Hidden Hierarchical Markov Fields for Image Modeling

by

Ying Liu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Random heterogeneous, scale-dependent structures can be observed from many image sources, especially from remote sensing and scientific imaging. Examples include slices of porous media data showing pores of various sizes, and a remote sensing image including small and large sea-ice blocks. Meanwhile, rather than the images of phenomena themselves, there are many image processing and analysis problems requiring to deal with *discrete-state* fields according to a labeled underlying property, such as mineral porosity extracted from microscope images, or an ice type map estimated from a sea-ice image. In many cases, if discrete-state problems are associated with heterogeneous, scale-dependent spatial structures, we will have to deal with complex discrete state fields. Although scale-dependent image modeling methods are common for continuous-state problems, models for discrete-state cases have not been well studied in the literature. Therefore, a fundamental difficulty will arise which is how to represent such complex discrete-state fields.

Considering the success of hidden field methods in representing heterogenous behaviours and the capability of hierarchical field methods in modeling scale-dependent spatial features, we propose a Hidden Hierarchical Markov Field (HHMF) approach, which combines the idea of hierarchical fields with hidden fields, for dealing with the discrete field modeling challenge. However, to define a general HHMF modeling structure to cover all possible situations is difficult. In this research, we use two image application problems to describe the proposed modeling methods: one for scientific image (porous media image) reconstruction and the other for remote-sensing image synthesis.

For modeling discrete-state fields with a spatially separable complex behaviour, such as porous media images with nonoverlapped heterogeneous pores, we propose a Parallel HHMF model, which can decomposes a complex behaviour into a set of separated, simple behaviours over scale, and then represents each of these with a hierarchical field.

Alternatively, discrete fields with a highly heterogeneous behaviour, such as a sea-ice image with multiple types of ice at various scales, which are not spatially separable but arranged more as a partition tree, leads to the proposed Tree-Structured HHMF model. According to the proposed approach, a complex, multi-label field can be repeatedly partitioned into a set of binary/ternary fields, each of which can be further handled by a hierarchical field.

Acknowledgements

First, I wish to thank my supervisor Paul Fieguth for his support and encouragement throughout my PhD study. His keen insight, expert guidance, and tireless efforts help make the research and writing of this thesis be possible.

I would like to thank the members of my examining committee Prof. Arsen Hajian, Prof. David Clausi, Prof. Zhou Wang, and Prof. Robert Dony for their careful reading of the thesis and suggestions for improvements.

I am also very grateful for the support and assistance of my fellow graduate students in the VIP lab. Their talent, devotion and thoughtfulness made me know what qualities are necessary to make productive, high value researchers. Especially, thanks for Alex Wong and Azadeh Mohebi for our enjoyable collaborations.

Last, but not least, I would like to thank my family for their constant support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis will describe research work regarding discrete-state field models for image processing applications. This introduction presents a general discussion about the motivations and contributions of this thesis work.

1.1 Motivation

Nowadays, large sets of image data are obtained from many imaging sources, such as microscopy, MRI, satellites, and video cameras. To automatically deal with large amounts of data, image models are created, which offer a convenient way to represent, code, and analyze many problems in image processing and computer vision. Their primary concerns are how to define a correct representation for a problem and how to find its optimal solution. Markov Random Fields (MRFs) [26, 64, 113] and their associated algorithms have provided successful approaches to address modeling concerns. The MRF models have been widely applied in solving image and vision problems [64, 113].

The MRF modeling technique itself has experienced significant developments. In the simplest case, a single MRF has been employed to represent global statistical phenomena [25, 27, 92, 118]. However, many scenes, which might be distorted, blurred, or have multiple underlying behaviours, are hardly represented by a single MRF with a simple neighborhood structure. Such modeling problems lead to the proposed classical Hidden Markov Field (HMF) [13, 22, 28, 39]. The basic idea of the HMF is to simplify an image

modeling problem by introducing a hidden field to capture the underlying characteristics of the observed image, so that, based on the hidden field, the image can be decomposed to two simpler random processes: the observation and the underlying prior. In the classical HMF, the prior is assumed to be a MRF.

However, a local, stationary MRF prior model is too strict to represent complex scenes faced by common applications, since in many cases real images do not have global stationarity. In particular, there are many problems in texture analysis, remote sensing and scientific imaging where the underlying discrete fields of observed images possess complex, non-stationary, scale-dependent, spatial structures. For example, the microscopic sample of porous media shown in Fig. 1.1(c) exhibits pores (black) at multiple scales. Moreover we can see that those scale-dependent pores have an uneven spatial distribution. In Fig. 1.1(d), the underlying label field of a SAR sea-ice sample has complicated multi-label states and all states display heterogenous behaviours. Although a single MRF prior model (Fig. 1.3(a)) is good at modeling homogenous behaviours at a single scale [36], they cannot handle such complex phenomena as shown in Fig. 1.1(c,d). MRFs with complicated neighbourhood structures may work, however the costs for training models with high-order neighbourhood systems are very expensive [24, 113]. Therefore, to modeling complex discrete fields, in particular with scale-dependent spatial structures as Fig. 1.1(c,d), brings a challenge to the classical MRF framework.

To deal with this challenge, more flexible methods need to be considered. Instead of using a single MRF, methods have been proposed to employ multiple separated MRFs [10, 57, 59] or joint MRF fields [7, 77, 89] to model complex situations such as mixed densities, non-stationary states and multi-sensor states. For example, we can apply multiple fields to capture the piece-wise, heterogenous behaviours of pores in Fig. 1.1(c), such as one for the large-scale pores (Fig. 1.2(a)), one for the small-scale, high density distributed pores (Fig. 1.2(b)), and one for the small-scale, low-density distributed pores(Fig. 1.2(c)), where each field represents a relative homogeneous behaviour.

However, for more complicated situations where spatial structures are highly dependent and have heterogenous behaviours at multiple scales, as shown Fig. 1.1(d), such discrete fields cannot be well-represented by two or three simple, stationary fields. Moreover, the multiple field methods can only process spatial structures at the finest scale, which makes computational costs very high for asserting large-scale structures. This computational inefficiency limits the application of the method.

At the same time, there are hierarchical MRF field methods [2, 43, 53, 60, 79], which can decompose scale-dependent characteristics at a set of scale spaces. In particular, a Frozen State Hierarchical Field (FSHF) method [20] has been proposed with attractive computational complexity. According to this FSHF method, a large-scale structure will be asserted hierarchically from the coarsest scale, and only those pixels, whose states cannot be determined at coarser scales, need to be processed at the finest scale. As a result, the computational cost is significantly reduced. However, a single hierarchical MRF cannot model heterogenous behaviours existing across the scales. For example, as we hierarchically downsample a discrete field such as Fig. 1.1(c), at many scales the downsampled fields are still heterogenous and cannot be modeled well by a stationary MRF.

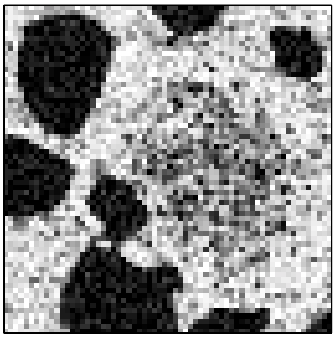
Since the existing MRF based methods have some limitations in modeling complex, heterogeous, scale-dependent behaviors, the work of this thesis is motivated to address this challenge by developing new methods with a more powerful modeling capability.

1.2 Contribution

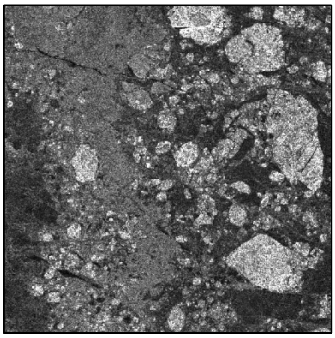
This thesis proposes the Hierarchical Hidden Markov Field (HHMF) method in modeling scale-dependent discrete fields, as well its application to real problems, such as porous media images and remote sensing imagery. Even though multi-scale, continuous-state modeling methods have been studied by many researchers [30, 73, 92], the modeling difficulty for discrete-state fields with structures on more than one or two scales has not been well solved. The proposed models in the literature are for relatively simple situations [2, 20, 80, 82]. In contrast, the HHMF approach provides a powerful and efficient way to handle complex, non-stationary discrete-state fields which possess structures at multiple scales.

In summary, with respect to modeling different scale-dependent situations, the proposed HHMF approach has four inter-related stages:

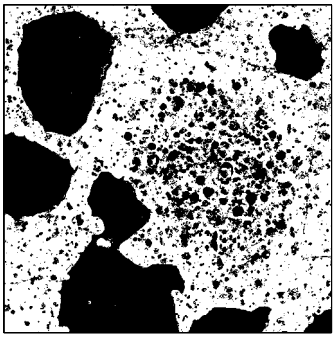
- To model discrete fields with piecewise, heterogenous, scale-dependent spatial structures, we apply the existing idea of multiple separated MRFs (Fig. 1.3(b)) to handle the non-stationarity, where each field is used to capture a simple stationary behaviour. For example, for a discrete field shown in Fig. 1.1(c), we can employ three MRFs to separately represent the behaviours of large-scale pores, high-density small-scale



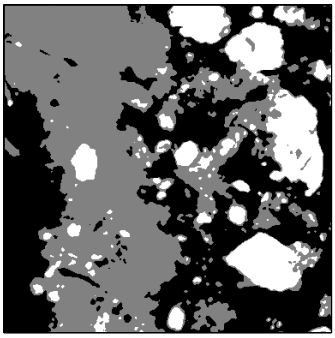
(a) Noisy LR measurement sample



(b) Sea-ice texture sample



(c) True microscopic sample of (a)

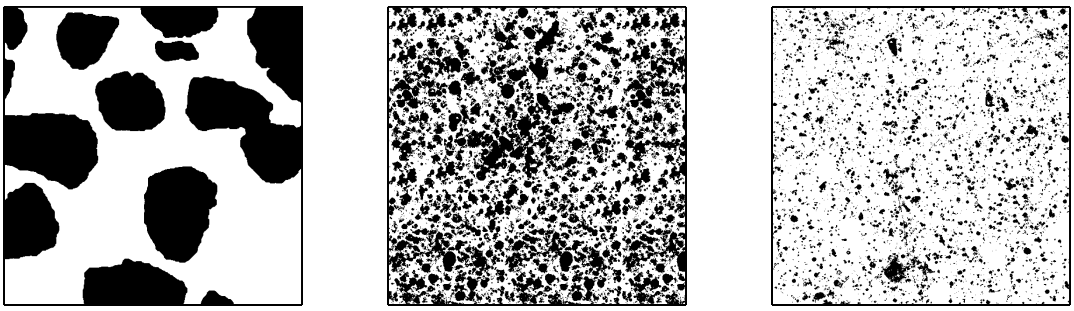


(d) The underlying label field of (b)

Figure 1.1: Both a microscopic sample (a) and a RADARSAT-1 SAR sea-ice texture sample (b) have complicated underlying label maps with multi-scale structures as in (c) and (d) respectively.

pores, and low-density small-scale pores. In doing so, a complex, non-stationary modeling problem is addressed by multiple simple models. However, considering the computational cost, in the thesis we only apply two separated MRFs to solve a reconstruction problem with two scale-dependent behaviours and discuss it in Section 5.2.

- The existing ideas of hierarchical fields (Fig. 1.3(c)) provide a natural way to capture spatial structural features at different scales, especially since the FSHF naturally offers computational efficiency. To efficiently capture the scale-dependent, spatial features of an underlying behaviour, we take the advantage of the efficient modeling



(a) Large-scale pores

(b) Small-scale, high density pores

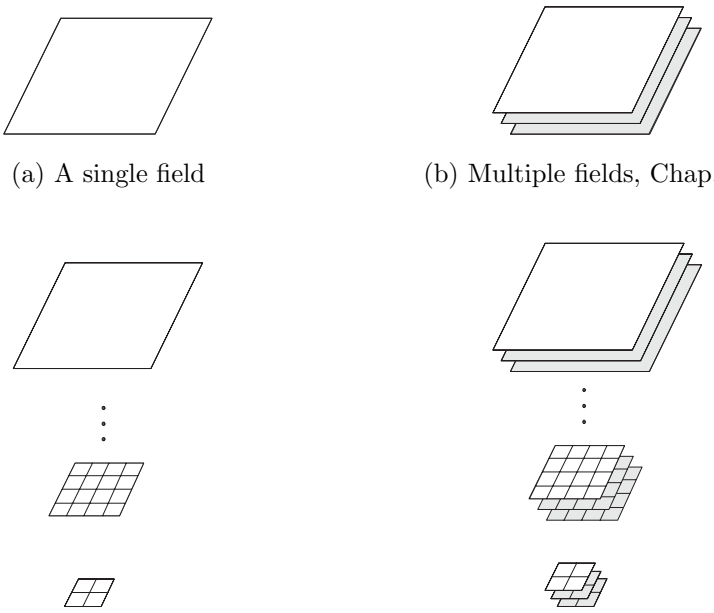
(c) Large-scale, low density pores

Figure 1.2: Multiple behaviours underlying a microscopic sample shown in Fig. 1.1(c).

capability of the FSHF, and apply it as the hidden layer to model the structural characteristics of an underlying behaviour. For example, the spatial structure of high-density, small-scale pores in Fig. 1.1(c) can be modeled by a hidden FSHF. In this thesis, the hidden FSHF method is discussed and applied to reconstruct porous media images in Section 5.2, and to synthesis remote sensing images in Section 6.3.

- To provide a powerful modeling tool with computational efficiency, scale-dependent modeling and non-stationary modeling capability, the ideas of hidden field, multiple fields, and hierarchical field are combined together to generate a Hidden Hierarchical Markov Field (HHMF). As for modeling a discrete field with piecewise, heterogenous behaviours, we introduce parallel HHMFs (Fig. 1.3(d)). For the example, as shown in Fig. 1.1(c), in applying the parallel HHMF method the different behaviours of pores are modeled separately and hierarchically. In the thesis, this parallel HHMF model is discussed and applied to porous media image reconstruction in Section 5.3.

- However, there are many discrete fields with spatially dependent, non-stationary structures that cannot be separated in a parallel way, such as shown in Fig. 1.1(d). To address this problem, a tree-structured HHMF (Fig. 1.3(e)) is proposed by combining a partition-tree modeling structure with hidden hierarchical fields. According to the proposed approach, a complex, multi-label, non-stationary field (Fig. 1.1(d)) can be repeatedly partitioned to yield a set of region-oriented, binary/ternary fields, each of

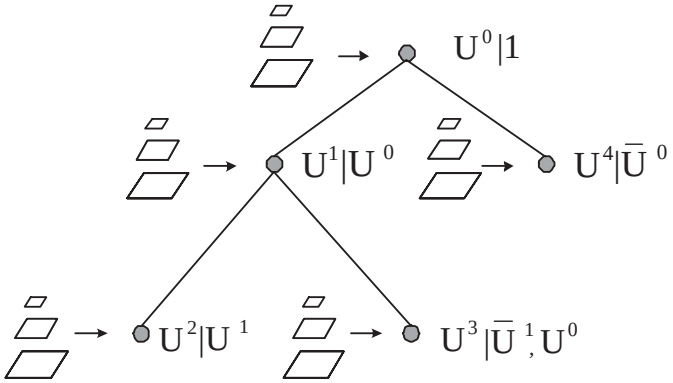


(a) A single field

(b) Multiple fields, Chap. 3, 5

(c) A single hierarchical field, Chap. 3, 5

(d) A hierarchical field with multiple hidden hierarchies, Chap. 5



(e) Tree-structured hidden hierarchies, Chap. 6

Figure 1.3: Possible structures for modeling scale-dependent multi-model behaviour.

which can be further handled by a resolution-oriented hierarchy. The tree-structured model is discussed and applied to remote-sensing image synthesis in Section 6.4.

1.3 Thesis Organization

The organization of the remaining chapters is as follows: Chapter 2 provides the general background of image modeling; Chapter 3 is about the specific background of MRF modeling approaches; Chapter 4 describes the problem formulation and some preliminary approaches to the solution. In Chapters 5 and 6, the HHMF based modeling methods are proposed and applied in porous media image reconstruction and remote sensing image synthesis, respectively. Finally, Chapter 7 summarizes the thesis work and discusses possible future research directions.

Chapter 2

Background: Image Modeling

This chapter contains a brief introduction to statistical modeling techniques. The chapter starts by describing the Bayesian framework as well as estimation and sampling problems. Then, the Markov/Gibbs random fields are introduced, and simple probability models used in this research are discussed. Finally, classical techniques including MCMC sampling methods along with simulated annealing are reviewed.

2.1 Images and Random Fields

Modeling a natural image is difficult. Just investigating simple first- and second- order statistics is not enough to adequately represent natural image characteristics [113]. Instead, one of the most general approaches is to consider an image as a realization of a random field where each pixel value is a realization of a discrete random variable [113]. A discrete random field is a two dimensional random process and can be defined as follows:

Definition 2.1.1 (Random Field)
Let $S = \{s_1, s_2, \cdots, s_N\}$ be a finite site set. A random field X on S is $X = \{X_s : s \in S\}$.

Let $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_L\}$ define a finite set of state values, $X_s \in \Lambda$, $\forall s \in S$. Denote the configurations for the realization of x as

$$\Omega = \{x = (x_{s_1}, \cdots, x_{s_N}) : x_{s_i} \in \Lambda, 1 \leq i \leq N\}.$$

A random field can also be defined on a lattice:

$$X = \{X_{i,j} | 1 \leq i \leq n_1, 1 \leq j \leq n_2\}$$

where $X_{i,j}$ is a random variable corresponding to the pixel (i, j), and where n_1 and n_2 are the numbers of the row indexes and the column indexes, respectively. X can be modeled by a probability distribution $p(x)$ on Ω with $p(x) \geq 0$ and $\sum_{x \in \Omega} p(x) = 1$. As image samples x from X are measured, the set of measurement samples y can be denoted by a random field Y.

2.1.1 Bayesian Framework

A stochastic relation between X and Y can be defined as a joint distribution $p(x, y)$ that can be further interpreted by Bayes' theorem:

$$p(y)p(x|y) = p(y|x)p(x) \tag{2.1}$$

Re-ordering (2.1), we have

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \tag{2.2}$$

The Bayesian formula (2.2) provides a framework for inferring x from y. Given knowledge of the prior distribution $p(x)$ and the likelihood function $p(y|x)$, the random sample x is represented by the posterior distribution $p(x|y)$. For an estimation problem, an estimate $\hat{x}$ can be decided by a statistical decision criterion such as maximum a posterior (MAP) [112]

$$\hat{x} = \arg\max_x p(x|y) \tag{2.3}$$

If $p(y)$ is given, based on (2.2), the MAP estimator can be rewritten as

$$\hat{x} = \arg\max_x \{p(y|x)p(x)\} \tag{2.4}$$

Then, the MAP solution to $\hat{x}$ can be considered as the optimal compromise between a prior model $p(x)$ and a measurement model $p(y|x)$.

More straightforward than estimation, image synthesis is a pure sampling problem which picks a configuration at random from a prior model $p(x)$

$$\hat{x} \leftarrow p(x) \tag{2.5}$$

Since synthesized samples $\hat{x}$ do not depend on any measurement, the behaviour of the synthesized samples reflects only the properties of the prior model [36].

2.1.2 Markov Random Fields

Although images and observations can be represented by random fields and probability distributions, in practice it is almost impossible to specify an image model based on a joint distribution of all of the pixels directly. Consider a small 32×32 binary image: the number of possible configurations is $2^{32\times32} = 2^{1024} \approx 10^{100}$. Such a large configuration space makes it impossible for us directly model the image with a joint distribution. So certain assumptions need to be introduced to simplify random field models.

Most images exhibit spatial dependencies, i.e., the pixels close to each other tend to have similar values or features. This local characteristic of images allows a Markovian property to be a reasonable assumption to reduce the complexity of modeling.

Definition 2.1.2 *[Markov Chain] [36]*
A random process $x(t)$ is a Markov Chain, if the knowledge of the process at some time t_0 is x_0, x_0 decouples the "past" x_p and the "future" x_f:

$$p(x_f|x_0, x_p) = p(x_f|x_0), \quad and \quad p(x_p|x_0, x_f) = p(x_p|x_0) \tag{2.6}$$

where $x_p = \{x(t)|t < t_0\}$, $x_f = \{x(t)|t > t_0\}$.

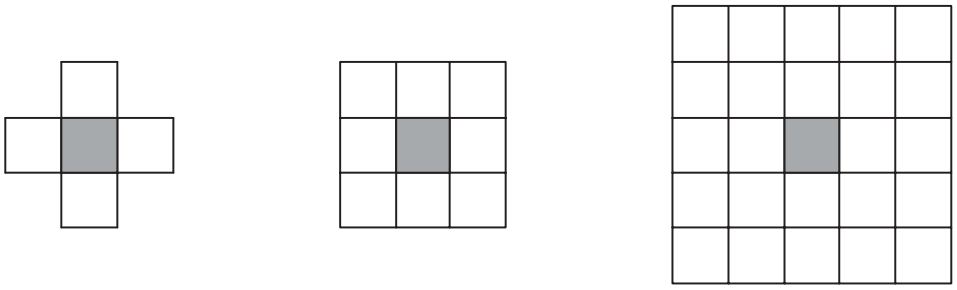
Whereas a Markov Chain (Definition 2.1.3) uses causal concepts of "past" and "future", a Markov Random Field (MRF, Definition 2.1.4) is noncausal. As such, it employs the variables $x_{\mathcal{N}_s}$ in the local neighborhood $\mathcal{N}_s$ to decouple a random variable x_s on site s from the rest of the domain $\{r \in S, r \neq s\}$. The neighborhood system is illustrated in Fig. 2.1 and defined as follows:

Definition 2.1.3 *[neighborhood system]*

$$A \text{ neighborhood system:} \quad \mathcal{N}_s \subset S \tag{2.7}$$

must have two properties:

1. *A site is not a neighbor of itself: $s \notin \mathcal{N}_s$.*
2. *A site and its neighbors have reciprocity: $s \in \mathcal{N}_r \Leftrightarrow r \in \mathcal{N}_s$.*



(a) First-order neighborhood (b) Second-order neighborhood (c) Fifth-order neighborhood

Figure 2.1: Neighborhood structures for MRF.

Definition 2.1.4 *[Markov Random Field]*
A random field X is Markov with respect to neighborhood system $\mathcal{N}$, if

$$p(x_s|x_r,\ r \in S, r \neq s) = p(x_s|x_r, r \in \mathcal{N}_s) \tag{2.8}$$

or in more compact notation as

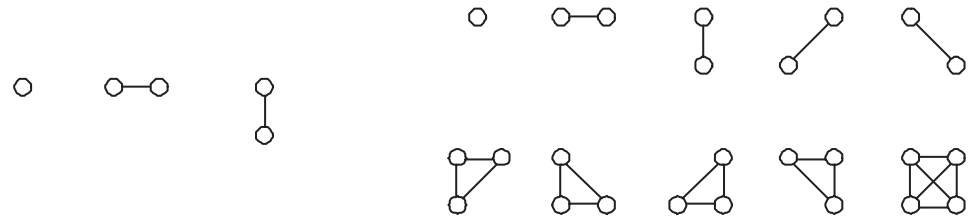
$$p(x_s|x_{S\setminus s}) = p(x_s|x_{\mathcal{N}_s}) \tag{2.9}$$

Therefore, given a neighborhood structure the MRF provides a convenient approach to represent spatial dependencies of a random field through a conditional marginal probabilistic distribution (2.9).

However, the remaining difficulty for MRFs is how to specify $p(x)$ or the conditional probabilities $p(x_s|x_{\mathcal{N}_s})$. This limitation of MRFs can be solved by GRFs which will be discussed in the next subsection.

2.1.3 Gibbs Random Fields

Gibbs Random Fields (GMF) are characterized by neighboring-site interactions [36], which were introduced into image modeling by Hassner & Sklansky [47], and popularized by Geman & Geman [39] and Besag [12].



(a) First-order cliques (b) Second-order cliques

Figure 2.2: Gibbs cliques for the first and second-order MRF neighborhoods.

Definition 2.1.5 *[Gibbs Random Fields]*
A random field X on S is a Gibbs Random Field if and only if its configurations obey a Gibbs distribution.

The Gibbs distribution is defined as

$$p(x) = \frac{e^{-H(x)/T}}{\mathcal{Z}} \tag{2.10}$$

where T is a temperature, $\mathcal{Z}$ is the partition function, and $H(x)$ is the energy function taking the form:

$$H(x) = \sum_{c \in \mathcal{C}} V(x_s, s \in c) \tag{2.11}$$

Here, c is a clique denoting a single site or a set of neighboring sites with $c \subset S$. $\mathcal{C}$ is the set of all cliques. $V(\cdot)$ is a clique potential which only depends on those x_s on the local sites $s \in c$.

The partition function $\mathcal{Z}$ is a normalizing constant

$$\mathcal{Z} = \sum_{x} e^{-H(x)/T} \tag{2.12}$$

The enormity of the possible configurations for x prevents $\mathcal{Z}$ to be evaluated but for the tiniest problems.

Since any pair of sites in c are neighbors to each other, the clique set can be associated with a neighborhood structure, for example, the clique types for the first and second order neighborhoods (Fig. 2.1(a)(b)) are shown in Fig. 2.2.

Actually, not only are $\mathcal{C}$ and $\mathcal{N}$ related, the MRF and the GRF have been proven to be equivalent [12, 45]:

Theorem 2.1.1 *[Hammersley-Clifford theorem] [36, 45]*

> *"X is a MRF with respect to $\mathcal{N}$ if and only if $p(x)$ is a Gibbs distribution with respect to $\mathcal{C}$, where $\mathcal{C}$ is the set of cliques with respect to neighborhood system $\mathcal{N}$."*

From Theorem 2.1.1, the local conditional distributions in a MRF model can be characterized by a Gibbs distribution, so $p(x)$ can be specified as

$$p(x) = \frac{1}{\mathcal{Z}} exp\Big\{ -\frac{1}{T} \sum_{c \subset \mathcal{C}} V_c \Big\} \tag{2.13}$$

Then, based on (2.13), the joint probability of a MRF can be specified by an energy function encoding the interactions of the local sites. Therefore, MRFs and GRFs provide an approach to simplify the modeling complexity of random fields. Their appealing properties for image and vision modeling can be summarized as follows [36, 64]:

1. They provide an effective mechanism for modeling spatial dependencies;
2. They offer a flexible way to describe enormously complicated probability functions by using relatively simple, intuitive energy functions, which will be discussed in Section 2.2.
3. They can be formulated within the Bayesian framework in Section 2.1.1, and can be easily estimated/sampled by existing statistical decision and estimation algorithms, as is shown in Section 2.3.

2.2 Common Prior Models

In Bayesian image analysis, the local and global characteristics of an image can be represented by random field models by setting up appropriate priors. In this section, some useful pixel-based models are reviewed including local and non-local models.

2.2.1 Ising Model

The Ising model (Ising 1925), based on the first-order neighborhood $\mathcal{N}_s$ for a pixel s, has been shown in Fig. 2.1(a). Although the Ising model is a very simple binary model, it exhibits a fundamental and typical local homogenous property shared by many complex systems: x_s tends to have the same value as its immediately adjacent neighbors $x_{\mathcal{N}_s}$. Hence it often acts as a test model for substantial problems about Markov fields.

Initially the Ising model was used by physicist E. Ising to explain ferromagnetism based on a crystal lattice of spin up or spin down dipoles [112]. In physics the energy $H(x)$ is formed as

$$H(x) = -\frac{1}{kT_a}\left(J \sum_{<s,r>\in\mathcal{C}_2} x_s x_r - mB\sum_s x_s\right) \tag{2.14}$$

where $\mathcal{C}_2$ represents all pairwise cliques of adjacent pixels. The first term of $H(x)$ represents the interaction energy of spin pairs and the second term represents the influence of an external field. As for the physical parameters, T_a is absolute temperature, k is Boltzmann's constant, the sign of J denotes whether spins are desired to be the same direction or not, B is the external field intensity, and m is a property of the material. If there is no external field, the model (2.14) is simplified as

$$H(x) = -\beta \sum_{<s,r>\in\mathcal{C}_2} x_s x_r \tag{2.15}$$

where, when $\beta > 0$ unequal neighbor pairs are penalized by high energy while the equal neighbor pairs are encouraged.

With only four nearest neighbors, the Ising model describes a variety of highly related random fields. The most favorable case is when the whole field is a homogenous region, which is much too simple to faithfully capture subtle structural characteristics. However, it is widely used in those cases where simplicity is pursued rather than searching for the best model, such as method testing [8, 111].

2.2.2 Potts Model

The Potts model (R.B. Potts, 1952) [93] is a natural generalization of the Ising model, where each site extends to have a L-state labels: $x_s \in \Lambda = \{0, 1, \cdots, L-1\}, s \in S$. Its

energy has the form

$$H(x) = -\beta \sum_{<s,r>\in\mathcal{C}_2} \delta(x_s, x_r) \tag{2.16}$$

where $\delta(x_a, x_b)$ is defined as:

$$\delta(x_s, x_r) = \begin{cases} 1 & \text{if } x_a = x_b \\ -1 & \text{if } x_a \neq x_b \end{cases} \tag{2.17}$$

Here, $H(x)$ is the number of unequal pairs minus equal pairs, thus measuring the degree of smoothness of X.

The model can also be formed as

$$\mathcal{H}(x) = \beta \sum_{<s,r>\in\mathcal{C}_2} (1 - \delta(x_s, x_r)) \tag{2.18}$$

Here, $\mathcal{H}(x)$ is equal to the length measuring the discontinuity of X.

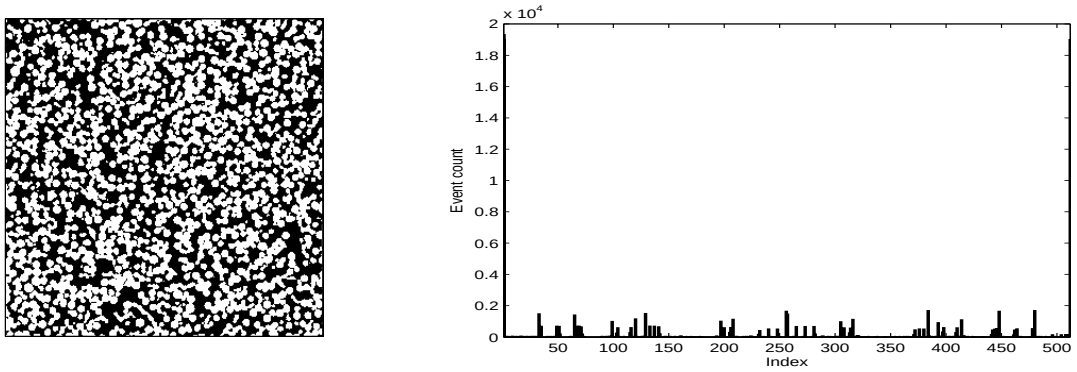
A common application for the Potts model is modeling images with more than two colors or images mixed with more than two kinds of textures, such as in [13] [83] [86].

2.2.3 Local Histogram Model

The local histogram model [1], Local Binary Pattern (LBP) [84] and Grey Level Co-occurrence (GLC) [46] are all non-parametric models defined on a neighborhood region, where both histogram and LBP accumulate the count of all the possible configurations about the whole neighborhood, whereas the GLC collects information about the joint probability of pixel pairs. In this thesis, the local histogram model will be used for all experiments.

To simplify the discussion, here we define a local histogram model on a binary random field X. For a 3×3 neighborhood as shown in Fig.2.1(b), the number of possible configurations will be 2^9. Target histograms ($\tilde{h}$) should be learned from the training data. An example of $\tilde{h}(n)$ is shown in Fig. 2.3(b).

An energy functions $H(x)$ of the local histogram model measures the consistency between the estimate histogram $h(n)$ and the target histogram $\tilde{h}(n)$. For example, for describing binary porous media the authors in [20] has proposed a histogram energy function



(a) An example image

(b) $h(n)$ of the example image

Figure 2.3: The histogram of an image example (a) is shown in (b).

as

$$H(n) = \sum_{n=0}^{N_h-1} \frac{|h(n) - \tilde{h}(n)|}{\nu(n) + \epsilon} \tag{2.19}$$

where N_h is the number of possible neighborhood configurations, ϵ is a small value to control the penalty for the unallowed local configurations, and ν normalizes coefficients, such as $\nu(n) = \sqrt{\tilde{h}(n)}$ defined in [20].

As a non-parametric, the model (2.19), well-defined in a small local neighborhood, can effectively model the nature of stationary structures as shown in Fig. 2.3(a), but is not suitable to capture large-scale heterogenous behaviours, such as in Fig. 1.1(c), whose configurations cannot be investigated locally [80]. However, it is computationally intractable to extend the local histogram model by employing a large neighbourhood, because the number of possible configurations increases exponentially with the size of the neighbourhood.

2.2.4 Chordlength Model

The chordlength model [105] is a standard model for handling dense structures in binary images. It describes the likelihood of finding black or white chords (line segments) as a function of length. If we only consider chords in the horizontal and vertical directions, and

denote the distribution of the counts of possible chord lengths as c^h and c^v, respectively, an energy function of the chordlength prior can be defined as

$$H_c(n) = \sum_{n=0}^{L_c-1} \|c^h(n) - \overline{c}^h(n)\| + \|c^v(n) - \overline{c}^v(n)\| \tag{2.20}$$

where $\|\cdot\|$ defines a norm function, $\overline{c}^h(n)$ and $\overline{c}^v(n)$ are the horizontal and vertical target chordlength distributions obtained from training data, and L_c denotes the longest chordlength. Fig. 2.4 shows an example of $c^h(n)$ for black chords.

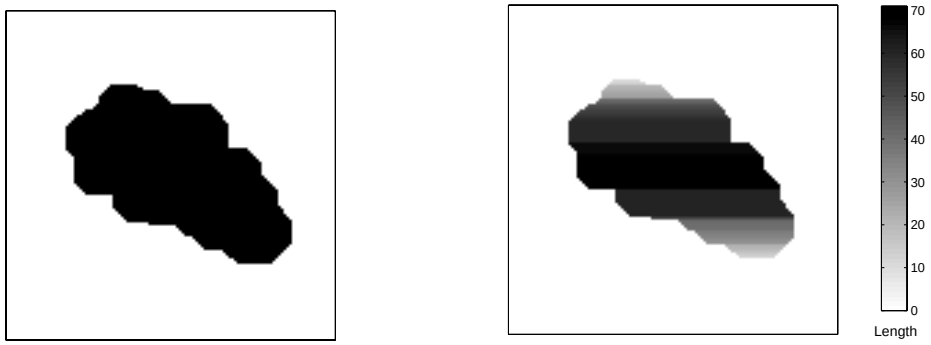
The chordlength model is widely used in porous media reconstructions [104, 105], which is effective at modeling structures at a single scale. However, it cannot effectively model complex, scale-dependent phenomena (Fig. 1.1(c)). Because the model does not define the correlation between the horizontal and vertical chords, in complex, multi-scale cases there are some unexpected configurations which are not forbidden by the model, such as those combining the horizontal chords of large-scale structures with the vertical chords from small-scale ones [81].

