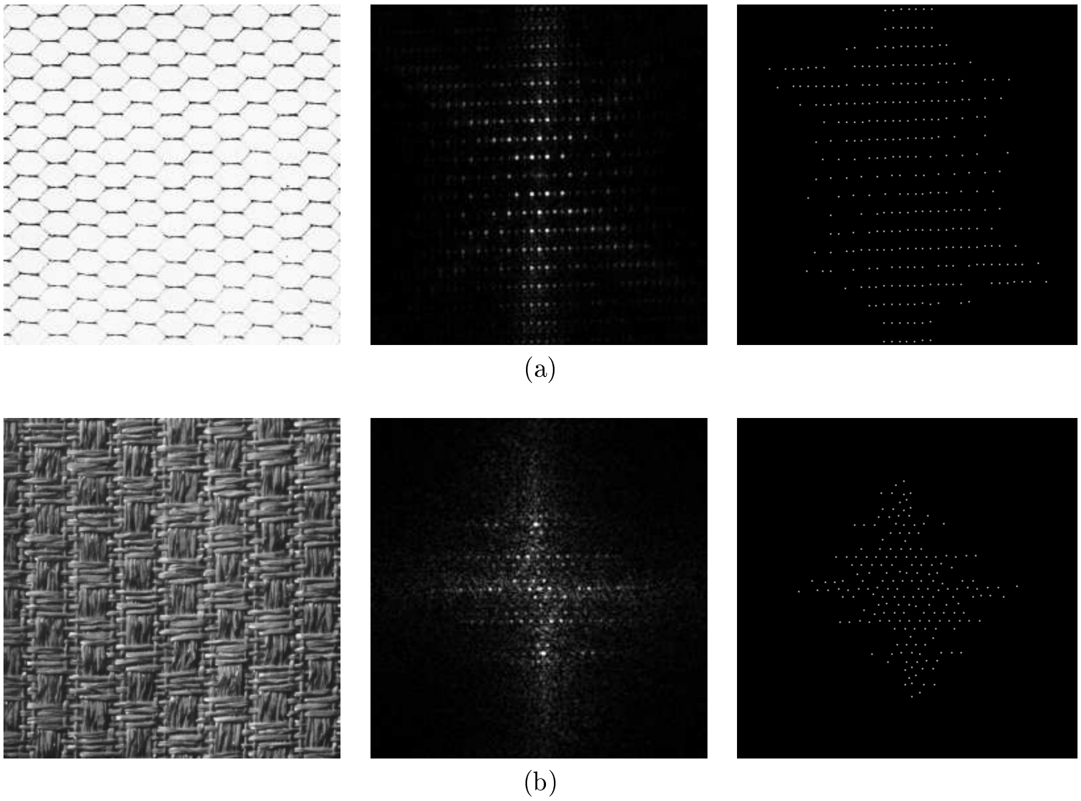


Figure 3-4: Examples of spectral harmonic peak detection. Left column: original image. Middle column: Fourier magnitudes. Right column: detected harmonic peak locations. (a) Brodatz texture D34: Netting. (b) Brodatz texture D82: Oriental Straw Cloth.

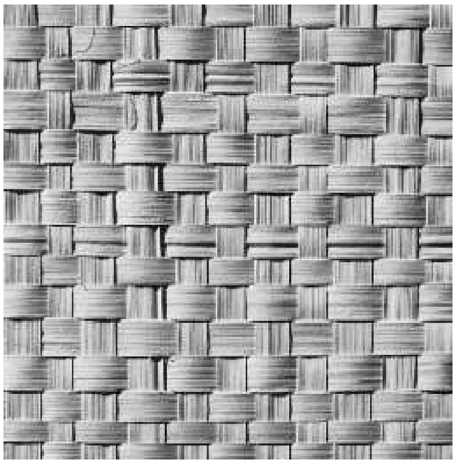
---

### Initializing Mask

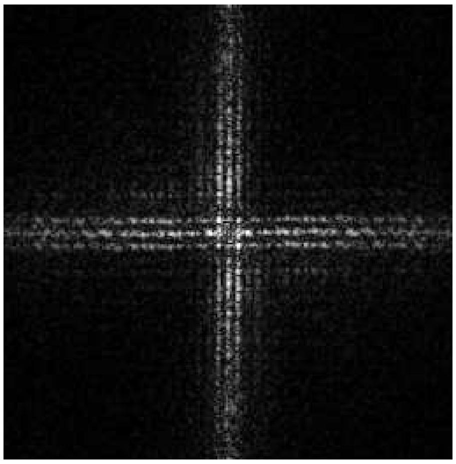
Before starting the iterative procedure, a mask image, which has the size of the Fourier magnitude image, is created and initialized to zero. The mask records the support regions estimated in each iteration. The frequency locations of the estimated harmonic peaks and evanescent lines are marked in the mask as the initial support regions. If a frequency on an evanescent line has a magnitude value less than 5% of the entire magnitude range, that frequency is not marked.

