

**Polygonal Random Fields**  
**and the Reconstruction of Piecewise Continuous Functions**

by

Nathan P. Judish

Submitted to the  
Department of Electrical Engineering and Computer Science  
in partial fulfillment of the requirements  
for the degree of

Master of Science in Electrical Engineering

at the

Massachusetts Institute of Technology

February, 1993

© Nathan P. Judish, 1993. All rights reserved.

The author hereby grants to MIT permission to reproduce and to  
distribute publicly copies of this thesis document in whole or in part.

Signature  
of Author

-  
Department of Electrical Engineering and Computer Science,  
February, 1993

Certified by

-  
Sanjoy K. Mitter  
Professor of Electrical Engineering  
Thesis Supervisor

Accepted by

\_\_\_\_\_  
Campbell L. Searle, Chairman  
Committee on Graduate Students

# **Polygonal Random Fields and the Reconstruction of Piecewise Continuous Functions**

by

Nathan P. Judish

Submitted to the  
Department of Electrical Engineering and Computer Science  
on January 28, 1993 in partial fulfillment of the  
requirements for the degree of  
Master of Science in Electrical Engineering

## **Abstract**

Polygonal random fields (PRFs) are random fields which map a region of the plane into a finite set such that the mapping's edges divide the region into polygons. Because PRFs provide a probability measure for the set of polygonal colorings of a region, they can be used to form a model for a Bayesian approach to the image segmentation problem. Previously, Markov random fields have been used in Bayesian formulations of the segmentation problem, but such formulations confine the problem to a discrete lattice.

In this thesis, PRFs are defined and their properties are discussed. Models are formulated to reconstruct piecewise constant functions and piecewise smooth functions. In each case, an algorithm is presented to generate sample realizations from the posterior distributions of the models. The algorithms are then implemented to demonstrate the performance of each model.

Thesis Supervisor: Sanjoy K. Mitter

Title: Professor of Electrical Engineering

## **Acknowledgements**

I thank Professor Sanjoy Mitter for his vision and guidance, his encouragement and patience. I thank Julia for her love and support.

I also thank the Office of Naval Research for the graduate fellowship which has supported my work here at MIT.

# Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
<b>2</b>	<b>Polygonal Random Fields</b>	<b>9</b>
2.1	Introduction . . . . .	9
2.2	The Poisson Line Process . . . . .	11
2.3	Polygonal Random Fields . . . . .	14
2.4	Conditional Distributions for Polygonal Random Fields . . . . .	18
2.5	The Markov Property for Polygonal Random Fields . . . . .	21
2.6	The Arak process . . . . .	24
<b>3</b>	<b>Reconstruction of Polygonal Random Fields</b>	<b>29</b>
3.1	Introduction . . . . .	29
3.2	A Model for Reconstructing Piecewise Constant Functions . . . . .	30
3.2.1	Prior Distribution . . . . .	31
3.2.2	Stochastic Measurement Model . . . . .	32
3.2.3	Posterior Distribution . . . . .	33
3.3	Obtaining Realizations of PRFs with Additive Potentials . . . . .	35
3.3.1	Monte Carlo Method Preliminaries . . . . .	36
3.3.2	Algorithm for the Generation of PRFs with Additive Potentials . . . . .	38
3.3.3	Analysis of the PRF Sample Coloring Algorithm . . . . .	39
3.4	Implementing the PRF Sample Coloring Algorithm . . . . .	44

3.5	Experimental Results . . . . .	49
3.5.1	Prior Potentials . . . . .	49
3.5.2	The Image Segmentation Problem . . . . .	51
<b>4</b>	<b>Reconstruction of Piecewise Continuous Functions</b>	<b>57</b>
4.1	Introduction . . . . .	57
4.2	Markov Random Fields . . . . .	58
4.3	A MRF Model for the Reconstruction of Piecewise Continuous Functions	60
4.4	A PRF–MRF Model for the Reconstruction of Piecewise Continuous Functions . . . . .	62
4.4.1	Prior Distribution . . . . .	63
4.4.2	Measurement Model and Posterior Distribution . . . . .	65
4.5	Monte Carlo Algorithm for the PRF–MRF Model . . . . .	66
4.5.1	Algorithm for the Generation of PRF–MRF Sample Realizations	67
4.5.2	Analysis of the PRF–MRF Algorithm . . . . .	69
4.6	Experimental Results . . . . .	71

# Chapter 1

## Introduction

To segment an image is to partition the image into coherent regions; the purpose of an image segmentation problem is to produce such partitions. A physical example is the problem of dividing an image into regions corresponding to objects in the region. A political segmentation problem, marked by internecine solutions, is the partition of land masses into nations. In this thesis, we will be concerned only with segmentation of an image based on intensity. Specifically, given a set of noisy measurements of the intensity of an image, we are concerned with reconstructing a piecewise constant or piecewise smooth approximation of the image. The principal contribution of this thesis is to address this problem through Bayesian models based on the use of polygonal random fields (PRFs).

Image segmentation problems have previously been addressed both through Markov Random Field (MRF) formulations and through variational formulations. Variational formulations are based on the minimization of energy functionals. For example, Mumford and Shah [13] proposed a variational formulation based on minimizing a cost functional over a set of boundaries with smooth functions within the boundaries. Under this formulation, if the image is on a region  $T$  and the observation is  $g$ , then a reconstructed image  $f$  and its edges  $\Gamma$  are found by minimizing

$$E(f, \Gamma) = c_1 \int_T (f - g)^2 dA + c_2 \int_{T \setminus \Gamma} \|\nabla f\|^2 dA + c_3 L(\Gamma)$$

where  $L(\Gamma)$  is the length of  $\Gamma$  and  $c_1, c_2$ , and  $c_3$  are constants.

A MRF approach to the reconstruction of piecewise continuous functions is discussed in Geman and Geman [7] and in Marroquin [11]. This formulation is described more fully in chapter 4. Unlike the variational formulation, it is a probabilistic approach and takes place in a discrete setting. The MRF formulation is Bayesian. The prior distribution is modeled with two coupled MRFs: one MRF represents the function intensity, and the other represents the presence or absence of edge segments between adjacent sites of the function intensity lattice. The prior distribution is then a Gibbs distribution with a potential which includes terms which reflect the cost of edge segments and the cost of discontinuities in the intensity at adjacent sites not separated by an edge. Suppose the measurement of an image at the sites of the intensity lattice is modeled by the image's intensity plus some Gaussian error term. Then the posterior distribution is also a Gibbs distribution; its potential includes the terms of the prior distribution plus a term which reflects the cost of differing with the measurements. A segmentation of the image can be obtained by sampling from or by minimizing the posterior distribution.

Both the variational and the MRF formulation are based on functionals which include three terms. One is a fidelity term which ensures that segmentations will reflect the data. One is a smoothing term which requires that segmentations be piecewise smooth off the set of edges. The final term limits the edges included in the segmentation. The PRF formulation of this thesis will also include these three terms.

Polygonal random fields are random fields which map a region of the plane into a finite set such that the mapping's edges divide the region into polygons. A large class of PRFs possess a spatial Markov property. Because PRFs provide a probability measure for the set of polygonal colorings of a region, they can be used in image segmentation problems in a manner similar to the use of MRFs described above. However, unlike the discrete MRFs, PRFs are defined as continuous functions on the plane. They can model discontinuities in arbitrary directions, and they can be chosen

to be invariant under translations and rotations. In some situations, these features make PRFs more natural models for segmentation problems than MRFs.