2.3 Markov Chain Monte Carlo Methods

As the configuration space of a random field X is extremely large, in particular the computation of the partition function $\mathcal{Z}$ of the Gibbs distribution is prohibitive, so directly sampling of the distribution p is impossible [113]. Therefore, dynamic Markov Chain Monte Carlo (MCMC) Methods [96] are prevailingly applied as the alternative approaches to simulate samples from random fields.

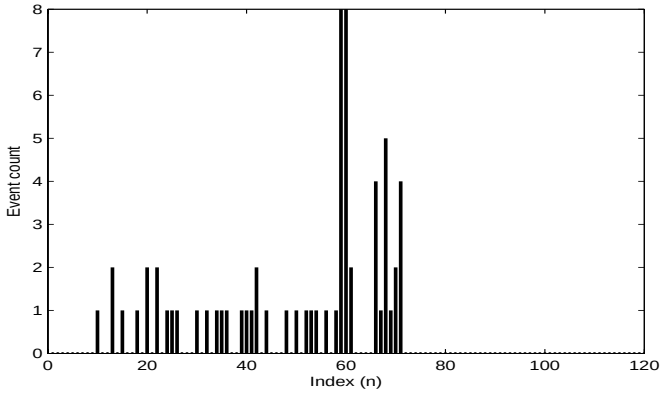
Dynamic Monte Carlo methods convert a problem of spatial complexity to a problem of temporal complexity. Starting from an initial configuration $x^{(0)}$, the algorithms update $x^{(0)}$ in subsequent steps under some proposed probabilistic rule which only depends on the current configuration and the number of updates. With a long running time, Dynamic Monte Carlo methods construct a sequence $\{x^{(0)}, x^{(1)}, \cdots, x^{(n-1)}\}$ to converge to x:

Theorem 2.3.1 *[Convergence Theorem] [112]:*
"Suppose P, $P_{i,j} = p\big(x^{(n)} = j | x^{(n-1)} = i\big)$, is an irreducible and aperiodic transition probability on a finite space with an invariant probability distribution p. Then uniformly



(a) An example image x

(b) The horizontal chord length map for black pixels in x



(c) $c^h(n)$ of x for black pixels

Figure 2.4: A chordlength example of event count distribution of all possible configurations with black pixels in the horizontal direction.

in all initial distributions $p^{(0)}$, $(p^{(0)} \cdot P^n) \to p$ *as* $n \to \infty$, *where* P^n *is the n-step transition probability."*

Theorem (2.3.1) indicates that for a Markov chain with ergodicity and steady-state probability, $p^{(n)}$ should be close to p after sufficient time. Although here only the convergence of homogenous chains is discussed, the inhomogenous cases also hold a similar theorem but need more assumptions to converge [112].

In MCMC methods, the Gibbs and the Metropolis samplers are the two most commonly used in solving image and vision problems, hence they are briefly reviewed below.

2.3.1 Gibbs Sampler

The Gibbs sampler [39] is an algorithm that simulates samples of a target Gibbs distribution by sampling its conditional marginal distribution. The conditional marginal distribution can be derived as

$$\begin{aligned} p(x_s|x_{\mathcal{N}_s}) &= \frac{p(x)}{\sum_{\lambda_i\in\Lambda} p(x)|_{x_s=\lambda_i}} \\ &= \frac{\frac{1}{\mathcal{Z}}\exp\{-H(x)/T\}}{\frac{1}{\mathcal{Z}}\sum_{\lambda_i\in\Lambda}\exp\{-H(x)/T\}|_{x_s=\lambda_i}} \\ &= \frac{\exp\{-H(x)/T\}}{\sum_{\lambda_i\in\Lambda}\exp\{-H(x)/T\}|_{x_s=\lambda_i}} \end{aligned} \tag{2.21}$$

where $\mathcal{Z}$ is canceled out. Representing the energy term $H(x)$ in terms of clique potentials $V(\cdot)$, (2.21) can be rewritten as

$$p(x_s|x_{\mathcal{N}_s}) = \frac{\exp\{-\sum_{c\subset\mathcal{C},s\in c} V_c\}}{\sum_{\lambda_i\in\Lambda}\exp\{-\sum_{c\subset\mathcal{C},s\in c} V_c\}|_{x_s=\lambda_i}} \tag{2.22}$$

For the Gibbs sampler, the canceling of $\mathcal{Z}$ is a crucial idea that helps the conditional marginal distribution $p(x_s|x_{\mathcal{N}_s})$ to be formulated in a feasible form as (2.22).

During the sampling process, the Gibbs sampler only changes one site value at each time and chooses a new value with a probability conditionally independent of the past choices. That means, taking a configuration $(x^{(t)})$ at time t and supposing the visiting site to be s, the new configuration $x^{(t+1)}$ will be:

$$x_r^{(t+1)} = \begin{cases} \text{a sample from} \quad p(x_s^{(t+1)}|x_{\mathcal{N}_s}^{(t)}) & r = s \\ x_r^{(t)} & r \neq s \end{cases} \tag{2.23}$$

The sampling process is shown in Algorithm 2.3.1.

According to (2.22), a marginal conditional distribution of the Ising model (2.15) at a site s can be written as

$$p(x_s|x_{\mathcal{N}_s}) = \frac{exp\{-\beta x_s \sum_{r\in\mathcal{N}_s} x_r\}}{exp\{-\beta\sum_{r\in\mathcal{N}_s} x_r\} + exp\{\beta\sum_{r\in\mathcal{N}_s} x_r\}} \tag{2.24}$$

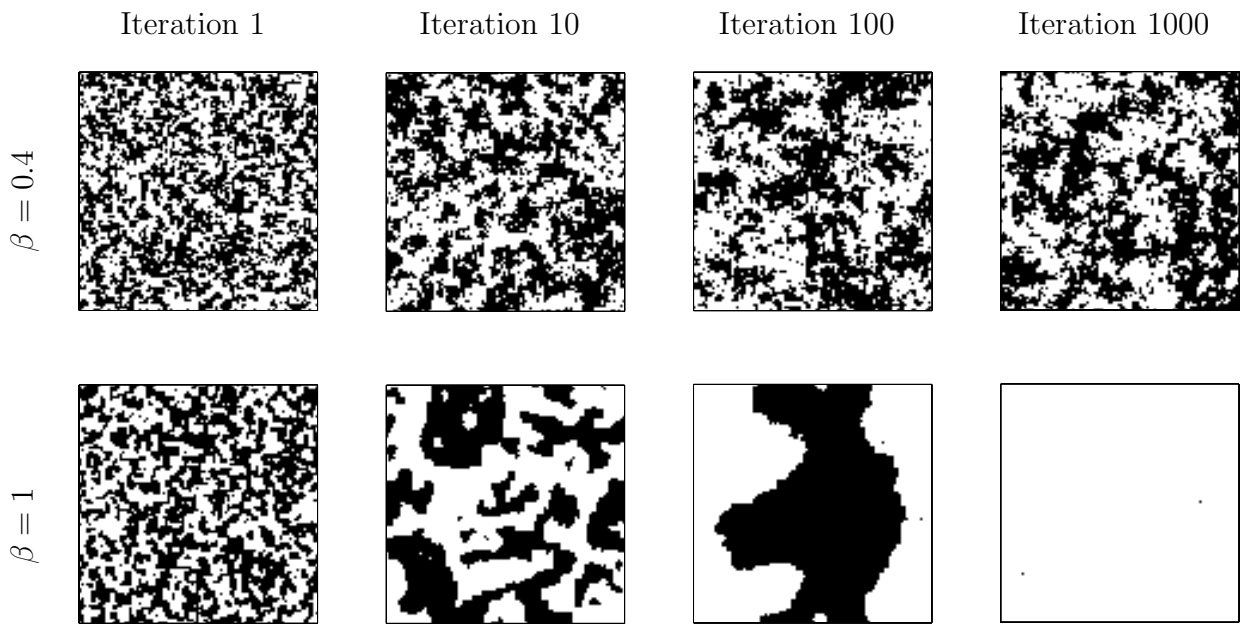


Figure 2.5: Realizations sampled from an Ising model by the Gibbs sampler. Notice how the samples are smoother for high β than for low β. That is because for high β adjacent pixels are more strongly coupled, discouraging irregular rough samples.

Algorithm 2.3.1: Gibbs Sampler

1: **repeat**

2: for all $s \in S$ do

3: compute all $p\Big(x_s^{(t+1)} = \lambda_i | x_{\mathcal{N}_s}^{(t)}\Big)$, $\lambda_i \in \Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_L\}$, and accumulate $P_{G_i} = P_{G_{i-1}} + p\Big(x_s^{(t+1)} = \lambda_i | x_{\mathcal{N}_s}^{(t)}\Big)$

4: uniformly draw a number ξ from $[0, 1]$, find $\lambda_{i^*} \in \Lambda$ with $P_{G_{i^*-1}} < \xi \leq P_{G_{i^*}}$, set $x_s^{(t+1)} = \lambda_{i^*}$

5: end for

6: **until** convergence

An example using the Gibbs algorithm to sample an Ising model is shown in Fig 2.5, where the samples are generated from (2.24) at both low and high inverse temperature ($\beta = 1/T$).

2.3.2 Metropolis Sampler

The Metropolis sampler [48] operates by choosing a state at random and then testing whether to accept or reject. In a single site case, at time t, for a site s randomly pick a new value λ. If the energy of this new configuration is lower than the former energy, then accept the new value, otherwise accept the new state λ with a probability P_M:

$$P_M = exp\Big(H(x^{(t)}) - H(x^{(t+1)})\Big) \Rightarrow x_s^{(t+1)} = \lambda.$$

We can see that, to avoid being trapped in local minima, the Metropolis sampler allows the acceptance of a new value λ which leads a higher energy, then even if $x^{(t)}$ is in a local minima, there is a change of P_M for $x^{(t+1)}$ to jump out of it. The sampling process of the sampler is shown in Algorithm 2.3.2, where the configuration change is decided by a ratio of the new configuration probability $p(x^{(t+1)})$ to the former one $p(x^{(t)})$, so there is no need to compute the normalization factor $\mathcal{Z}$.

Moreover, different from the Gibbs sampler, the performance of the Metropolis sampler depends crucially on the visiting scheme. It is irreducible only with random visiting

Algorithm 2.3.2: Metropolis sampler

repeat

  randomly pick a site $s \in S$

  hypothesize a value $\lambda \in \Lambda$ for x_s, which is corresponding to a configuration $x^{(t+1)}$

  compute $P_M = \min\left\{1, \frac{p\left(x^{(t+1)}\right)}{p\left(x^{(t)}\right)}\right\}$, and uniformly draw a number ξ from $[0, 1]$

  if $P_M > \xi$, then $x_s^{(t+1)} = \lambda$

  else $x_s^{(t+1)} = x_s^{(t)}$

until convergence

schedule [112], so the systematic sweep strategies, such as a chequerboard scan, may lead to some configurations that cannot be reached by the sampler.

2.3.3 Simulated Annealing

Simulated Annealing (SA), proposed by Kirkpatrick [58], is a stochastic optimization algorithm which provides an approach to reach the global minimum/maximum of non-convex energy functions. In image analysis problems, SA seeks the maximal modes of the (posterior) probability distribution on an image space.

For a Gibbs field, SA can be applied to a MCMC sampler to search the most probable samples by defining the parameter T of the distribution (2.10) as a function of the index of a sequence ($T = T(n)$). At the beginning, one sets T to be a large value that leads $p(x)$ to approach a uniform distribution on the configuration space. Then one slowly cools down T based on an annealing schedule $\{T_n\}$. As $T \to 0$, $p^{(T)}(x)$ concentrates on the peaks of $p(x)$ with a uniform distribution on the space of maximal modes. In essence, under $\{T_n\}$ and a visiting scheme $\{q_1, \cdots, q_N\}$, the method replaces the homogeneous Markov chains defined by the previous samplers to a non-homogeneous Markov chain $\{x^{\{(q_1),(T_1)\}}, \cdots, x^{\{(q_N),(T_1)\}}, \cdots, x^{\{(q_1),(T_n)\}}, \cdots, x^{\{(q_N),(T_n)\}}\}$. Algorithm 2.3.3 shows this process. The convergence theorem for SA (Theorem 2.3.2) has been proved by Geman & Geman [39].

Algorithm 2.3.3: Simulated Annealing

1: Initial configuration $x^{(0)}$
2: $i \leftarrow 0$
3: **repeat**
4: $\quad \beta^{(i)} = 1/T^{(i)}$
5: $\quad$ Apply the Gibbs/Metropolis sampler with $\beta(i)$ and get $x^{\{(i),\beta^{(i)}\}}$ from $x^{\{(i-1),\beta^{(i-1)}\}}$
6: $\quad i \leftarrow i + 1$
7: **until** $T \to 0$

Theorem 2.3.2 *[Convergence Theorem for SA] [39]:*
"Let T_n $(n \geq 2)$ be a cooling schedule decreasing to zero with

$$T_n \geq \frac{N\Delta}{\ln n} \tag{2.25}$$

then uniformly for any initial distribution $p^{(0)}$,

$$\lim_{n\to\infty} p^{(0)} P_1 \cdots P_n(x) = \begin{cases} |M|^{-1} & \text{if } x \in M \\ 0 & \text{otherwise} \end{cases}$$

where $|M|$ denotes the number of maximal modes M, and Δ is the absolute maximum change in Gibbs energy when only a single site changes its value."

The key idea of a cooling schedule is that, for a sufficiently slow cooling, the final energy $H(x^{(0)})$ should be close to the global minimum energy H_{min}. But a logarithmic schedule like Theorem 2.3.2 is too slow to be practical. Therefore, many other schemes are used such as geometric cooling [58]:

$$T_n = \rho^n \cdot T_0 \quad (0 < \rho < 1) \tag{2.26}$$

Although geometric schemes (2.26) have practical advantages, their convergence is not guaranteed. Therefore, we need a trade-off between the quality of convergence and the convergence speed. In practice, T_0 is set high enough in order to explore the whole configuration space. When T_n is cooled with rate ρ, the constraints defined by energy functions are more and more emphasized. As $T_n \to 0$, whether a new state value will be accepted only depends on energy reductions, so configuration searching process is not stochastic any more.

For SA applied to a random field, the conflict between the computational cost and the convergence to a global minimum becomes more serious in modeling large-scale phenomena than small-scale ones, since large structures composed of many pixels are hard to change by single-site samplers. As discussed in [42], for a random walk process the required number of Monto Carlo updates (the relaxation time τ) can be roughly approximated as a quadratic function of the correlation length ξ of the process:

$$\tau \approx \xi^2. \tag{2.27}$$

Thus, annealing a random field with very large spatial structures will lead to an enormous computational demand [35]. Clearly the methods, which can represent large-scale structures with reduced pixel numbers and local correlations, are motivated. Thus, hierarchical representation methods are intuitively attractive, where large spatial structures in an image can be decomposed into local ones at coarse scales by repeatedly downsampling the original image. As we hierarchically sample a field, an annealing process is allowed to start from a coarse scale where the field is related to few pixels, and create pixels over scales. During the annealing process, for coarse scale pixels whose state values are converged or close to converged, at fine scales their corresponding pixels may not need a re-sampling [20] or may not require a highly relaxed stochastic re-sampling [79]. Therefore, the annealing cost can be significantly reduced by either decreasing the number of pixels to be sampled or decreasing the relaxation time.

According to the efficiency of hierarchical annealing, many methods with various structures and schemes have been proposed in the literature, such as a pyramid with multi-temperature annealing [53], a pyramid with frozen-state annealing [20], a quadtree structure [60], a hybrid structure combining pyramid and quadtree [18], and a hierarchical region growing [109].

Moreover, except for hierarchical methods, various other acceleration techniques have been proposed in the literature, such as causal Markov model [51, 62], and clustering [5, 103, 106].

Chapter 3

Background: Markov Random Field Modeling

In this chapter MRF based modeling methods are reviewed. We start with a discussion of the single MRF field and the classical Hidden Markov Field, and then present some popular extended MRF methods. Finally, hierarchical MRF methods are introduced. Those reviewed materials cover three important ideas: hidden fields, multiple/joint MRF fields, and hierarchical fields, which will be applied later in this thesis to create new approaches for addressing challenges of modeling scale-dependent, non-stationary structures.

3.1 Single Markov Random Fields

The simplest approach to MRF modeling is to represent an image with a single random field $X = \{X_s : s \in S\}$ which has been successfully applied in image related problems such as modeling [105], classification [25] and synthesis [2] [27].

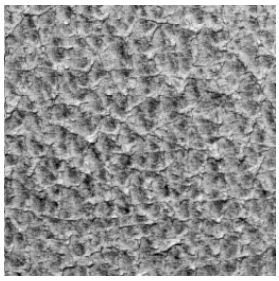
3.1.1 Gaussian Markov Random Fields

One particularly common MRF is the Gaussian-Markov Random Field (GMRF) [24] which changes the conditional independence of the Markov definition from (2.9):

$$p(x_s|x_r,\ r \in S, r \neq s) = p(x_s|x_r, r \in \mathcal{N}_s)$$

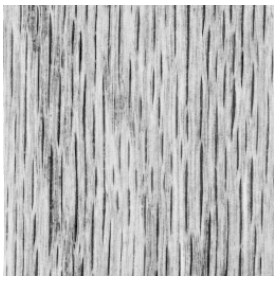
0	-0.0773	0.2031	-0.0576	0
-0.0308	0.2308	-0.6233	0.1963	-0.0232
0.0674	-0.3844		-0.3844	0.0674
-0.0232	0.1963	-0.6233	0.2308	-0.0308
0	-0.0576	0.2031	-0.0773	0

Grass



0	-0.0333	0.0869	-0.0264	0
-0.0087	0.1396	-0.3612	0.1111	0.0074
0.0709	-0.4863		-0.4863	0.0709
0.0074	0.1111	-0.3612	0.1396	-0.0087
0	-0.0264	0.0869	-0.0333	0

Pigskin



0	-0.0025	0.0592	-0.0018	0
0.0104	0.1112	-0.5573	0.1072	0.0111
-0.0208	-0.2166		-0.2166	-0.0208
0.0111	0.1072	-0.5573	0.1112	0.0104
0	-0.0018	0.0592	-0.0025	0

Wood

Figure 3.1: Textures and their estimated GMRF model parameters corresponding to a 4th-order MRF neighborhood structure. Textures are taken from the Brodatz Texture Database. The method used to learn the model parameters is from [24].

to conditional decorrelation, we get a Weakly Markov Field [44]:

$$E(x_s|x_r,\ r \in S, r \neq s) = E(x_s|x_r, r \in \mathcal{N}_s) \tag{3.1}$$

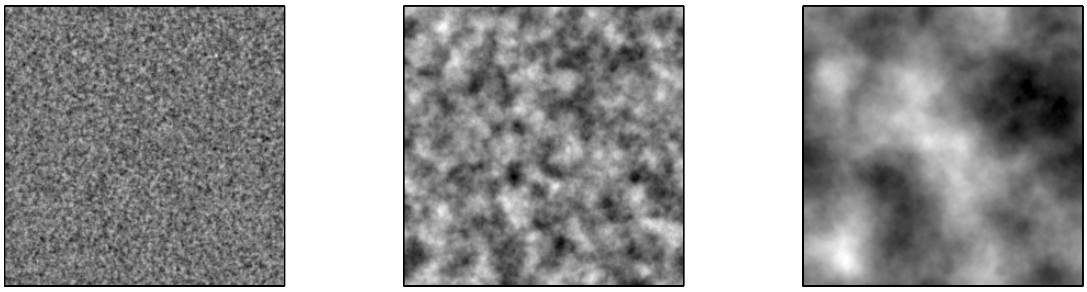
In the Gaussian case, (3.1) is equal to the best linear estimate of x_s [36], so the GMRF can be written as

$$x_s = \sum_{r \in \mathcal{N}_s} a_{s,r} x_r + \omega_s \tag{3.2}$$

Here $\theta_s = \{a_{s,r}, r \in \mathcal{N}_s\}$ is a model parameter set which can be used to capture the stationary characteristics of x, and ω is zero mean Gaussian noise with $E(x_r \omega_s) = 0$ for all $r \neq s$.

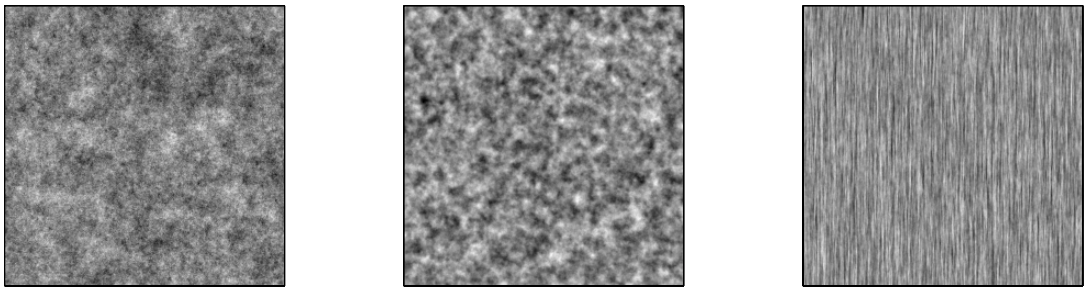
Three examples of applying a GMRF to model textures are shown in Fig. 3.1, where θ_s is based on a 4th-order neighborhood structure and the different estimated θ_s reflect the different features of those textures.

In general, the complexity of a MRF model is related to the size of neighbourhood (Fig. 3.3), whereas the scale characteristic of a MRF model, which is determined by model coefficients, actually has no relationship with the size of neighbourhood — a local neigh-



(a) small scale (b) medium scale (c) large scale

Figure 3.2: A given order MRF model can model wide range scales, for example, a 3rd-order thin-plate model generates (a-c).



(a) Membrane (b) Thin-plate (c) Tree-bark

Figure 3.3: The complexity of a MRF model is related to the size of neighbourhood, such as (a) a 1st-order membrane model, (b) a 3rd-order thin-plate model, and (c) a 4th-order tree-bark model [36].

bourhood can define a small or a large scale process (Fig. 3.4(a,b)), but cannot define a complex multi-scale process (Fig. 3.4(c)) [36].

At the same time, a single MRF parameter set, corresponding to a stationary model, is unable to represent a nonstationary scene with multiple textures (Fig 3.5(b)). Generally, a simple linear MRF model is also not suitable to describe scenes which are distorted, blurred, or contaminated, since the statistics of those scenes are usually not driven by a single random process associated with the true scene. To model those cases, another method — Hidden Markov Fields — is considered.

(a) small scale (b) large scale (c) (a)×(b)

Figure 3.4: Thresholded binary samples from a thin-plate model and a simple operation. A local thin-plate model can represent small scale structure (a) or large scale structure (b), but is unable to represent a two-scale structure (c) which is generated from the simple operation (a)×(b) [36].

3.1.2 Classical Hidden Markov Fields

Since the work of Geman & Geman [39] and Besag [13], the Hidden Markov Field (HMF) has made significant contributions to MRF modeling. The main reason for its success is that the HMF extends the modeling flexibility of MRF by using two coupled fields which allow modeling a broader range of features than a single field on its own.

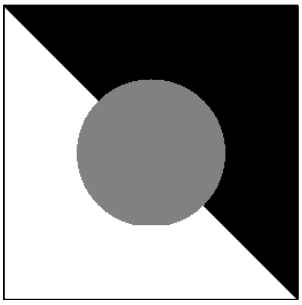
Basically, the classical HMF models a scene as a two-layer random field: an observable field Y and a hidden field X to be estimated. Here, Y can be considered as a deterministic or random transformation of X. A general measurement y of a random sample x can be written as

$$y = g(x) + \nu \tag{3.3}$$

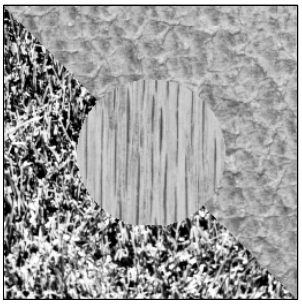
where y is a realization of Y, $g(\cdot)$ denotes a forward operation, and ν denotes measurement noise. Examples in Fig. 3.5 show a ground truth $\overset{*}{x}$ and several observations corresponding to different forward models.

Therefore, the classical HMF has two random processes: one is driven by ν, which constrains the inconsistency between $g(x)$ and y; the other is a MRF which constrains the statistical features of x based on the prior knowledge of $\overset{*}{x}$.

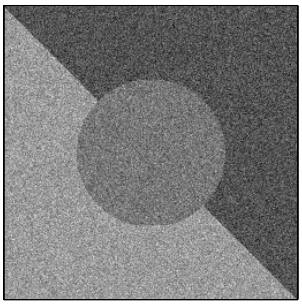
The general problem for a HMF is how to generate a sample notation or an estimate $(\hat{x}(y))$ from y. In general, there are many possible x which can match a given y, so the



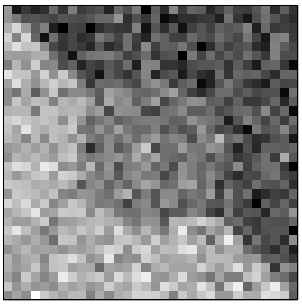
(a) Ground truth, $\overset{*}{x}$, $\overset{*}{x}_s \in \{1,2,3\}, s \in S$



(b)Mixed texture image, $y = g(\overset{*}{x})$, $g(i)$ is a texture decided by labels $i \in \{1,2,3\}$



(c) Noisy observation, $y = \overset{*}{x} + \nu$



(d) Noisy downsampled observation, $y = \Downarrow (\overset{*}{x}) + \nu$

Figure 3.5: A ground truth $\overset{*}{x}$ and several of its possible observations resulting from different $g(\cdot)$. The observations correspond to different applications of HMF models such as (b) texture segmentation, (c) restoration, and (d) reconstruction. Specifically, (c) can be modeled by the classical HMF, (d) can be modeled by a relaxed classical HMF with a loose assumption (A.1) only based on (3.6), and (b) can be modeled by the Double MRF (Section 3.2.1).

HMF is usually approached as an optimization problem which can be addressed under a Bayesian framework and solved by a MAP estimator:

$$\begin{aligned} \hat{x} &= \arg\max_{x} p(x|y) \\ &= \arg\max_{x} \{p(y|x)p(x)\} \end{aligned} \tag{3.4}$$

However, as discussed in Section 2.1.2, except for uninterestingly small scenes the configuration space of x is too enormous to permit (3.4) to be solved directly. To make the above estimation tractable, $p(x|y)$ is desired to be Markovian so that existing algorithms, such as Markov Chain Monte Carlo (MCMC) methods (Section 2.3) can be used to solve the problem. For the classical HMF, two assumptions are required to have this Markovianity [7, 13, 39]:

Assumption (*A*.1) : *Y is conditionally independent, that is*

$$p(y|x) = \prod_{s \in S} p(y_s|x_s) \tag{3.5}$$

which implies two sub-assumptions:

1. Given X, $\forall\ s \in S$, y_s has conditional independence:

$$p(y|x) = \prod_{s \in S} p(y_s|x) \tag{3.6}$$

2. Given X, $\forall\ s \in S$, the conditional probability of y_s only depends on its corresponding x_s:

$$p(y_s|x) = p(y_s|x_s) \tag{3.7}$$

(3.6) implies that the measurement model should be driven by independent noise, and (3.7) implies that $g(\cdot)$ should not allow multiple variables of x corresponding to a single variable of y, such as a blurring operation.

Assumption (*A*.2) : *X is a MRF with distribution* $p(x)$.

Therefore, according to (A.1), the Markovianity of $p(x)$ can lead $p(x|y)$ to be Markovian:

$$p(x|y) \propto \Big[\prod_{s \in S} p(y_s|x_s)\Big] p(x) \tag{3.8}$$

However, it is easy to be aware of the drawback of the classical HMF: it maintains the Markovianity of $p(x|y)$ by adding a strict assumption on $p(y|x)$. The problem is that (A.1) is hard to be satisfied in many complex situations, for example when the observation is corrupted by a correlated noise, or when it is a blurred version of the true scene. In those cases, the classical HMF will introduce approximations, which results in the deterioration of the estimates.

3.2 Multiple Markov Random Fields

3.2.1 Extended Hidden Markov Fields

To improve the modeling ability of the classical HMF, many models with extended HMFs have been proposed. A double MRF structure, shown in Fig. 3.6(c), has been introduced in [77]. A scene is modeled as a joint distribution for X and Y:

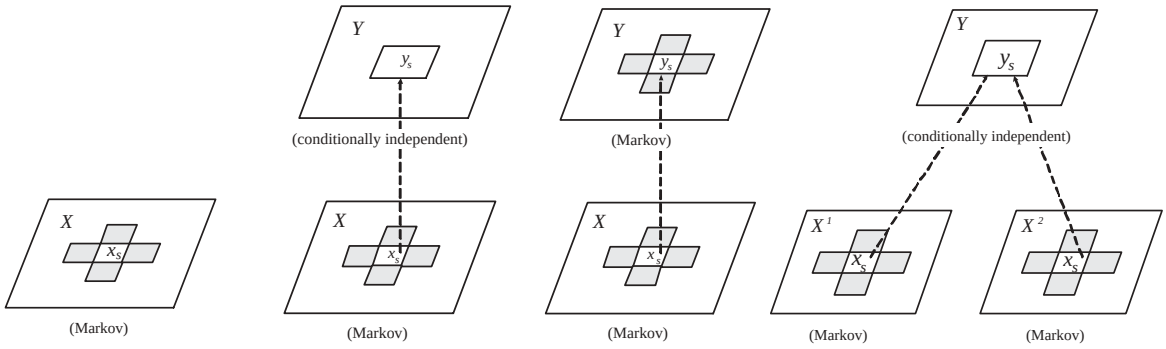
$$p(x|y) \propto p(y|x)p(x)$$

Here, $p(x)$ is Markov, $p(y|x)$ is conditional Markov, and their spatial dependency can be represented as

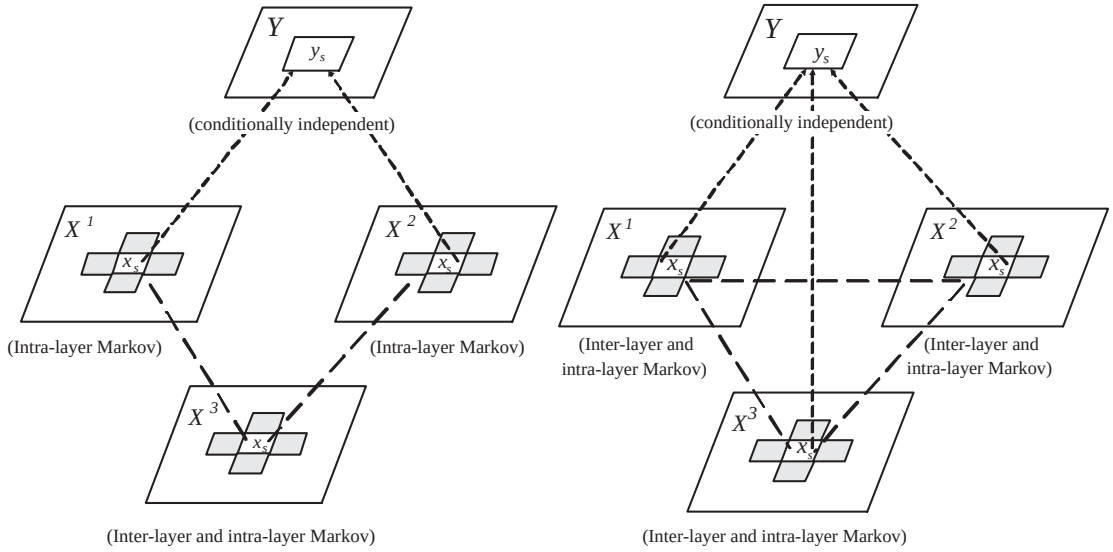
$$p(y_s|y_{S\setminus s}, x) = p(y_s|y_{\mathcal{N}_s^{y|x}}, x) \tag{3.9}$$
$$p(x_s|x_{S\setminus s}) = p(x_s|x_{\mathcal{N}_s})$$

where $\mathcal{N}_s^{y|x}$ and $\mathcal{N}_s^x$ are the neighborhoods for the MRFs of $Y|X$ and X, respectively, at site s.

Compared to the classical HMF, the double MRF model provides more powerful modeling capability, since in many cases observations have locally dependent characteristics. For examples of remote-sensing data, in a landmass image the noise may be related to land cover types, in an lake image the reflectance of adjacent pixels may be very different from the pixels far away, and in a forest image the density of tree distribution may be different from one area to another. In addition, the mixed-texture modeling problem (Fig. 3.5(b)) mentioned in Section 3.1.2 can now be addressed by a double MRF model where, given a field X labeling the type of texture the observation, Y only contains a single texture and then $Y|X$ can be modeled by a GMRF (Section 3.1.1).



(a) Classical MRF [24] (b) Classical HMF [39] (c) DoubleMRF[77] (d) FMRF [57]



(e) Multi-MRF [54] (f) Three-layer MRF [10]

Figure 3.6: Some examples of extended HMF models. The classical HMF (b) has been extended from a two-layer random field with a single MRF to multi-layer, multi-MRF structures, such as (c)(d)(e)(f). Those extended HMF models relax the conditional independency assumption of the classical HMF. All HMFs are extensions of the classical MRF (a).

Although the classical HMF model performs well in dealing with low-level vision problems [63], the model seems inadequate to handle high-level vision problems without extensions, since most high-level vision problems are related to much more complicated situations such as multiple feature spaces, multiple measurement channels, disconnected spatial regions, etc. For example: in a multi-object recognition problem, where images show a wooden table with three objects on it: a newspaper, a cup, and a basket of strawberries, to successfully label the items on the table we need to model the objects, the background, and their relations. In [57], a factorial Markov Random Fields (FMRF) model was proposed, which is based on layer representations [107] to represent scenes with overlapped objects. This model assumes an observation (Y) based on a set of independent hidden fields $\{X^l, l \in \mathcal{L}\}$. Here $\mathcal{L}$ is a set of layer labels, and at each layer has $X^l = \{X_s^l : s \in S^l\}$, where S^l is the site set at layer l. Given y, a FMRF can be modeled as

$$p(\mathbf{x}|y) \propto p(y|\mathbf{x})p(\mathbf{x}) = \prod_{s \in S} p(y_s|\mathbf{x}_s) \prod_{l \in \mathcal{L}} p(x^l) \tag{3.10}$$

where all the hidden fields are assumed to be Markov with a neighborhood $\mathcal{N}^l$. An example of a typical FMRF structure is shown in Fig. 3.6(d)). For some complex vision problems such as motion recognition and pose detection, more than decomposing a scene into independent layers, features captured by different layers need to combine together to contribute to a final conclusion. For example, to detect a human body pose, statistical features of relevant body parts, such as head, arms and the legs, need be integrated to make a final pose decision. For those cases requiring to model inter-layer relations, multi-MRF models have been proposed and applied [10] [54]. An example structure is shown in Fig. 3.6(e), where

$$Y = Y^1 \cup Y^2. \tag{3.11}$$

Y could been decomposed into two independent feature fields Y^1 and Y^2. Y^1 and Y^2 respectively correspond to two independent hidden fields X^1 and X^2. In addition, the third hidden field X^3, which combines the information of X^1 and X^2, is defined as

$$X^3 = \{X_s^3 : s \in S^1 \cup S^2 \cup S^3\}. \tag{3.12}$$

X^3 has intra-layer and inter-layer spatial dependencies.