### Gaussian Fitting

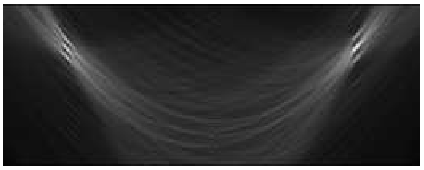
A 2-D Gaussian surface is used to coarsely model the indeterministic component in the Fourier magnitude image. The indeterministic component corresponds to the relatively smooth "background" of the Fourier spectra. The purpose of the Gaussian fitting is not to parameterize the indeterministic component, but to facilitate the determination of the harmonic peak and evanescent line support regions. This will become clear in the discussion that follows.



(a) Original



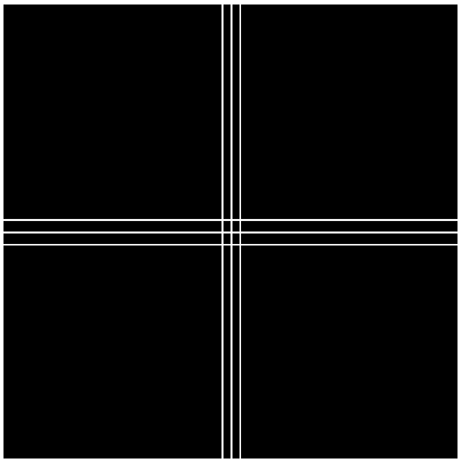
(b) Fourier Magnitudes



(c) Hough Transform

| $d$ | $\phi$ | Hough Value |
|---|---|---|
| 127.0 | 90.0 | 9220.0 |
| 128.0 | 0.0 | 7673.0 |
| 123.0 | 0.0 | 6648.0 |
| 133.0 | 0.0 | 6648.0 |
| 134.0 | 90.0 | 5951.0 |
| 120.0 | 90.0 | 5951.0 |

(d) Hough Local Maxima



(e) Detected Lines

Figure 3-5: Example of spectral evanescent line detection. (a) Brodatz texture D64: Handwoven Oriental Rattan. (b) Fourier magnitudes of (a). (c) Hough transform of the magnitude image. The abscissa is the normal angle $\phi$, and the ordinate the normal length $d$. (d) List of large local maxima in (c). (e) Lines corresponding to the local maxima.

The equation of the Gaussian surface is

$$M_s\, g_s(\mathbf{f}) = M_s\, e^{-\frac{1}{2}(\mathbf{f}-\bar{\mathbf{f}})^T \boldsymbol{\Sigma}_{\mathbf{f}}^{-1}(\mathbf{f}-\bar{\mathbf{f}})}, \tag{3.11}$$

where $T$ denotes the transpose and $\mathbf{f} \in \mathcal{D}$. Vector $\bar{\mathbf{f}}$, matrix $\boldsymbol{\Sigma}_{\mathbf{f}}$, and scaler $M_s$ are the parameters to be estimated from the Fourier magnitude data. The least-squares estimation of $\bar{\mathbf{f}}$ and $\boldsymbol{\Sigma}_{\mathbf{f}}$ is a difficult nonlinear problem. An alternative is to regard the magnitude data as a histogram of Gaussian-distributed random frequency samples and estimate the parameters $\bar{\mathbf{f}}$ and $\boldsymbol{\Sigma}_{\mathbf{f}}$ by using a maximum likelihood estimator. The bin size of the magnitude "histogram" in this case is unity in both dimensions. Denote the integer part of the Fourier magnitude value $|Y(k,l)|$ by $N_{(k,l)}$. The histogram interpretation of the magnitude image is that there are $N_{(k,l)}$ random samples observed at frequency $(k,l)$. The maximum likelihood estimate of the vector $\bar{\mathbf{f}}$ and matrix $\boldsymbol{\Sigma}_{\mathbf{f}}$ in (3.11) are (see Appendix A.2 for derivations)

$$\bar{\mathbf{f}} = (\bar{k}, \bar{l}), \qquad \boldsymbol{\Sigma}_{\mathbf{f}} = \begin{bmatrix} \sigma_{kk}^2 & \sigma_{kl}^2 \\ \sigma_{lk}^2 & \sigma_{ll}^2 \end{bmatrix}, \tag{3.12}$$