This thesis explores the use of PRFs in a Bayesian formulation of the segmentation problem to reconstruct piecewise smooth functions. In chapter 2, we define PRFs and discuss their important mathematical properties. We show how the probability measure for PRFs is constructed from the Poisson line process, and we show that a large class of PRFs has the spatial Markov property. In chapter 3, we formulate a Bayesian model for the image segmentation problem to reconstruct piecewise constant functions. We then present an algorithm to generate sample realizations of PRFs, and we discuss the problems involved in implementing the algorithm. We then use the algorithm to generate sample realizations of PRFs, including the PRFs of the piecewise constant segmentation problem. In chapter 4, we formulate a Bayesian model for the reconstruction of piecewise smooth functions. This model couples a PRF which represents the edges of an image with a MRF which represents the intensity of the image at a set of sites on a lattice. We then present an algorithm to generate sample realizations for this model. Finally, we present experimental results of the algorithm to reconstruct piecewise continuous functions.

The mathematical development of PRFs is due to Arak and Surgailis [2], and Clifford and Middleton [4] suggested a Monte Carlo type algorithm for the generation of PRFs. A different approach to the segmentation problem can be found in Mitter and Zeitouni [12]. The principal contribution of the thesis is the dual PRF-MRF model for the reconstruction of piecewise continuous functions and the algorithm to reconstruct piecewise continuous functions. Other contributions here include a more detailed discussion of the issues involved in the implementation of Monte Carlo algorithms for the generation of PRFs and the presentation of experimental results of PRF generation algorithms.



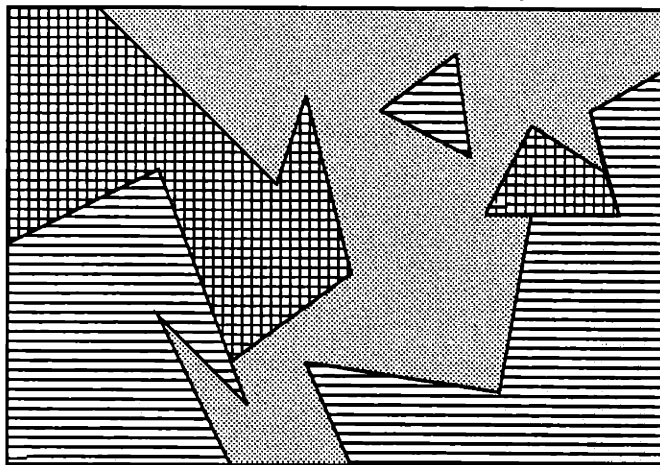
# Chapter 2

## Polygonal Random Fields

### 2.1 Introduction

Polygonal random fields are random fields which map a region of the plane  $T$  into a finite set  $J$  such that the mapping's discontinuities divide the region into polygons. The elements of  $J$  can be thought of as colors. A sample realization of a PRF is then a partition of  $T$  into colored polygons; hence, realizations are generally referred to as colorings. Figure 2.1 shows a possible coloring of a PRF which takes on three values.

Figure 2.1: Sample coloring of a PRF taking on three values.



PRFs provide a probability measure for the set of polygonal colorings of a region, and they can be used as a model for reconstruction problems in a manner analogous

to MRF models. In fact, PRFs share many of the features which make MRFs useful for reconstruction problems:

1. The probability density for PRFs can be specified explicitly. This feature makes PRFs analytically tractable, so it is possible to apply Bayes' theory to PRF reconstruction problems.
2. Because the probability density can be specified explicitly, there is a clear relationship between the model and the characteristics of the sample colorings of a PRF. For example, the density can include terms which directly affect the length of a coloring's edges, the angles formed by the coloring's edges, and the coloring's particular color at individual points.
3. The class of PRFs is broad enough to model a wide variety of situations.
4. Sample colorings from PRFs can be generated through Monte Carlo procedures.
5. PRFs have a spatial Markov property which makes parallel implementation of PRF algorithms possible.

Polygonal random fields also have characteristics which distinguish them from Markov random fields. PRFs are defined on the plane; MRFs are defined on the nodes of a graph. For certain kinds of images, the continuous formulation of PRFs seems to be a more natural model than the discrete formulation of MRFs. For example, images containing artificial objects often have discontinuities along line segments with arbitrary orientations. To model such images with MRFs, the images must be discretized into pixels, and the images' discontinuities can only be modeled in a few fixed directions. Certain intrinsic properties of an image, like the total length of its discontinuities, are lost through this discretization process. In contrast, the PRF model includes such discontinuities directly, and the total length of an image's discontinuities is preserved. Of course, modeling image segmentation problems with PRFs also has disadvantages. The continuous nature of PRFs can make them difficult

– though certainly not impossible – to handle computationally. We will discuss the computational problems of PRFs in the next chapter, when we begin to discuss the use of PRFs in reconstruction problems.

The development of PRFs is largely due to the work of Arak and Surgailis. Arak[1] first noted that the trajectories in space-time of a system of particles moving in one dimension could yield a partition of a two-dimensional region with a spatial Markov property. Such a system is called an Arak process, and we will discuss it at the end of this chapter. Expanding on this notion, Arak and Surgailis[2] developed the mathematical basis for more general PRFs. Clifford[5] and Clifford and Middleton[4] suggested an algorithm for the generation of sample colorings of PRFs.

In this chapter, we define PRFs and explore their important mathematical properties. We begin by reviewing the Poisson line process, the fundamental building block from which PRFs are constructed. Next, we define PRFs, and then we discuss conditional probabilities and the Markov property for PRFs. Finally, we examine Arak’s particle model for the generation of PRFs. Although the particle model does not play an important role in reconstruction problems, it will provide insight into the generation of PRFs and their Markov property.

## 2.2 The Poisson Line Process

The Poisson line process is the basic process through which PRFs are defined and their sample colorings are constructed. A Poisson line process generates random sets of lines which intersect a given region. Because the Poisson line process is so important to the development of PRFs and the generation of sample colorings, we will review its characteristics and describe a procedure to simulate it.

Throughout this chapter, let  $T$  be a bounded convex region in the plane. The homogeneous Poisson line process on  $T$  with intensity  $\lambda$  has the following properties:

1. Given a convex region  $S \subset T$  with perimeter  $P_S$ , the number of lines intersecting

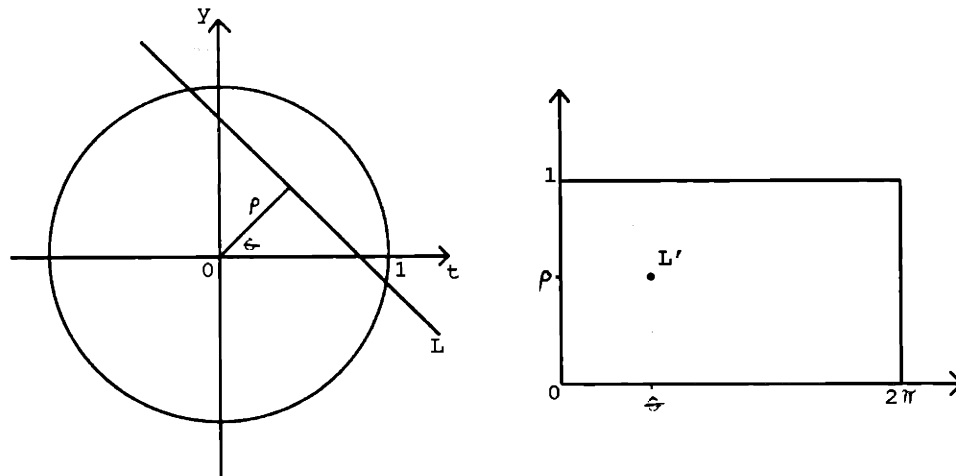
$S$  has a Poisson distribution with mean  $\lambda P_S$ .