Therefore, given y, the posterior distribution of this model can be formulated as

$$\begin{aligned} p(x|y) &\propto p(y^1, y^2|x^1, x^2, x^3)p(x^1, x^2, x^3) \\ &= p(y^1|x^1)p(y^2|x^2)p(x^1, x^2|x^3)p(x^3) \\ &= p(y^1|x^1)p(y^2|x^2)p(x^1|x^3)p(x^2|x^3)p(x^3) \end{aligned} \tag{3.13}$$

where conditioned on a hidden field X^3, a complex joint field model $p(x|y)$ is decomposed as several independent models. Generally, those independent models are simpler than $p(x|y)$ and are relatively easy to deal with.

In more general cases the hidden layers associated with the different feature fields could be related to each other, so the information of X^l not only depends on Y but also depends on other hidden layers. A three-layer MRF model was proposed in [10] to ensure the connection of different feature layers. A modeling structure is shown in Fig. 3.6(f). Given $Y = \{Y^l, l \in \mathcal{L}\}$, the energy function $H(x^l|y^l)$ of the posterior distribution for $p(x^l|y^l)$ can be written as

$$H(x^l|y^l) = C_m H_m(y^l|x^l) + C_i H_i(x^l) + C_o H_o(x^l|x^j, j \in \mathcal{L}, j \neq l) \tag{3.14}$$

where C_m, C_i, and C_o are modeling parameters. All the hidden layers are assumed to be Markov, $H_i(x^l)$ models intra-layer features, and the prior $H_o(x^l|x^j, j \in \mathcal{L}, j \neq l)$ reflects the influences from other layers.

Although the double MRF, the FMRF, the multi-MRF, and the three-layer MRF models were introduced for addressing different image and vision issues, all of them extend the classical HMF by relaxing its strict assumption of a conditional independency, and provide more powerful capabilities for capturing complex statistical characteristics.

3.2.2 Pairwise Markov Fields and Triplet Markov Fields

To remedy the drawbacks of the classical HMF, and in contrast to the extended HMF models discussed in the previous section, the Pairwise Markov Fields (PMF) [89] and the Triplet Markov Fields (TMF) [7] directly assume the joint fields to be Markov instead of assuming single fields or conditional fields to be so. The PMF (Fig. 3.7(b)) asserts (X, Y) to be pairwise Markov, so given Y or given X, the conditional distributions $p(x|y)$ or $p(y|x)$ are guaranteed to be Markovian, respectively. But X and Y individually are not

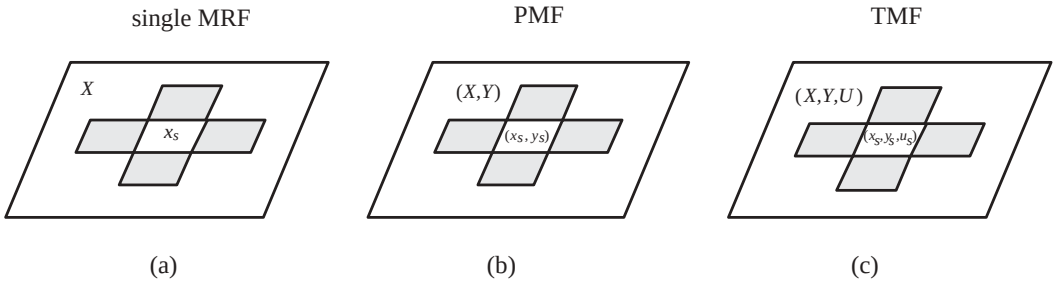


Figure 3.7: The structure graph of single and joint MRF models. Here, models are separated as joint fields (a)(b)(c). Among them the TMF (c) is the most general one. The TMF can be the PMF (b) by calculating marginal distribution, and similarly (b) can be simplified as the single MRF (a).

necessarily Markov. Therefore, the PMF offers a more general modeling capability than other two-field Markov models, such as the double MRF (3.2.1), since the PMF not only allows modeling complex noise processes based on $p(y|x)$, but also estimates underlying properties based on $p(x|y)$, for example, given observation Y we can estimate an underlying label field X for segmentation [8].

A local dependency for the PMF can be described as

$$p(x_s, y_s|x_{S\setminus s}, y_{S\setminus s}) = p(y_s|x_s, x_{\mathcal{N}_s^x}, y_{\mathcal{N}_s^y})p(x_s|x_{\mathcal{N}_s^x}, y_{\mathcal{N}_s^y}) \tag{3.15}$$

where $\mathcal{N}_s^x$ defines a neighborhood structure in X, and $\mathcal{N}_s^y$ defines a neighborhood structure in Y. From (3.15), we can see that, compared to the classical HMF, the PMF relaxes assumption (3.5) without requiring $p(y|x)$ to be conditionally independent. Therefore, (3.6) is relaxed so that a measuring process could be driven by correlated noise, and (3.7) is also relaxed so that $g(\cdot)$ is allowed to determine y_s based on a subregion of x: $x_{w_s} = \{x_s, s \in w_s \subset S\}$. Moreover, the model for X is also relaxed and is not necessarily to be Markov.

Generally, the distribution of the PMF can be written as

$$p(x, y) = \frac{e^{-H(x,y)/T}}{\mathcal{Z}} \tag{3.16}$$

Further, by introducing an auxiliary random field $U = \{U_s : s \in S\}$, the PMF can be extended to a TMF (Fig. 3.7(c)) which is capable of handling more complex situations,

such as nonstationary regions, by asserting (X, Y, U) to be triplet Markov:

$$p(x, y, u) = \frac{e^{-H(x,y,u)/T}}{\mathcal{Z}} \tag{3.17}$$

For a TMF, if we assert (X, Y) to be pairwise fields and U to be an underlying field labeling heterogenous regions, $p(x, y|u)$ will be PMF. Taking advantage of an auxiliary field U, the TMF can provide a broader representation than the PMF, for examples, the TMF is able to model a nonstationary image with correlated noise, or segment a nonstationary, noisy image [7, 8]. Moreover, the TMF offers a more general relation between an image field X and its observation Y: if we assume (X, U) to be pairwise hidden fields, then not only $p(x)$ and $p(y)$ but also $p(x|y)$ and $p(y|x)$ are not necessarily Markovian.

At the same time, similar to the PMF, the TMF does not need any conditionally independence assumption or the Markov assumption for individual fields. Thus, it is more general than other three-field Markov models, such as the FMRF (Section 3.2.1).

3.3 Hierarchical Markov Random Fields

There are many problems in texture analysis, remote sensing and scientific imaging where observed images possess highly scale-dependent structure, for example, imagery of damaged woolen fabrics with holes of various sizes, imagery of an archipelago including large and small islands, and imagery of rock samples possessing pores at multiple scales. Although such structures can, in principle, be represented with sufficiently complex models, the development of such models is a difficult task and leads to computationally intractable algorithms if executed on a single, fine scale. Instead, hierarchical methods [43, 60, 62, 111] by constructing fields at a set of sequential scales offer an approach to scale-dependent modeling.

3.3.1 Hierarchical Random Fields

In general, hierarchical models define X via a sequence of fields $\{X^k,\ k \in K = (0, 1, \cdots, M)\}$, where $k = 0$ defines the finest scale. At each scale k, X^k is defined on site space S^k and results from the downsampling of X^0.

However, the fact is that a downsampled Markov field is no longer Markov, in general [49], to which there are two main responses: one is to formally model the downsampled fields properly based on group normalization [41], the other is to construct an approximation and assert the downsampled fields to be Markov. The latter MRF approximation gives us a simple, convenient modeling environment with a scale dependent model, where the hierarchy very nearly obeys the statistics of the finest scale. Therefore, at each scale a single MRF (X^k) can be used to capture the features local to that scale, inherently allowing scale-dependent structure.

Many researchers have been working on hierarchical MRF modeling [2, 20, 43, 53, 60, 79] to provide a more natural and efficient way to deal with label fields. Kato et al [53] proposed a hierarchical MRF model with a 3D neighborhood system for modeling discrete states, but with considerable computational cost. A MRF model based on a quad-tree structure was discussed by Laferté et al [60], but does not model the interactions within scales.

Thus, in defining a hierarchical model two issues need emphasizing: the inter-scale context, and the computational complexity. To model the spatial context, Mignotte et al [79] proposed a Markov chain in scale $p(x^k|x^{K\backslash k}) = p(x^k|x^{k+1})$, where the intra-scale relation is a MRF $p(x_s^k|x_{S\backslash s}^k) = p(x_s^k|x_{\wp(s)}^{k+1}, x_{\mathcal{N}_s^k}^k)$, where $\wp(s)$ denotes the parent site of s at the parent scale and $\mathcal{N}_s^k$ defines a local neighborhood. Although this hierarchical model is computationally cheaper than one with a 3D neighborhood, a single-site sampler still needs to scan all of the pixels at every scale.

3.3.2 Frozen State Hierarchical Field

To achieve computational efficiency, a Frozen State Hierarchical Field (FSHF) was presented in [20] to synthesize binary images. In that work, a given HR field ($x = x^0$) can be represented by a hierarchical field $\{x^k\}$ (Fig. 3.8) where $x^k = \Downarrow^k(x^0)$, and $\Downarrow^k(\cdot)$ denotes a downsampling operator.

At coarse scales ($k > 0$), x^k is defined with a ternary state $x^k(s) \in \{0, 1, \frac{1}{2}\}$, where $0, 1$ (black, white) are determined states, and $\frac{1}{2}$ (grey) is undetermined. In terms of *modeling*,

a fine to coarse representation can be derived as

$$x_s^k = \begin{cases} 1 & \text{if } x_q^{k-1} = 1, \quad \forall q \in \Re^{k-1}(s) \\ 0 & \text{if } x_q^{k-1} = 0, \quad \forall q \in \Re^{k-1}(s) \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{3.18}$$

where $\Re^{k-1}(s)$ is the set of sites in scale $k-1$ corresponding to location s in scale k. Then, for *synthesis*, the key idea of the FSHF model is that, at each scale, only the sites which are undetermined need to be sampled, with those sites determined by the parent scale fixed (or frozen):

$$p(x_s^k | x_{S\setminus s}^k) = \begin{cases} \delta_{x_s^k, x_{\wp(s)}^{k+1}} & \text{if } x_{\wp(s)}^{k+1} \in \{0,1\} \quad \longleftarrow \text{Frozen} \\ p(x_s^k | x_{\mathcal{N}_s^k}^k) & \text{if } x_{\wp(s)}^{k+1} = \frac{1}{2} \quad \longleftarrow \text{Sampled} \end{cases} \tag{3.19}$$

With the frozen state, large scale features captured at the coarse scale are frozen and maintained to the fine scale, regardless of annealing schedule or sampling method. Since the "grey" interface between black and white determined regions represents only a small fraction of most images (Fig. 3.9), this approach offers a huge reduction in computational complexity relative to standard, full-sampling hierarchical techniques. The site sampling strategy corresponding to (3.19) is

$$\hat{x}_s^k = \begin{cases} \hat{x}_{\wp(s)}^{k+1} & \text{if } \hat{x}_{\wp(s)}^{k+1} \in \{0,1\} \\ \text{a sample from } \; p(x_s^k | x_{\mathcal{N}_s^k}^k) & \text{if } \hat{x}_{\wp(s)}^{k+1} = \frac{1}{2} \end{cases} \tag{3.20}$$

where $\hat{x}^k$ is the sampled (estimated) random field at scale k.

With a determined inter-scale relationship, the idea of FSHF modeling [20] inherits the advantage of hierarchical methods, such as the method of [79], which can provide a good capability to model scale-dependent structures by using simple models. At the same time, the FSHF brings significant computational benefits. Since a single-site sampler only needs to scan the interface between determined regions which represents a small fraction of most images, as shown in Fig. 3.9, the method offers a huge reduction in computational complexity relative to standard, full-sampling hierarchical techniques.

3.3.3 Modeling and Sampling

A hierarchical modeling and sampling approach as proposed in [2, 20, 72] will be applied in this thesis, which is shown in Fig. 3.11. This approach use a bottom-up modeling process

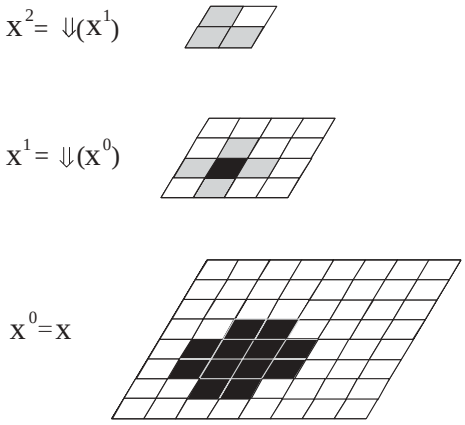


Figure 3.8: An example of ternary hierarchical subsampling [20]: a given field x^0 is coarsified by repeated 2×2 subsampling $\Downarrow(\cdot)$. All-white and all-black regions are preserved, with mixtures labeled as uncertain (grey).

starting from the finest scale, and a corresponding top-down sampling process starting at the coarsest scale.

Specifically, a hierarchical model, such as the FSHF, can be described by a Gibbs distribution (Section 2.1.3) at each scale k

$$p(x^k) = \frac{e^{-H^k(x^k)/T}}{\mathcal{Z}^k} \tag{3.21}$$

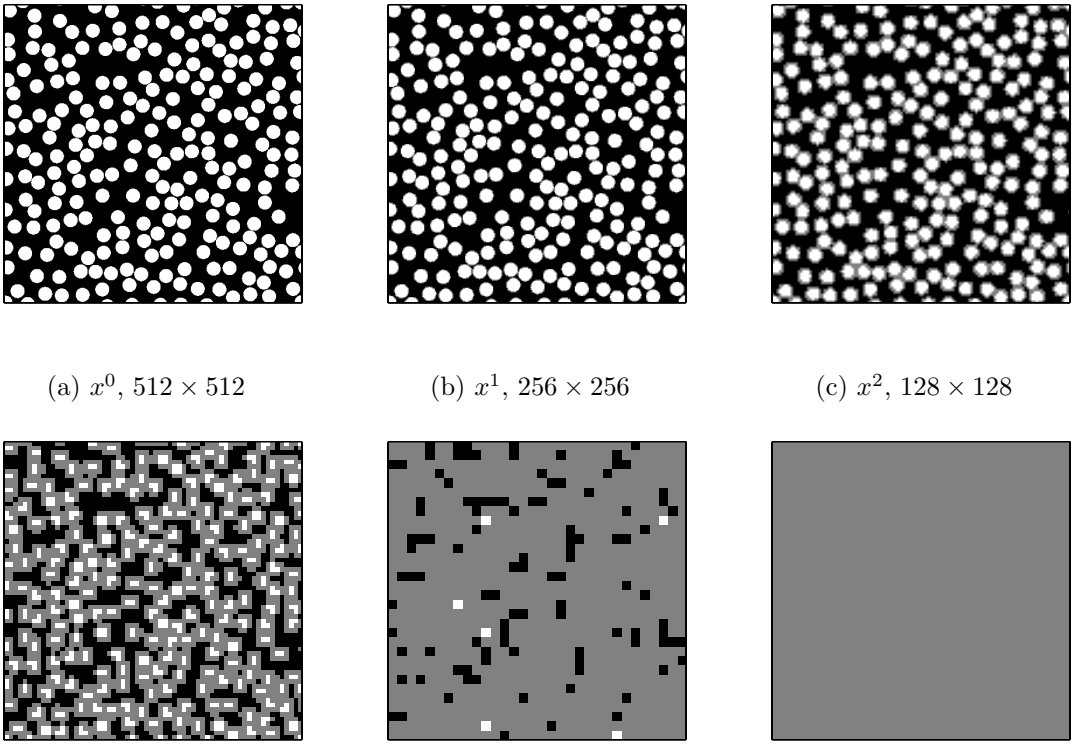
where T is the temperature, $\mathcal{Z}^k$ is the partition function, and $H^k(x^k)$ is the energy function at scale k.

A variety of methods [2, 104, 105] can be applied to define binary energy functions to model complex structures. In this thesis, we apply a local histogram model (Section 2.2) to set up the target and sample models. The energy function at each scale k is then defined as

$$H^k(x^k) = \|h^k(x^k) - H^k\| \tag{3.22}$$

where H^k is the target histogram learned from the training samples, and h^k is the sample histogram corresponding to x^k.

Once a model has been built up, random samples can be generated through MCMC samplers with simulated annealing as we have described in Section 2.3.



(a) x^0, 512×512 (b) x^1, 256×256 (c) x^2, 128×128

(d) x^3, 64×64 (e) x^4, 32×32 (f) x^5, 16×16

Figure 3.9: A frozen state down-sampling example. A binary image (a) with determined states (black, white) is repeatedly downsampled (b-f) until reaching a scale with all pixels undetermined (f). The uncertain label dominates at coarse scales (e,f), but is present only around boundaries at fine scales (b,c). Thus, under a frozen state hierarchy the number of undetermined pixels is only a small fraction of the original image size, 19% for this example.

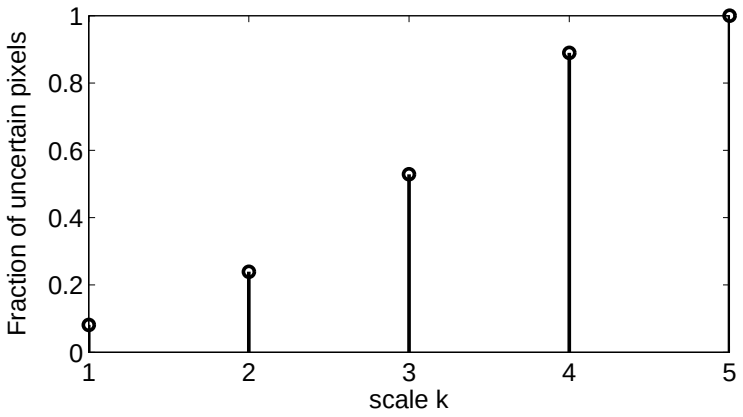


Figure 3.10: A plot of the fraction of uncertain pixels as a function of scale k. For the down-sampling example shown in Fig. 3.9, the portion of uncertain (grey) pixels drops as the scale k goes to finer scales.

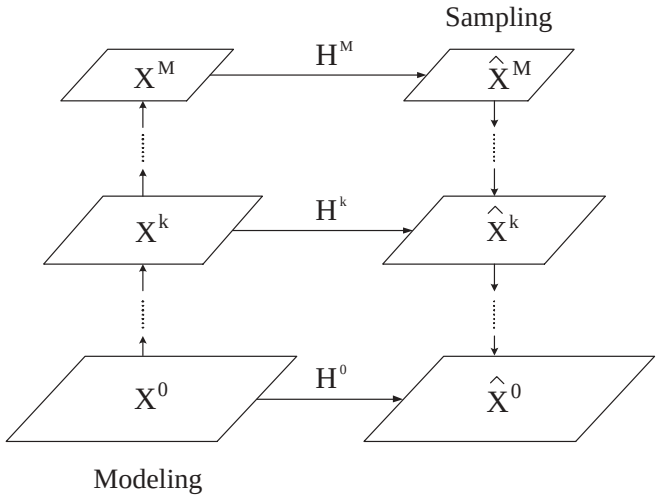


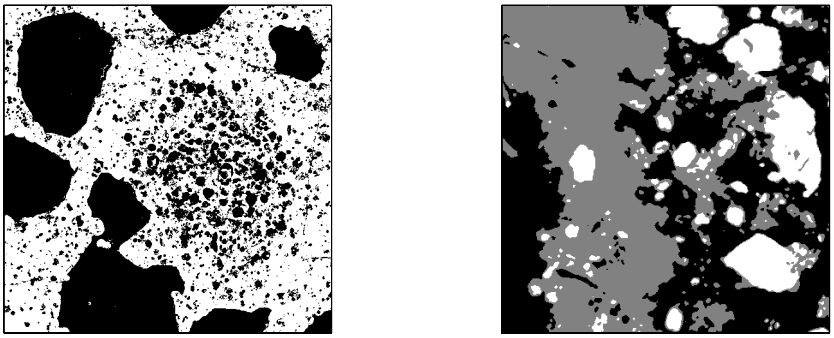
Figure 3.11: The hierarchical modeling approach is a bottom-up process, started at the finest scale to infer a model H^k at each scale k. Hierarchical sampling, on the other hand, is a top-down process starting at the coarsest scale, then sampling at each scale k, constrained by the model H^k and parent state $\hat{X}^{k+1}$.

Chapter 4

Problem Formulation

Random heterogeneous, scale-dependent structures are omnipresent and can be observed from many image sources, especially from remote sensing [17, 40, 100, 114] and scientific imaging [20, 50, 82]. Examples include a slice of microscope data showing cell nuclei with different sizes, a remote sensing image of a forest area with scattered small and large fires, or a telescope image of the observable universe with stars and galaxies having various forms and scales. Meanwhile, rather than the images of phenomena themselves, there are many image processing and analysis problems requiring to deal with *discrete-state* fields according to a labeled underlying property. For examples, the porosity information extracted from human bone images helps to analyze age-related bone changes; a labeled map of an infrared image indicates the distribution of noise sources; a remote-sensing landmass image segmented according to types of plants is for plant distribution study. In many cases, if discrete-state problems are associated with heterogeneous, multi-scale structures, we will have to deal with complex discrete state fields, such as shown in Fig. 4.1. Therefore, a fundamental difficulty will arise: how could we represent such complex, scale-dependent discrete-state fields? The focus of this thesis is to answer this question.

Certainly, for continuous-state problems scale-dependent modeling methods are common in image representation, analysis and processing, such as the widespread application of wavelets [29, 30, 74, 91], Gaussian and Laplacian pyramid methods [19, 85], and quadtree based models [6, 36]. In addition, multi-fractal analysis has been used to characterize self-similarity to study the statistics of natural images [23], to synthesize textures [23], and as a prior to regularize reconstruction problems [67]. Moreover, nonparametric example-



(a) A microscopic sample of a complex porous media image

(b) The underlying label field of a sea-ice sample

Figure 4.1: Both a microscopic sample (a) and the underlying label field of a RADARSAT-1 SAR sea-ice texture sample (b) have complicated multi-scale structures.

based methods have been proposed and successfully used in texture synthesis or enhancement [37, 65], which have the ability to handle spatial inhomogeneities by introducing patch-based representations.

On the other hand, the modeling problem for scale-dependent, discrete-state fields, which is emphasized in this research, has not been well studied in the literature — many related works are limited to modeling relatively simple situations [2, 20, 80]. However, in many cases the characteristics of discrete-state fields, such as porous media images (Fig. 4.1(a)) and the label fields underlying remote-sensing images (Fig. 4.1(b)), are diverse, complex, scale-dependent, and require further study.

In Chapters 5 and 6, we will describe our proposed modeling approaches in detail through a reconstruction of porous media images and a synthesis of remote-sensing images. In this chapter, we intend to formulate the scale-dependent modeling problem for discrete-state fields by briefly discussing existing methods, challenges and possible approaches.

4.1 Modeling Scale-dependent Structures

In Chapter 3, we reviewed MRF-based methods which can be used for representing scale-dependent discrete state fields. For a stationary field, no matter whether on a small

(Fig. 4.2(a)) or a large scale (Fig. 4.2(b)), the statistical characteristics of the field can be captured by a single field X with a stationary local model, such as a MRF with a small neighborhood (Section 3.1). However, a local stationary model cannot capture both small and large scales at the same time (Fig. 4.2(c)).

A nonlocal model, such as the MRF with a very large neighborhood, may be able to model such multi-scale structures, however, the training costs for learning models with high-order neighbourhood systems are expensive [24, 113], and how to simplify a complex neighborhood to a relatively simple form is an unsolved problem.

The widely used nonlocal discrete-state models, such as correlation [105] and chordlength distributions (Section 2.2.4), are more powerful than simple, local discrete Markov models. However although these models are effective at modeling microstructural information at a single scale, nevertheless they perform rather poorly in modeling phenomena at multiple scales [80].

Recently, a powerful modeling structure — the TMF/PMF (Section 3.2.2)— has been proposed to model complex nonstationary situations by using ternary/pairwise joint fields [8, 88]. This method can be applied to represent scale-dependent structures. For example, for a two-scale problem shown in Fig. 4.2(c), we can introduce two random fields X^1 and X^2 to describe the small and the large scale structures, respectively. Then, Z is defined as a joint field

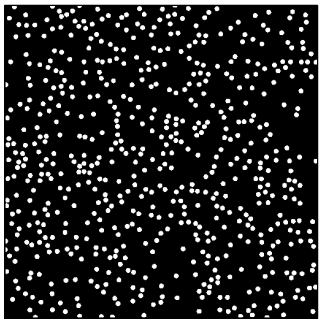
$$X \equiv f(X^1, X^2) \tag{4.1}$$

where, although X is scale-dependent, both X^1 and X^2 only have one scale and can be modeled by local stationary models.

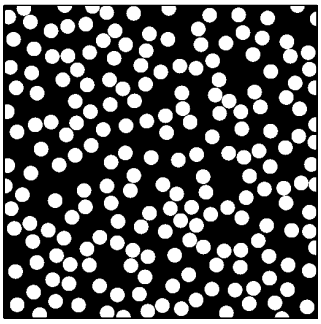
However, the TMF has its own problems: the modeling structure is limited to triplet fields, and the joint correlation $f(\cdot)$ is difficult to infer [16]. To develop a more practical method, we also consider introducing a modeling structure with N parallel fields, each of which has a local stationary neighborhood structure to model a single scale. In this case, we assume that X can be decomposed as a set of independent parallel fields $X = \{X^i\}$. Then, we have

$$p(x) \equiv \prod_{i=1}^{N} p(x^i) \tag{4.2}$$

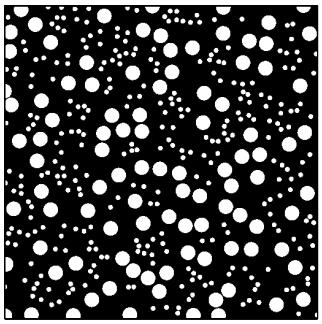
where $p(x^i)$ can be defined by simple local or nonlocal models. However, this approach also has limitations:



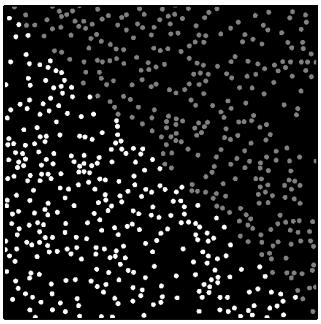
(a) Stationary, small scale



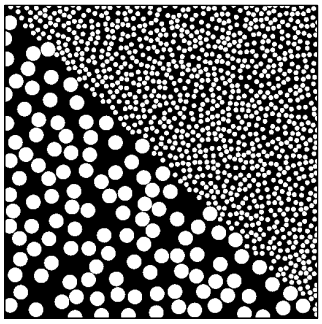
(b) Stationary, large scale



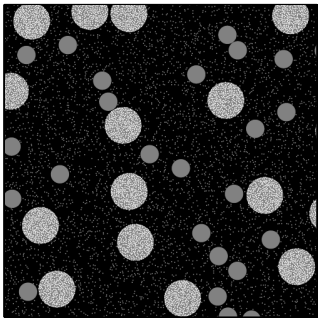
(c) Stationary, scale-dependent



(d) Nonstationary, small scale



(e) Scale-dependent, nonstationary



(f) Scale-dependent, multi-label, nonstationary

Figure 4.2: Synthetic image examples of discrete fields with different scale-dependent structures.

- The multiple parallel field method is based on an assumption that scale-dependent structures of a field are spatially independent of each other. However, for those cases where the multi-scale structures are spatially correlated (parallelly unseparable), this method will have a model approximation [70].

- Generating large-size image fields is computationally expensive when using common single-site samplers (Section 2.3.1 and 2.3.2) to update all of the pixel sites, and to allow large-scale spatial structures to converge demands a long relaxation time for an annealing process (Section 2.3.3). Therefore, the computational cost of this multiple parallel field method will be very high for sampling a set of large fields, especially for those fields carrying large-scale structures.

To reduce the modeling approximation in using multiple parallel fields, a more natural approach for modeling scale-dependent structures is to gradually model their spatial features decomposed at different scales. The idea can be realized by setting up hierarchical fields at a series of scales from large to small and using a local model to capture features at each scale [2, 20]. A hierarchical model, Markov in scale, with $M+1$ scale levels can be written as

$$p(x) \quad = \quad \Big[\prod_{k=0}^{M-1} p(x^k | x^{k+1}) \Big] \cdot p(x^M) \tag{4.3}$$

Further, to reduce computational cost an efficient approach is to apply the Frozen State Hierarchical Field (FSHF) model (Section 3.3.2), which is defined as

$$p(x_s^k | x_{S \setminus s}^k) = \begin{cases} \delta_{x_s^k, x_{\wp(s)}^{k+1}} & \text{if } x_{\wp(s)}^{k+1} \in \{0, 1\} \quad \longleftarrow \text{Frozen} \\ p(x_s^k | x_{\mathcal{N}_s^k}^k) & \text{if } x_{\wp(s)}^{k+1} = \frac{1}{2} \qquad \longleftarrow \text{Sampled} \end{cases}$$

For the FSHF, at a given scale only the states undetermined by the coarser level are allowed to be sampled and forces the rest pixels to be fixed (frozen). Based on the FSHF, large-scale structures are modeled at a sequence of scales, where in general the proportions of undetermined pixels are large at coarse scales, but gradually narrowed down at finer scales (Section 3.3.2). Since the number of pixels to update for a hierarchical field is smaller than that for a flat field, the computational cost will be significantly reduced, especially for those fields with large scale structures.

4.2 Challenges and Proposed Approaches

In general, scale-dependent structures may possess other complex attributes, such as having spatially nonstationary behaviour. As a simple synthetic example, an image X is shown in Fig. 4.2(d), where balls are at the same scale, but have two nonstationarily distributed label states. A hidden field (Section 3.1.2) model can easily handle this nonstationary image by introducing an additional underlying field U. Defining U to capture the spatial distribution of the label state of balls, X obtains a conditional stationarity given U. Then, a modeling problem about (X, U) (Fig. 4.2(d)) can be addressed by modeling a stationary field $X|U$ and a prior U:

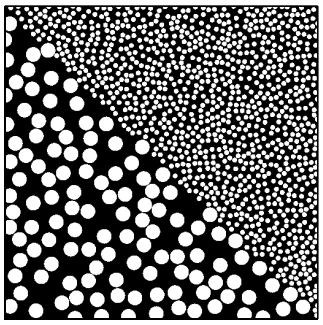
$$p(x, u) = p(x|u)p(u) \tag{4.4}$$

where both $X|U$ and U can be described by simple local/nonlocal models as discussed in Section 2.2.

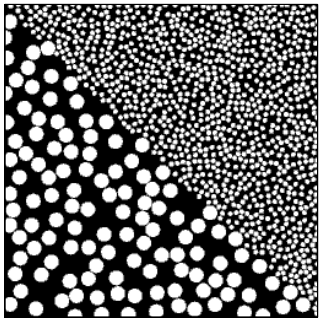
However, for a discrete field X with scale-dependent, nonstationary structures, it cannot be appropriately represented by either by a hidden field or by hierarchical methods based on simple models. For example, for a synthetic two-scale field with separately distributed large-small structures as shown in Fig. 4.2(e), even given a hidden field U to capture the label nonstationarity, $X|U$ still has large and small structures which cannot be represented by a simple local/nonlocal model. Moreover, when we define this two-scale nonstationary field with a FSHF $\{X^k\}$ (Section 2.3.3 and 3.3.2), at each scale the field X^k is still nonstationary, as demonstrated in Fig. 4.3(b-e), which also cannot be well modeled with simple local/nonlocal models.

Further, in more complicated cases, instead of only having a binary state, scale-dependent structures may have different label states associated with multi-model behaviors as shown in Fig. 4.2(f). In general, it would not be a trivial step to extend the existing local histogram model for binary microstructure modeling (Section 2.2.3) to a multi-label model. Since the number of possible configurations is decided by an exponentiation with the number of states as the base, a multi-label field will lead to a much larger configuration space than the binary one. Therefore, more sophisticated and practical approaches are desired to overcome the modeling challenges of complex scale-dependent behaviours.

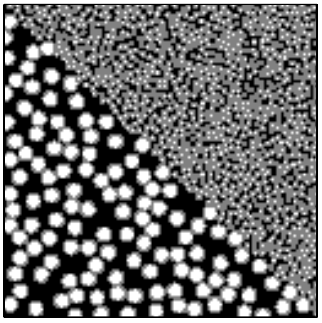
In this thesis, we propose the Hidden Hierarchical Markov Field (HHMF) model by combining the idea of a hierarchical field (Section 3.3) with a hidden field (Sections 3.1



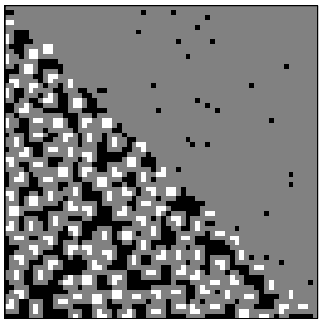
(a) x^0, 512×512



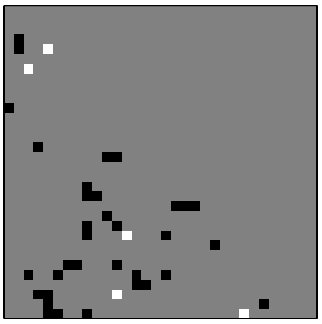
(b) x^1, 256×256



(c) x^2, 128×128



(d) x^3, 64×64



(e) x^4, 32×32

Figure 4.3: A two-scale synthetic example under a frozen state down-sampling.

and 3.2). The considerable success of hidden field models for representing complex nonstationarities is highly motivating in the sense that they allow more complex models than would otherwise be possible. Meanwhile, the computational efficiency of the FSHF approach (Section 3.3.2) strongly motivates the use of hierarchical discrete fields. Therefore, the proposed HHMF aims to take the advantages from both of them and offers a modeling approach with a capability to capture scale-dependent, nonstationary, spatial structures by using simple local models through a hierarchy.

At the same time, the proposed HHMF approaches are based on such an assumption that a complex scale-dependent problem can be simultaneously or repeatedly partitioned to yield a set of simple problems. Since many complicated phenomena are actually composed of simple ones, such as the synthetic problems in Fig. 4.2(e,f), this assumption can hold in many situations.

However, to implement a HHMF model is not straightforward. The most important problem is how to assert the HHMF modeling structure to make a complex problem simple again. Moreover, is it good sufficient to use a single HHMF to handle all the possible modeling situations? If not, how many HHMFs do we need? And how could we infer the relationship among those HHMFs? Further, how to infer the relationship between HHMFs and observable fields? It is significant that the answers for those questions are highly problem-dependent, for example, to well handle the synthetic samples in Fig. 4.2(e) and (f) need different HHMF modeling structures. It is hard to define a general structure to cover all possible situations. Thus, in the following two chapters we are going to illustrate the proposed HHMF modeling approaches through two application problems: porous media image reconstruction and remote-sensing image synthesis.