where

$$\bar{k} = \frac{1}{N_t} \sum_{(k,l)\in\mathcal{D}} k\, N_{(k,l)} \tag{3.13}$$

$$\bar{l} = \frac{1}{N_t} \sum_{(k,l)\in\mathcal{D}} l\, N_{(k,l)} \tag{3.14}$$

$$\sigma_{kk}^2 = \frac{1}{N_t} \sum_{(k,l)\in\mathcal{D}} k^2\, N_{(k,l)} - \bar{k}^2 \tag{3.15}$$

$$\sigma_{kl}^2 = \sigma_{lk}^2 = \frac{1}{N_t} \sum_{(k,l)\in\mathcal{D}} kl\, N_{(k,l)} - \bar{k}\,\bar{l} \tag{3.16}$$

$$\sigma_{ll}^2 = \frac{1}{N_t} \sum_{(k,l)\in\mathcal{D}} l^2\, N_{(k,l)} - \bar{l}^2 \tag{3.17}$$

$$N_t = \sum_{(k,l)\in\mathcal{D}} N_{(k,l)}. \tag{3.18}$$

Since the Fourier magnitude image is symmetric with respect to $\mathbf{f}_o$, it is expected that $\bar{\mathbf{f}} \approx \mathbf{f}_o$.

After the parameters $\bar{\mathbf{f}}$ and $\boldsymbol{\Sigma}_{\mathbf{f}}$ are estimated, the magnitude $M_s$ of the Gaussian surface can be obtained by solving a least-squares problem (see Appendix A.3 for derivations):

$$M_s = \arg\min_{M_s} \sum_{(k,l)\in\mathcal{D}} \left[\, |Y(k,l)| - M_s\, g_s(k,l)\, \right]^2 \tag{3.19}$$

$$= \frac{\sum_{(k,l)\in\mathcal{D}} g_s(k,l)\, |Y(k,l)|}{\sum_{(k,l)\in\mathcal{D}} g_s^2(k,l)} \tag{3.20}$$

#### Growing Support Regions

The support regions are grown in each iteration from the existing support areas that are marked in the mask. The Gaussian fitting residual, which is the difference between the magnitude image and its fitted Gaussian surface, is first computed. Then the local standard deviations of the residual image are estimated at each frequency using typically a $5 \times 5$ estimation window. The candidates of new support frequencies are found by detecting large positive residual values that are at least as large as the local standard deviation. These "outliers" become part of the peak support if they are adjacent to the existing support frequencies. The new support frequencies are then recorded in the mask and the values of these frequencies in the Fourier magnitude image are replaced by the corresponding values of the Gaussian surface.

#### Terminating Iterations

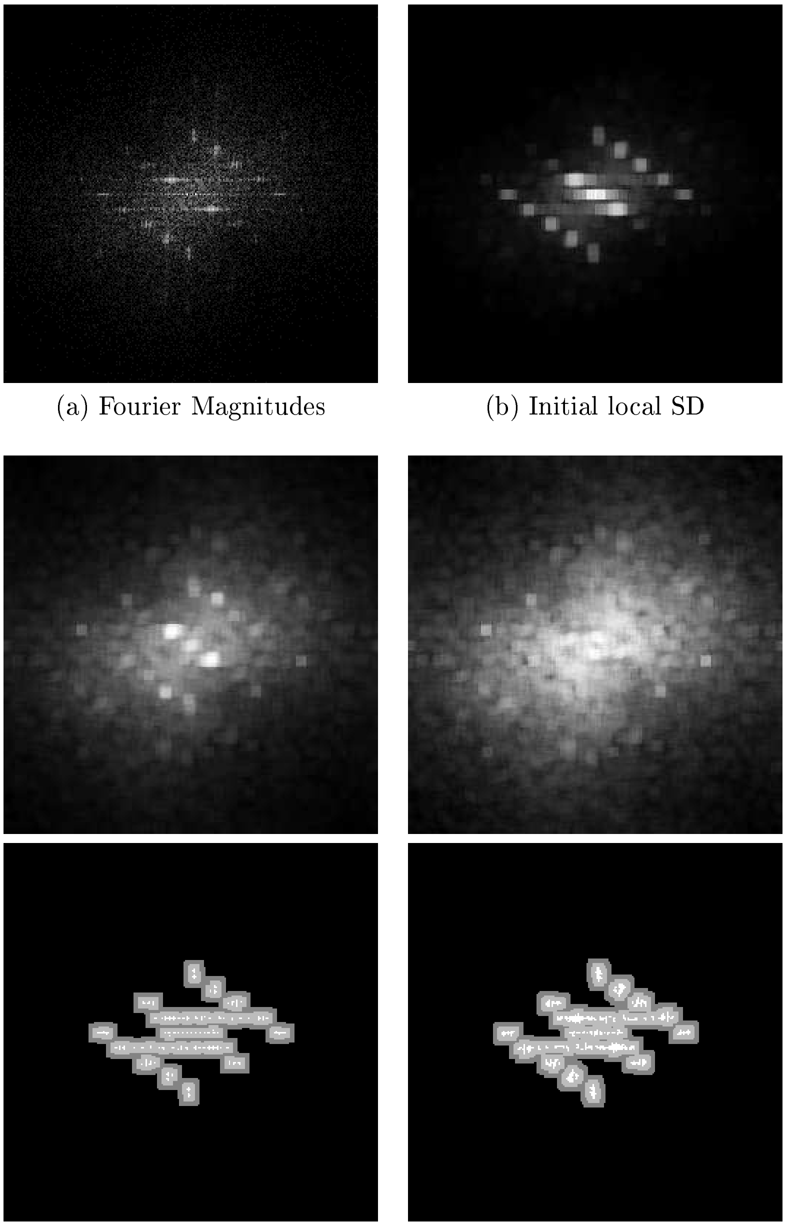
The iterative process should terminate when the large spectral peaks are removed and the remaining Fourier magnitude image becomes "smooth". However, this smoothness is only to the global scale; the Fourier magnitude image is generally very noisy. Hence, any applicable smoothness measure has to be statistical.