2. The distribution of the lines is invariant under Euclidean transformations.

To choose random lines on  $T$ , we will parameterize the lines which intersect  $T$ ; each line will then correspond to a point in a two-dimensional parameter space. Choosing random points from the parameter space is thus equivalent to choosing random lines which intersect  $T$ . The two properties above will then follow from results of geometric probability theory.

First, fix some point  $O$  in  $T$  as the origin. Any line  $l$  which intersects  $T$  can be parameterized  $l = l(\rho, \theta)$ , where  $\rho \geq 0$  and  $\theta \in [0, 2\pi)$  are the polar coordinates of the intersection between the line and the perpendicular to the line through the origin. Given the region  $T$ , this parameterization produces a set  $T' = \{(\rho, \theta) : l(\rho, \theta) \text{ intersects } T\}$  in the  $\rho - \theta$  parameter space, where each point in  $T'$  corresponds to a line which intersects  $T$ . For example, if  $T$  is an open disc of radius  $r$  with its center at the origin, then  $T'$  is the rectangle  $T' = \{(\rho, \theta) : \rho \in [0, r), \theta \in [0, 2\pi)\}$ . Figure 2.2 illustrates the correspondence between  $T$  and  $T'$ .

Figure 2.2: A unit disc and the corresponding parameter space. The line  $L$  intersecting the disc corresponds to the point  $L'$ .



The region  $T$ .

The region  $T'$  in the parameter space.

We wish to select lines intersecting  $T$  such that properties 1 and 2 above are

satisfied. From geometric probability theory (Kendall and Moran[9], p. 56),  $d\rho d\theta$  is the area measure in  $T'$  which yields a distribution of lines in  $T$  which is invariant under Euclidean transformations. Hence, the distribution of lines intersecting  $T$  is invariant under Euclidean transformations if we sample uniformly from  $T'$ . Geometric probability theory (Kendall and Moran[9], p. 58) also yields another important result: for any convex region  $S$ , the area in the parameter space of  $S'$  is

$$\int_{S'} d\rho d\theta = P_S.$$

These results suggest that we run a Poisson point process with constant intensity  $\lambda$  on  $T'$  in the parameter space. This means that from a subset  $S'$  of  $T'$  with area  $A$ , the number of points chosen has a Poisson distribution with mean  $\lambda A$ , and that those points are distributed as if they were sampled independently from a uniform distribution over  $S'$ . Thus, the distribution of the corresponding lines is invariant under Euclidean transformations, and property 2 is satisfied. Furthermore, given a convex  $S \subset T$ , the area of  $S'$  is  $P_S$ . Thus, the number of lines intersecting  $S$  has a Poisson distribution with mean  $\lambda P_S$ , so property 1 is satisfied. Therefore, this sampling procedure produces a homogeneous Poisson line process on  $T$ .

Using this procedure, it is easy to simulate a Poisson line process of intensity  $\lambda$  on a disc of radius  $r$ . First, generate  $n$ , the number of lines to be drawn, by sampling from a Poisson distribution with mean  $2\pi r\lambda$ . Then, for each of the  $n$  lines, sample independently from a uniform distribution on the corresponding rectangle  $\{(\rho, \theta) : \rho \in [0, r], \theta \in [0, 2\pi)\}$ . Finally, draw the lines corresponding to the points chosen in the parameter space.

To simulate a Poisson line process with intensity  $\lambda$  on an arbitrary convex bounded region  $T$  with perimeter  $P_T$ , again generate  $n$  by sampling from a Poisson distribution with mean  $\lambda P_T$ . The task is then to choose  $n$  points by sampling uniformly from the corresponding region  $T'$ .  $T'$  is no longer a rectangle, as it was when  $T$  was a disc. This problem is minor: simply sample from a uniform distribution on a larger rectangle containing  $T'$  and disregard points chosen which correspond to lines which do not

intersect  $T$ . Again, draw the lines corresponding to the  $n$  points chosen from  $T'$ .

The Poisson line process induces a measure on finite sets of lines intersecting  $T$ , and this measure is the basis for the construction of PRFs. Let  $L$  be the set of all lines intersecting  $T$ , and let  $\ell = \{l_1, l_2, \dots, l_n\}$  be a finite set of distinct lines with each  $l_i \in L$ . Let  $L_T$  be the set of all  $\ell$ , so  $L_T$  contains all finite sets of lines intersecting  $T$ . The empty set of lines is also considered to be an element of  $L_T$ . Then the Poisson line process with intensity  $\lambda$  on  $T$  induces a measure  $\mu$  on  $L_T$ ; we call  $\mu$  the Poisson line measure. For example,  $\mu(\emptyset) = e^{-\lambda P_T}$ .

## 2.3 Polygonal Random Fields

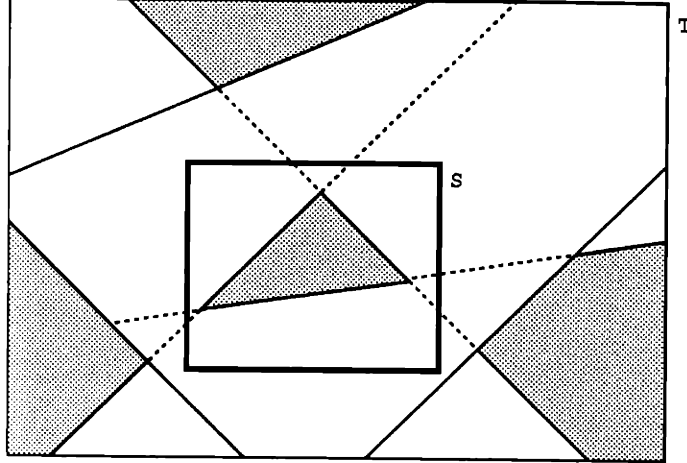
A Poisson line process on  $T$  divides  $T$  into convex polygons (except along the boundary of  $T$ , which does not necessarily consist of line segments). Assigning an element of a finite set  $J$  to each polygon creates a coloring of  $T$ . This assignment can be made in many different ways, and each method induces a probability distribution on the set of colorings of  $T$ . A simple method would be to color each polygon independently: coloring each one black or white with probability  $1/2$ , for example.

Although this independent coloring process has a Markov property along lines crossing  $T$  (Switzer [14]), it does not have a spatial Markov property. The problem arises because a given line may contribute more than one edge to the coloring, leaving the line “invisible” between the segments. (A line segment which forms a discontinuity of a coloring is called an edge). Given a region  $S \subset T$ , a line which is visible on the exterior of  $S$  may not be visible on the boundary of  $S$ . Such lines will of course affect the coloring on the interior of  $S$ . Figure 2.3 illustrates this situation.

The problem of invisible line segments can be avoided by requiring that each line contribute only one edge to the coloring, for then there can be no invisible gaps between two edges from one line. This requirement is fundamental to PRFs. But, as we shall see, this requirement makes the development of PRFs more complicated

than the simple independent coloring process described above.

Figure 2.3: An independent coloring process does not have the spatial Markov property. The dashed lines are invisible on the boundary of  $S$ .



Recall that PRFs are random fields which map a convex region of the plane into a finite set  $J$  such that the mapping's discontinuities partition the region into polygons. To define PRFs, some notation must be introduced. Here, we largely follow the development of Clifford and Middleton[5], which is somewhat simpler than that of Arak and Surgailis[2].