Chapter 5

Parallel Hidden Hierarchical Fields for Multi-scale Reconstruction

Porous media images are typically complex, nonstationary and possessing structures at multiple scales. Because Markov random fields are poor at representing other than single-scale phenomena, the classic Hidden Markov Model is a relatively poor fit to complex scenes such as porous media, and a single hierarchical field with local models cannot handle scale-dependent heterogeneous behaviours. Consequently this chapter proposes a Parallel Hidden Hierarchical Markov Field method for images having structure on more than one scale based on a prior model with multiple hidden hierarchical fields. We illustrate the effectiveness of the model with the reconstruction of porous media from low-resolution measurements.

5.1 Introduction

Porous media are materials containing heterogeneous pores, where the examples include cement, concrete, cartilage, bone, wood, and soil. Their application covers many science and engineering areas, such as geosciences, biology, material science, petroleum engineering, construction engineering,and environmental engineering. In the study of porous media, scientific imaging plays a significant role, especially with the supporting from more and

This chapter is written based on the papers: [68], [69] and [70].

more imaging tools including magnetic resonance imaging (MRI), micro-CT, X-ray, confocal microscopy, and so on.

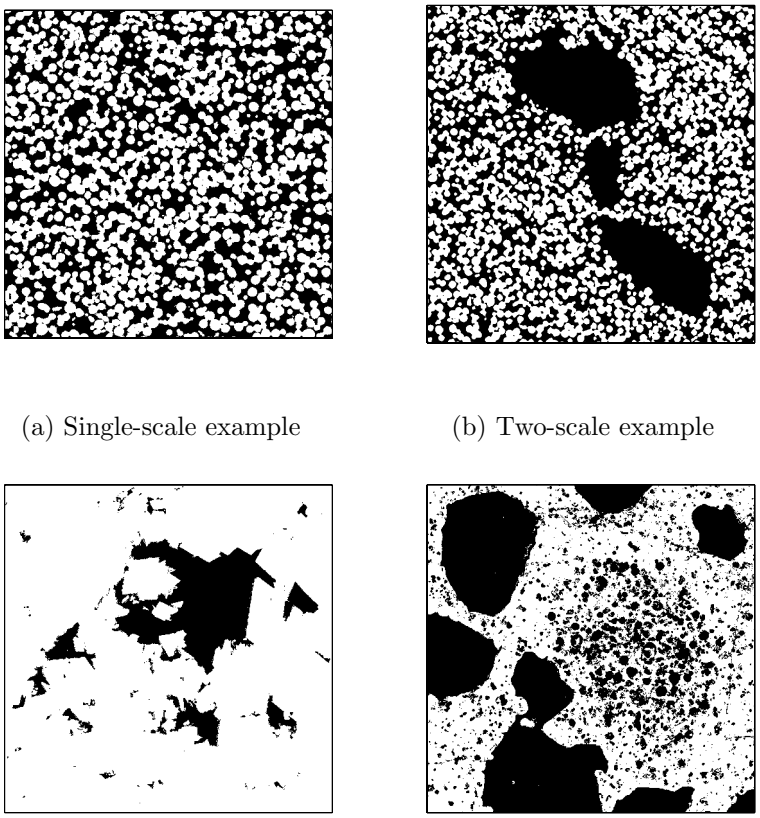
However, to obtain High Resolution (HR) binary images (Fig. 5.1) of internal structures from a porous media sample is difficult, which requires to physically process the sample with cutting, polishing, etc. [80]. Alternatively, Low Resolution (LR) images can be noninvasively observed by MRI 3D imaging, but the resolution is not enough to resolve small-scale pores (Fig. 5.2). Therefore, researchers have published papers on synthesizing [2, 20, 104] or reconstructing HR image from LR measurements [80, 82, 116]. In this chapter, the work is focused on the reconstruction. Since porous media images are typically complex, nonstationary and possessing structures at multiple scales (Fig. 5.1), to generate their HR samples is challenging.

Hidden Markov Fields (HMFs) are widely used in image restoration and resolution enhancement [39, 80, 105], however because local MRF/Gibbs models and the widely used non-local models, which are discussed in Chapters 2 and 3, can only describe structures at a single scale, most methods fail to produce convincing fine-scale and coarse-scale reconstructions.

The Frozen State Hierarchical Field (FSHF) method, as discussed in Section 3.3.2, has attractive computational complexity and scale-dependent modeling. The goal of this chapter is the extension of FSHF to allow hidden fields.

Generally, a single hierarchy with a scale-dependent model can capture a stationary structure (Fig. 5.1(a)), whereas many random fields have some sort of nonstationary piecewise multi-model behaviour which requires additional hidden fields (Section 3.2). Although multiple hidden fields are routinely used in Markov modeling, asserting a hierarchical context creates additional subtleties. Recently, Scarpa et al. [101] proposed a hierarchical texture model which represents texture at the region level with a superimposed finite-state hierarchical model. Their approach has some similarities with ours, but focuses on unsupervised model inference, whereas our approach requires more accurate, supervised models, with an emphasis on computational tractability for large problems.

In this chapter, we explore the combination of hierarchical and hidden fields to perform reconstruction for complex, nonstationary problems. We have chosen to apply our methods to reconstruct scientific images of porous media, such as the ones shown in Fig. 5.1, since the images include multiple challenging behaviours, with fractal-like scale-dependent



(a) Single-scale example

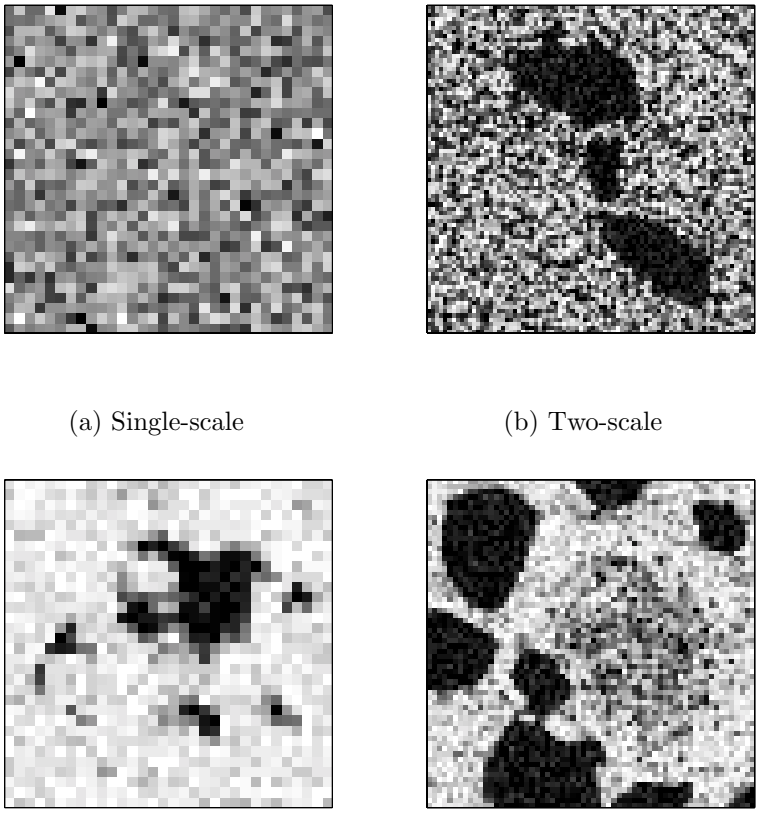
(b) Two-scale example

(c) Multi-scale example

(d) Complex, multi-scale, multi-model

Figure 5.1: Excerpts from microscopic images of physical porous media. A variety of behaviours can be observed, such as (a) single-scale stationary behaviour, (b) two-scale nonstationary behaviour, (c) near-fractal multi-scale behaviour, (d) multi-scale, multi-model behaviour. To correctly model these complex multi-scale behaviours in general poses a significant modeling challenge.

structures.



(a) Single-scale

(b) Two-scale

(c) Multi-scale

(d) Complex, multi-scale, multi-model

Figure 5.2: Noisy low resolution measurements for the porous media images shown in Fig. 5.1.

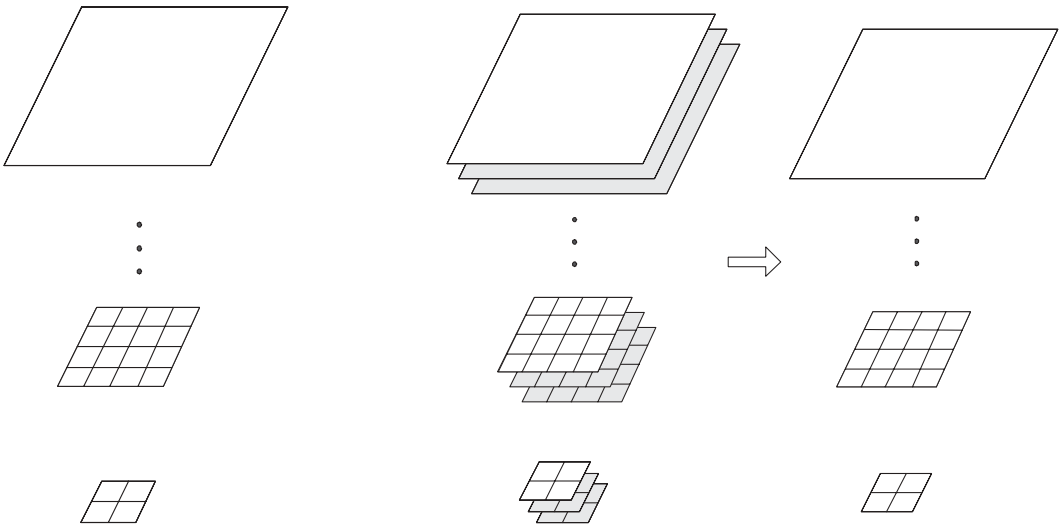
5.2 Markov Random Field Models

Based on the classical HMF framework (Section 3.1.2), image reconstruction can be achieved by estimating a hidden random field X from an observed field Y, where $Y = \{Y_s : s \in S_L\}$ is defined on a LR grid space S_L, and $X = \{X_s : s \in S_H\}$ is defined on a HR grid space S_H. The relationship between X and Y is expressed by a forward model $Y = g(X) + \nu$, where ν denotes the measurement noise. If ν is i.i.d, the classical HMF is written as

$$p(x|y) \propto \prod_{s \in S_L} p(y_s|x)p(x) \tag{5.1}$$

(a) A single field

(b) Multiple fields



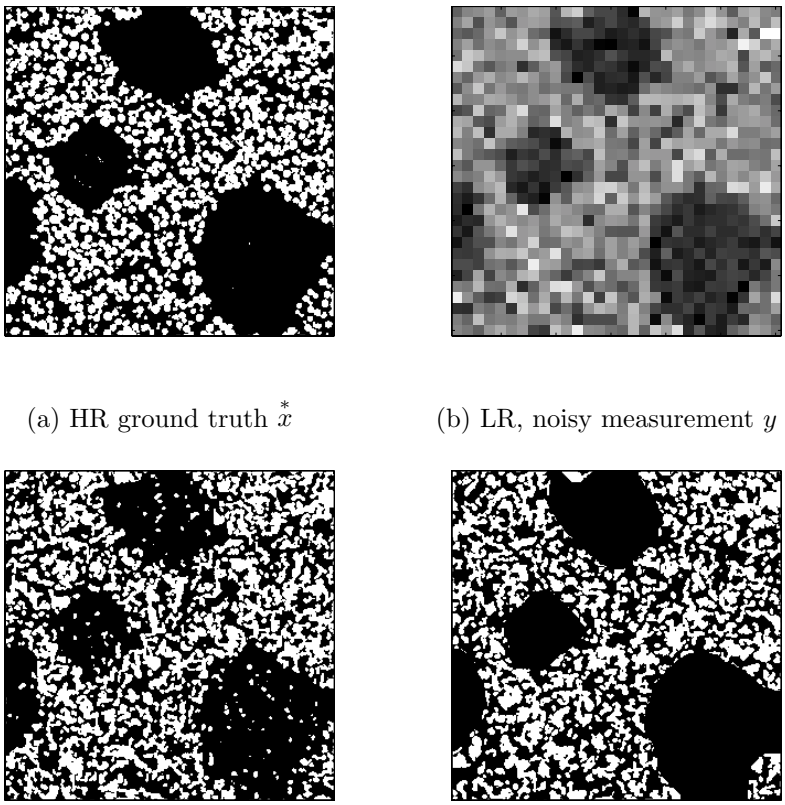
(c) A single hierarchical field

(d) A hierarchical field (right) with multiple hidden hierarchies (left)

Figure 5.3: Possible structures for modeling scale-dependent multi-model behaviour. This chapter focuses on case (d).

where X is assumed to be MRF. However, a single local MRF (Fig. 5.3(a)) cannot perform well in modeling a multi-scale nonstationary X, as seen in Fig. 5.4.

The TMF (Section 3.2.2) is a more general framework, assuming a joint triplet random field (Y, X_1, X_2) to be Markov. We can apply the joint field (X_1, X_2) to capture the multi-model behaviour present in a scene, such that both X_1 and X_2 can be stationary. The idea can be extended, in principle, to multiple hidden fields $(X_1, \cdots, X_{N_f})$, where we can use multiple flat fields (Fig. 5.3(b)) to model multi-scale phenomena. Then, the reconstruction



(a) HR ground truth $\overset{*}{x}$

(b) LR, noisy measurement y

(c) HR estimate $\hat{x}$ from classical HMF [80]

(d) HR estimate $\hat{x}$ from multiple parallel MRFs [70]

Figure 5.4: Suppose we have a two-scale structure (a) with low resolution measurement (b). A flat Markov model is poor at simultaneously modeling the large black pores and tiny spaces between beads, therefore the HR result (c) using a classical HMF model cannot strongly assert the presence of large scale structures. Instead, (d) shows that multiple flat MRFs lead to an improved result with this two-scale case.

can still be addressed under the HMF framework:

$$p(x|y) \propto \prod_{s \in S_L} p(y_s|x_1, \cdots, x_{N_f}) \prod_{i=1}^{N_f} p(x_i) \tag{5.2}$$

where each X_i is assumed to be Markov. The modeling improvement from (5.1) to (5.2) is obvious (e.g., comparing Fig. 5.4(d) with Fig. 5.4(c)). However, this multiple MRFs

method has its limitations: it is only tractable for cases with a limited number of scales (N_f), it ignores any dependencies among the different scale structures, and flat MRFs will have difficulty in modeling large nonlocal structure.

Under the HMF framework, if we use a hierarchical field (Fig. 5.3(c)) discussed in Section 3.3, to model a HR field X, a reconstruction can be achieved by

$$p(x|y) \propto \Big[\prod_{s \in S_L} p(y_s|x) \Big] \cdot \Big[\prod_{k=0}^{M-1} p(x^k|x^{k+1}) \cdot p(x^M) \Big] \tag{5.3}$$

where $k = M$ denotes the coarsest scale of X and the prior is defined as FSHF.

A FSHF can work well in modeling stationary scale-dependent structures, however not with nonstationary, piece-wise behaviour, because conditioned on X^{k+1}, X^k will still have nonstationary features which cannot be captured by a single model.

5.3 Hidden Hierarchical Markov Fields

The considerable success of hidden field models (Section 3.1.2 and 3.2) for modeling complex nonstationarities is highly motivating in the sense that they allow more complex models than would otherwise be possible. At the same time, the computational efficiency of the FSHF (Section 3.3.2) approach strongly motivates the use of hierarchical discrete fields. However, the combination of these two ideas has not yet been explored; in this chapter, we propose a Hidden Hierarchical Markov Field (HHMF) (Fig. 5.3(d)) model to model complex scale-dependent behaviours.

5.3.1 Single Hidden Hierarchy

We start with the simplest possible hidden model, in which the nonstationarity in X can be entirely attributed to a single binary hidden field U. Both X and U are hierarchical, asserted to be Markov in scale, so the joint relationship can be written as

$$p(x,u) = p(x|u)p(u) = \left[\prod_k p(x^k|x^{k+1},u^k) \right] \left[\prod_k p(u^k|u^{k+1}) \right] \tag{5.4}$$

We select the coarsest scales of X and U, k_x and k_u respectively, at which determinable state structure appears, such that (5.4) becomes

$$p(x,u) = \left[\prod_{k=0}^{k_x-1} p(x^k|x^{k+1},u^k)\right] p(x^{k_x}|u^{k_x}) \left[\prod_{k=k_x}^{k_u-1} p(u^k|u^{k+1})\right] p(u^{k_u}) \tag{5.5}$$

Since U describes the model behaviour in X, the determinable state in X is expected to vanish at a finer scale than in U ($k_x < k_u$).

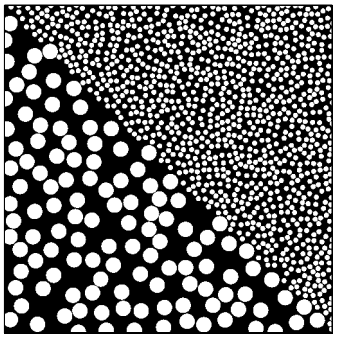
As illustrated in Fig. 5.5(e) versus Fig. 5.5(c,d), the introduction of a hidden field allows for a superior behaviour separation. However two clear issues remain: first, a great many problems (e.g., Fig 5.1(d)) cannot be represented by a single binary hidden field; second, in Fig. 5.5(e) the hidden field U was given or known, whereas in practice it must be estimated. These issues form the basis for the following two sections.

5.3.2 Multiple Hidden Hierarchies

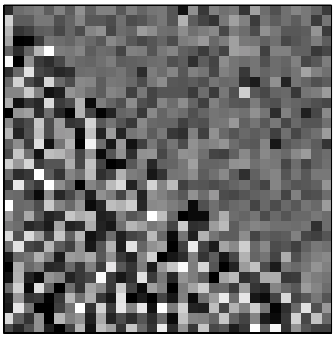
In most cases, the behaviour of a random field X will be determined by more than one spatial variable, such that X remains nonstationary when conditioned on a single binary field U. The obvious solution to this problem is to define U as a multi-label field; for example, the behaviour of Fig. 5.6(a) is determined by two binary variables of scale and shade, corresponding to a quad-label hidden field.

Although multi-label models (e.g., Potts [93]) do exist, in practice the established set of spatial binary models is much richer, and the size N_c of the configuration space of a local joint model grows terribly large, $N_c = N_l^{n^2}$ for N_l labels in an $n \times n$ two-dimensional neighbourhood. However more troubling is the representation of uncertainty in the frozen hierarchy, since N_l labels imply $\binom{N_l}{2}$ pairwise uncertainties, $\binom{N_l}{3}$ three-way uncertainties, etc.

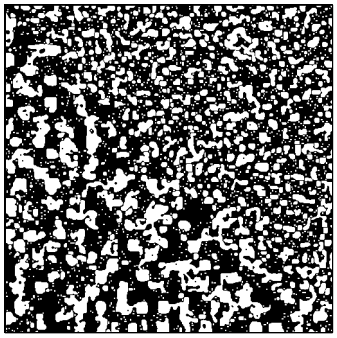
To address the challenge of multi-label modeling, we extend the HHMF with a single hidden hierarchical field to a more general parallel HHMF. The key idea of the parallel HHMF is to maintain tractability by introducing multiple binary hidden label fields $U = \{U_i, i \in (0, \cdots, N_{v-1})\}$, such that each field U_i is on a hierarchy $\{U_i^k\}$, where each hierarchy $\{U_i^k\}$ is used to capture a single binary structure or model behaviour. The hidden fields



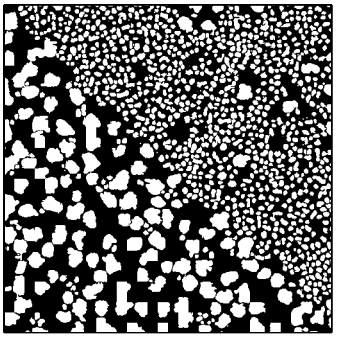
(a) Ground truth, $\overset{*}{x}$, 512×512



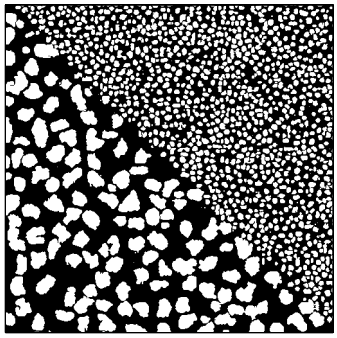
(b) LR, noisy y, 32×32



(c) HR estimate $\hat{x}$ based on single MRF, 512×512



(d) HR estimate $\hat{x}$ based on FSHF, 512×512

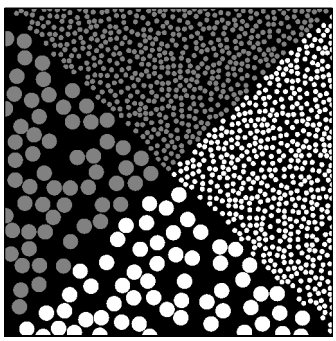


(e) HR estimate $\hat{x}$ based on HHMF, 512×512

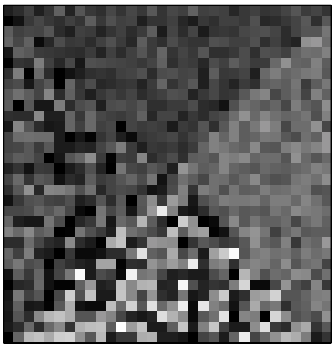
Figure 5.5: The reconstruction of a two-scale image (a) from low resolution measurements (b) with different Markov field frameworks. The clear scale separation of the result from the hidden hierarchy [69] (e) should be compared to the results from a single flat MRF model [80] (c) and a single hierarchical model [20] (d). In (e), the hidden field u is given, and not estimated.

remain Markov in scale, so (5.4) becomes

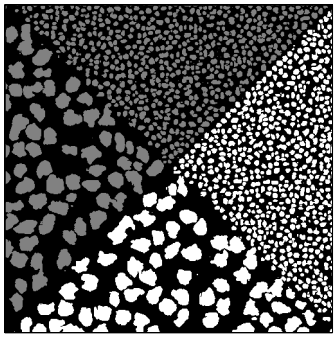
$$p(x,u) = p(x|u)p(u) = \left[\prod_k p(x^k|x^{k+1},u^k)\right]\left[\prod_k p(u_0^k,\cdots,u_{N_v-1}^k|u^{k+1})\right] \qquad (5.6)$$



(a) HR image with two spatial variables



(b) LR, noisy y, 32×32



(c) HR estimate $\hat{x}$ from HHMF with two binary hidden hierarchies, 512×512

Figure 5.6: To reconstruct a two-scale image with multiple intensity states (a) from a low resolution measurements (b) requires two hidden fields: one for scale and one for shading. The modeling capability of a hierarchical model with two independent hidden fields is demonstrated in (c), showing that the reconstructed result has a clear scale and shade separation.

As in (5.5) we limit the scales for X and U, so (5.6) becomes

$$p(x,u) = \left[\prod_{k=0}^{k_x-1} p(x^k|x^{k+1},u^k)\right] p(x^{k_x},u^{k_x}) \left[\prod_{k=k_x}^{k_u-1} p(u_0^k,\cdots,u_{N_v-1}^k|u^{k+1})\right] p(u^{k_u}) \qquad (5.7)$$

The key question is what may be assumed regarding the hidden fields $\{u_i\}$. In rare cases (e.g., Fig. 5.6) the hidden fields may be independent, however normally (e.g., Fig. 5.1(d))

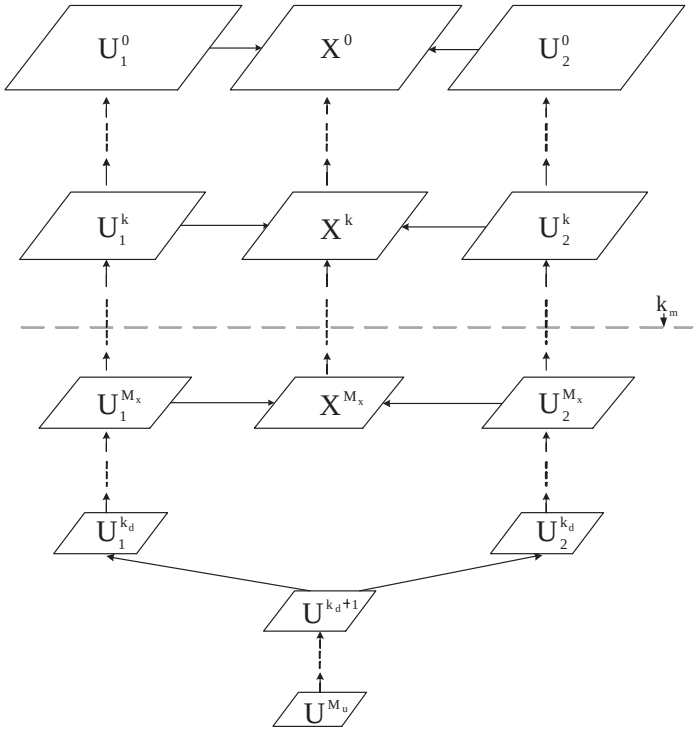


Figure 5.7: An example of the proposed parallel Hidden Hierarchical Markov field model with two hidden fields U_1, U_2, such that the hidden label field U is a joint field only at coarse scales. As the features of different model behaviour become separable at some scale k_d, U is decoupled to multiple parallel hierarchical fields. At scales coarser than k_x the entire random field X of interest is uncertain, and so only U is represented to scale k_u.

that will not be the case. However, very likely that the fields are conditionally independent at sufficiently fine scales. That is, if the coupling between hidden fields (e.g., a spatial exclusion) is asserted by the coarser scales, with finer scales primarily serving to refine the hidden-state boundaries, then the fine scales $u^k, k < k_d$ become conditionally independent:

$$p(u_0^k, \cdots, u_{N_v-1}^k | u^{k+1}) = \prod_i p(u_i^k | u_i^{k+1}) \qquad k < k_d \tag{5.8}$$

Joint fields $\{U^k, k > k_d\}$ may be needed at coarse scales $k > k_d$, however the limited number of state elements at coarse scales make such a joint model feasible:

$$p(u) = \left[\prod_{k=0}^{k_d} \prod_i p(u_i^k | u_i^{k+1}) \right] \left[\prod_{k=k_d+1}^{k_u-1} p(u^k | u^{k+1}) \right] p(u^{k_u}) \tag{5.9}$$

Fig. 5.7 shows an example with two hidden variables ($N_v = 2$). At finer scales, the hidden fields are spatially decoupled, binary, and simply modeled. The complex, joint hidden structure appears only at very coarse scales, where the small number of state elements allows such a structure to be computationally tractable.

This approach simplifies modeling in three significant ways: first, the parallel HHMF consists entirely of simple models, both local and stationary. Specifically, although X^k and U^k may have complex, non-local behaviour, the conditional residuals $(U^k|U^{k+1})$, $(X^k|X^{k+1}, U^k)$ are local, by virtue of the fact that all non-local matters have been absorbed into the conditioned (coarser) scale.

Second, in general the hidden fields will be dependent upon each other. However, they can be *conditionally* independent at finer scales, because the dependency among the fields will be captured at coarse scales. Therefore, conditioned on those coarse scales, the hidden fields at finer scales are in fact conditionally independent. At coarse scales, $k > k_d$ where only few pixels exist, it is computationally tolerable to assert a joint model for U^k, where the joint model is needed to allow the hidden models to interact (Fig. 5.8(b)). In most problems, empirically, the assumption of conditionally independent fine-scale hidden field is therefore reasonable.

Third, because $\{X^k\}$ and $\{U_i^k\}$ are modeled using simple, binary models, $\{X^k\}$ and $\{U_i^k\}$ are easily defined as hierarchical frozen states, leading to the computational cost of the parallel HHMF being linear in the number of hidden fields N_v, except at coarse scales.

5.3.3 Reconstruction

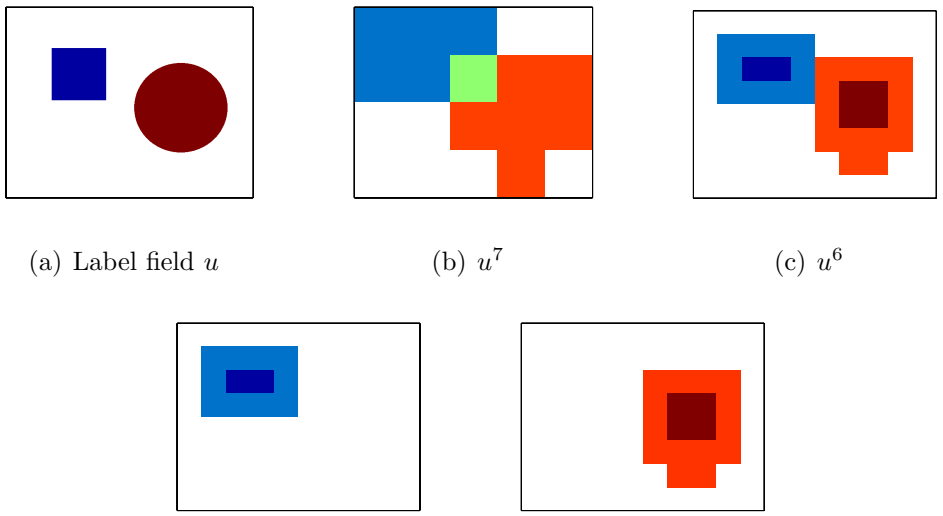
For image estimation, the hidden fields are invisible to the measurements, therefore $p(y|x^k, u^k) = p(y|x^k)$, and so the reconstruction model illustrated in Fig. 5.7 can be written as

$$p(x, u|y) \propto \prod_{s \in S_L} p(y_s|x) p(x, u) \tag{5.10}$$

where $p(x, u)$ is the parallel HHMF prior, defined as in (5.7) and (5.9), and where the measurements $p(y_s|x)$ are taken at some scale k_m.

Given measurements Y contaminated by i.i.d. noise, the posterior distribution of (X, U, Y) can be represented as a Gibbs distribution

$$p(x, u|y) \propto exp\big(-\frac{1}{T} H(x, u|y)\big) \tag{5.11}$$



(a) Label field u (b) u^7 (c) u^6

(d) u_1^6 (e) u_2^6

Figure 5.8: An example of label field U with two dependent features (a). At some coarse scale the features of different behaviours interact and lead to complex joint states (b), however at some finer scale the two behaviour in U become separable (c), and can be decoupled to multiple independent ternary fields (d,e).

where T is the temperature, such that H is the energy function implying the probability density p. Finding a good estimate $\hat{x}$ therefore corresponds to maximizing $p(x, u|y)$, correspondingly minimizing H, the sum of hidden joint $H_u^k(u^k|u^{k+1})$, decoupled $H_{u_i}^k(u_i^k|u_i^{k+1})$, visible $H_{x|u}^k(x^k|x^{k+1}, u^k)$, and measurement $H_m(y|x)$.

The prior models can be learned in many ways, since a huge number of Gibbs priors have been developed (Section 2.2). For simplicity, and to limit the assumptions made, we have selected as prior model a nonparametric joint local distribution, the exhaustive joint distribution of a local 3×3 neighbourhood of ternary state elements. The models are learned separately on each scale, based on downsampled training data $\tilde{x}^k = \Downarrow^k (\tilde{x}^0|\tilde{u}^0)$, $\tilde{u}_i^k = \Downarrow^k (\tilde{u}_i^0)$. The resulting energy function is the least-squares difference between the model and observed joint histograms (Section 2.2.3).

The measurement energy function is inferred from the given forward model $g()$. As with the prior model, a variety of measurements could be defined, depending on the measuring instrument, however in this chapter we focus on reconstruction from low-resolution images,

making $g()$ a downsampling operator.

To minimize $H(x, u|y)$, we need to anneal (Section 2.3.3) on each scale k in each field X, U_i, with consequent open questions: whether to minimize hidden states separately or jointly with the observable state, whether to minimize the scales in parallel or sequentially, and whether to have scale-dependent annealing schedules.

Normally the solution to the estimation problem, found by minimizing (5.11), is straightforward in principle. Here the minimization is much more subtle, in that (5.11) is defined over a hierarchy of multiple scales, in principle requiring the joint, simultaneous minimization over all fields and all scales. Because finer scales are highly dependent on coarser ones, joint minimization over scales is difficult, and the minimization almost certainly needs to proceed sequentially from coarse to fine scales. However at a coarse scale, it is not possible to interpret $p(x|u)$ if the hidden field is uncertain ($u = \frac{1}{2}$). Therefore, we propose, somewhat at odds with conventional practice in hidden fields, to first estimate U over all scales, such that the estimated hidden field is definitive (not uncertain) at every point, and then to estimate X.

The above argument, together with empirical testing, suggests that a constant annealing schedule, with sequential minimization over scales, and sequential minimization from hidden (U) to visible states (X), lead to a reliable and robust reconstruction. When estimating the hidden field U, in which case X is unknown, a model for $p(y|u)$ is required, and needs to be inferred empirically.

5.4 Computational Complexity

In this section, we discuss the computational complexity of our proposed model, from the perspective of both computational time cost and storage cost. The complexity of processing the proposed HHMF model is composed of the complexity of processing the image field X and the hidden field U.

Since we proceed estimations sequentially from U to X, the total processing time cost is

$$C_{HHMF} = C_X + C_U \tag{5.12}$$

where C_X and C_U denote the time cost of processing X and U respectively. C_U is dominated by the time of processing binary hidden label fields $\{C_{U_i}\}$

$$C_U \simeq \begin{cases} \max\{\, C_{U_i}, \forall i \in (0, \cdots, N_\nu)\,\} & \text{if parallel computing} \\ \sum_{i=0}^{N_\nu - 1} C_{U_i} & otherwise. \end{cases} \tag{5.13}$$

For each of those fields, the computational time is the time for processing a hierarchical field from coarse to fine scale, such as

$$C_X = \sum_{k=0}^{M_x} C_X^k \tag{5.14}$$

In this work, at a scale $k > 0$ an estimated field is represented as a ternary field with frozen state (3.18), and is generated from the posterior model (5.10) by Gibbs sampler (Algorithm 2.23) with Simulated Annealing (SA) (Algorithm (2.3.3). According to the strategy of the frozen state hierarchical annealing (Section 3.3.2), at a given scale only the pixels with undetermined state ($\frac{1}{2}$) are processed and others are frozen . Consider α_g^k to be the fraction of undetermined pixels, so that a sampler need to scan $N_g^k = \alpha_g^k \cdot N^k$ pixels instead of a total of N^k ($N^k = \frac{1}{4^k} N^0$) pixels at k scale. Therefore, the computational time per scale C^k (C_X^k or C_U^k) is determined by the number of SA updates τ^k, the fraction of pixels for sampling N_g^k, and the processing time for each pixel t_s, then we have

$$C^k = \tau^k \cdot N_g^k \cdot t_s \tag{5.15}$$

where t_s is determined by the measurement energy (H_m), the prior energy ($H_{u|x}^k$,H_u^k), and the number of label states (N_l^k) in a field. Consider the time of calculating energies for a label state l at site s is t_{sl}, then we have

$$t_s = \begin{cases} 3 \cdot t_{sl} & \text{for } X_s \in \{0, \frac{1}{2}, 1\} \\ 2 \cdot t_{sl} & \text{for } X_s \in \{0, 1\} \end{cases} \tag{5.16}$$

Now the only thing that remains unclear is the actual number of iterations required for SA convergence τ^k. In fact, τ^k is highly problem dependent and cannot be discussed quantitatively in the computational complexity. However, it has been demonstrated that the number of Monte Carlo updates (τ) can be approximated by a quadratic function of the correlation length (ξ^2) (2.27), which has been validated for Gaussian model [42]. Thus, for each scale we have

$$\tau^k \sim (\xi^k)^2 \tag{5.17}$$

Table 5.1: A comparison of the storage complexity

Model	**Single MRF**	**FSHF**	**Proposed HHMF**
Number of parameters	$2 \cdot 2^8 = 512$	$2 \cdot (M_x \cdot 3^9 + 2^8) = M_x \cdot 13122 + 512$	$2 \cdot (M_x \cdot 3^9 + 2^8) + 2 \cdot N_\nu \cdot (M_u \cdot 3^9 + 2^8) = (M_x + N_\nu \cdot M_u) \cdot 13122 + (N_\nu + 1) \cdot 512$
Size of domain	N	N	$(N_\nu + 1) \cdot N$

At same time, consider a fraction $\alpha_\xi^k = \frac{\xi^{k+1}}{\xi^k}$, then the number of SA updates is approximately reduced $(1 - (\alpha_\xi^k)^2)$ after subsampling.