The local standard deviation (SD) is used in the algorithm to construct a smoothness measure. At the end of each iteration, the local SD of the Fourier magnitude image is computed. An example is given in Figure 3-6 to show how the local SD of the Fourier magnitude image becomes smoother over the iterations (image (b) and the middle row). At the beginning, the local SD values are large in areas near the spectral peaks. When more frequencies are extracted as peak support, the local SD values adjacent to the peak support regions become similar to those further away from the support regions. In the figure, the local SD images are individually scaled to the display range $[0, 255]$.

The Fourier magnitude smoothness measure used in the decomposition algorithm is constructed as follows. First, two kinds of regions surrounding the spectral peaks are found. One is the **adjacent** area, which includes all frequencies within distance $d_{adj}$ from the estimated peak support frequencies. The other is the **ambient** area, which includes all frequencies within distance $d_{amb}$ from the frequencies in the adjacent area. Examples of the adjacent and the ambient areas are shown in the bottom row of Figure 3-6, with $d_{adj} = d_{amb} = 3$. The **smoothness measure** is defined as the ratio between the averaged local SD in the adjacent areas, $SD_{adj}$, and the averaged local SD in the ambient areas, $SD_{abm}$. Figure 3-7 shows the local SD ratio $SD_{adj}/SD_{abm}$ of 13 Brodatz textures in the first 7 iterations. For all 13 images, the local SD ratio tapers off after first few iterations. In the algorithm, the iterative processing terminates when the change of local SD ratio between iterations drops to below value 0.1.

### 3.3.6 Decomposition

The decomposition of a homogeneous random field is based on the decomposition of its spectral Wold components. When the peak support determination program is terminated, the mask contains the spectral frequencies of the deterministic component. Denote these frequencies as $\mathcal{D}_v$. The DFT



(a) Fourier Magnitudes (b) Initial local SD

(c) Iteration 1 (d) Iteration 4

Figure 3-6: Examples of Fourier magnitude local SD ($5 \times 5$ estimation window) and the adjacent and the ambient areas (3 pixel width) for local SD ratio computation. (a) Fourier magnitudes of Brodatz texture D11 (the original is shown in Figure 3-1 (a)). (b) Initial magnitude local SD. (c)-(d) Shown for iteration 1 and 4: top image: magnitude local SD; bottom image: estimated peak support (white) and corresponding adjacent (light gray) and ambient areas (dark gray). The local SD images are individually scaled to the display range $[0, 255]$.