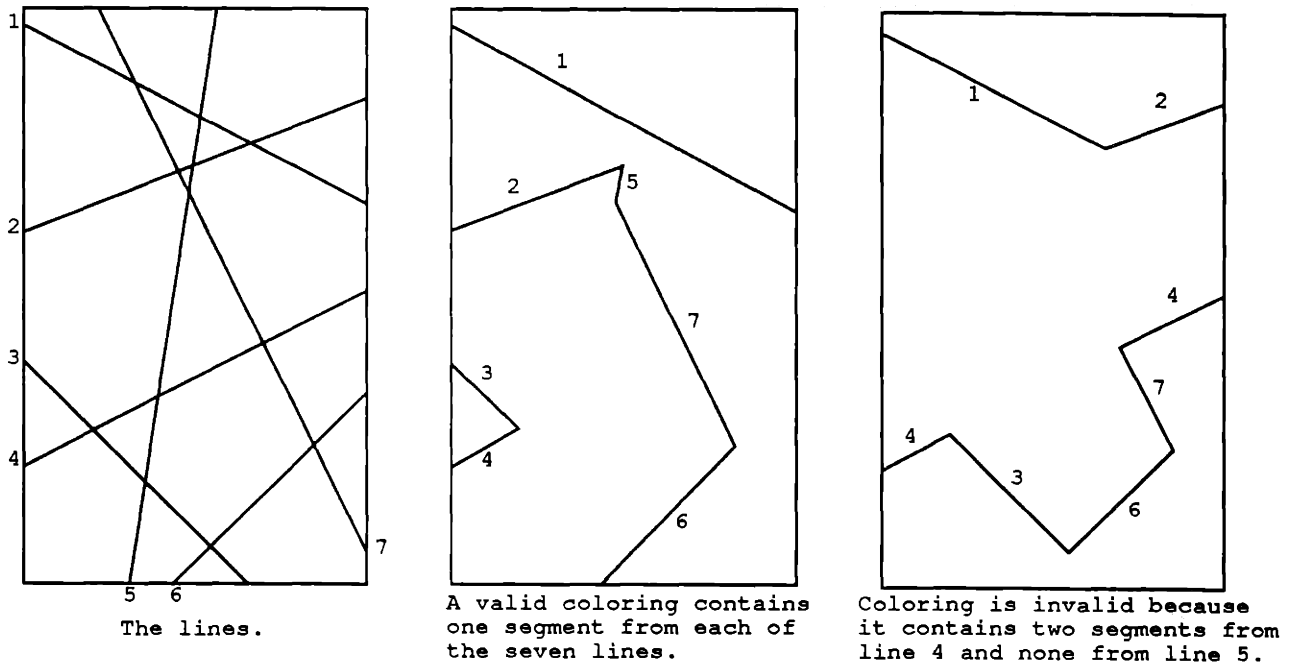
Again, let  $T$  be a convex bounded region of the plane, and let  $J$  be a finite set. A coloring of  $T$  is a function of the form  $\chi : T \rightarrow J$ . For a given line  $l_i$ , a segment of  $l_i$  is denoted by  $[l_i]$ . Also recall that  $L_T$  is the set of all  $\ell$ , where each  $\ell$  is a finite set of lines which intersect  $T$ . For each  $\ell \in L_T$ , with  $\ell = \{l_1, \dots, l_n\}$ , let  $\Omega_T^\ell$  be the set of colorings  $\chi$  satisfying:

1. For each  $i = 1, 2, \dots, n$ , there is a segment  $[l_i] \subset l_i$  such that  $[l_i]$  is an edge of  $\chi$ .
2.  $\bigcup_{i=1}^n [l_i]$  is the set of discontinuities of  $\chi$ .

With the requirement that each line of  $\ell$  must contribute exactly one edge to the coloring,  $\Omega_T^\ell$  is the set of polygonal colorings on  $T$  given a finite set  $\ell$  of lines. The union of all  $\Omega_T^\ell$ , for all  $\ell \in L_T$ , is called  $\Omega_T$ . It is the set of all polygonal colorings on  $T$ . Figure 2.4 shows a set of lines  $\ell$ , a coloring in  $\Omega_T^\ell$ , and a coloring which is not in

$\Omega_T^L$ . (Actually, it shows the discontinuities of a coloring in  $\Omega_T^L$ , not the coloring itself. For convenience, this is how we will usually display sample colorings of PRFs. In general, we will be concerned only with binary PRFs, for which the cardinality of  $J$  is two. For these PRFs, there are exactly two possible colorings corresponding to each set of discontinuities.)

Figure 2.4: A set of lines, a valid coloring, and an invalid coloring.



Technically, the above definition does not uniquely define a polygonal coloring along its edges. For example, an edge in a black and white coloring could be black, white, or a mixture of both. To ensure the uniqueness of colorings, the elements of  $J$  could be numbered, and the value of a coloring at a point of discontinuity could be defined to be the *lim inf* at the point. This is a minor detail, and it will not be addressed further.

We are now ready to define a probability measure  $P_T$  for PRFs. For simplicity, the probability measure is defined through a two step process.

First, let  $A \subset \Omega_T$ . For example,  $A$  could be the set of colorings of  $T$  containing



two colors and seven edges. Define the polygonal coloring measure  $\gamma$  by

$$\gamma(A) = \int_{L_T} |A \cap \Omega_T^\ell| \mu(d\ell) \quad (2.1)$$

where  $|\cdot|$  is the number of elements in a set. The polygonal coloring measure can be interpreted as the expected number of colorings in set  $A$  which can be formed from the lines obtained from a realization of the Poisson line process on  $T$ .

The probability measure for a PRF has a density of the form  $e^{-F(x)}$  with respect to  $\gamma$ , where the function  $F : \Omega_T \rightarrow \mathfrak{R} \cup \infty$  is called the potential. The probability measure for a PRF will thus have the form:

$$P_T(A) = \frac{\int_A e^{-F(x)} \gamma(d\chi)}{\int_{\Omega_T} e^{-F(x)} \gamma(d\chi)}. \quad (2.2)$$

Note that the density function simply adds a weighting term  $e^{-F(x)}$  to the polygonal coloring measure. To summarize this development, the Poisson line process induces a measure on sets of lines intersecting  $T$ . Each finite set of lines  $\ell$  determines a set of allowable colorings  $\Omega_T^\ell$ , and the density term gives a coloring  $\chi \in \Omega_T^\ell$  an additional weight  $e^{-F(x)}$ . Thus, the measure reflects the likelihood of obtaining sets of lines weighted by both the number of colorings which can be created from those lines and the value of the density function for those colorings.

In order for  $P_T$  to be a probability measure, the denominator of equation 2.2 must be finite. Therefore, the definition of PRFs is restricted to potentials  $F(\chi)$  which yield a finite denominator. Unfortunately, it is not known precisely which potentials satisfy this criterion. However, Arak and Surgailis[2] show that the PRF induced by the Arak process (see section 2.6) does have a finite denominator. For the Arak process,  $F(\chi)$  is infinite if any two edges cross each other, so there are no T-shaped, X-shaped, or more complex intersections of edges. Otherwise  $F(\chi) = 2\Gamma$ , where  $\Gamma$  is the total length of an image's discontinuities. This result is sufficient to guarantee the existence of a great number of PRFs: if a potential  $F$  is greater than or equal to the potential for the Arak process, the denominator of equation 2.2 will be finite,

and the PRF will exist. In this thesis, all potentials will satisfy this criterion, so all the PRFs here will be well-defined.