In practice, because the number of SA updates is highly variable from model to model and from image to image, the total time cost of the proposed HHMF method C_{HHMF} is difficult to predict and only can be obtained experimentally. For example, according to our implementation with matlab and running in windows xp, it took an intel core 2 Duo 2.4GHz processor about 5 hours to obtain the result shown in Fig. 5.11(d), close to 2 hours to get the reconstruction shown in Fig. 5.13(d).

The storage cost of the proposed HHMF method is mainly decided by the size of the HR estimated fields N, the number of the HR fields $N_\nu + 1$, the number of prior model parameters, such as the parameter number of a local histogram model (2.19). In details, the major storage cost is listed and is compared with other MRF based methods, as shown in Tabel 5.1.

5.5 Experimental Results

In this section, we present the results of the proposed HHMF reconstruction in three experiments based on the measurement sets shown in Fig. 5.10. In order to assess its advantage in modeling complex multi-scale behaviour, the proposed approach is compared with both discrete-state and continuous-state methods. Since the HR images are discrete, reconstructions from the continuous-state methods are thresholded.

To reconstruct a scale-dependent, near fractal, piece-wise nonstationary image such as the porous medium in Fig. 5.1(c) is a major modeling challenge. The image in Fig. 5.1(c) displays three types of behaviour: large-scale pores, regions of high density, and background

areas of low density. We therefore propose the ternary hidden field U to be decoupled into two parallel binary hierarchies, where $\{U_1^k\}$ identifies the presence of large pores, and $\{U_2^k\}$ identifies regions of high density. From Fig. 5.1(c) we can see that, clearly, U_1 and U_2 are not independent, since a porous state is not permitted to assert in both fields at the same time.

The relationship $p(y|u)$ is found empirically and is modeled as shown in Fig. 5.9, which plots the empirical distribution $p(y)$ as a function of the nine possible joint relationships in U_1 and U_2. Because the hidden fields are decoupled, four of the nine joint relationships are inadmissible (shown as shaded, in the figure), and are modeled as uniform, with a low marginal probability.

We will be comparing the proposed HHMF against both related Markov methods and against conventional and state-of-the-art methods in wavelet reconstruction. In terms of Markov methods, we will compare to a single, non-hierarchical MRF [80] and the frozen state (FSHF) approach [20], hierarchical but with no hidden state.

For the same noisy observation, estimates are also obtained by thresholding the results from several continuous-state methods: zero-pad wavelet interpolation (ZPWT) [74], wavelet based parametric texture model (WTPM) [91], example-based super-resolution (EBSR) [37], and wavelet based example-driven parametric model (WTEDPM) [65]. For the ZPWT and the WTPM cases, we first de-noise based on biorthogonal wavelet filters (bior4.4) at the measurement scale (k_m), then enhance the resolution to the finest scale. For the ZPWT, a fine scale image is directly obtained from the inverse wavelet transform (WT) by zero-padding the fine scale bands. For the WTPM, parametric models are learned from the training sample ($\tilde{x}$) to constrain interpolation. When using the EBSR, we reduce the LR noise by matching patches of y with the training sample at the measurement scale $\tilde{x}^{k_m}$ and build $\hat{x}^{k_m}$ based on the best matching patches: at finer scales we estimate $\hat{x}^k$ by cubic interpolation and use it to get matches from $\tilde{x}^{k-1}$ to build the finer estimate $\hat{x}^{k-1}$ [37]. In the WTEDPM [65], every best matching patch is only taken as the initial driving data, then the WTPM follows to get the final finer scale patch.

In the first experiment, we reconstruct a two-scale porous media sample (Fig. 5.11(a)) from its LR noisy observation y (Fig.5.10(a)), with MRF and wavelet reconstructions shown in Fig. 5.11 and Fig. 5.12. The proposed HHMF with a single hidden hierarchy is used in this case. In comparing the estimated results ($\hat{x}$) with the true sample, we see that

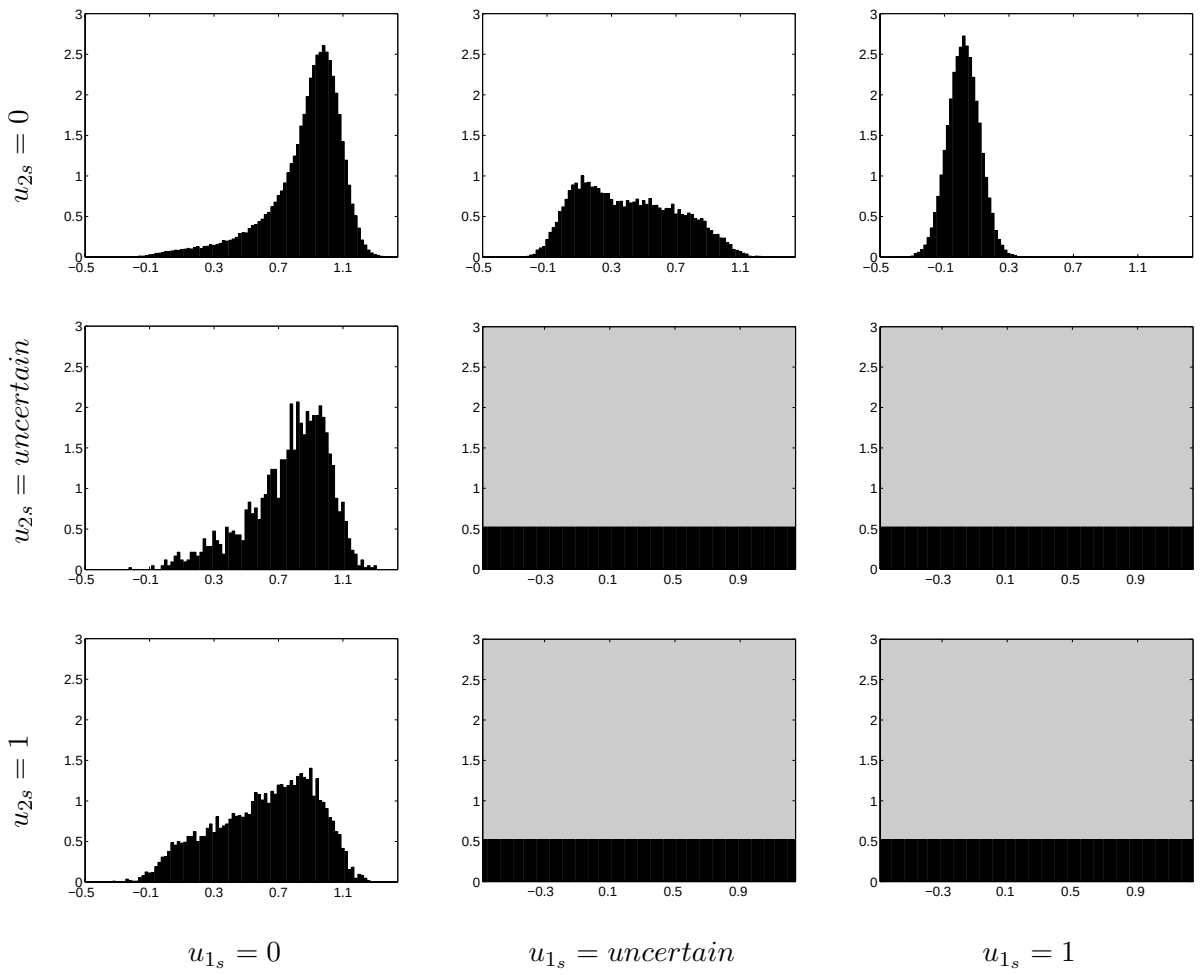
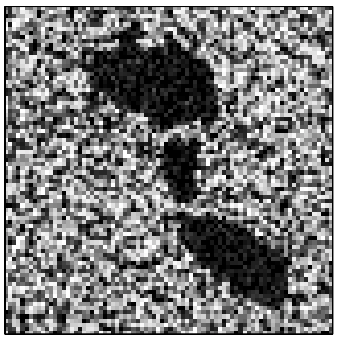
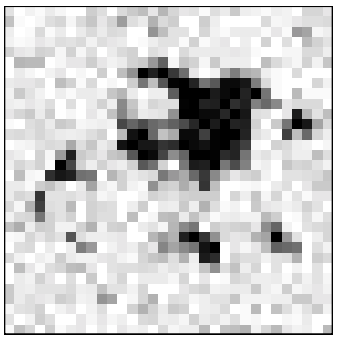


Figure 5.9: An example of the conditional target measurement histograms of the training measurements $\tilde{y} = g(\tilde{x})$ corresponding to Fig. 5.10(c) for decoupling a joint field U^{k_d} into two simpler fields $U_1^{k_d}$, $U_2^{k_d}$. Each panel shows the distribution of $\tilde{y}$ for one of nine possibilities on U_1, U_2. Since the hidden fields are asserted to be decoupled, those cases where both fields are asserted (shaded distributions) are never observed, and so are assigned a uniform distribution with low marginal probability. To the extent that the joint state configuration of $(U_{1_s}^{k_d}, U_{2_s}^{k_d})$ relates to distinguishable model behaviour in $\tilde{y}$, we expect the hidden fields to be estimatable.

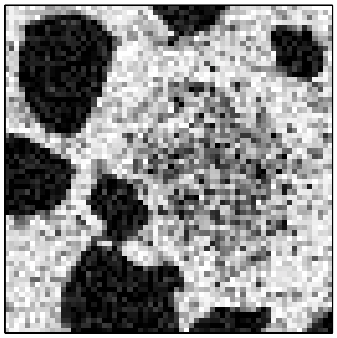
the HHMF (Fig. 5.11(d)) achieves better large-scale structure than the local flat MRF (Fig. 5.11(b)), and better small-scale details than the FSHF (Fig. 5.11(c)). Even the most superficial examination of the non-MRF results in Fig. 5.12 shows that the example-wavelet



(a)



(b)



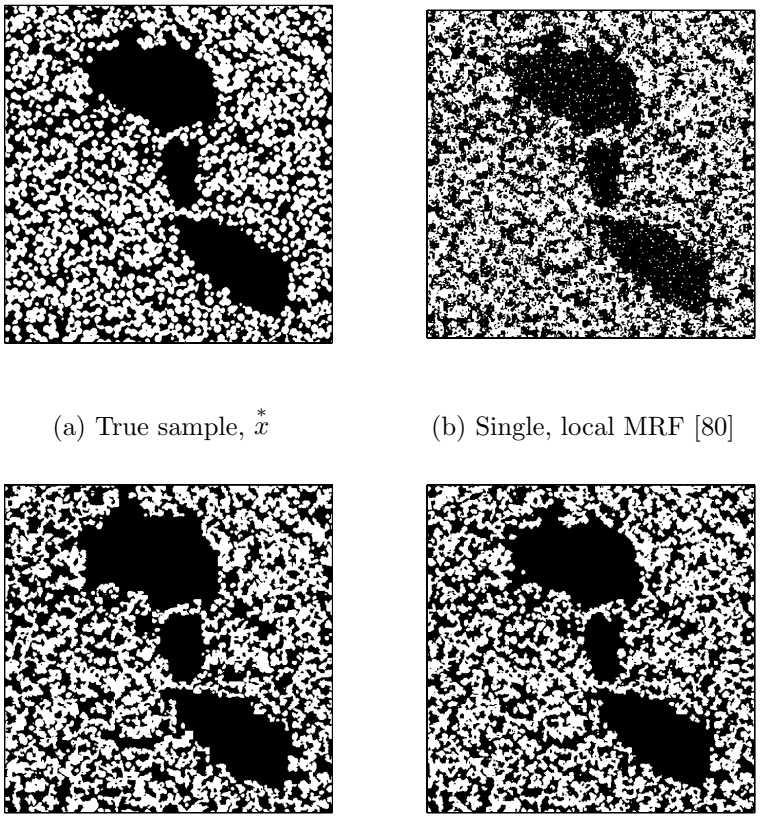
(c)

Figure 5.10: Noisy LR measurements $y = g(\overset{*}{x}) + \omega$ are generated from three microscopic samples ($\overset{*}{x}$) in Fig. 5.1 by downsampling $g()$, with added Gaussian noise ω (σ=0.1, SNR=14dB).

approach of the WTEDPM greatly outperforms the three other implemented methods — ZPWT, WTPM, and EBSR.

In the second experiment, we reconstruct a near-fractal multi-scale porous media sample (Fig. 5.13(a)) from its LR noisy observation (Fig.5.10(b)). Here, we use the HHMF with a single hidden hierarchy to get the reconstruction (Fig. 5.13(d)). Although the visible piecewise-flat facets in $\overset{*}{x}$ are not able to be fully reproduced, $\hat{x}$ from the HHMF in recovering multi-scale characteristics is clearly better than the other MRF methods (Fig. 5.13(b)(c)).

The reconstructions based on the continuous state methods are shown in Fig. 5.14,



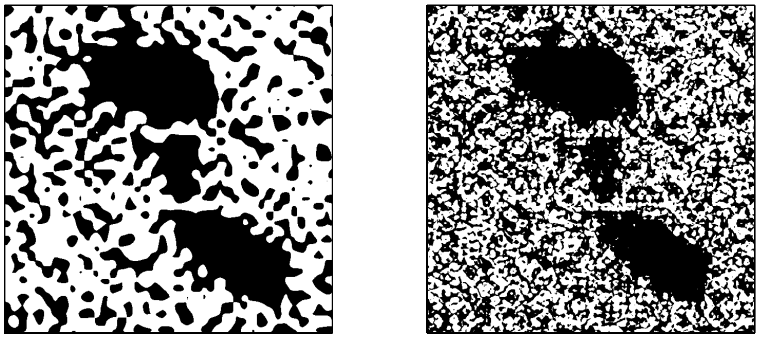
(a) True sample, $\overset{*}{x}$ (b) Single, local MRF [80]

(c) FSHF [20] (d) Proposed HHMF

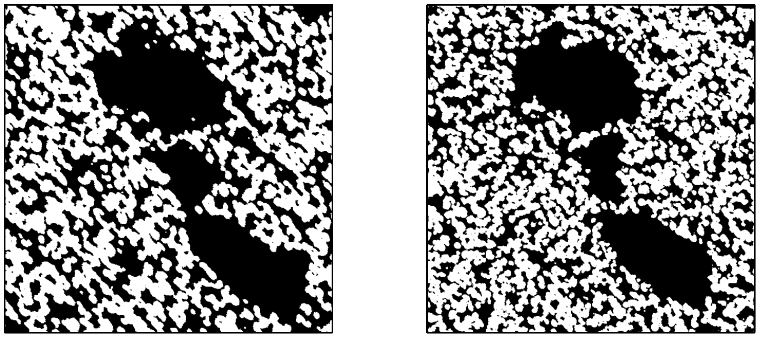
Figure 5.11: Reconstructing a two-scale porous media sample from the measurements shown in Fig. 5.10(a). Although (c) provides a good reconstruction, (d) is more faithful to $\overset{*}{x}$ in the details (e.g., the shapes of the large pores and the small-scale structures along pore edges).

where the $\hat{x}$ based on the WTEDPM performs the best and provides a good recovery of facets. However, many small-scale structures fail to be recovered by the WTEDPM since those structures do not appear at the measured scale and cannot be reproduced by enhancement.

In the third experiment, we reconstruct the complex, multi-model porous media sample (Fig. 5.15(a)) from its LR noisy observation (Fig. 5.10(c)). For this difficult problem, we build a HHMF with two parallel hidden hierarchies (U_1, U_2). During the reconstruction,



(a) ZPWT (b) WTPM [91]



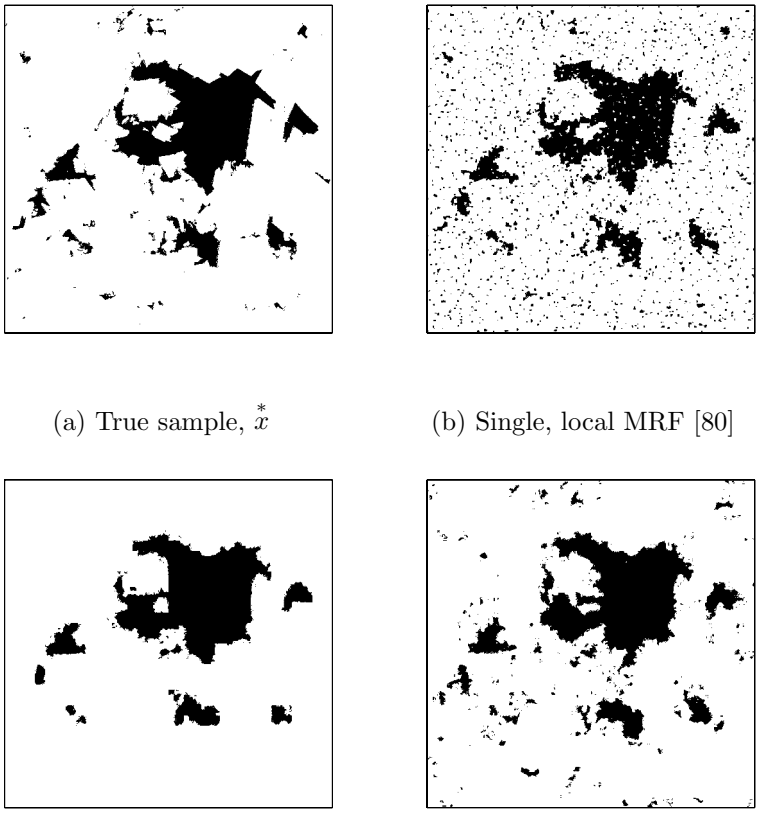
(c) EBSR [37] (d) WTEDPM [65]

Figure 5.12: Reconstructing a two-scale porous media sample (Fig. 5.10(a)) with non-MRF methods. Only (d) provides a credible result.

first we estimate U_1, U_2 and then X, as proposed. The parallel HHMF based estimates $\hat{x}$, $\hat{u}_1$ and $\hat{u}_2$ are shown in Fig. 5.15(d-f). The performance of the HHMF with two hidden hierarchies is clear from the comparison to other MRF methods (Fig. 5.16(b-d)). The reconstructions based on the continuous state methods are shown in Fig. 5.17.

5.6 Evaluations

To evaluate the performance of different reconstruction methods, we consider three metrics. First, to study the statistical similarity between estimate $\hat{x}$ and ground truth $\overset{*}{x}$, we evaluate

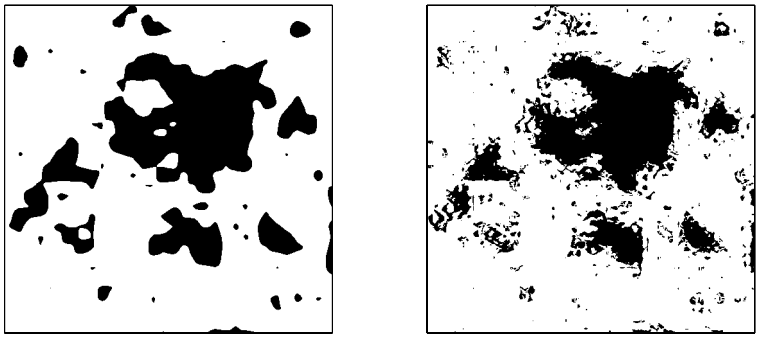


(a) True sample, $\overset{*}{x}$ (b) Single, local MRF [80]

(c) FSHF [20] (d) Proposed HHMF

Figure 5.13: Reconstruction from Fig. 5.10(b) by MRF methods. Although the flat facets in $\overset{*}{x}$ are not fully reproduced in (d), the improvement in recovering the multi-scale characteristics are clearly better than in (b) and (c).

the inconsistency in terms of chordlength models (Section 2.2.4) learned from $\hat{x}$ and $\overset{*}{x}$, the chordlength inconsistency measured as count differences $|c_{\hat{x}}(n) - c_{\overset{*}{x}}(n)|$. A second, related metric is to compare on the basis of porosity, the fraction of black pixels. Third, to demonstrate the structural consistency between $\hat{x}$ and $\overset{*}{x}$, the correlation coefficient $\rho(\overset{*}{x}, \hat{x})$ is calculated as a function of scale, which is defined as the average number of decimations leaving a pixel value unchanged (see appendix A for details). Clearly, very tiny structures fail to exert much influence on the measurements, therefore the correlation ρ is expected to decrease at finer scales.



(a) ZPWT (b) WTPM [91]

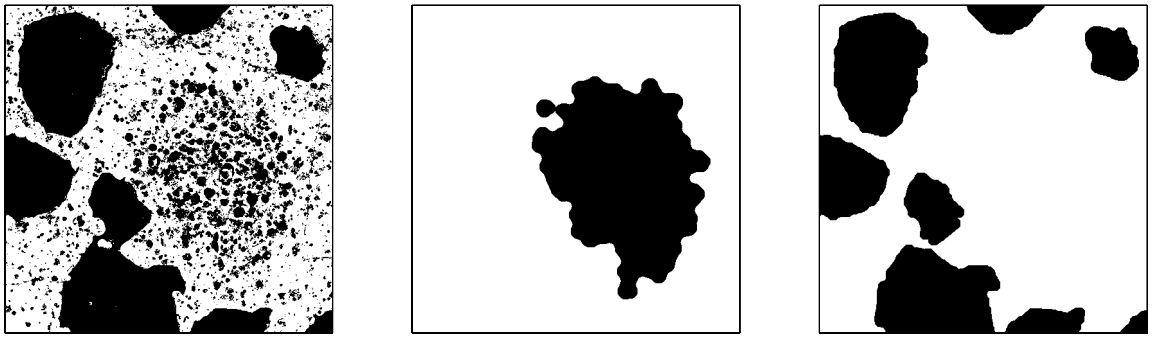
(c) EBSR [37] (d) WTEDPM [65]

Figure 5.14: Reconstruction from Fig. 5.10(b) by non-MRF methods. Method (d) provides the best result, however it fails to recover many small-scale structures.

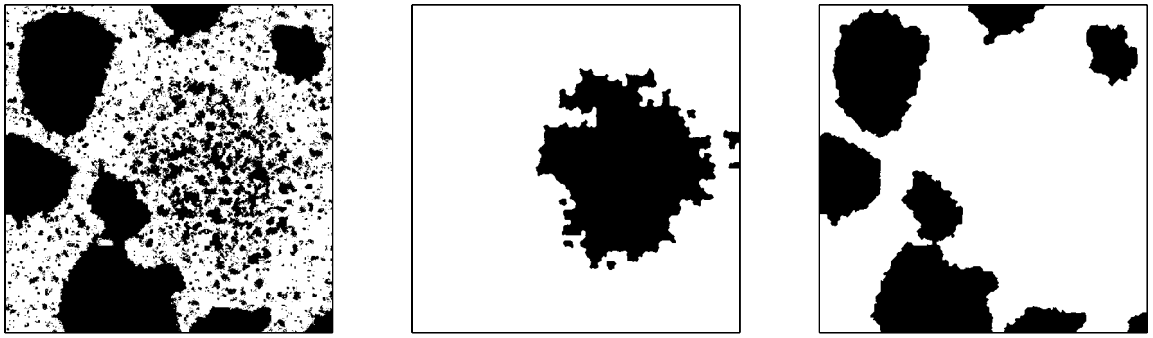
Because the proposed HHMF and the competing WTEDPM quite obviously visibly outperform all other implemented methods, for the quantitative tests our focus will be on the comparison of HHMF and WTEDPM.

Fig. 5.18 plots the correlation coefficient ρ_k between truth and reconstruction as a function of scale k. The results reveal a significant correlation between estimates and truth one to two scales finer than measured. The proposed method performs similarly to or outperforms the WTEDPM.

As an independent test of the reconstruction, Fig. 5.19 plots the chordlength inconsistency $|c_{\hat{x}}(n) - c_{\overset{*}{x}}(n)|$ between estimated field $\hat{x}$ and true field $\overset{*}{x}$, measured by the prevalence



(a) True sample, $\overset{*}{x}$ (b) True label field, $\overset{*}{u}_1$ (c) True label field, $\overset{*}{u}_2$

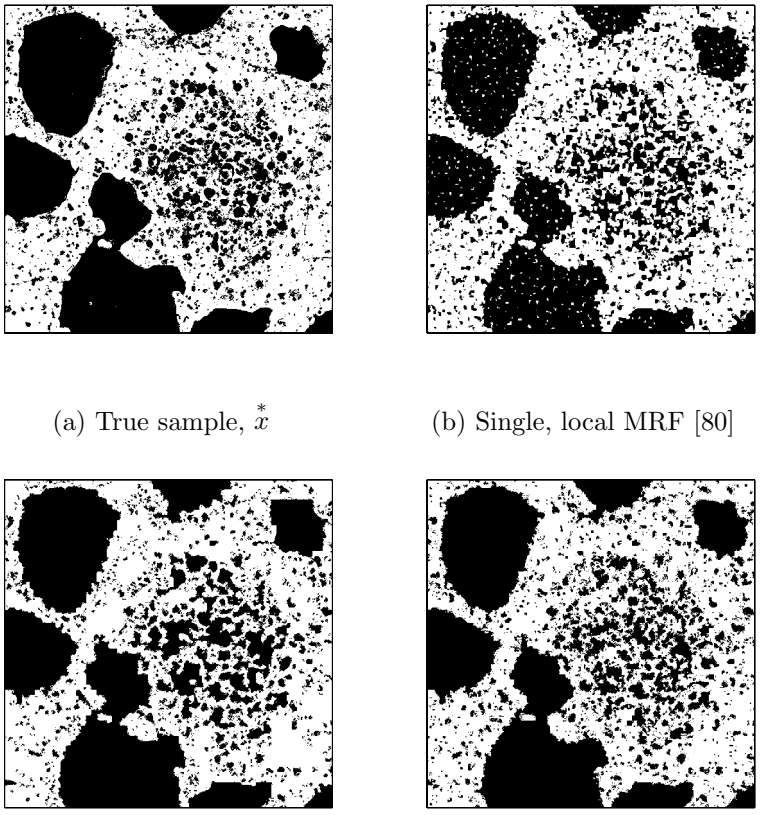


(d) HR estimate, $\hat{x}$ (e) Label field Estimate, $\hat{u}_1$ (f) Label field Estimate, $\hat{u}_2$

Figure 5.15: A sample of a complex porous media reconstruction with the proposed HHMF method, for the measurements in Fig. 5.10(c). Although (d) is not able to fully reconstruct some subtle structures (e.g., a line-like structure connecting two large pores at the top right), the improvement in relevant detail of (d) over Fig. 5.10(c) is stunning. The HHMF here demonstrated its capability in modeling multiple spatial nonstationarities.

of a given chordlength, where the chordlength is a statistic not captured explicitly in any of the wavelet or Markov models. The better consistency of the proposed HHMF is striking, especially at shorter lengths (finer scales) in (b) and (c).

Finally, researchers in porous media are very sensitive to the preservation of aggregate behaviour, such as overall porosity, thus a final comparison is made in Table 5.2. In all three experiments the aggregate porosity is better preserved in the HHMF, compared to the



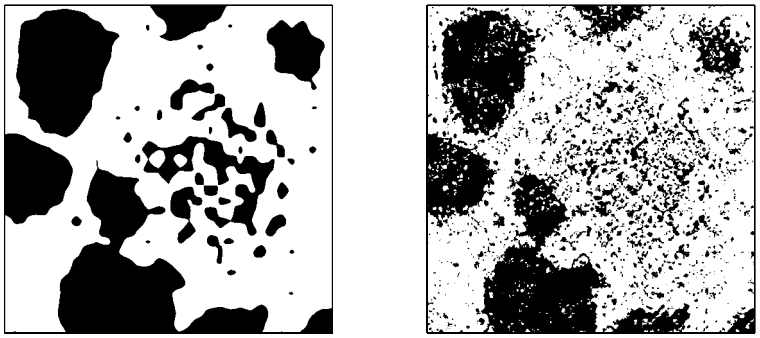
(a) True sample, $\overset{*}{x}$ (b) Single, local MRF [80]

(c) FSHF [20] (d) Proposed HHMF

Figure 5.16: Reconstructing complex porous media image Fig. 5.10(c) by MRF methods. With two parallel hidden hierarchies, (d) provides the best reconstruction among the results, although there are some small-scale structures in the highly dense regions and the boundaries along large pores which are lost.

wavelet WTEDPM. Most significant is the consistency in the most complex reconstruction of Fig. 5.10(c), where the absence of an explicit hidden model greatly compromises the WTEDPM relative to the HHMF.

For the first experiment, the correlation coefficients are plotted in Fig. 5.18(a); the chordlength inconsistencies is plotted in Fig. 5.19(a). Corresponding to different behaviour region (smalls/large scale structure) in Fig. 5.11(a), the proportions of black pixels are calculated for reconstructions and the true sample, and listed in Table 5.2(a), where the



(a) ZPWT (b) WTPM [91]

(c) EBSR [37] (d) WTEDPM [65]

Figure 5.17: Reconstructing a complex porous media image (Fig. 5.15(a)) from Fig. 5.10(c) by non-MRF methods. The fine-scale statistics in (d) are not matched as well as those in the HHMF (Fig. 5.16(d)).

proportions of the proposed HHMF method is closer to the true proportion than the WTEDPM method. As checking all those evaluations, although Fig. 5.19(a) shows that the WTEDPM works better than the HHMF in reconstructing the small detail structures ($chordlength < 5$), we can see the overall superiority of the HHMF is still overwhelming.

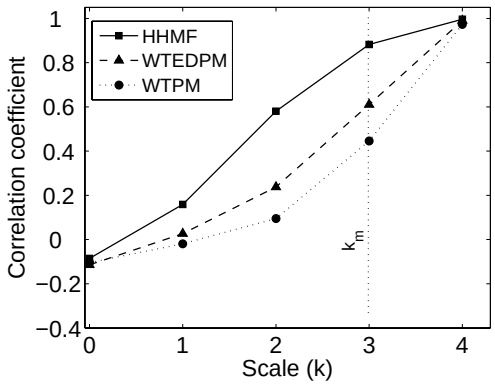
For this experiment, the performance evaluations are plotted in Fig. 5.18(b) and Fig. 5.19(b) respectively. For the structure based correlation, the WTEDPM based $\hat{x}$ is more consistent with $\overset{*}{x}$ than the HHMF except at fine scales. For the similarity of the statistical characteristics, the HHMF outperforms with the WTEDPM at most significant chordlength entries.

Table 5.2: Porosity comparison of ground truth and reconstruction

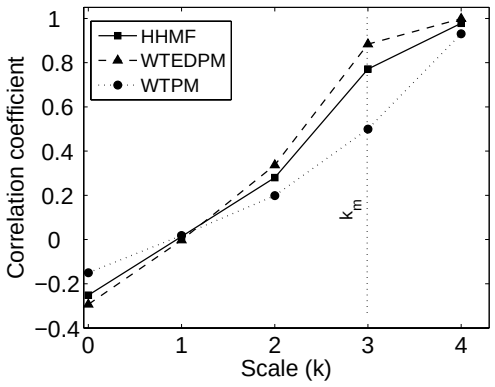
Fig. 5.10(a):	**True sample**	**WTEDPM**	**Proposed HHMF**
Background	0.448	0.401	0.459
Large pores	1.000	0.973	0.988
Fig. 5.10(b):	**True sample**	**WTEDPM**	**Proposed HHMF**
Whole image	0.140	0.108	0.115
Fig. 5.10(c):	**True sample**	**WTEDPM**	**Proposed HHMF**
High density background	0.428	0.278	0.445
Low density background	0.140	0.074	0.164
Large pores	1.000	0.960	0.992

From this experiment, we see that even a phenomenon with numerous scale-dependent behaviours creates a challenge for a HHMF with limited number of hidden hierarchies, the reconstruction is still comparable to the WTEDPM based result.

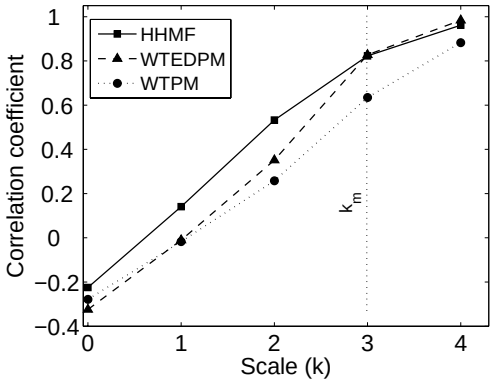
Whereas, the HHMF exhibit its stronger capability to handle the piece-wise multi-model behaviours in Fig. 5.15, as demonstrated in Table 5.2(c). Moreover, the evaluation plots both from a correlation aspect (Fig. 5.18(c)) or a chordlength model respect (Fig. 5.19(c)) also indicates the HHMF outperforms with the WTEDPM in this experiment. Besides, it is worthwhile to clarify that although Fig. 5.19(c) indicating the MTPM possesses smallest inconsistency at some short chordlength entries, the actual performance of the MTPM is doubtable due to the strong artifacts caused by the intrinsic limitation of the wavelet transform.



(a) (a)



(b) (b)



(c) (c)

Figure 5.18: The correlation ρ_k between ground truth $\overset{*}{x}$ and reconstruction $\hat{x}$ as a function of structure scale k for (a) a two-scale reconstruction from Fig. 5.10(a), (b) a fractal multi-scale reconstruction from Fig. 5.10(b), (c) a multi-model behaviour reconstruction from Fig. 5.10(c), where k_m denotes the resolution scale of the measurement, where k_m denotes the measurement scale. The proposed HHMF performs similarly to or better than the competing wavelet methods.