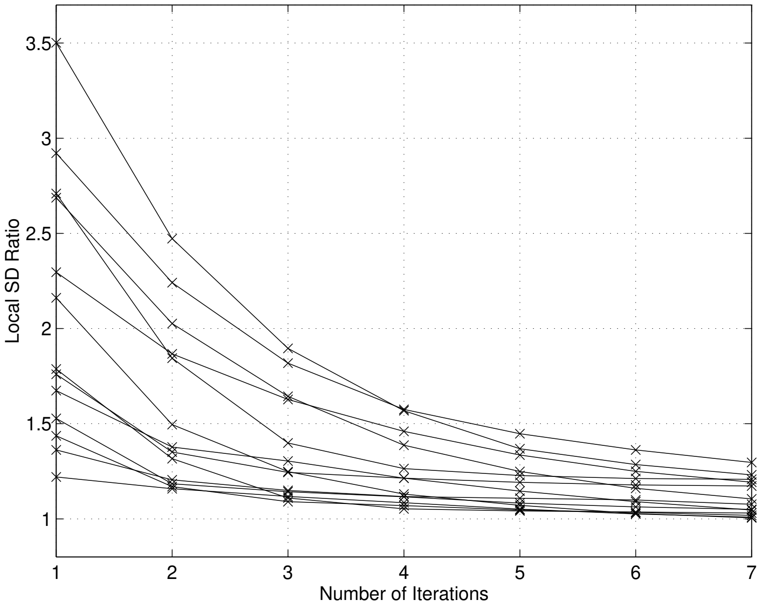



Figure 3-7: Fourier magnitude local SD ratio $SD_{adj}/SD_{abm}$ of 13 Brodatz textures in the first 7 iterations. This ratio is used to measure the global smoothness of the Fourier magnitude images.

---

of the random field is then decomposed into the deterministic component

$$V(k,l) = \begin{cases} Y(k,l), & (k,l) \in \mathcal{D}_v \\ 0, & otherwise \end{cases} \tag{3.21}$$

and the indeterministic component

$$W(k,l) = \begin{cases} Y(k,l), & (k,l) \in \mathcal{D},\ (k,l) \notin \mathcal{D}_v \\ 0, & otherwise \end{cases} \tag{3.22}$$

The deterministic field $v(m,n)$ and the indeterministic field $w(m,n)$ are obtained by computing the inverse DFT of $V(k,l)$ and $W(k,l)$, respectively.

Note that the mask can also record whether an extracted peak frequency belongs to a harmonic peak or an evanescent line. Consequently, the deterministic field can be decomposed further into the harmonic and the generalized evanescent components by using a spectral decomposition procedure similar to the one presented above.

| Iteration | D11 | Sweater | D78 |
|---|---|---|---|
| 1 | 1.410867 | 1.579968 | 1.331581 |
| 2 | 1.183042 | 1.308202 | 1.215476 |
| 3 | 1.140345 | 1.199033 | 1.172281 |
| 4 | 1.121880 | 1.147715 | 1.145736 |

Table 3.1: Fourier magnitude local SD ratio of textures D11, Sweater, and D78 at the end of processing iteration 1 to 4.

### 3.3.7 Examples

Three decomposition examples are shown in Figures 3-8, 3-9, and 3-10. The images in the top rows are the original, the Fourier magnitudes of the original, and the detected harmonic peak central frequencies or evanescent lines. The Fourier magnitudes are computed after the original images are zero-meaned. Shown in row 2 to row 5 of each figure are the mask images, the Fourier magnitude images, the harmonic components, and the indeterministic components obtained at the end of iterations 1 to 4. The Fourier magnitude images are individually scaled to the display range $[0, 255]$.

The Fourier magnitude local SD ratio at the end of each iteration are shown in Table 3.1. For the three examples, the automatic termination of the iterative peak support estimation occurs after iteration 3, 4, and 3, respectively.

## 3.4 Discussion

Spectral peak support estimation is an important issue for a spectral decomposition algorithm since in practice the spectral singularities seldom appear as pure 1-D or 2-D impulses in estimated spectra. The algorithm presented here uses a non-parametric peak support estimation method, which is based on the local variance of image Fourier magnitudes. The main advantages of this algorithm are its robustness and computational efficiency. A parametric approach, such as fitting 2-D Gaussian models to spectral evanescent and harmonic peaks, can also be considered for robust peak support estimation. For fast computation of model parameter estimation, an efficient fitting algorithm, such as the one presented in Section 3.3.5, can be used.

## 3.5 Summary

In this chapter, a spectral 2-D Wold decomposition algorithm for homogeneous or near homogeneous random fields is presented. This algorithm relies on the fundamental-harmonic relationship among spectral peaks to identify the harmonic frequencies, and uses Hough transformation to detect spectral evanescent components. A local variance based procedure is developed to determine the spectral peak support. Comparing to the existing global thresholding scheme and maximum-likelihood parameter estimation, this algorithm is more robust and flexible for the large variety of natural textures, as well as computationally more efficient than the maximum-likelihood method.