The spatial Markov property is extremely important for reconstruction problems involving PRFs, but we have not yet shown that PRFs have this property. In fact, not all PRFs have a spatial Markov property. For example, for a black and white PRF, let

$$F(\chi) = 2\Gamma \quad \text{if there is exactly one black triangle in } \chi \\ = \infty \quad \text{otherwise.}$$

Clearly, for this potential, the coloring outside a subregion  $S \subset T$  affects the coloring inside the subregion. If there is a black triangle outside of  $S$ , there cannot be one inside  $S$ . Furthermore, the information on the boundary of  $S$  is not sufficient to determine whether or not there is such a triangle outside  $S$ . Therefore, this PRF is not Markov.

There is, however, a large class of PRFs which do have the spatial Markov property. We will address this issue in the next two sections. First, we develop the notion of conditional probability for PRFs, and then we show that PRFs with additive potentials have the Markov property.

## 2.4 Conditional Distributions for Polygonal Random Fields

Conditional probability is fundamental to a Markov property, so we must discuss conditional probability for PRFs before we can discuss their spatial Markov property. In addition, conditional probability plays an important role in our reconstruction problems, as we will use our ability to sample from local conditional probabilities to generate colorings on the whole of  $T$ . Our development here of conditional probability is similar to our development of PRFs: first, we define a conditional polygonal coloring measure, and then we define a conditional probability measure. Both have essentially

the same form as the polygonal coloring measure and the probability measure of equations 2.1 and 2.2, but they add the requirement that the colorings be consistent with the conditional information given.

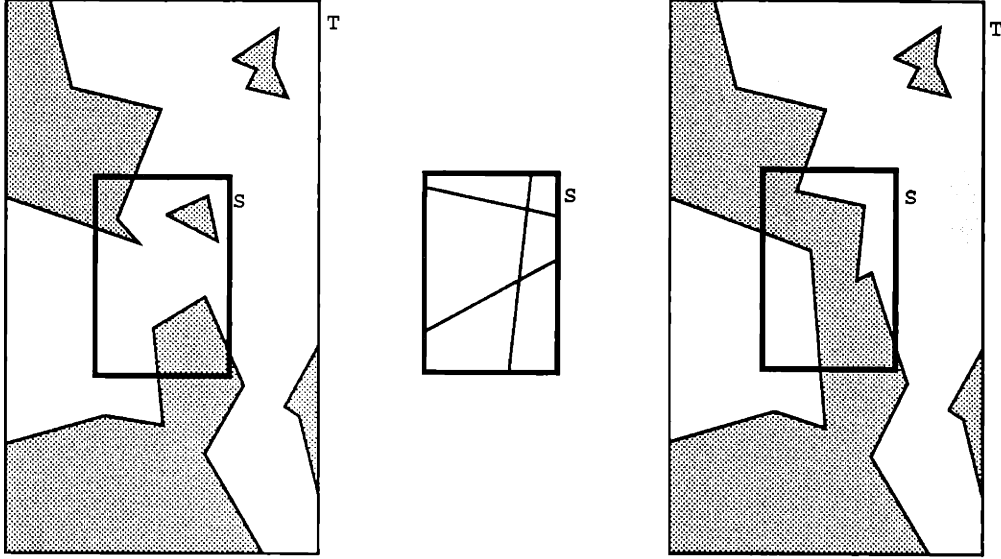
Additional notation is needed to define conditional distributions, though much of it is similar to the notation already introduced to define PRFs. For any  $U \subset T$ , let  $\chi^U : U \rightarrow J$  be a coloring of  $U$ , let  $\mu_U$  be the Poisson line measure on  $U$ , and let  $L_U$  be the set of all finite sets of lines intersecting  $U$ . These terms are, of course, analagous to  $\chi$ ,  $\mu$ , and  $L_T$ .

Let  $S$  be an open convex subset of  $T$ , and let  $\partial S = \bar{S} \setminus S$ .  $\Omega_S$  is the set of polygonal colorings of  $S$ , defined as  $\Omega_T$  was defined for  $T$ . Given a finite set of lines  $\ell = \{l_1, \dots, l_n\}$  intersecting  $S$ , let  $\Omega_S^\ell(\chi^{T \setminus S})$  be the set of polygonal colorings  $\chi^S$  satisfying:

1.  $\chi^S$  is consistent with the coloring of  $\chi^{T \setminus S}$  on  $\partial S$ .
2. For each  $i = 1, 2, \dots, n$ , there is a segment  $[l_i] \subset l_i$  such that  $[l_i]$  is an edge of  $\chi^S$ .
3. Let  $\{l'_1, \dots, l'_m\}$  denote the lines of  $\chi^{T \setminus S}$  which are visible on  $\partial S$ . For each  $l'_j$  there is a segment  $[l'_j] \subset l'_j$  such that  $[l'_j]$  is an edge of  $\chi^S$ .
4.  $\bigcup_{i=1}^n [l_i] \cup \bigcup_{j=1}^m [l'_j]$  is the set of discontinuities of  $\chi^S$ .

$\Omega_S^\ell(\chi^{T \setminus S})$  is the set of polygonal colorings of  $S$  consistent with  $\chi^{T \setminus S}$  and using the lines of  $\ell$ . The additional edges  $\{[l'_j]\}_{j=1}^m$  must be added to  $\chi^S$  to enable it to be consistent with  $\chi^{T \setminus S}$ , because the edges of  $\chi^{T \setminus S}$  which intersect  $\partial S$  must continue into  $S$  along these segments until they leave  $S$  or until they intersect some other edge. Figure 2.5 gives an example of a coloring  $\chi$ , a set of lines  $\ell$  intersecting  $S$ , and a new coloring of  $\chi$  in which the old coloring of  $S$  is replaced by an element of  $\Omega_S^\ell(\chi^{T \setminus S})$ . In the figure, the element of  $\Omega_S^\ell(\chi^{T \setminus S})$  contains seven edges. Three come from the lines of  $\ell$ , and four come from the lines necessary to retain consistency with the coloring on  $\partial S$ .

Figure 2.5: A coloring of  $T$ , a set of lines intersecting  $S$ , and a consistent polygonal coloring using segments from the set of lines.



Now, let  $A_S \subset \Omega_S$ . For example,  $A_S$  could be colorings of  $S$  with exactly three edges. Define the conditional polygonal coloring measure  $\gamma_S$  by

$$\gamma_S(A_S | \chi^{T \setminus S}) = \int_{L_S} |A_S \cap \Omega_S^l(\chi^{T \setminus S})| \mu_S(dl). \quad (2.3)$$

This measure can be interpreted as the expected number of colorings  $\chi^S$  in set  $A_S$  which can be formed from the lines obtained from a Poisson line process on  $S$ , with the condition that the colorings be consistent with the coloring  $\chi^{T \setminus S}$ .

The conditional polygonal coloring measure can now be used to define the conditional probability measure. Let

$$P_S(A_S | \chi^{T \setminus S}) = \frac{\int_{A_S} e^{-F(\chi)} \gamma_S(d\chi^S | \chi^{T \setminus S})}{\int_{\Omega_S} e^{-F(\chi)} \gamma_S(d\chi^S | \chi^{T \setminus S})}. \quad (2.4)$$

Here, the potential  $F(\chi)$  is a function of the entire coloring  $\chi$  on  $T$ .  $\chi$  is composed of the partial colorings  $\chi^{T \setminus S}$  and  $\chi^S$ . Note that  $\chi^{T \setminus S}$  remains constant, while  $\chi^S$  varies over  $A_S$  and  $\Omega_S$ . This measure has essentially the same form as the probability measure for PRFs, except that the integrals are restricted to colorings of  $T$  which agree with the given coloring  $\chi^{T \setminus S}$  on  $T \setminus S$ .