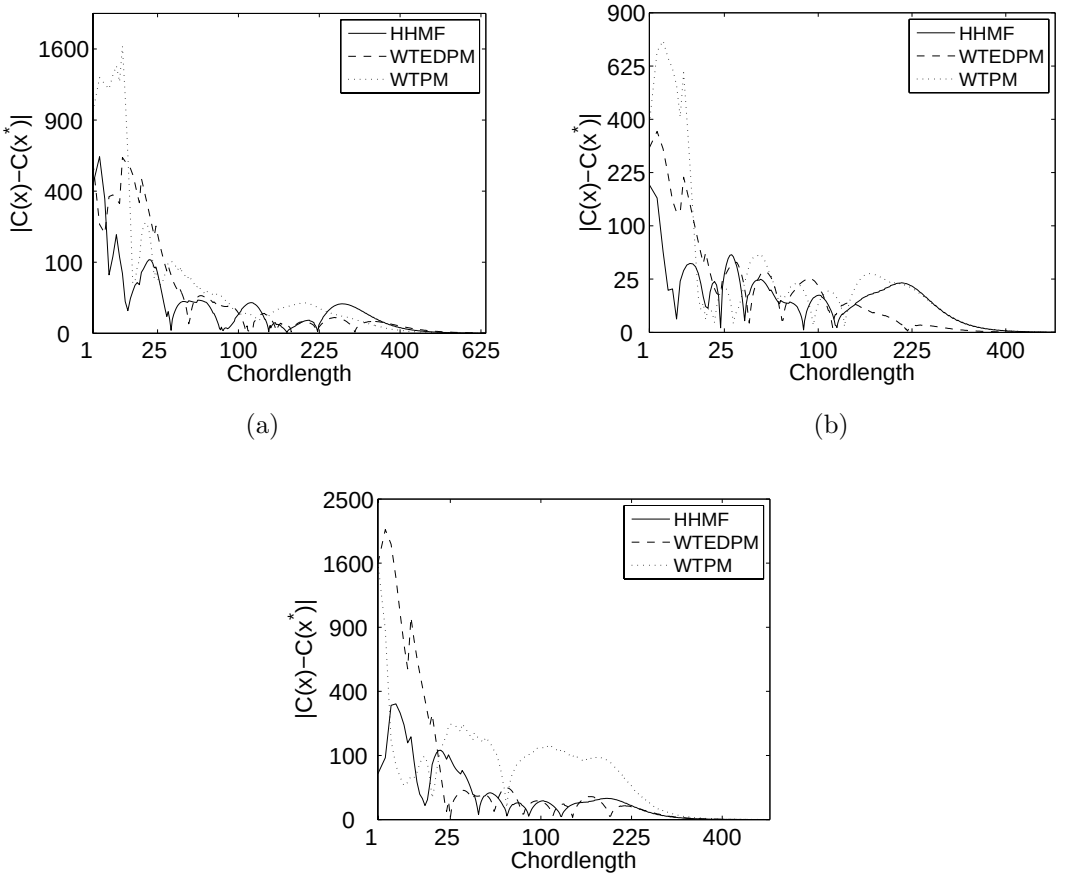


Figure 5.19: The inconsistency of the reconstruction result $\hat{x}$ with ground truth $\overset{*}{x}$ assessed in terms of a chordlength model [105], plotted as a function of chordlength. With the exception of long chordlengths in (b), the inconsistency of the HHMF is consistently better (lower).

Chapter 6

Tree-structured Hidden Hierarchical Markov Fields for Remote Sensing Image Synthesis

The systematic evaluation of data analysis tools, such as segmentation and classification algorithms for geographic information systems (GIS), is difficult given the unavailability of ground-truth data in most cases. Therefore, testing is typically limited to small sets of pseudo-ground truth data collected manually by trained experts, or primitive synthetic sets composed of simple geometries. In this chapter, to address this issue, we will propose a more substantial approach on the basis of hidden hierarchical fields to synthesize remote sensing data for use as a reliable evaluation test-bed. Given the scale-dependent, non-stationary nature of remotely sensed data, a new modeling approach that combines a resolution-oriented hierarchical method with a region-oriented binary tree structure is introduced to synthesize such complex data in a realistic manner. Experimental results using operational RADARSAT SAR sea-ice image data and SIR-C/X-SAR land-mass image data show that the proposed hierarchical approach can better model complex, nonstationary scale structures than local MRF approaches, thus making it well-suited for synthesizing

This chapter is written based on the paper: Y. Liu, A.Wong and P. Fieguth, Synthesis of remote sensing label fields using a tree-structured hierarchical model, accepted by *IEEE Trans. on Geoscience and Remote Sensing*, Nov. 28, 2010.

remote sensing data.

6.1 Introduction

The use of aerial and satellite remote sensing imagery has become an integral part of terrestrial ecological studies and environmental monitoring, ranging from sea-ice monitoring in polar regions [117] and land-use and land-cover change analysis [56, 61], to flood risk and damage assessment [33]. Given the large volume of high-resolution remote sensing data being acquired on a daily basis and the time consuming nature of manual data manipulation, considerable research effort in the design of geographic information systems (GIS) has been spent on the development of tools for analyzing remote sensing data in an automated fashion. Two classes of automatic data analysis tools that have great importance to GIS are automatic segmentation [52, 94, 99, 102, 117] and classification algorithms [11, 33, 55, 78, 87].

A major challenge in the design of automatic segmentation and classification algorithms for the purpose of remote sensing image analysis is the reliable, systematic evaluation of algorithmic performance for assessing the potential for transition from the research stage to real-world operational use in GIS. While a plethora of quantitative assessment metrics are available for evaluating the performance of such automatic data analysis tools [76], it is not feasible in the case of polar regions to acquire ground-truth segmentation and classification information pertaining to the remote sensing data. As such, the evaluation of automatic segmentation and classification techniques have been limited to the use of small sets of pseudo-ground truth data collected manually by trained experts in a time-consuming manner, or primitive synthetic sets composed of simple geometric shapes [117]. The reliability of performance assessment using pseudo-ground truth data is limited not only by the small set of test data available, but also by the limited time and accuracy of trained experts who are able to produce manual segmentations and classifications on a pixel level. The performance assessment using primitive synthetic sets is more reliable than that using pseudo-ground truth data given the large amount of test data available and pixel-level accurate ground truth. However, such primitive synthetic tests are a poor representation of real remote sensing imagery and as such do not provide a realistic testing scenario for evaluating the operational potential of an automatic data analysis algorithm.

To address these issues associated with the evaluation of automatic analysis algorithms, we propose an alternative approach where the systematic evaluation of analysis algorithms is performed using realistic-looking remote sensing data, generated from real data, and corresponding synthesized ground truth. This approach allows for the generation of large test sets that are representative of real-world operational scenarios and have *known* ground-truth. Furthermore, the randomness associated with the synthesis process improves the reliability of testing by reducing bias towards algorithms tuned to work well with specific test data.

Much of the research literature in remote sensing image synthesis deals with model-based texture synthesis [14, 15, 21, 115]. However, such methods are designed to capture and generate textural characteristics only and as such are ill-suited for generating realistic-looking remote sensing data as they do not capture the complex structural characteristics found in operational settings. More recent general nonparametric texture synthesis methods (e.g., [34, 66, 108]) are able to better capture both textural and structural characteristics, but exhibit two main limitations. First, while such methods are able to capture small-scale structural characteristics, they are ill-suited for capturing large-scale structural characteristics, which will be illustrated later in the experimental results. Second, and crucially, no texture method provides a corresponding label field, which is the necessary ground truth in the evaluation of data analysis tools.

In this chapter, we aim to address the issues faced by existing image synthesis methods in generating realistic-looking remote sensing data by decoupling the synthesis of texture and structure into two different modeling structures. That is, we will first explicitly synthesize the discrete-state label field, which hidden under the remote sensing image contains the complex structural characteristics of the image, then separately synthesize the textures of remote sensing images using a modification of the nonparametric texture synthesis strategy proposed by Efros and Leung [34]. We introduce a practical approach to synthesizing underlying multi-label discrete fields by combining a resolution-oriented hierarchy with a region-oriented hierarchy. Multi-resolution representations have a long history and the method of representing an image or a label map with binary partition trees has been previously discussed [32, 90, 98]. Indeed, recently there has been growing interest in a generalization of hierarchical partition fields [38, 100] for segmentation or hidden hierarchical fields [68]. However, the idea of using partition trees and hierarchical models for synthesis is novel.

6.2 Related Work

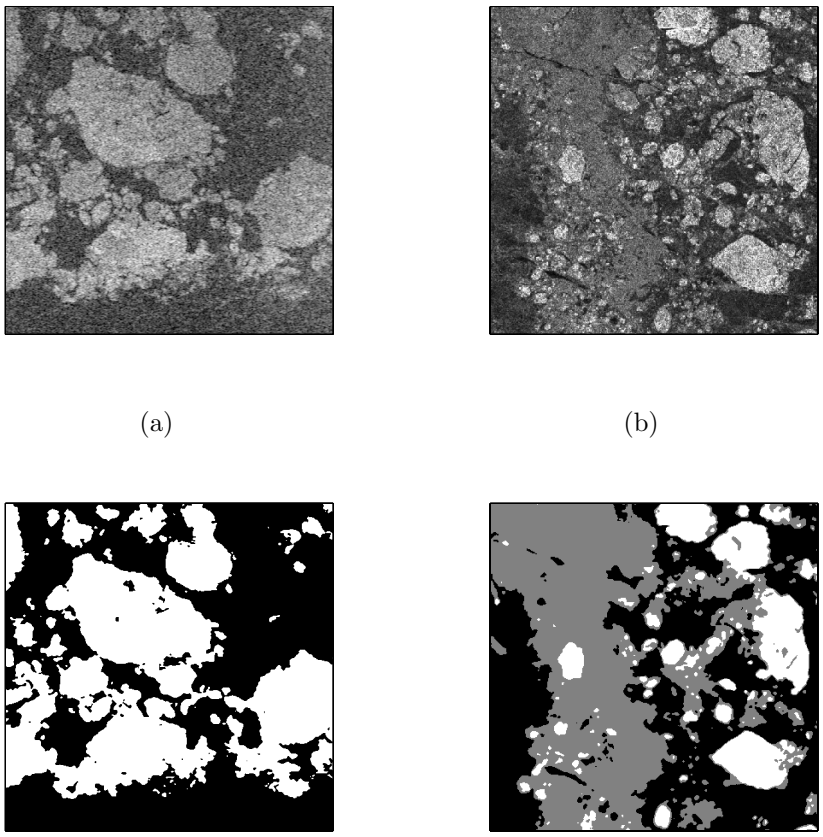
While texture synthesis approaches for remote sensing data have been proposed in previous research literature [14, 15, 21, 115], little attention has been paid to synthesizing remote sensing data with complex structural and textural characteristics. One can view the problem of generating structures for synthetic remote sensing data as a label synthesis problem, where a label corresponds to a particular class of structure or feature (e.g., ice type, vegetation type).

While there is a large research literature [24, 39] on texture classification and processing, in most cases the problem involves comparatively simple labels or lies at a single scale. In particular, many approaches utilize blob-like priors that enforce boundary smoothness and as such assert little in terms of subtle structures and complexity in the simulation of the field. Therefore for synthesizing complex label fields, such as those in Fig. 6.1(c,d) illustrating structures in remote sensing imagery, a more subtle model is required.

The simplest method to improving the modeling of subtle structures is through the use of a Fourier basis for a Gaussian Markov Random Field (GMRF) model kernel (Section 3.1.1) method with threshold. Unfortunately, this approach is able to give only a stationary binary field and, even more problematic, only at a single scale. To capture complex scale structures, we are unable to uese single-scale models such as local MRF (Section 3.1.1), chordlength and local histogram (Section 2.2.3).

On the other hand, creating a prior based on a hierarchical structure provides a more natural way to introduce scale-dependent models. The Frozen State Hierarchical Field (FSHF) model, as discussed in Section 3.3.2, defines an inter-scale relation for computational efficiency with only annealing relatively small structures at each scale. Although it is possible to apply the FSHF to directly model a binary label map (Fig. 6.1(c)), this approach cannot be used directly in the cases with more than two labels or with nonstationary behavior (Fig. 6.1(d)), both of which are common in remote sensing imagery.

In Chapter 5, we have proposed a Hidden Hierarchical Markov Field (HHMF) model with parallel hidden hierarchies to capture heterogenous behaviours. Although this model can be adapted to synthesize multi-label fields, an assumption has to be satisfied: the multi-label behaviours have to be either independent to each other or conditionally independent. However, in the remote-sensing case, the label field may have highly spatially



(a) (b)

(c) (d)

Figure 6.1: Sea-ice texture samples (a,b) and their underlying label maps (c,d). Many remote sensing textures have underlying label maps with multi-scale structures which can be binary (c) or multi-labeled (d). The scale-dependent behavior in (c, d) will usually not be well captured by a single random field.

dependent structures (Fig. 6.1(d)) which can be well represented by two or tree independent or conditional independent stationary fields.

There is an established literature on the use of partition trees to decompose multi-label problems in image segmentation, compression, and synthesis [4, 32, 90, 98]. In this approach, a multi-label problem is repeatedly partitioned to yield a set of simple problems. In the past, the simple partitioned problem were solved using simple models, such as a single

MRF. However, a tree of simple random fields does not gives the modeling flexibility to represent complex structures. For example, a sea-ice label map (Fig. 6.1(d)) is set apart into a set of simple component fields (Fig. 6.2), however some partitioned label fields, even with binary states, may still possess sufficiently complex scale-dependent structures which are not modeled well by single random field, such as Fig. 6.2(b). Thus, to model those partitioned fields requires a sophisticated model, such as the FSHF.

In this chapter, we are proposing to combine the ideas from the two preceding paragraphs, combining hierarchical modeling with partitioning methods. This modeling approach will allow us to reliably capture multi-label structures as a function of scale, which is very important for the realistic synthesis of complex data.

6.3 Hierarchical Markov Fields

6.3.1 Hidden Label Field Synthesis

In this chapter, we propose to synthesize remote sensing images on the basis of first synthesizing the label field. There are a variety of approaches to synthesize label fields and, in particular, we are using methods taken from MRFs (Chapter 3). In modeling a given underlying label field U, a MRF (Section 2.1.2) characterized by a *local* neighbourhood $\mathcal{N}_s$,

$$p(u_s|u_{S\setminus s}) = p(u_s|u_{\mathcal{N}_s}),$$

cannot assert the presence of structures on more than one scale (Section 3.1), whereas learning a huge *nonlocal* model which can, in principle, learn such structures, is prohibitive to learn and to use. Instead, to model a label field U having multi-scale structure, we would propose using scale-dependent modeling, such that U is defined via a sequence of fields $\{U^k,\, k \in K = (0, 1, \cdots, M)\}$, where $k = 0$ denotes the finest scale and $k = M$ the coarsest. At each scale k, U^k is defined on site space S^k and results from the downsampling of $U \equiv U^0$. A hierarchical model can be written as

$$p(u^0, \cdots, u^M) \;=\; \big[\prod_{k=0}^{M-1} p(u^k|u^{k+1})\big] \cdot p(u^M) \tag{6.1}$$

The advantage of hierarchical modeling is that nonlocal large-scale features become local at a sufficiently coarse scale, therefore at each scale a single MRF can be used to capture

the features local to that scale, inherently allowing for scale-dependent structures. We will define u_s^k to be the label state at site s on scale k, with an associated local neighbourhood $\mathcal{N}_s^k$ and parent $u_{\wp(s)}^{k+1}$ on the next coarser scale. As synthesizing binary images, a label field U can be represented by FSHF model with a ternary state $u^k(s) \in \{0, 1, \frac{1}{2}\}$ at coarse scales $(k > 0)$.

6.3.2 Multi-state Hierarchy

In general, extending beyond binary modeling leads to rather complex representations and models, as well as to significant computational complexity. Although the FSHF model is effective in binary modeling, extending the method to multi-state modeling is not a trivial step [20, 68]. The problem related to modeling all pairwise, triplet-wise, etc. label interactions at coarser scales is quite complicated even for the ternary case.

However, for some ternary-state phenomena, if there is an intermediate medium acting as a physical separation or layer between two others, we have a particularly convenient context for modeling the ternary phenomenon. We can change the state definition of the FSHF by letting grey $(\frac{1}{2})$ denote the intermediate layer or undetermined, with the effect that the FSHF method can be directly applied and the intermediate state will lead to a ternary rather than binary field at the finest scale. Noticeably, here we insist a spatial decoupling assumption that the intermediate state conditionally separates the other two states. In most cases it will be simpler to decompose a complex multi-label structure into a set of simpler components as discussed in the Section 6.4.

6.4 Tree-Structured Hidden Hierarchical Markov Fields

The FSHF method discussed in Section 6.3 offers a compelling approach to modeling, which is computationally highly efficient, and admits a scale dependent model for the synthesis of binary label maps. However there are two obvious issues that need to be addressed for the synthesis of more complex label fields as encountered in remote sensing: first, we generally have to solve a multi-label problem. Second, the label maps may be nonstationary, meaning that there are different behaviours in different parts of the image, which cannot be well modeled by a single hierarchy: forcing a single hierarchy to learn the

variability of a nonstationary behaviour leads to an averaging effect, so we need more than one model.

There is an existing literature on partition trees [32, 98] which allows a given image or label map to be partitioned into pieces. The general idea behind the partition tree is that behaviours are split and successively subdivided until homogeneous portions of images are found. In general, such binary partition trees can be used in problems of classification. Here, we choose to use them equally suitably in image synthesis, as a proposed Tree-Structured Hidden Hierarchical Markov Field (Tree-Structured HHMF) for modeling underlying label field. The assumption is that a given multi-label image can be produced as a tree-structured conditional sequence of binary or ternary images, such that the dominant large-scale structure is produced first (the root node of the partition tree), then with further details inside and outside of this structure developed in the child nodes, a detailed example of which will be seen in the experimental results. The key idea is to use the existing method of partition trees to combine multiple hierarchical models to allow the nonstationary and nonbinary representation that we are seeking, and at the same time to preserve the scale-dependent computational efficiency of the hierarchical approach. The modeling structure of tree-structured HHMF is shown in Fig. 6.3.

In the proposed modeling approach, the structural components of U are progressively specified by a sequence of nodes in a binary tree (6.2) from mixed to pure labeled states.

$$T = \{U^i | \mathcal{Q}^i, 0 \leq i < N\} \tag{6.2}$$

Every node is defined as a conditional hierarchical field $U^i|\mathcal{Q}^i = \{U^{i,k}|\mathcal{Q}^i, k \in (0, 1, \cdots, M^i)\}$, where $\mathcal{Q}^i$ denotes the set of fields on which U^i depends. The partition tree starts at the root $U^0 = \{U^{0,k}|\mathcal{Q}^0 = 1\}$, used to capture the most significant structure of U.

The influence of U^i on the partition tree is mediated through the up to two children of U^i, conditional on U^i or $\tilde{U}^i$, such that binary field U^i controls the spatial extent of its children. The conditioning is encoded in $\mathcal{Q}^i$, which consists of one or more fields, such that

$$\begin{array}{rcll} \mathcal{Q}^i & = & U^a & \rightarrow U^i_s = 0 \quad \text{if } U^a_s = 0 \\ \mathcal{Q}^i & = & \tilde{U}^a, U^b & \rightarrow U^i_s = 0 \quad \text{if } U^a_s = 1 \text{ or } U^b_s = 0 \end{array} \tag{6.3}$$

etc. Since each node under T only models simple binary/ternary structures, each field $U^i|\mathcal{Q}^i$ can be well modeled by the FSHF, as discussed in Section 6.3, and each scale of each

field $U^{i,k}|\mathcal{Q}^i$ can be sampled as

$$\hat{U}^{i,k}|\mathcal{Q}^i \leftarrow p\left(U^{i,k}|\mathcal{Q}^i\middle|\,\hat{U}^{i,k+1}\right). \tag{6.4}$$

This process proceeds recursively, first over all scales in U^0, then over scales on fields further down the partition tree.

The process by which we infer a partition tree structure T from a given training label field $\tilde{u}$ is a creative one, requiring human input, and is highly problem dependent. The main example of this paper, the field shown in Fig. 6.1(d), has as its dominant large-scale structure the binary behaviour $u^0|1$ (Fig. 6.2(a)). Since both foreground and background in Fig. 6.2(a) correspond to mixed labels, the partition process needs to continue. The foreground is partitioned into two binary fields (Fig. 6.2(b,c)), whereas the background is decomposed into a ternary field (d) and two minor, residual binary ones (e,f). The original label field $\tilde{u}$ has thus been decomposed, represented as a partition tree as shown in Fig. 6.6(i).

Having specified a partition tree, the inverse step, the process of recombining of the synthesized conditional fields $\{U^i|\mathcal{Q}^i\}$ to get $\hat{u} = J(\,\{\hat{u}^i|\hat{\mathcal{Q}}^i\}\,)$, is straightforward.

Thus, the proposed tree-structured HHMF method synthesizes a label field in two ways: in terms of shape complexity, the structure is gradually refined hierarchically from coarse to fine resolution; in terms of label complexity, the states are specified through a partition tree from coarse to fine labeling. The proposed modeling approach, with both resolution-oriented and region-oriented hierarchies, provides a capability to model complex discrete fields using simple models while maintaining high computational efficiency. A hierarchical model on its own, such as the FSHF, can be considered as a special case of the tree-structured HHMF with only one region-oriented component.

Admittedly, one of the limitations of the proposed approach is its spatial decoupling assumption, which assumes that a multi-label field can be decomposed into multiple binary/ternary fields. In some cases in which the different label regions are highly interacting this assumption may not hold true, however our tests show that a variety of SAR and other data can indeed be modeled in this way.

The overall modeling process (Algorithm I) is therefore to select a partition tree, to find ground truth for each state in the partition, and to learn a histogram model H_k^i in (2.19) from the empirical histogram of the ground truth data for each scale k in model i.

Algorithm I: Tree-Structured HHMF Modeling

1: Initialize a training partition tree $T = \{U^i|\mathcal{Q}^i, 0 \leq i < N\}$;
2: Learn a hierarchical histogram model H_k^i at k scales for each conditional node $U^i|\mathcal{Q}^i$ in T;
3: Initialize each sampling hierarchy randomly at the coarsest scale;
4: Sample each hierarchical model from coarser to finer scales by using simulated annealing;
5: Combine the generated fields to the partition-tree structure and obtain the synthesized label field.

With modeling performed, the sampling process follows the dependency structure of the partition tree. Each hierarchy is randomly initialized at the coarsest scale, and sampled at progressively finer scales using simulated annealing. When all of the frozen state hierarchies have been sampled, the generated fields are combined according to the partition-tree structure to obtain the label synthesis.

6.5 Image Synthesis

The textured images in Fig. 6.1(a,b), having a complex, non-local, non-stationary behaviour are difficult to model. Therefore the direct synthesis

$$\hat{x} \leftarrow p(x) \tag{6.5}$$

is a complicated undertaking. On the other hand, because U represents the salient features of interest in X, what remains in X, given U, are the fine-scale details not of interest: noise, speckle, quantization, blurring, etc., all of which are comparatively simple and *local* textural phenomena. That is, the synthesis

$$\hat{x} \leftarrow p(x|\hat{u}) \tag{6.6}$$

is comparatively straightforward, therefore we are deliberately picking an existing texture synthesize method [34] to generate the fine-scale texture on top of $\hat{u}$.

The method in [34] is a sample based approach to synthesis, such that a pixel x_s is synthesized by comparing its neighbourhood X_{N_s} to all possible neighbourhoods in the

training data $\tilde{x}$, and selecting x_s at random from among the matching $\tilde{x}$ neighbourhoods. This allows a synthesis $\hat{x} \leftarrow p(x)$, as in (6.5), with the problem that a small neighbourhood $\mathcal{N}_s$ fails to reproduce large-scales structures, and large neighborhoods tend to memorize the structure of $\tilde{x}$, failing to generate a random sample.

We slightly modify the method of [34] to allow a synthesized texture $\hat{x}$ to be sampled from the conditional MRF $X|U$:

$$\hat{x}_s \leftarrow p(x_s|x_{\mathcal{N}_s}, u_s), \tag{6.7}$$

rather than directly from the texture field X. Given the conditioning on u_s, we now search for a set of closely matching patches in $\tilde{x}$, for which the training label $\tilde{u}$ also matches.

We will see in Section 6.6 that this simple texture synthesis approach leads to good results. There is nothing inherent necessitating the use of [34] with our approach; indeed, any other advanced texture synthesis method may be used as well.

6.6 Experimental Results and Evaluation

This chapter has two goals:

i) The synthesis of realistic remote sensing imagery.

ii) The synthesis of the underlying label field as ground truth.

To demonstrate the effectiveness of the proposed tree-structured HHMF model, the image synthesis approach described in Sections 6.4 and 6.5 was used to generate random SAR sea-ice imagery based on operational RADARSAT-1 SAR sea-ice imagery of the polar region provided by the Canadian Ice Service (CIS), as well as SIR-C/X-SAR land-mass imagery of Hong Kong, China provided by NASA JPL. The SAR sea-ice data used to learn the model for generating sea-ice imagery are acquired in the microwave band (C-band), with HH polarization, 100m pixel spacing, and three ice types. The sea-ice imagery of the polar region is difficult to model and synthesize given the complex sea-ice structures and formations, as well as non-homogeneous texture characteristics. The SAR land-mass data of Hong Kong, China used to learn the model for generating land-mass imagery are acquired in the microwave band (C-band). The sea-ice and landmass imagery are very different from one another, with the intent of illustrating that our method is not specialized to a single type of imagery.

6.6.1 Single Hierarchical Approach

A single hierarchical field model, such as the FSHF, can be considered as a special case of the tree-structured HHMF with only one component. As the first test for the proposed model, we apply a single FSHF model to synthesize a scale-dependent binary field. This initial test is undertaken to demonstrate the morphological modeling performance of a single hierarchy in modeling a binary field. The model is trained by the binary field $\tilde{u}$ shown in Fig. 6.1(c). Two synthesized samples $\hat{u}$ are shown in Fig. 6.4(a,b). We can see the structures in the synthesized fields essentially resemble the multi-scale phenomena of the training data.

The texture at each pixel $\hat{x}_s$ is sampled, as described in Section 6.5. In the synthesized texture samples, Fig. 6.4(c-d), we see that the created texture skin is consistent with the texture characteristics in the training data shown in Fig. 6.1(a).

6.6.2 Tree-Structured Hidden Hierarchical Approach

A more general test for the proposed tree-structured HHMF is the image shown in Fig. 6.1(b), with a corresponding label field in Fig. 6.1(d). Based on the tree-structured modeling representation of Section 6.4, a partition tree of binary or ternary component fields $\tilde{u}^i$ is constructed, as shown in Fig. 6.6(i), such that the hidden field is produced from the components as

$$\begin{aligned} \hat{u} &= J\left(\{\hat{u}^i|\hat{\mathcal{Q}}^i\}\right) \\ &= \hat{u}^0|1+\hat{u}^1|\hat{u}^0-\hat{u}^2|\hat{u}^0,\bar{\hat{u}}^1+\hat{u}^3|\bar{\hat{u}}^0-\hat{u}^4|\hat{u}^3,\bar{\hat{u}}^0+\hat{u}^5|\bar{\hat{u}}^3,\bar{\hat{u}}^0 \end{aligned} \tag{6.8}$$

where the partition tree structure is subjectively inferred, by hand, from the training data.

We thus obtain a complex, synthesized, multi-label field (Fig. 6.6(g)). By comparing Fig. 6.6(a-f) to Fig. 6.2(a-f), respectively, we can see that the synthesized components generally resemble the multi-scale structure in their corresponding training components. Since the structured features of each training component can be well captured by each node hierarchy, the final label field should possess similar statistical characteristics to the training data, comparing Fig. 6.6(g) to Fig. 6.1(b). Given the synthesized label field (Fig. 6.6(g)), the sea-ice texture may be generated (Fig. 6.6(h)), and can be compared with Fig. 6.1(b).

To emphasize the variability in the synthesized samples and the suitability of our approach in generating ground truth test data, three additional results based on the same training data are shown in Fig. 6.7. We can see that the synthesized label fields provide substantial variations, and yet share similar statistical characteristics.

The hierarchical model in the proposed modeling structure is evaluated, as plotted in Fig. 6.5, using a chordlength model (Section 2.2.4) from a large (2048×2048) binary sample (Fig. 6.5(a)). This test sample, a microscopic image of a physical porous medium, contains a wide variety of multiscale structures, exactly the sort of structure our proposed hierarchical model is expected to model. Comparing the chordlength plots between the synthesized and true fields (Fig. 6.5(b)) the chordlength model, which is unrelated to our model in (2.19), validates the consistency of the synthesized samples with each other and with the true sample. The sensitivity of the free parameter ϵ in (2.19) is assessed by generating samples as a function of ϵ; as shown in Fig. 6.5(c) the proposed algorithm is insensitive to small ϵ.

To illustrate the strength of the tree-structured HHMF, we compare our proposed method with other methods in label field modeling and texture synthesis. First, a single MRF is used to synthesize both binary and ternary fields based on Fig. 6.1(c,d). The synthesized label fields are shown in Fig. 6.8(a,b) where we see that the synthesized structures are local and stationary, rather than presenting the multi-scale structures appearing in the true label maps. In contrast, the FSHF and the tree-structured HHMF models exhibit their capabilities of capturing complex structures in the label fields by capturing the presence of scale-dependent behavior in Fig. 6.8(c,d).

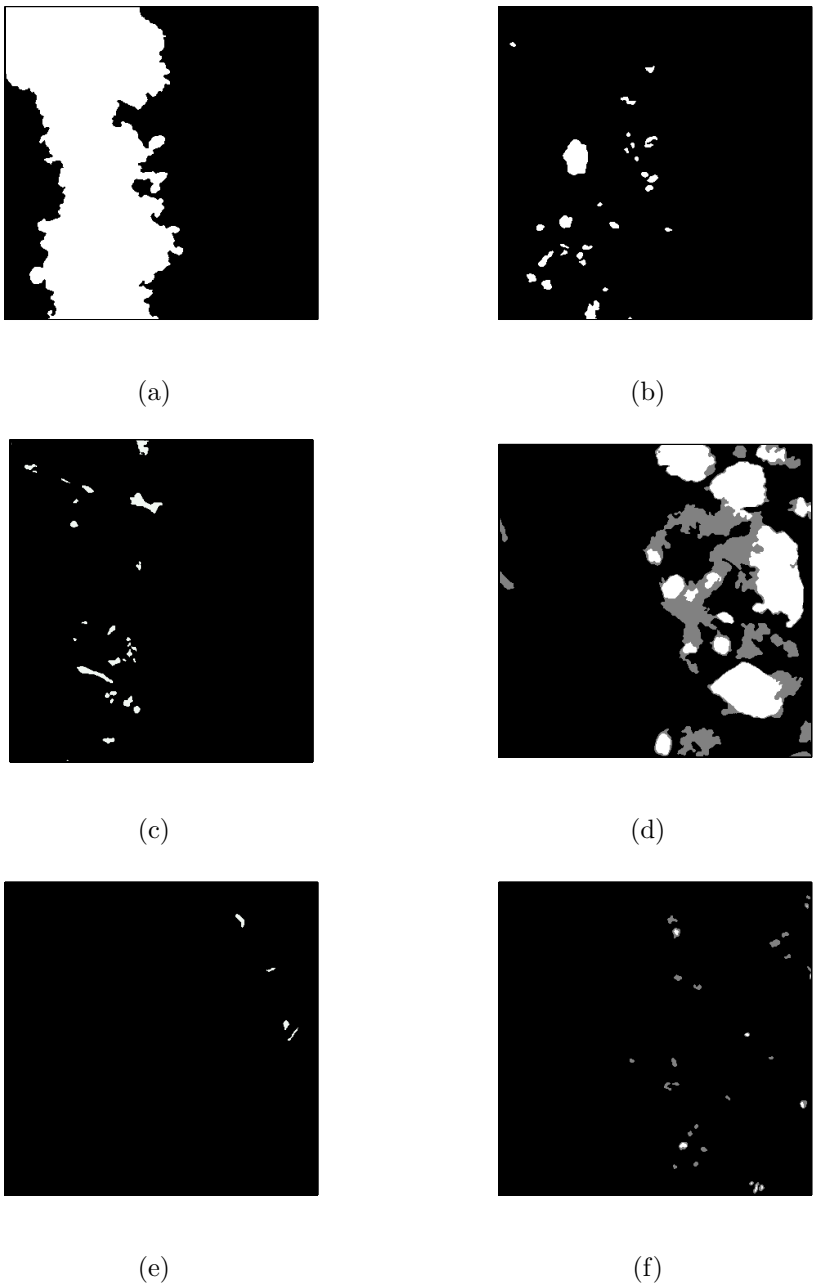
As a second comparison, we compare our proposed data synthesis method with two non-parametric texture synthesis methods: one a pixel-based sampling [34] and the other a patch-based sampling [66]. The basic idea of both non-parametric methods is to directly sample the given image by using self-similarity, a concept widely used in texture synthesis. For the relatively simple sea-ice training data in Fig. 6.1(a), the generated textures from the two comparison methods are given in Fig. 6.9(a,b), which provide a good reproduction of the training sample. Similarly, given training data with more complex structure in Fig. 6.1(b), the nonparametric methods also provide quite attractive results in Fig. 6.9(d,e) and demonstrate a good ability in structure representation.

However, there are three significant issues. First, the nonparametric methods are sensi-

tive to the synthesis staring seed, such that for certain seeds the synthesis may fail to sense certain significant structures present in the training data, as may be seen in comparing Fig. 6.9(e) with Fig. 6.1(b). Second, the nonparametric methods are sensitive to window size, such that a small window fails to sense large-scale structure, whereas a large window can lead to copying portions of the training image, as may be seen in Fig. 6.9(d), rather than random sampling. Finally, and most significantly, the texture synthesis methods synthesize *only* the texture, and have no notion of the underlying label field, which is essential for the testing of classification and segmentation algorithms.

A third experiment compares to a recent method [114] in which a nonparametric method *does* generate the label field as part of synthesis. Developed from the patch-based sampling method of [66], the method inherits the same advantages and disadvantages of patch-based methods. Because the focus of [114] was on texture synthesis, and not necessarily the quality of the underlying field, the synthesized ground-truth is relatively poor, as shown in Fig. 6.10. In particular, the synthesized label field is rather sensitive patch size (Fig. 6.10(a-c)), and suffers from blocky and repetitive artifacts (Fig. 6.10(c,d,f)).

As a final experiment, in contrast to the sea-ice imagery shown in previous examples, land-mass imagery in Fig. 6.11(a), with a corresponding label field in Fig. 6.11(d), are also used as training samples. The synthesized label field and texture are shown in Fig. 6.11(e,f). Compared to the results in Fig. 6.11(b,c) from the nonparametric methods [34, 66], our proposed approach shows more flexibility in producing random syntheses, particularly given the similar structures which appear in the nonparametric syntheses, copied from the training data.



(a) (b)

(c) (d)

(e) (f)

Figure 6.2: A complex multi-label map (training label field) can be manually decoupled as several binary or ternary fields with relatively simple structures. For example, the label map from Fig. 6.1(d) is decomposed here in (a-f). Although some fields, such as (a), (b) and (d), still contain structures at multiple scales, each decomposed field becomes much simpler than the original.