The conditional probability measure  $P_S$  has been defined separately from the PRF probability measure  $P_T$ . Therefore, we have not actually shown that  $P_S$  is the true conditional probability measure for  $P_T$ . However, this fact is intuitively clear from a comparison of the two probability measures in equations 2.2 and 2.4. Equation 2.4 is essentially the same as equation 2.2, except the integrals are restricted to colorings which are consistent with the given coloring  $\chi^{T \setminus S}$ . A rigorous proof that  $P_S$  is the proper conditional probability measure for  $P_T$  can be found in Arak and Surgailis[2].

## 2.5 The Markov Property for Polygonal Random Fields

A PRF has the spatial Markov property if the interior coloring of a region  $S \subset T$  is independent of the exterior coloring, given the information along the region's boundary. For PRFs with the Markov property, this localized probabilistic dependence makes the PRFs amenable to the computation necessary for image segmentation problems. In this section, we define additive potentials and the Markov property for PRFs, and we show that PRFs with additive potentials have the Markov property.

A potential  $F(\chi)$  is additive if its value for a coloring equals the sum of its values for the components of a partition of the coloring. That is, if  $S_1, S_2, \dots, S_n$  partition  $T$ , then  $F$  is additive if  $F(\chi) = \sum_{i=1}^n F(\chi^{S_i})$ . This property holds for a wide range of useful potentials, including the total length of a coloring's discontinuities and functions of a coloring's color at particular points. Because of the importance of the Markov property, we restrict ourselves from now on to the consideration of PRFs with additive potentials.

We have said that a PRF has the Markov property if the interior coloring of a region  $S \subset T$  is independent of the exterior coloring, given the information along the boundary. This boundary information must include both  $\chi^{\partial S}$  and  $\ell^{\partial S}$ , where  $\ell^{\partial S}$  is the set of lines which contribute discontinuity points to the coloring  $\chi^{\partial S}$ . There

is an intuitive explanation for why it is necessary to include  $\ell^{\partial S}$  in the boundary information. The coloring  $\chi^{T \setminus S}$  determines  $\chi^{\partial S}$ . But each discontinuity point of  $\chi^{\partial S}$  corresponds to an edge of  $\chi$  which intersects both  $\chi^{T \setminus S}$  and  $\chi^S$ . (This holds unless the edge ends on  $\partial S$ , an event with probability 0.) The direction of this edge can be determined from  $\chi^{T \setminus S}$  but not from  $\chi^{\partial S}$ . The lines of  $\ell^{\partial S}$  provide this information: they specify the direction of the edges which pass from  $T \setminus S$  into  $S$ .

Before we can define the spatial Markov property, we must define boundary conditional probability for PRFs with additive potentials. That is, we define  $P_S(A_S | \chi^{\partial S}, \ell^{\partial S})$ , where  $A_S \subset \Omega_S$ . As with our previous development of conditional probability, this definition is a two step process, and we first define a conditional polygonal coloring measure. Here we define the boundary conditional polygonal coloring measure  $\gamma_S(A_S | \chi^{\partial S}, \ell^{\partial S})$ . Recall that the conditional polygonal coloring measure  $\gamma_S(A_S | \chi^{T \setminus S})$  depends on  $\chi^{T \setminus S}$  only through  $\chi^{\partial S}$  and  $\ell^{\partial S}$ . This is because the conditional polygonal coloring measure is the expected number of colorings in  $A_S$  which can be formed from a Poisson line process on  $S$ , with the condition that the colorings be consistent with the coloring of  $\chi^{T \setminus S}$ . A coloring  $\chi^S$  is consistent with  $\chi^{T \setminus S}$  if and only if it is consistent with the boundary information  $\chi^{\partial S}$  and  $\ell^{\partial S}$ . Therefore, the boundary conditional polygonal coloring measure is the same as the conditional polygonal coloring measure. That is,

$$\gamma_S(A_S | \chi^{\partial S}, \ell^{\partial S}) = \gamma_S(A_S | \chi^{T \setminus S}). \quad (2.5)$$

Mirroring our previous development, we use the boundary conditional polygonal coloring measure to define the boundary conditional probability measure for additive functions. Let

$$P_S(A_S | \chi^{\partial S}, \ell^{\partial S}) = \frac{\int_{A_S} e^{-F(\chi^S)} \gamma_S(d\chi^S | \chi^{\partial S}, \ell^{\partial S})}{\int_{\Omega_S} e^{-F(\chi^S)} \gamma_S(d\chi^S | \chi^{\partial S}, \ell^{\partial S})}. \quad (2.6)$$

The form of this probability measure is similar to the conditional polygonal coloring measure of equation 2.4. The difference is that the potential here is a function only of the coloring on  $S$ , rather than of the entire coloring on  $T$ .

Two separate conditional probabilities have been defined:  $P_S(A_S | \chi^{T \setminus S})$  and  $P_S(A_S | \chi^{\partial S}, \ell^{\partial S})$ . A PRF has the spatial Markov property if, for  $A_S \subset \Omega_S$ ,

$$P_S(A_S | \chi^{T \setminus S}) = P_S(A_S | \chi^{\partial S}, \ell^{\partial S}). \quad (2.7)$$

This definition has the obvious explanation: a PRF has the Markov property if the interior coloring of  $S$  is independent of the exterior coloring, given the information on the boundary.

It remains to show that PRFs with additive potentials have the Markov property. From equation 2.4, recall that:

$$P_S(A_S | \chi^{T \setminus S}) = \frac{\int_{A_S} e^{-F(x)} \gamma_S(d\chi^S | \chi^{T \setminus S})}{\int_{\Omega_S} e^{-F(x)} \gamma_S(d\chi^S | \chi^{T \setminus S})}.$$

Note that  $\chi^S$  and  $\chi^{T \setminus S}$  partition  $\chi$ . Therefore, if  $F$  is additive,  $F(\chi) = F(\chi^S) + F(\chi^{T \setminus S})$ , so that:

$$P_S(A_S | \chi^{T \setminus S}) = \frac{\int_{A_S} e^{-F(\chi^S) - F(\chi^{T \setminus S})} \gamma_S(d\chi^S | \chi^{T \setminus S})}{\int_{\Omega_S} e^{-F(\chi^S) - F(\chi^{T \setminus S})} \gamma_S(d\chi^S | \chi^{T \setminus S})}. \quad (2.8)$$

Both the integrals in equation 2.8 are over colorings of the subregion  $S$ . Therefore,  $e^{-F(\chi^{T \setminus S})}$  is a constant term in both integrals, so it cancels. This cancellation reduces equation 2.8 to:

$$P_S(A_S | \chi^{T \setminus S}) = \frac{\int_{A_S} e^{-F(\chi^S)} \gamma_S(d\chi^S | \chi^{T \setminus S})}{\int_{\Omega_S} e^{-F(\chi^S)} \gamma_S(d\chi^S | \chi^{T \setminus S})}.$$

Finally, since  $\gamma_S(A_S | \chi^{T \setminus S}) = \gamma_S(A_S | \chi^{\partial S}, \ell^{\partial S})$ , it follows from equation 2.6 that  $P_S(A_S | \chi^{T \setminus S}) = P_S(A_S | \chi^{\partial S}, \ell^{\partial S})$ . Thus, PRFs with additive potentials have the Markov property.

It is instructive to note why PRFs without additive potentials do not necessarily have the Markov property. The interaction between the coloring on  $S$  and the coloring on  $T \setminus S$  is simple for additive potentials: they contribute the independent factors  $e^{-F(\chi^S)}$  and  $e^{-F(\chi^{T \setminus S})}$  to the density. However, for non-additive potentials, the

interaction can be more complex, and the coloring of  $T \setminus S$  can influence the potential on  $S$ . In other words, non-additive potentials do not necessarily have localized probabilistic dependence.

## 2.6 The Arak process

In the previous sections, we have defined PRFs and discussed some of their properties. However, we have not generated any sample colorings from PRFs, nor have we mentioned any procedures to generate them. Naturally, such a procedure is essential to reconstruction problems. Unfortunately, the general problem of producing sample colorings from PRFs with arbitrary additive potentials is somewhat complex. It requires a Monte Carlo simulation procedure, and it will not be introduced until the next chapter. In this section, we discuss the Arak process, a simple process which generates sample colorings from a PRF with potential  $F$  proportional to  $2\Gamma$ , where  $\Gamma$  is the total length of the discontinuities of the partition.

The Arak process is discussed here because it demonstrates the existence of many PRFs, and because it provides insight into the generation of sample colorings of PRFs and their Markov property. Also, it is historically important: its discovery by Arak[1] in 1982 led to the development of PRFs. However, because the Arak process and other closely related processes are only useful in generating realizations of PRFs with very specific potentials, they do not play an important role in the reconstruction problems we address in this thesis. Thus, we will not analyze the Arak process in depth here. A more complete treatment of the process can be found in Arak and Surgailis[2].

The Arak process is based on the evolution of a system of particles moving with piecewise constant velocity in one dimension. Arak noted that the space-time trajectories of such a system could give rise to a Markov polygonal partition of a convex region of the plane. Here, we will describe the system and show that it clearly yields a polygonal partition with the Markov property.



First, consider a single particle moving in one dimension with piecewise constant velocity  $v$ . Its path can be plotted as position  $y$  versus time  $t$ , where the slope of the line is  $v$ .

Let the particle makes jumps in velocity such that the length of each line segment in space-time is exponentially distributed. That is, given that a particle is traveling with velocity  $v$ , the length of time between jumps is exponentially distributed with parameter proportional to  $\sqrt{(1 + v^2)}$ . Furthermore, when the particle's velocity changes, let the new velocity  $v'$  be distributed such that  $\tan(v')$  has uniform distribution between  $-\pi/2$  and  $\pi/2$ . This means that the direction of each line segment in space-time will be uniformly and independently distributed. Figure 2.6 shows the possible evolution of a particle.

Figure 2.6: Particle moving with piecewise constant velocity. The length of each segment is exponentially distributed.



The evolution of this particle is Markov: if the position and velocity of the particle at some time is given, its future is independent from its past. This follows from the fact that the time between jumps is exponentially distributed based on the velocity, and the fact that the jumps in velocity are independent and identically distributed.

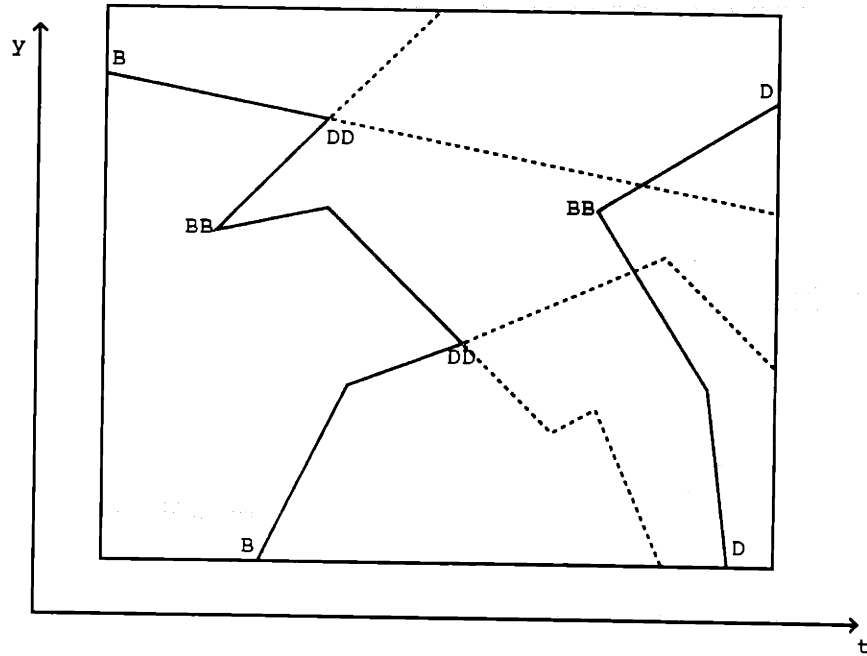
The Arak process generates a Markov polygonal partition of a convex bounded region  $T$  through the birth, evolution, and death of a system of particles, with each

particle evolving like the particle above. We must also describe the “birth” and “death” of the particles. Particles may be born either on the boundary of  $T$  or in the interior of  $T$ . Particles are born on the boundary of  $T$  according to the measure induced by a Poisson line process with parameter  $\lambda$  on  $T$ . The number of lines induced will thus have a Poisson distribution with mean  $\lambda P_T$ , where  $P_T$  is the perimeter of  $T$ . For each line intersecting  $T$ , the initial position of a particle is the leftmost intersection of the line with  $T$ , and let the initial velocity be the slope of the line. Pairs of particles are born on the interior of  $T$  according to the measure on the intersection of two lines induced by the product of two Poisson line processes with parameter  $\lambda$  on  $T$ . From probabilistic geometry, the number of particles born on the interior of  $T$  will have a Poisson distribution with mean proportional to  $2\pi\lambda^2 A$ , where  $A$  is the area of  $T$ . The initial position of each particle will be the point of intersection of the lines, and the initial velocity of each particle will be the slope of the corresponding line.

Particles are created and evolve as described above. It remains to be noted that they “die” at the moment of their first intersection with the boundary or with another particle. They are then annihilated. Figure 2.7 shows an example of the birth, evolution, and death of a system of particles.

The Arak process creates a polygonal partition of the region  $T$ . Let the number of colors in  $J$  be two, so that there are only two possible colorings for a given partition. Assigning equal probabilities to the two colorings induces a measure on  $\Omega_T$ , the set of polygonal partitions of the region. Furthermore, the process can intuitively be seen to have the spatial Markov property. Let  $S \subset T$ . The coordinates of any particles “born” in  $S$  are independent of the birth coordinates of particles external to  $S$ , since all births are induced by Poisson line processes. Also, the evolution of each particle is Markov, with the state including both its position and velocity. Furthermore, the state of all particles entering and exiting  $S$  can be determined from  $\chi^{\partial S}$  and  $\ell^{\partial S}$ .