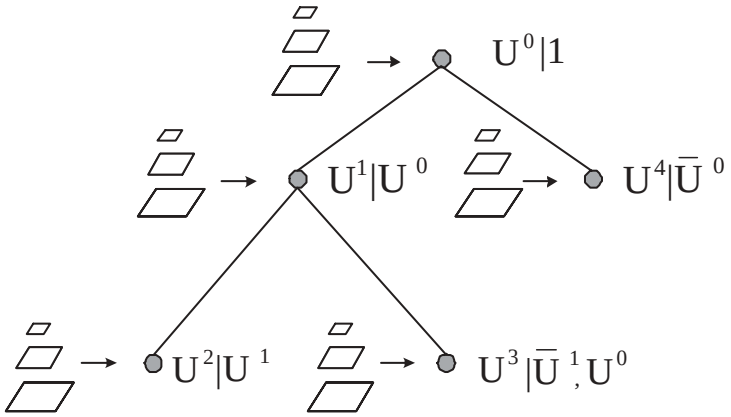
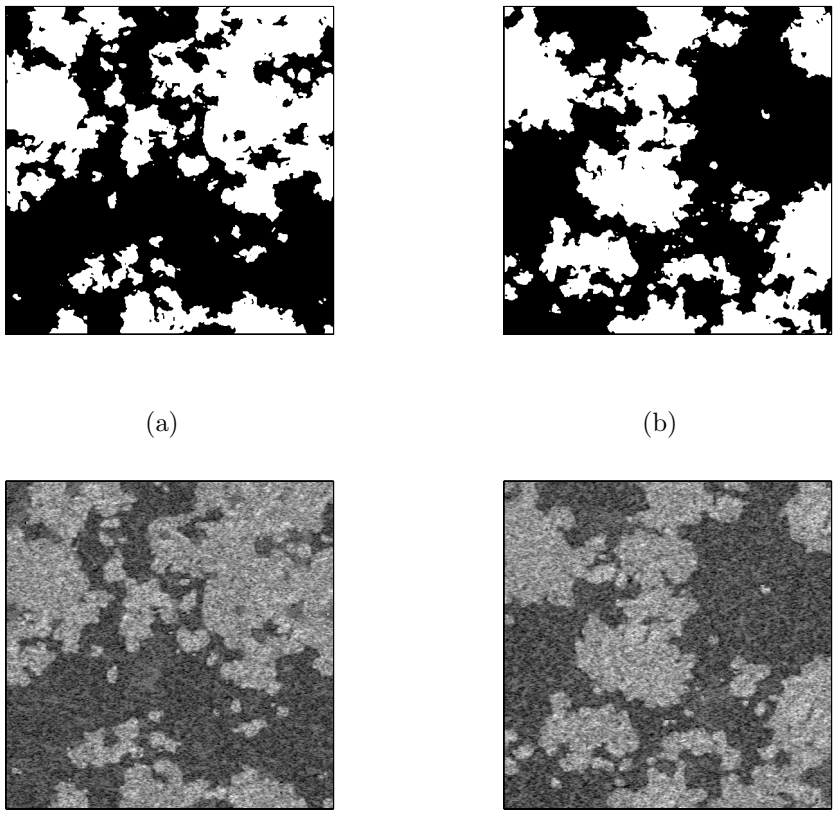


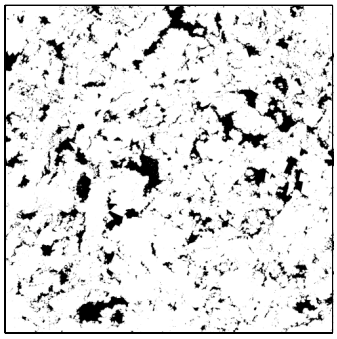
Figure 6.3: The modeling structure of tree-structured HHMF. The partition tree has a hierarchical field at each node, where the field U^i is conditioned on the behaviour of its parent, or both parent and grandparent.



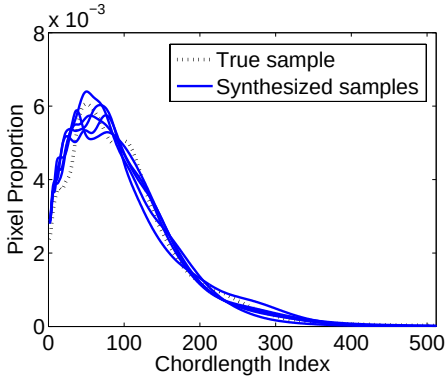
(a) (b)

(c) (d)

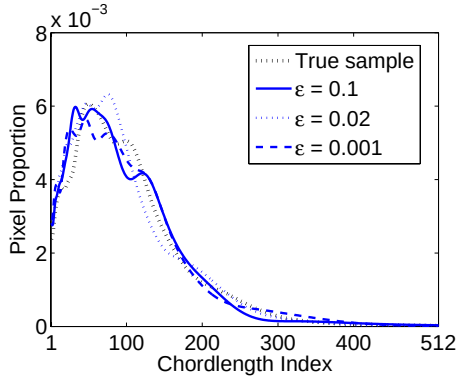
Figure 6.4: Binary-label sea-ice samples synthesized using a frozen state hierarchical model. Trained by the label sample in Fig. 6.1(c), the synthesized label maps (a) and (b) not only maintain similar structure statistics of the training sample, but have significant variations. From (a) and (b), it is straightforward to conditionally sample the textures shown in (c) and (d), which are comparable to the true sea-ice sample of Fig. 6.1(a).



(a) Binary test sample

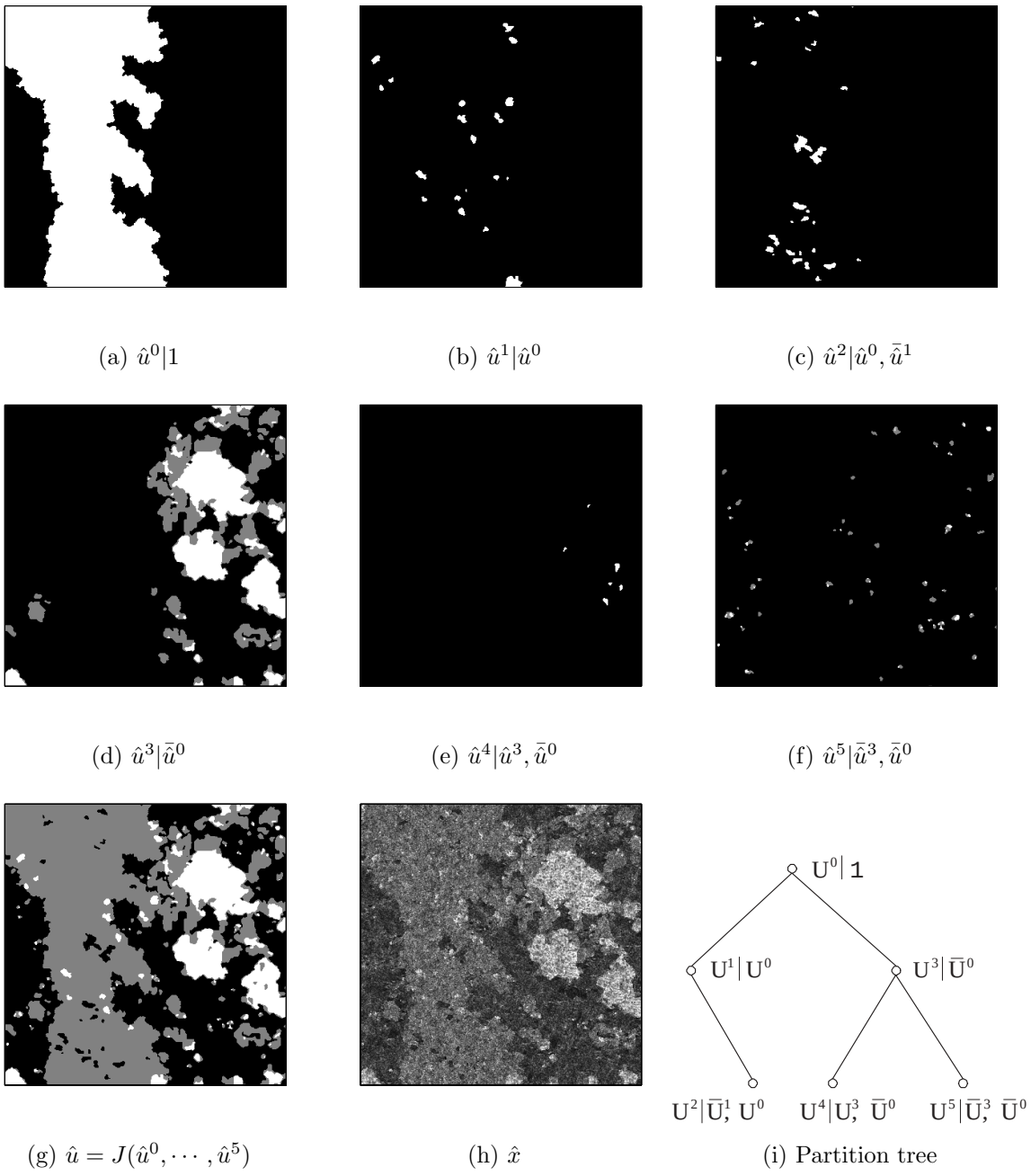


(b) Chordlength of syntheses with multiple runs



(c) Chordlength of syntheses as a function of ϵ,showing limited sensitivity.

Figure 6.5: The hierarchical model of the proposed modeling structure is evaluated using a chordlength model [104]: (a) A large binary microscopic excerpt (2048×2048) for model evaluation. (b) Chordlength distributions from multiple synthesis runs, (c) Chordlength as a function of parameter ϵ in (2.19).

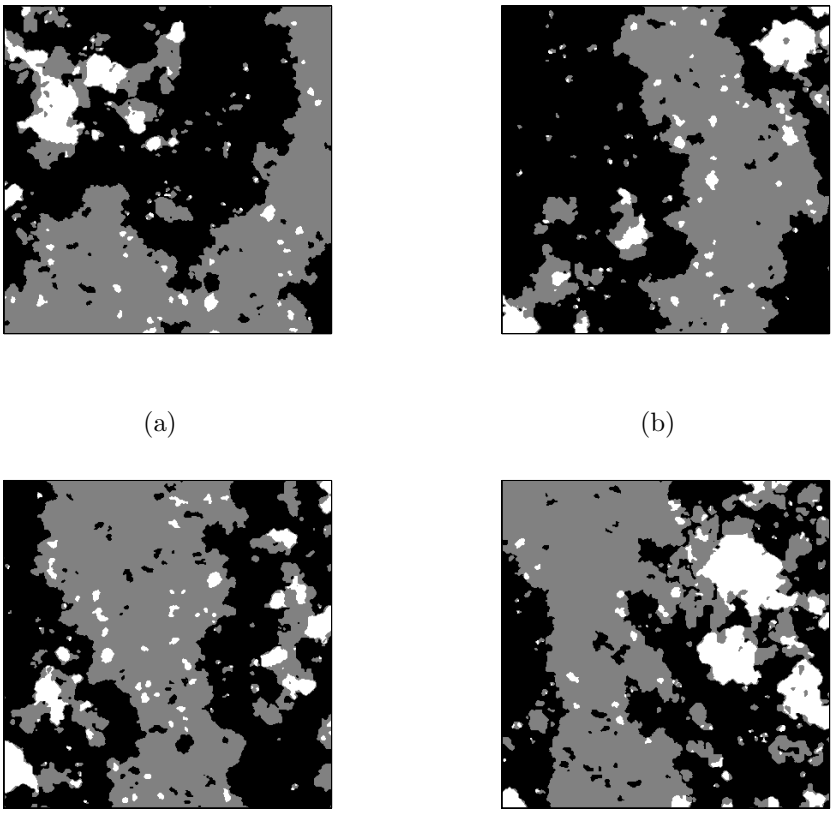


(a) $\hat{u}^0|1$ (b) $\hat{u}^1|\hat{u}^0$ (c) $\hat{u}^2|\hat{u}^0, \bar{\hat{u}}^1$

(d) $\hat{u}^3|\bar{\hat{u}}^0$ (e) $\hat{u}^4|\hat{u}^3, \bar{\hat{u}}^0$ (f) $\hat{u}^5|\bar{\hat{u}}^3, \bar{\hat{u}}^0$

(g) $\hat{u} = J(\hat{u}^0, \cdots, \hat{u}^5)$ (h) $\hat{x}$ (i) Partition tree

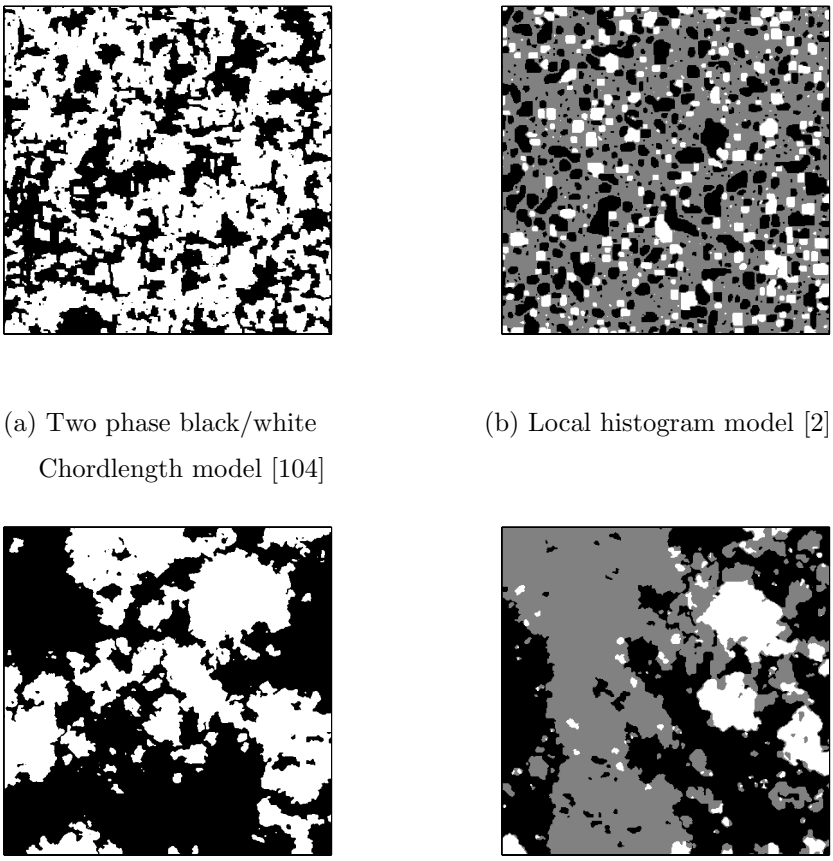
Figure 6.6: A multi-label synthesis with the proposed tree-structured hidden hierarchical model. The synthesized component fields are shown (a-f), corresponding to the training samples shown in Fig. 6.2. The simple binary / ternary fields can be combined, based on the tree structure (i), to achieve the final label map (g), which is clearly similar to the true label map in Fig. 6.1(d). Given (i), a synthesized sea-ice texture sample (h) is produced.



(a) (b)

(c) (d)

Figure 6.7: Multi-label sea-ice samples synthesized with multiple runs, based on the same training samples from Fig. 6.2 and the same tree structure in Fig. 6.6(i).



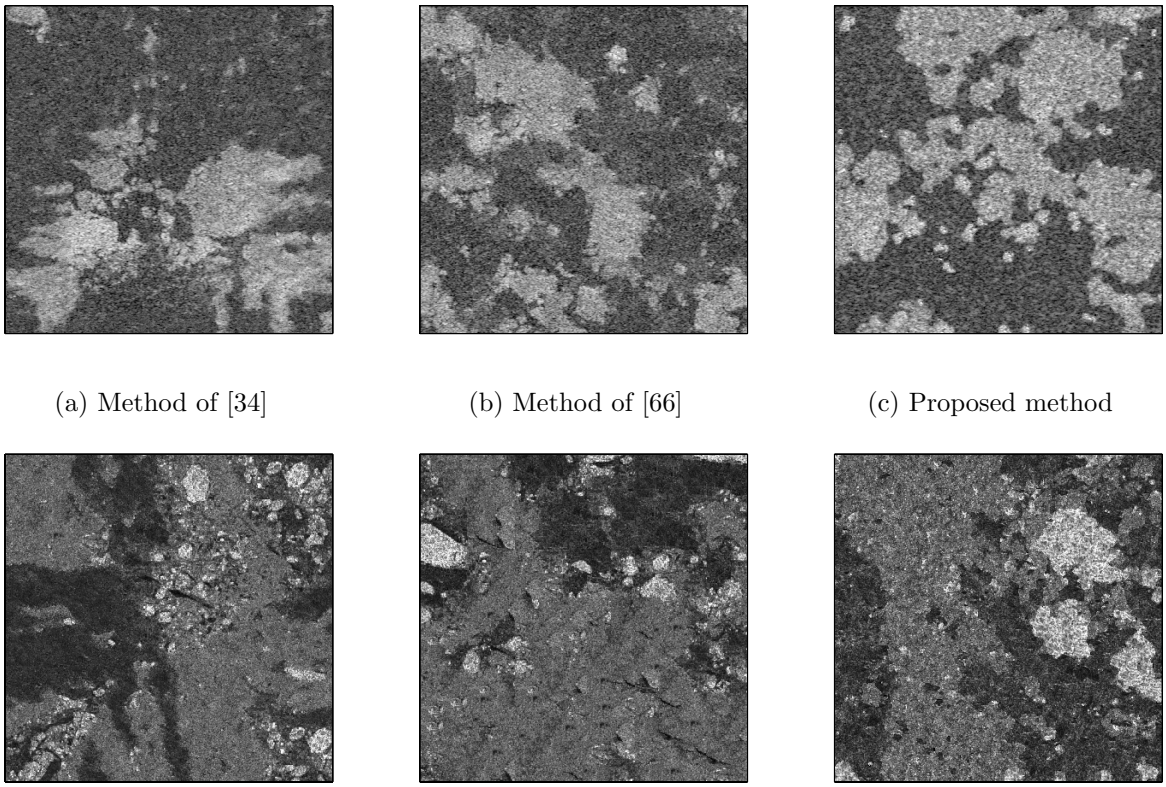
(a) Two phase black/white Chordlength model [104]

(b) Local histogram model [2]

(c) FSHF

(d) proposed method

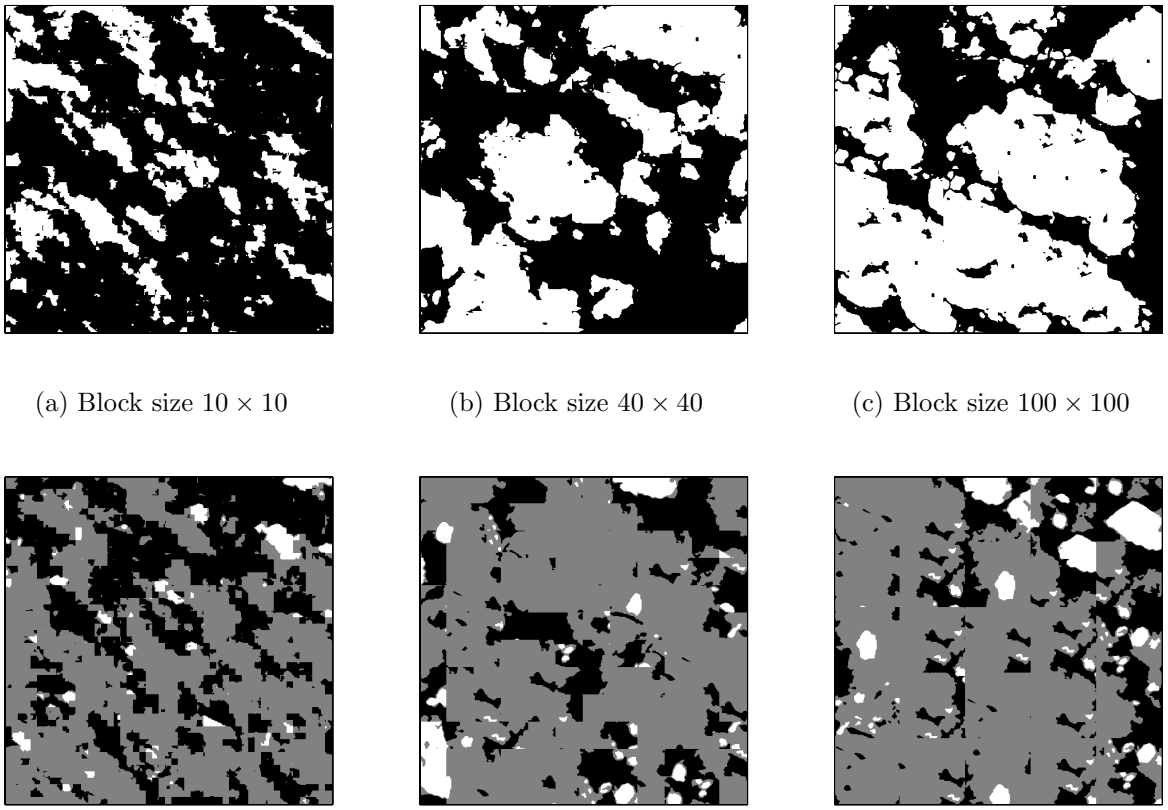
Figure 6.8: Sea-ice label map synthesis comparison. Panels (a,b) show label fields resulting from single Markov fields, based on chordlength and local-histogram models, whereas panels (c,d) show the label fields from the scale-dependent FSHF and the tree-structured HHMF, respectively. It is clear that the single Markov models can only provide stationary fields, with structure on one scale, as opposed to the nonstationary and scale-dependent structures possessed by the real label maps in Fig 1(c,d), and which are well captured by the multi-scale models in the FSHF and the tree-structured HHMF.



(a) Method of [34] (b) Method of [66] (c) Proposed method

(d) Method of [34] (e) Method of [66] (f) Proposed method

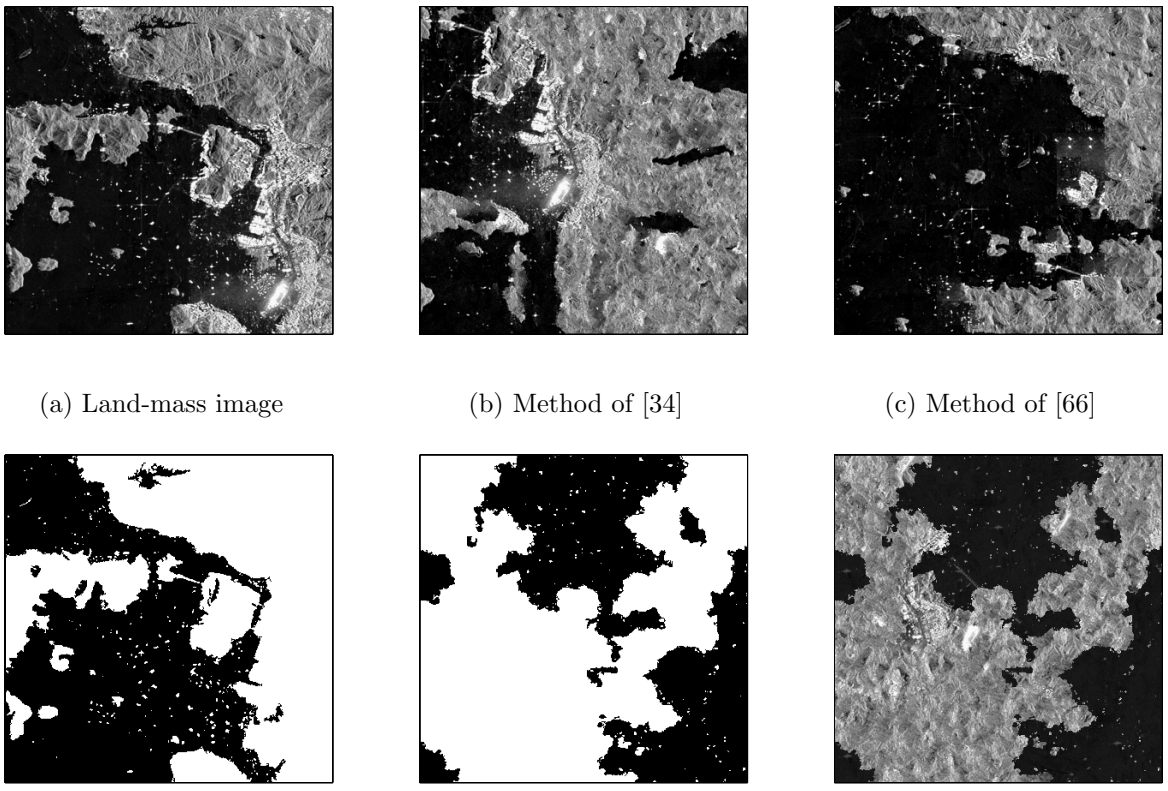
Figure 6.9: Sea-ice texture synthesis comparison, based on the pixel-based non-parametric sampling method [34] (a,d), patch-based sampling method [66] (b,e), and our texture synthesis method proposed in Section 6.5 (c,f). The top row shows two-label synthesis results, and should be compared to Fig. 6.1(a). The bottom row shows three-label syntheses, compared to Fig. 6.1(b).



(a) Block size 10×10 (b) Block size 40×40 (c) Block size 100×100

(d) Block size 10×10 (e) Block size 40×40 (f) Block size 100×100

Figure 6.10: Sea-ice label map synthesized by the IceSynth2 method [114]. Based on the binary and ternary label samples of Fig. 6.1(c,d), samples (a-c) and (d-f) are synthesized with the stated patch sizes. The sensitivity of the result to patch size is clear, as is the blockiness and repetitive artifacts in (c,d,f).



(a) Land-mass image

(b) Method of [34]

(c) Method of [66]

(d) True label field of (a)

(e) Synthesized label field from proposed method

(f) Synthesized texture from (e)

Figure 6.11: Land-mass imagery (a) with its given label field (d). Both the synthesized label field (e) and texture (f) resemble the true label map (d) and texture (a) well.

Chapter 7

Conclusions and Future Directions

7.1 Summary

The contributions of this thesis research are in both methodologies and applications. In methodology, HHMF models are proposed (Chapter 5, 6) for addressing significant challenges of modeling scale-dependent, heterogeneous structures. In application, the proposed HHMF models have been successfully applied to porous media image reconstruction (Chapter 5) and remote-sensing image synthesis (Chapter 6).

The basic idea of the HHMF is to combine hierarchical field modeling with multiple hidden field partitioning methods, so that the heterogenous behaviors of the original discrete field can be decomposed into several simpler fields, at the same time, the computational complexity can be taken care of by an efficient hierarchical modeling method, such as the FSHF. Thus, the proposed methods provide a practical and reliable way for dealing with multi-model behaviours, multi-label states, and scale-dependency. However, how to infer HHMF modeling structures is highly problem-dependent. Through two real problems with complex data we have illustrated the design and implementation of the proposed HHMF models. From both reconstruction and synthesis results, we can see that the proposed HHMF methods have exhibited positive effects in handling different complex scale-dependent situations. In summary, in this thesis we have achieved the following accomplishments:

- A Parallel HHMF model has been proposed to handle complex scale-dependent struc-

tures with nonstationary, piece-wise heterogeneous behaviours. Given joint fields at coarser scales, the proposed model represents multi-model behaviours by a set of conditionally independent fields at finer scales, so that each of those parallel fields contains a simpler behaviour and can be captured hierarchically by simple, local models. This proposed modeling approach has been applied to porous media image reconstruction [68, 69, 70], where the significant reconstruction results show the flexibility and capability of the proposed model in dealing with complex, multi-scale structures.

- Alternatively, a highly heterogeneous arrangement of hidden fields, which is not parallelly separable but arranged more as a partition tree, leads to the proposed Tree-Structured HHMF model. This modeling approach integrates a region-oriented binary tree structure with a resolution-oriented hierarchical approach to allow for complex, multiscale structure modeling while maintaining high computational efficiency. Although inferring the tree structure makes a forward problem with measurements difficult, such a tree structure can be easily used to synthesize highly complex hidden fields. Given the scale-dependent, non-stationary nature of remotely sensed data, we have applied the proposed model to remote sensing image synthesis [71].

7.2 Future Research Directions

To extend the work of this thesis, potential research directions can be considered from both methodology and application points of view:

- Setting up an efficient model for representing multi-label states, especially for the discrete field, is still a challenge. For example, in a four-label case to represent a joint relationship among the eight neighboring pixels surrounding a central pixel requires a model with $4^8 = 65536$ parameters per scale. On the basis of this parameter explosion, the local histogram model used in the FSHF modeling suffers a critical problem, since a multi-label state, rather than binary, will bring too many degrees of freedom in defining the unfrozen-state. To avoid increasing parametrization and to keep the computational efficiency of the FSHF, the proposed HHMF approaches apply parallel or tree partition to decompose a multi-label field into multiple binary/ternary fields.

However, this modeling simplification is based on a spatial decoupling assumption which may not hold true for regions with highly spatially interrelated structures. Therefore a further study for efficient, multi-label modeling is required.

- Another prospective direction is to generalize to 3D discrete-state models. This is interesting topic that has been little touched but is related to many unsolved research problems. For example, how could we model a 3D discrete-state field with scale-dependent structures? Could we address a 3D modeling problem by a 2D HHMF sequence? Or, how could we develop a 3D hierarchical model with a tractable computational complexity? Besides, more challenges will rise in the situations where we need to associate a 3D prior with measurements. For example, issues related to 3D porous media image reconstruction may include how to match a high resolution 3D prior with a low resolution 2D sequence, how to train the 3D prior model with high resolution 2D images, how to reduce the computational complexity of 3D sampling, and so on.

- Although the HHMF model has offered a powerful, flexible modeling approach to handle the heterogenous behaviours of scale-dependent structures, in this thesis inferring a modeling structure still requires human input. An automated algorithm is desired for this purpose. Since the modeling structure is highly dependent on the behaviours and properties of the image in question, there is no general automatic way to go about this, but at least for some particular, actual data, an automated algorithm for model learning is worth a study.

- For the image application problems we have solved in this thesis, there is still a lot of further work to do. For example, for the problem of HHMF based remote sensing image synthesis, more than synthesizing samples, its impact on classification processes would be an interesting problem to study, such as the differences between the synthesized ground-truth and the true ground-truth in training a classifier.

Although in this thesis the proposed HHMF has been demonstrated to reconstruct porous media images and synthesize remote-sensing images, it should not limited in those two applications. In the future work, we would like to work on extending the HHMF to solve other image processing and analysis topics, such as denoising, segmentation, and feature extraction, where scale-dependent, discrete-state fields are associated with problem solving.

Appendix A

Correlation Evaluation

The quantification of our results is challenging because we are seeking to reconstruct the actual measured sample at a finer resolution, and not just to synthesize a statistically similar sample. Standard statistical comparisons, such as comparing the overall porosity or chordlength distributions, assesses the statistical similarity of the true sample $\overset{*}{x}$ and the reconstruction $\hat{x}$, but fails to assess whether the reconstruction is actually revealing the structures of the original sample. The mean-squared-error between $\overset{*}{x}$ and reconstructed field $\hat{x}$ could be computed, however the MSE fails to quantify the ability to reconstruct large-scale as opposed to small-scale features. Here a correlation method proposed by Dr. Paul Fieguth for evaluating discrete field reconstructions, to assess structural consistency between $\hat{x}$ and $\overset{*}{x}$ as a function of scale.

The scale of a feature is defined to be the average number of scales n_k over which a pixel value x_s is unchanged under repeated decimation. The decimation $q^k(x)$ at the kth scale for binary field $x_s \in \{-1, 1\}$ is a majority-vote decimator, returning zero in the case of a tie. The decimation can be a function of the dyadic origin o, therefore we refer to the decimation $q_o^k(x)$ as a function of o. The scale associated with the feature at pixel s is therefore

$$n_k(s) = \text{average}_o \left\{ \min \left\{ k \ni \{q_o^k(x)\}_s = x_s \right\} \right\} \tag{1}$$

With a feature scale defined, we can therefore correlate $\hat{x}$ and $\overset{*}{x}$, calculated over a subset Q as a function of feature scale n_k:

$$\rho_{n_k}(\overset{*}{x}, \hat{x}) = \text{correlation}(\overset{*}{x}_Q, \hat{x}_Q) \;\text{ over }\; Q = \left\{ s \ni |n_k(s) - n_k| < \frac{1}{2} \right\} \tag{2}$$

Appendix B

Example Reconstruction Procedure

The overall procedure of the proposed parallel hidden hierarchical field method for porous media image reconstruction (Chapter 5) includes model training, learning, and sampling.

Training data preparation

The training data used for a reconstruction example is a large, high resolution binary image $\tilde{x}$ (4096×4096) for a particular type of porous media. For example, to reconstruct a two-scale porous media image (X) (Fig. 5.1(b)) from its observation Y (Fig. 5.2(b)), an excerpt of training image is used and shown in Fig. 1(b). The number of model behaviours N_ν in the training sample $\tilde{x}$ is inferred manually, such as $N_\nu = 1$ for Fig. 1(b). Then, according to N_ν, $\tilde{x}$ is segmented as a label field $\tilde{u}$ under human supervision.

For example: the training image $\tilde{x}$ (Fig. 1(b)) is first segmented by a morphological method to close small scale pores (black pixel). Then, the small pores are further merged by thresholding the region size, where the threshold is set empirically. Finally, a binary segmentation result $\tilde{u} = \{\tilde{u}_i, i = 0\}$ is obtained, such as shown in Fig. 1(c), which has been used as the training data for modeling a hidden label field U.

For a more complicated type of training image field $\tilde{x}$ where we are not able to get an appropriate segmentation easily, we label the image manually to get a multi-label field $\tilde{u}$

All the porous media data are provided by Professor M. Ioannidis, Dept. of Chemical Engineering, University of Waterloo for providing sample porous media images.

and decompose it to a set of binary label fields $\{\tilde{u}_i, i \in (0, \cdots, N_{\nu-1})\}$, where

$$\tilde{u}_{i,s} = \begin{cases} \tilde{u}_s & \text{if } \tilde{u}_s = i, \quad \forall s \in S \\ 0 & otherwise. \end{cases} \tag{3}$$

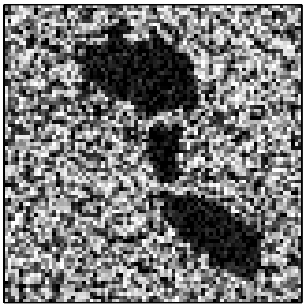
Given the label field $\tilde{u}$ and its decomposed binary fields $\{\tilde{u}_i\}$, we can have conditional training data $\tilde{x}|\tilde{u}$ and $\{\tilde{x}|\tilde{u}_i\}$ which are used to train image models for a particular hidden state or every model behaviour.

Model learning

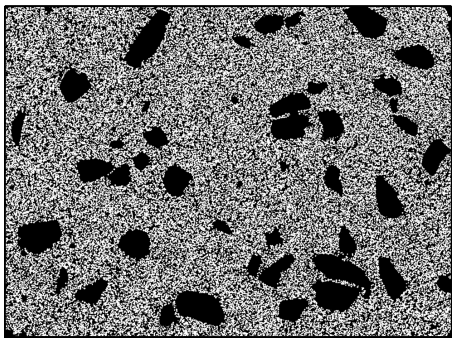
After we obtain the training data, we can start setting up a hidden hierarchical field model similar shown in Fig. 5.7. At first, to find the coarsest scales (M_x, M_u) for each hierarchy, we keep down-sampling training image $\Downarrow^k(\tilde{x})$, and training hidden field $\Downarrow^k(\tilde{u})$ with a ternary state representation $\{0, 1, \frac{1}{2}\}$ (3.18). At some scale k, as the majority states of $\tilde{x}^k$ become undetermined, we set the coarsest scale for hierarchical field $\{X^k\}$ to be $M_x = k$. Similarly, $\{U^k\}$ can find its coarsest scale M_u under the same procedure. Meanwhile, if at some scale k the different behaviour labels of $\tilde{u}^k$ start touching, we record this scale k as k_d, then $\{U^k, k > k_d\}$ needs to be modeled as a joint-label field; otherwise $\{U^k\}$ can always be decoupled as multiple independent fields $\{U_i^k\}$.

Then, given the coarsest scales M_x, M_u and the decomposition scale k_d, we represent training samples ($\{\tilde{x}|u_i\}$, $\{\tilde{u}\}$, $\{\tilde{u_i}\}$) as frozen state hierarchical fields (3.18) and correspond them to a modeling structure, such as Fig. 5.7. Then, at each scale target histograms $\tilde{h}^k_{x|u}$, $\tilde{h}^k_u$, and $\{\tilde{h}^k_{u_i}\}$ are defined by a local histogram (Section 2.2.3), and learned from hierarchical training image field $\tilde{x}^k|(\tilde{x}^{k+1}, \tilde{u}^k)$, training hidden field $\tilde{u}^k|\tilde{u}^{k+1}$, and training decoupled hidden field $\{\tilde{u}_i^k|\tilde{u}_i^{k+1}\}$.

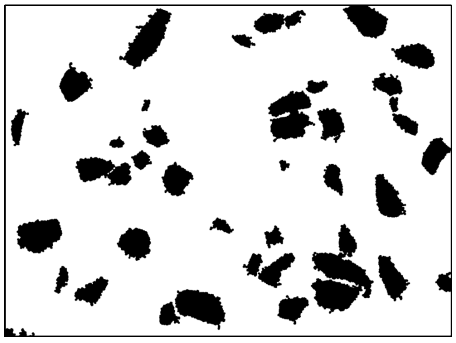
The relationship $p(y|u)$ is defined by histogram and learned from the training data $\tilde{x}$ and $\tilde{u}$. The measurement histogram $\tilde{h}^k_{y|u}$ is learned at each scale from training data $\tilde{y} = g(\tilde{x}^k)$ given every possible configuration of a joint field $\tilde{u}^k = (\tilde{u}_0^k, \cdots, \tilde{u}_{N_\nu-1}^k)$, such as shown in Fig. 5.9. Here, $g(\cdot)$ denotes a forward operation which down-samples $\Downarrow^k(\cdot)$ (as a bottom-up process for Fig.3.8) or up-samples $\Uparrow^k(\cdot)$ (as a top-down process for Fig.3.8) a field to the measurement scale ($k = k_m$).



(a) A noisy low resolution sample shown in Fig. 5.2(b)



(b) An excerpt train data



(c) The hidden label field of (b)

Figure 1: For reconstructing a noisy low resolution sample (a), the high resolution training data (b) is selected with the same porous media image type, and its corresponding hidden label field of (b) is inferred and shown in (c).

The estimation process

In estimation, we first estimate hidden field U over all scales, such that the estimated hidden field $\hat{u}$ is definitive (not uncertain) at every point, and then estimate an image field X. This estimation procedure suggests that a constant annealing schedule, with sequential minimization over scales from coarse ($k = M_u, M_x$) to fine ($k = 0$), and sequential minimization from hidden (U) to visible states (X), lead to a reconstruction $\hat{x}$.

When estimating the hidden field U, in which case X is unknown, U is estimated according to a measurement model $p(y|u)$ and prior $p(u)$. The hidden field U is estimated as joint fields $\{u^k\}$ only at some coarse scales $k > k_d$

$$p(u^k|u^{k+1}, y) \propto \prod_{s \in S_L} p(y_s|u^k) \cdot p(u^k|u^{k+1}), \quad k > k_d, \tag{4}$$

Then, as we learned from the training hidden field $\tilde{u}$, different model behaviour become separable at some scale k_d and U is estimated separately as multiple independent hierarchical fields $\{u_i\}$

$$p(u_i^k|u^{k+1}, y) \propto \prod_{s \in S_L} p(y_s|u_i^k) \cdot p(u_i^k|u^{k+1}), \quad k = k_d, \tag{5}$$

$$p(u_i^k|u_i^{k+1}, y) \propto \prod_{s \in S_L} p(y_s|u_i^k) \cdot p(u_i^k|u_i^{k+1}), \quad k < k_d, \tag{6}$$

After U is estimated, image field X is estimated given the estimated hidden hierarchical fields $\{\hat{u}_i\}$, therefore at each scale we have

$$p(x^k|x^{k+1}, u_0^k, \cdots, u_{N_\nu - 1}^k, y) \propto \prod_{s \in S_L} p(y_s|x^k) \cdot (x^k|x^{k+1}, u_0^k, \cdots, u_{N_\nu - 1}^k), \quad k < M_x, \tag{7}$$

At each scale, the above posterior distributions (4)-(7) are represented as a Gibbs distribution (2.10). Finally, at each scale $\{X^k\}$ and $\{U_i^k\}$ can be estimated by minimizing their posterior energies using MCMC method (Section 2.3).

In this work, all the prior energy functions are defined by local histogram model (2.19)

$$H(n) = \sum_{n=0}^{N_h - 1} \frac{|h(n) - \tilde{h}(n)|}{\nu(n) + \epsilon}$$

where $h(n)$ is the estimated local histogram, $\tilde{h}(n)$ is the target histogram learned from training data, N_h is the number of possible neighborhood configurations, ϵ is empirically

set as 0.001 to control the penalty for the unallowed local configurations, and ν normalizes coefficients, such as $\nu(n) = \sqrt{\tilde{h}(n)}$ [20].

Measurement energies $E_m(y|x^k)$ evaluates the inconsistency between the hierarchical reconstruction x^k and the observed field y, and is defined as

$$H_m(y|x^k) = \|y - g(x^k)\| \tag{8}$$

Measurement energies of $\{p(y|u_i^k)\}$ and $\{p(y|u^k)\}$ are inferred empirically from the training data.

Model sampling

With modeling performed, we first sample hidden fields $\{\hat{u}_i\}, i \in (0, \cdots, N_\nu)$, then sample image field X, given the hidden estimates. For each estimated hierarchical field, the sampling process follows the structure of frozen state hierarchical field (3.20) from the coarsest scale to the finest scale. Each hierarchy is randomly initialized, and sampled at progressively finer scales using Gibbs sampler (Algorithm 2.23) with Simulated Annealing (SA)(Algorithm 2.3.3).

The SA process is started at a high temperature ($T_0 = 100$), then we decrease the temperature slowly according to a geometric scheme: $T_n = \rho(T_n)T_{n-1}$, where n is the number of the iteration, and $\rho(T_n)$ is separately set to be $0.996, 0.993, 0.99, 0.98, 0.96, 0.93$ as T_n in the different interval: $[100\ 30], [30\ 10], [10\ 3], [3\ 0.1], [1e-1\ 1e-2], [1e-2\ 1e-4]$.

When all of the hidden hierarchical field have been sampled $\{\hat{(u)}_i\}$, the image field X is sampled based on (7) to obtain the reconstruction $\hat{x}$.

Appendix C

Example Synthesis Procedure

Remote sensing image synthesis (Chapter 6) includes two steps:

- Synthesize hidden label field U as ground truth (Section 6.4).
- Synthesize realistic imagery X given the given hidden label field U (Section 6.5).

The implementation procedure of the two steps is described as follows.

Training data preparation

To train the models for representing both label field and texture field of a remote sensing image, the data need to be prepared for three purposes:

- to infer a tree-structure T (6.2) which is used to combine multiple binary/ternary node fields to be the hidden label field U,
- to train frozen state hierarchical field models (6.4) which represent the multiple binary/ternary fields,
- to train a conditional nonparametric model (6.6) which describes the image field X conditioning on the state of an underlying label field U.

The sea-ice images used in this work are provided by Professor D. Clausi, and their corresponding true label fields are provided by former Ph.D. student A. Wong, both from Department of Systems Design Engineering, University of Waterloo.

In our preparation, first a true label field sample (with L label states), such as shown in Fig. 6.1(d), is used as training data $\tilde{u}$ for inferring a partition tree structure T. After the partition process, the complex training label field $\tilde{u}$ has been decomposed as a set of simple binary/ternary fields $\{\tilde{u}^i|\tilde{\mathcal{Q}}^i, 0 \leq i < N\}$, such as shown in Fig. 6.2. Then, we just take $\{\tilde{u}^i|\tilde{\mathcal{Q}}^i\}$ as training data for learning the conditional hierarchical fields $\{U^i|\mathcal{Q}^i\}$. Meanwhile, since the true label $\tilde{u}$ of the training image $\tilde{x}$ is known, for every label state $\{l, l \in (0, \cdots, N-1)\}$ its corresponding training image$\tilde{x}|\tilde{u}_{|l}$ ($\tilde{u}_{|l} = \{\tilde{u}_s|\tilde{u}_s = l, \forall s \in S\}$) can be easily obtained, which is taken as training data for learning the image's texture behaviour, such as shown in Fig. 6.1(b), given every label state l in its label field $\tilde{u}$.

Inferring a partition tree

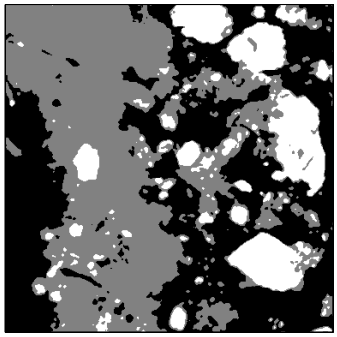
The process for inferring a partition tree structure T from a given training label field $\tilde{u}$ requires human input. We manually partition a training label field from a multi-label, multi-scale field to a set of binary/ternary fields and each partitioned field contains spatial structures at similar scales. A successful tree inference requires that the training label field is spatially separable at each partition.

An inference example (Fig. 6.2) has been discussed in Section 6.4, however we would repeat it here to make the implementation step complete. For a training label sample $\tilde{u}$ (Fig. 2(a)) which has a dominant large-scale structure, we manually extract the dominant structure and represent its spatial behaviour by a binary field $\tilde{u}^0|1$ (Fig. 2(b)). Then, given $\tilde{u}^0|1$ the original label is partitioned as two label fields, as shown in Fig. 2(c)(d), where the scales of the spatial structures are less variant than the original field (Fig. 2(a)).

Since both foreground and background still contains mixed labels and multiple scale structures, the partition process continues until $\tilde{u}$ is decomposed as binary/ternary fields, such as shown in (Fig. 6.2). Finally the original training label field $\tilde{u}$ has been represented as a partition tree T as shown in Fig. 6.6(i) and the relationship $J(\cdot)$ among the partitioned fields is recorded, such as (6.8).

Conditional hierarchical field learning

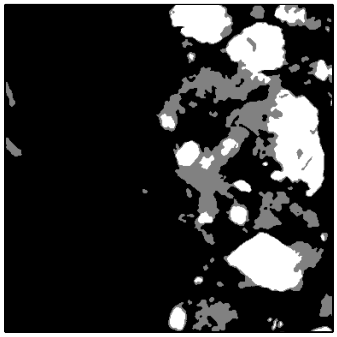
After a partition tree structure has been inferred, to train the conditional hierarchical fields at each node is straightforward. The partitioned binary/ternary label fields $\{\tilde{u}^i|\tilde{\mathcal{Q}}^i\}$



(a) Training Label field shown in Fig. 6.1(d)

(b) Dominant large-scale structure of (a) shown in Fig. 6.2(a)

(c) Foreground field given (b)



(d) Background field given (b)

Figure 2: An example of partition tree inferring. A complex label field (a) is manually partitioned as two simpler fields (c) and (d) according to its dominant large-scale structure (b).

are used as training data. Those partitioned fields are represented as frozen state hierarchical fields, where the coarsest scale M^i for each hierarchy is learned by down-sampling $\tilde{u}^{i,k}|\tilde{\mathcal{Q}}^{i,k} = \Downarrow^k(\tilde{u}^i|\tilde{\mathcal{Q}}^i)$ until the majority of label states in $\tilde{u}^{i,k}|\tilde{\mathcal{Q}}^{i,k}$ become undetermined ($\frac{1}{2}$), then set $M = k$. For all of the training hierarchies, at each scale a target histogram $\tilde{h}^k_{u^i|\mathcal{Q}^i}$ is defined by a local histogram (Section 2.2.3).

Hidden label field sampling

With modeling performed, the sampling process follows two steps sequentially: First, a set of binary/ternary fields $\{\hat{u}^i|\hat{\mathcal{Q}}^i\}$ are sampled from the learned conditional hierarchical fields separately. Second, according the learned tree-structure T the synthesized fields $\{\hat{u}^i|\hat{\mathcal{Q}}^i\}$ are combined together to be the final synthesized label field $\hat{u}$.

In the sampling process, each hierarchy is randomly initialized at the coarsest scale, and sampled at progressively finer scales using Gibbs sampler (Algorithm 2.23) with simulated annealing (Algorithm 2.3.3). The cooling scheme is a geometric scheme: $T_n = \rho(T_n)T_{n-1}$, where $\rho(T_n)$ is set to be $0.996, 0.993, 0.99, 0.98, 0.96, 0.93$ separately as T_n in the different interval: $[100\ 30], [30\ 10], [10\ 3], [3\ 0.1], [1e-1\ 1e-2], [1e-2\ 1e-4]$.

When all of the conditional hierarchies $\{\hat{u}^{i,k}|\hat{\mathcal{Q}}^{i,k}\}$ have been sampled, recombining the synthesized conditional fields $\{\hat{u}^i|\hat{\mathcal{Q}}^i\}$ to obtain $\hat{u}$ is straightforward based on the specified partition tree T with a recorded relationship of partitioned fields $j(\cdot)$, such as the example shown in (6.8).

Image synthesis

For image synthesis, we slightly modify an existing method [34] to allow generating the fine-scale $\hat{x}$ on top of a synthesized underlying label field $\hat{u}$ (6.6)

$$\hat{x} \leftarrow p(x|\hat{u})$$

where $p(x|\hat{u})$ is assumed to be a MRF. The sampling process (6.7) generates a pixel x_s based on its neighbourhood $x_{\mathcal{N}_s}$ and its corresponding hidden label state $\hat{u}_s$:

$$\hat{x}_s \leftarrow p(x_s|x_{\mathcal{N}_s}, \hat{u}_s).$$

The MRF can be specified nonparametrically by an empirical histogram. In our implementation, same as [34] we directly sample image field X from the training image sample $\tilde{x}$ instead of constructing a model. However, differently from [34] we search for a set of matching patches in $\tilde{x}$ given the conditioning on u_s rather than just based on an image field matching. The distance between two image neighbourhoods given a hidden label state u_s is defined as

$$d\left(\mathcal{N}_{x_s}, \mathcal{N}_{\tilde{x}_{s'|u_s}}\right) = \left\|G\left(\mathcal{N}_{x_s} - \mathcal{N}_{\tilde{x}_{s'|u_s}}\right)\right\|_2 \tag{9}$$

where G is a Gaussian kernel (7×7) and used to emphasize the local structure [34]. The size of matching neighbourhoods ($\mathcal{N}_{x_s}$ and $\mathcal{N}_{\tilde{x}_{s'}|u_s}$) has been set as 7×7 for our experiments in Chapter 6.

After comparing a matching neighbourhood $x_{\mathcal{N}_s}$ to all possible neighbourhoods $\{\mathcal{N}_{\tilde{x}_{s'}|u_s}\}$ in a conditional training subset $\tilde{x}|u_s$, we can find the minimum distance

$$d_{min} = \arg\min_{d} \; d\left(\mathcal{N}_{x_s}, \mathcal{N}_{\tilde{x}_{s'}|u_s}\right), \quad \forall s' \in S. \tag{10}$$

Then, a group of close matches $\Lambda_{\tilde{x}_{s'}}$ are selected by thresholding d

$$\Lambda_{\tilde{x}_{s'}} = \{\tilde{x}_{s'} \mid d\left(\mathcal{N}_{x_s}, \mathcal{N}_{\tilde{x}_{s'}|u_s}\right) < (1+\tau) \cdot d_{min}\} \tag{11}$$

where τ is a relaxation coefficient and is empirically set to be 0.3 in our experiments.

Finally, the texture of the image is sampled randomly from $\Lambda_{\tilde{x}_{s'}}$

$$\hat{x_s} \leftarrow \Lambda_{\tilde{x}_{s'}}. \tag{12}$$

Bibliography

[1] S. K. Alexander, P. Fieguth, M. Ioannidis, and E. R. Vrscay. Hierarchical annealing for random image synthesis. *In Proc. of Energy Minimization Methods in Computer Vision and Pattern Recognition 2003, LNCS, Springer*, (2683):357–378, 2003.

[2] S. K. Alexander, P. Fieguth, M. Ioannidis, and E. R. Vrscay. Hierarchical annealing for synthesis of binary porous media images. *Mathematical Geoscienes*, 41:357–378, 2009.

[3] S. K. Alexander, P. Fieguth, and E. R. Vrscay. Parameterized hierarchical annealing for scientific models. *International Conference on Image Analysis and Recognition 2006, LNCS, Springer*, pages 236–243, 2006.

[4] C. B. Atkins, C. A. Bouman, and J. P. Allebach. Tree-based resolution synthesis. *In Proc. of the Image Processing, Image Quality, Image Capture Systems Conference*, pages 405–410, 1999.

[5] A. Barbu and S. C. Zhu. Generalizing Swendsen-Wang to sampling arbitrary posterior probabilities. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(8):1239–1253, 1994.

[6] M. Basseville, A. Benveniste, K. C. Chou, S. A. Golden, R. Nikoukhah, and A. S. Willsky. modeling and estimation of multiresolution stochastic processes. *IEEE Trans. on Information Theory*, 38(3):766–784, 1992.

[7] D. Benboudjema and W. Pieczynski. Unsupervised image segmentation using triplet Markov fields. *Computer Vision and Image Understanding*, 99(3):476–498, 2005.

[8] D. Benboudjema and W. Pieczynski. Unsupervised statistical segmentation of non-stationary images using triplet Markov fields. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(8):1367–1378, 2007.

[9] D. Benboudjema and W. Pieczynski. Unsupervised statistical segmentation of non-stationary images using triplet Markov fields. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(8):1367–1378, 2007.

[10] C. Benedekand, T. Sziranyiand, Z. Kato, and J. Zerubia. A multi-layer MRF model for object-motion detection in unregistered airborne image-pairs. *IEEE International Conference on Image Processing*, 6:VI – 141–VI – 144, 2007.

[11] J. Benediktsson, M. Pesaresi, and K. Amason. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Trans. on Geoscience and Remote Sensing*, 41(9):1940–1949, 2003.

[12] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society Series B (Methodological),*, 36(2):192–236, 1974.

[13] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 48(3):256–302, 1986.

[14] D. Blacknell. A new method for the simulation of K-distribution clutter. *IEEE Proc. on Radar Sonar Navigation*, 141:53–58, 1994.

[15] D. Blacknell, A. Blake, P. Lombardo, and C. Oliver. A comparison of simulation techniques for correlated gamma and K-distributed images for SAR applications. *In Proc. of IEEE International Geoscience and Remote Sensing Symposium*, 4:2182–2184, 1994.

[16] J. Blanchet and F. Forbes. Triplet matkov field for the classification of complex structure data. *research report RINRIA*, 2007.

[17] L. Bombrun, J. M. Beaulieu, G. Vasile, J. P. Ovarlez, F. Pascal, and M. Gay. Hierarchical segmentation of polarimetric sar images using heterogeneous clutter models. *IEEE International Geoscience and Remote Sensing Symposium 2009*, pages 5–8.

[18] C.A. Bouman and M. Shapiro. A multiscale random field model for Bayesian image segmentation. *IEEE Trans. Image Processing*, 3(2):162–177, 1994.

[19] B. Burt and E. Adelson. The laplacian pyramid as a compact image code. *IEEE Trans. on Communication*, 31(4):532–540, 1983.

[20] W.R. Campaigne, P. Fieguth, and S.K. Alexander. Forzen-state hierarchical annealing. *International Conference on Image Analysis and Recognition 2006, LNCS, Springer*, pages 41–52.

[21] H. Cantalloube. Texture synthesis for SAR image simulation. *Proc. SPIE*, 3497:242–250, 1998.

[22] M. Ceccarelli. A finite Markov random field approach to fast edge-preserving image recovery. *Image and Vision Computing*, 25:792–C804, 2007.

[23] P. Chainais. Infinitely divisible cascades to model the statistics of natural images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29:2015–2119, Dec. 2007.

[24] R. Chellappa. Two-dimensional discrete Gaussian Markov random field models for image processing. *In L. N. Kanal and A. Rosenfeld, editors, Progress in Pattern Recognition, Springer-Verlag*, 2:79–112, 1985.

[25] R. Chellappa and S. Chatterjee. Classification of textures using Gaussian Markov random fields. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 33(4):959–963, 1985.

[26] R. Chellappa and A. Jain. *Markov Random Fields: Theory and Applications*. Academic Press, 1993.

[27] R. Chellappa and R. L. Kashyap. Texture synthesis using 2-d noncausal autoregressive models. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 33(1):194–203, 1985.

[28] J. Chen and A. Kundu. Unsupervised texture segmentation using multichannel decomposition and hidden Markov models. *IEEE Trans. on Image Processing*, 4:603–619, 1995.

[29] H. Choi and R.G. Baraniuk. Multiscale image segmentation using wavelet-domain hidden Markov models. *IEEE Trans. on Image Processing*, 10(9):1309–1321, 2001.

[30] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk. Wavelet-based statistical signal processing using hidden Markov models. *IEEE Trans. on Signal Processing*, 46:886–902, 1998.

[31] P. Davis. *Circulant Matrices*. Wiley-Interscience, 1979.

[32] C. D'Elia, G. Poggi, and G. Scarpa. A tree-structured Markov random field model for Bayesian image segmentation. *IEEE Trans. on Image Processing*, 6(6):721–741, 2003.

[33] C. Van der Sande, S. De Jong, and A. De Roo. A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment. *International Journal of Applied Earth Observation and Geoinformation*, 4(3):217–229, 2003.

[34] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. *IEEE International Conference on Computer Vision, 1999*, pages 1033–1038.

[35] P. Fieguth. *The multigrid and Swendsen-Wang algorithms in statistical physics, Technical Report*. Massachusetts Institute of Technology, 1994.

[36] P. Fieguth. Hierarchical posterior sampling for images and random fields. *IEEE Trans. on Image Processing*, accepted in 2008.

[37] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Computer Graphics and Applications*, 22(2):56–65, 2002.

[38] R. Gaetano, G. Scarpa, and G. Poggi. Hierarchical texture-based segmentation of multiresolution remote-sensing images. *IEEE Trans. on Geoscience and Remote Sensing*, 47(7):2129–2141, 2009.

[39] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[40] M. G.Eramian, R. A. Schincariol, L. Mansinha, and R. G. Stockwell. Generation of aquifer heterogeneity maps using two-dimensional spectral texture segmentation techniques. *Mathematical Geology*, 31(3):327–348.

[41] B. Gidas. A renormalization group approach to image processing problems. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(2):164–180, 1989.

[42] J. Goodman and A. Sokal. Multigrid Monte Carlo method. conceptual foundations. *Physical Review*.

[43] C. Graffigne, F. Heitz, P. Pérez, F. Prêux, M. Sigelle, and J. Zerubia. Hierchical Markov random field models applied to image analysis: a review. *SPIE*, 2568:2–17, 1994.

[44] P. Kelly H. Derin. Discrete-index Markov-type random processes. *Proc. of the IEEE*, 77(10):1485–1510, 1989.

[45] J. M. Hammersley and P. Clifford. Markov field on finite graphs and lattices. *unpublished*, 1971.

[46] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics*, 3(6):610–621, 1973.

[47] M. Hassner and J. Slansky. The use of Markov random field as models of texture. *Computer Graphics Image Processing*, 12:357–370, 1980.

[48] W.K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–108, 1970.

[49] F.C. Jeng. Subsampling of Markov random fields. *Journal of visual communication and image representation*, 3(3):225–229, 1992.

[50] Y. Jiao, F. H. Stillinger, and S. Torquato. modeling heterogeneous materials via two-point correlation functions. *Physical Review*, E 76(031110):1–15, 2007.

[51] D. Joshi, J. Li, and James Z. Wang. A computationally efficient approach to the estimation of two- and three-dimensional hidden Markov models. *IEEE Trans. on Image Processing*, 15(7):1871–1886.

[52] J. Karvonen. Baltic sea ice SAR segmentation and classification using modified pulse-coupled neural networks. *IEEE Trans. on Geoscience and Remote Sensing*, 42(7):1566–1574, 2004.

[53] Z. Kato, M. Berthod, and J. Zerubia. A hierarchical Markov random field model and multitemperature annealing for parallel image classificaion. *Graphical Models and Image Processing*, 58(1):18–37, 1996.

[54] Z. Kato, T. C. Pong, and G. Q. Song. Unsupervised segmentation of color textured images using a multi-layer MRF model. *IEEE International Conference on Image Processing*, 1:961–964, 2003.

[55] R. Kettig and D. Landgrebe. Classification of multispectral images by extraction and classification of homogeneous objects. *IEEE Trans. on Geoscience and Remote Sensing*, 14:19–26, 2004.

[56] L. Kiage, K. Liu, N. Walker, N. Lam, and O. Huh. Recent land-cover/use change associated with land degradation in the Lake Baringo catchment, Kenya, East Africa: evidence from Landsat TM and ETM+. *International Journal of Remote Sensing*, 28(19):4285–4309, 2007.

[57] J. Kim and R. Zabih. Factorial Markov random fields. *European Conference on Computer Vision*, pages 321–334, 2002.

[58] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:671C–680, 1983.

[59] H. I. Koo and N. I. Cho. Prior model for the MRF modelling of multi-channel images. *International Conference on Acoustics, Speech and Signal*, 1:713–716.

[60] J. M. Laferté, P. Pérez, and F. Heiz. Discrete Markov image modeling and inference on the quadtree. *IEEE Trans. Image Processing*, 9(3):390–404, 2001.

[61] T. Landmann, C. Herty, S. Dech, and M. Schmidt. Land cover change analysis within the GLOWA Volta basin in West Africa using 30-meter landsat data snapshots. *In Proc. of IEEE International Geoscience and Remote Sensing Symposium*, pages 5298–5301, 2007.

[62] J. Li, R. M. Gray, and R. A. Olshen. Multiresolution image classification by hierarchical modeling with two-dimensional hidden Markov models. *IEEE Trans. on Information Theory*, 46(5):1826–1841.

[63] S. Z. Li. Markov random field models in computer vision. *European Conference on Computer Vision*, (2683):361–370, 1994.

[64] S. Z. Li. *Markov Random Field modeling in Computer Vision*. Springer-Verlag, 1995.

[65] X. Li. Image resolution enhancement via data-driven parametric models in the wavelet space. *EURASIP Journal on Image and Video Processing*, 2007.

[66] L. Liang, C. Liu, Y. Xu, B. Guo, and H. Shum. Real-time texture synthesis by patch-based sampling. *ACM Trans. on Graphics*, 20(3):127–150.

[67] I. Y. Liao, M. Petrou, and R. Zhao. A fractal-based relaxation algorithm for shape from terrain image. *Computer Vision and Image Understanding*, 109:227–243, 2008.

[68] Y. Liu and P. Fieguth. Parallel hidden hierarchical fields for multi-scale reconstruction. *In Proc. of Energy Minimization Methods in Computer Vision and Pattern Recognition 2009, LNCS, Springer*, Aug. 2009.

[69] Y. Liu and P. Fieguth. Image resolution enhancement with hierarchical hidden fields. *International Conference on Image Analysis and Recognition*, Jul. 2009.

[70] Y. Liu, A. Mohebi, and P. Fieguth. modeling of multiscale porous media using multiple Markov random fields. *In Proc. of the 4th BIOT Conference on Poromechanics, New York, US, DEStech Publications*, pages 435–440, June, 2009.

[71] Y. Liu, A. Wong, and P. Fieguth. Remote sensing image synthesis. *International Geoscience and Remote Sensing Symposium 2010*.

[72] A. Lorette, X. Descombes, and J. Zerubia. Texture analysis through a Markovian modelling and fuzzy classification: Application to urban area extraction from satellite images. *International Journal Computer Vision*, 36(3):221–236, 2000.

[73] S. W. Lyu and E. P. Simoncelli. Statistical modeling of images with fields of Gaussian scale mixtures. *Advances in Neural Information Processing Systems*, 19:886–902, May.

[74] S. G. Mallat. A theory for multiresolution signal decomposition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[75] J. L. Marroquin, E. A. Santana, and S. Botello. Hidden Markov measure field models for image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(11):1380–1387, 2003.

[76] G. Meinel and M. Neubert. A comparison of segmentation programs for high resolution remote sensing data. *International Archives of the ISPRS*, 35:1097–1105, 2004.

[77] D. E. Melas and S. P. Wilson. Double Markov random fields and Bayesian image segmentation. *IEEE Trans. on Signal Processing*, 50(2):357–365, 2002.

[78] F. Melgani and L. Bruzzone. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. on Geoscience and Remote Sensing*, 42(8):1778–1790, 2004.

[79] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy. Sonar image segmentation using an unsupervised hierarchical Mrf model. *IEEE Trans. on Image Processing*, 9(7):1216–1231, 2000.

[80] A. Mohebi and P. Fieguth. Posterior sampling of scientific images. *International Conference on Image Analysis and Recognition 2006, LNCS, Springer*, pages 339–350, 2006.

[81] A. Mohebi and P. Fieguth. Statistical fusion and sampling of scientific images. *IEEE International Conference on Image Processing 2008*, pages 1312–1315, 2008.

[82] A. Mohebi, Y. Liu, and P. Fieguth. Hierachical sampling with constraints. *International Conference on Image Analysis and Recognition 2010, LNCS, Springer*, pages 23–32.

[83] H. Noda, M. N. Shirazi, and E. Kawaguchi. An MRF model-based method for unsupervised textured image segmentation. *In Proc. of the 13th International Conference on Pattern Recognition*, 33(4):765–769, 1996.

[84] T. Ojala, M. Pietikainen, and D. Harwood. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29:51–59, 1996.

[85] H. Olkkonen and P. Pesola. Gaussian pyramid wavelet transform for multiresolution analysis of images. *Graphical Models and Image Processing*, 58(0032):394–398, Jul. 1996.

[86] H. Permutera, J. Francosb, and I. Jermync. A study of Gaussian mixture models of color and texture features for image classication and segmentation. *Pattern Recognition*, 39:695–706, 2006.

[87] M. Pesaresi and J. Benediktsson. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Trans. on Geoscience and Remote Sensing*, 39:309–319, 2001.

[88] W. Pieczynski and D. Benboudjema. Multisensor triplet Markov fields and theory of evidence. *Image and Vision Computing*, 24:61–69, 2006.

[89] W. Pieczynski and A.N. Tebbache. Pairwise Markov random fields and segmentation of textured images. *Machine Graphics and Vision*, 9(3):705–718, 2000.

[90] G. Poggi, G. Scarpa, and J. B. Zerubia. Supervised segmentation of remote sensing images based on a tree-structured MRF model. *IEEE Trans. on Geoscience and Remote Sensing*, 43(8):1901–1911, 2005.

[91] J. Portilla and E. P. Simoncelli. Parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1):49–71, 2000.

[92] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. *Image denoising using Gaussian scale mixtures in the wavelet domain*. Computer Science Technical Report, Courant Institute of Mathematical Sciences, New York University, 2002.

[93] R. B. Potts. Some generalized order-disorder transformations. *Proceedings of the Cambridge Philosophical Society*, 48(2):106–109, 1952.

[94] Q. Redmund, D. Long, and M. Drinkwater. A finite mixture algorithm for finding proportions in SAR images. *In Proc. of IEEE International Geoscience and Remote Sensing Symposium*, 4:1976–1978, 1998.

[95] M. Rivera and J. C. Gee. Two - level MRF models for image restoration and segmentation. *The British Machine Vision Conference 2004*, pages 809–818, 2004.

[96] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods.* second edition, Springer, 2005.

[97] H. Rue and L. Held. *Gaussian Markov Random Field: Theory and Applications.* CRC Press, 2005.

[98] P. Salembier and L. Garrido. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. on Image Processing*, 9(4):561–576, 2000.

[99] R. Samadani. A finite mixture algorithm for finding proportions in SAR images. *IEEE Trans. on Image Processing*, 4(8):1182–1186, 1995.

[100] G. Scarpa, R. Gaetano, M. Haindl, and J. Zerubia. Hierarchical multiple Markov chain model for unsupervised texture segmentation. *IEEE Trans. on Image Processing*, 18(8):1830–1843, 2009.

[101] G. Scarpa, M. Haindl, and J. Zerubia. A hierarchical texture model for unsupervised segmentation of remotely sensed images. *SCIA*, pages 303–312, 2007.

[102] L. Soh, C. Tsatsoulis, D. Gineris, and C. Bertoia. ARKTOS: An intelligent system for SAR sea ice image classification. *IEEE Trans. on Geoscience and Remote Sensing*, 42(1):229–248, 2004.

[103] R. H. Swendsen and J. S. Wang. Nonuniversal critical dynamics in Monte Carlo simulation. *Physical Rev. Letters*, 58(2):86–88, 1987.

[104] M.S. Talukdar, O. Torsaeter, and M.A. Ioannidis. Stochastic reconstruction of particulate media from two-dimensional images. *Journal of Colloid and Interface Science*, 248(2):419–428.

[105] S. Torquato. *Random heterogeneous materials: microstructure and macroscopic properties*. Springer-Verlag, 2002.

[106] J. S. Wang. Stochastic relaxation on partitions with connected components and its applications to image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(6):619–636, 1998.

[107] Y. A. Wang and E. H. Adelson. Representing moving images with layers. *IEEE Trans. on Image Processing*, 3(5):625–638, 1994.

[108] L. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. *In Proc. of the 27th annual conference on Computer graphics and interactive techniques*, pages 479–488, 2000.

[109] S. Wesolkowski and P. Fieguth. Hierarchical regions for image segmentation. *In Proc. of International Conference on Image Analysis and Recognition 2004, LNCS, Springer*, 3211:9–16, 2004.

[110] S. B. Wesolkowski. *Stochastic Nested Aggregation for Images and Random Fields, PhD thesis*. University of waterloo, 2007.

[111] R. Wilson and C. T. Li. A class of discrete multiresolution random fields and its application to image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(1):42–56, 2002.

[112] G. Winkler. *Image analysis, Random Fields, and Markov Chain Monte Carlo Methods*. Springer-Verlag, second edition, 2003.

[113] C.S. Won and R.M. Gray. *Stochastic Image Processing*. Kluwer Academic/Plenum Publishers, New York.

[114] A. Wong, P. Yu, W. Zhang, and D. A. Clausi. IceSynth II: Synthesis of SAR sea-ice imagery using region-based local conditional posterior sampling. *IEEE Geosciences and Remote Sensing Letters*, pages 4285–4309, 2009.

[115] Y. Wu, C. Wang, H. Zhang, X. Wen, and B. Zhang. Statistical analysis and simulation of high-resolution SAR ground clutter data. *In Proc. of IEEE International Geoscience and Remote Sensing Symposium*, 4:2182–2184, 2008.

[116] C. L. Y. Yeong and S. Torquato. Reconstructing random media. ii. three-dimensional media from two-dimensional cuts. *Physical Review*, E 58(1):224–233, 1998.

[117] Q. Yu and D. Clausi. SAR sea-ice image analysis based on iterative region growing using semantics. *IEEE Trans. on Geoscience and Remote Sensing*, 45(12):3919–3931, 2007.

[118] D. X. Zhou. Texture analysis and synthesis using a generic Markov-Gibbs image model. *Phd Thesis,University of Auckland*, Feb., 2006.