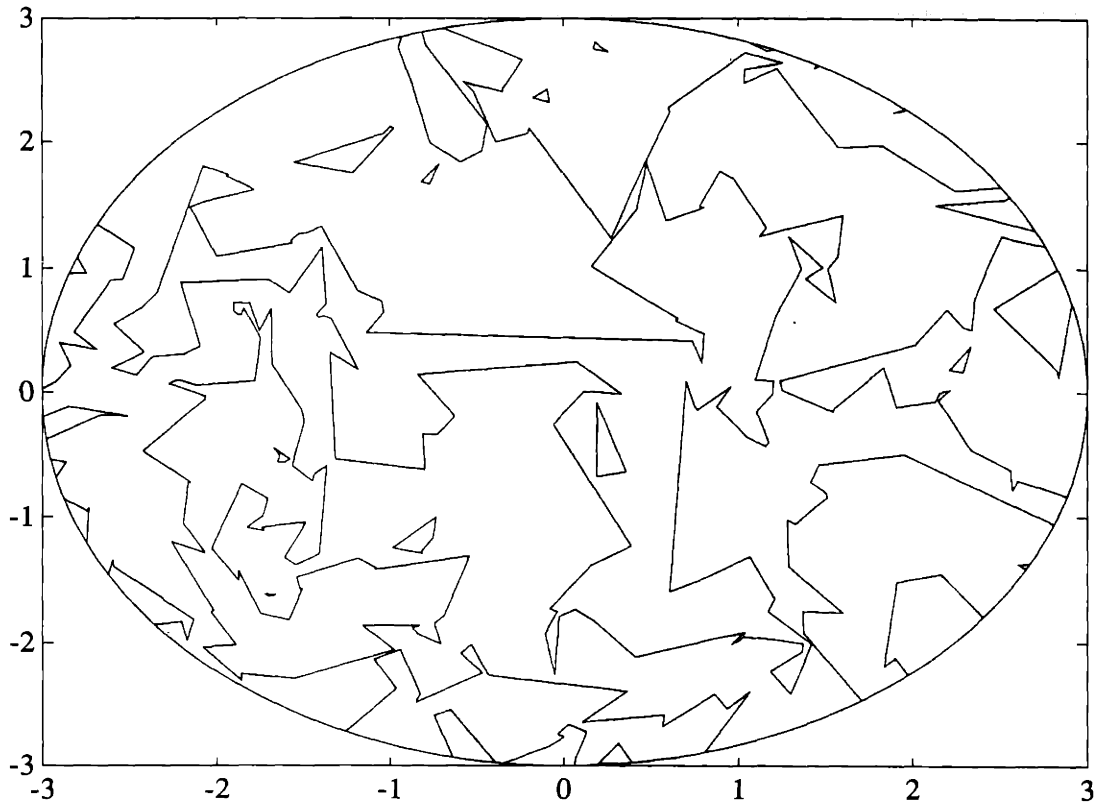
Figure 2.7: The birth, evolution, and death of a system of particles. Solid lines indicate the position of a particle  $y$  vs time  $t$ . Dashed lines indicate the system's evolution if the death of the particles is neglected. 'B' indicates a birth; 'D' indicates a death.



Thus,  $\chi^{\partial S}$  and  $\ell^{\partial S}$  provide as much information about the coloring  $\chi^S$  as does the coloring  $\chi^{T \setminus S}$ . Therefore, the partition will have the Markov property.

The partitions created by the Arak process never have more than two edges meet at a point, so the potential  $F_A(\chi)$  of the Arak process is infinite for colorings with nodes of degree higher than two. Arak and Surgailis[2] show that the potential  $F_A$  is otherwise proportional to  $2\Gamma$ , where  $\Gamma$  is the total length of the coloring's discontinuities. Figure 2.8 shows a computer generated partition for an Arak process. Because the reconstructions discussed in the following chapters do not rely on the Arak process, we will not review this analysis. However, two addition points should be mentioned now. First, particle models with more complex interactions between the particles also yield PRFs with the spatial Markov property, and this issue is addressed in Arak and Surgailis[3].

Figure 2.8: Partition of a circle of radius 3 produced by an Arak process.



More important for our purposes, the Arak process shows that the denominator of equation 2.2 is finite for the potential  $F_A$ . Therefore, for any potential  $F$  satisfying  $F(\chi) \geq F_A(\chi)$  for all  $\chi \in \Omega_T$ , the denominator of equation 2.2 will also be finite. Therefore, such PRFs will be properly defined. In order to use this result, the PRFs which we consider must have  $F(\chi) = \infty$  for all colorings with T-shaped, X-shaped, or more complex intersections of edges. From now on, we will make this assumption for all of our potentials.

# Chapter 3

## Reconstruction of Polygonal Random Fields

### 3.1 Introduction

A Bayesian formulation for reconstruction problems requires a number of elements: a prior probability distribution for the functions to be reconstructed, a probabilistic model for the observations, computation of the posterior distribution for the functions to be reconstructed, and methods to produce samples from the posterior distribution. PRFs provide a probability distribution on the set of polygonal colorings of a region. They are therefore a somewhat natural basis for a Bayesian approach to partitioning a region. In the next two chapters, we examine the use of PRFs in Bayesian models for image segmentation problems. That is, we will be attempting to restore and segment an image degraded by noise. In this chapter, we will formulate a model for partitioning the plane into piecewise constant regions. In the next chapter, we will use a coupled PRF and MRF model to reconstruct piecewise smooth functions.

We will begin this chapter by formulating a stochastic model for the Bayesian image segmentation problem, where we use PRFs in the model to reconstruct piecewise constant functions. A PRF with an additive potential will serve as a prior distribu-

tion for this model, and the measurement process will consist of noisy observations of an image taken at a finite number of locations. Under these circumstances, the resulting posterior distribution is also a PRF with an additive potential.

Since both prior and posterior distributions are PRFs with additive potentials, we must then develop a method to produce sample realizations from PRFs with arbitrary additive potentials. In the second section of this chapter, we discuss a Monte Carlo type algorithm which generates such samples, and then we discuss the issues involved in implementing the algorithm. In this chapter's final section, the measurement and reconstruction of images are carried out through computer simulation using the algorithm we have developed, and we discuss the results of these simulations.

It should be stressed that the problems in this and the next chapter are closely related. In the next chapter, we will need to partition a region and then interpolate over the subregions to reconstruct a piecewise smooth function. The algorithm we develop here to partition a region will be a necessary part of the algorithm developed for that problem.

## 3.2 A Model for Reconstructing Piecewise Constant Functions

In this section, we formulate a model for reconstructing piecewise constant images with polygonal discontinuities. That is, we are reconstructing polygonal colorings  $\chi : T \rightarrow J$ , where again  $T$  is a bounded convex region and  $J$  is a finite set. One potential application for such a model is object recognition in an industrial environment. The problem formulation has three components: the prior distribution, the stochastic measurement model, and the posterior distribution. We will begin with an artificially created piecewise constant image on  $T$ , and we will take noisy measurements of the image at a set of sites in  $T$ . Then, we will attempt to reconstruct the initial image from the measurements. The model here has been suggested in Clifford[5] and Clifford

and Middleton[4].

For the sake of clarity, we will discuss a specific example in detail throughout this section. For the example, let  $T$  be a square with side of unit length, and let  $J = \{-1, 1\}$ . That is, we will reconstruct a binary image in a square region. Fix  $\lambda = 2$ , where  $\lambda$  is the intensity of the Poisson line process which induces the Poisson line measure  $\mu$ . (Recall that the polygonal coloring measure  $\gamma$  of equation 2.2 is constructed from  $\mu$ . Although  $\lambda$  does not appear elsewhere, it is an important parameter in all reconstruction problems here, as it affects the scale of the reconstructions. For a Poisson line process on  $T$ , doubling  $\lambda$  is equivalent to doubling the side length of  $T$ .)

### 3.2.1 Prior Distribution

PRFs with additive potentials shall serve as prior distributions for our model. The prior potential will be denoted by  $F(\chi)$ , where  $F$  is an additive function. Equation 2.2 therefore gives the probability measure for the prior distribution. The PRFs of the prior distribution are required to have additive potentials because of the importance of the Markov property for reconstruction problems. However, it should be stressed that this is not an especially restrictive requirement: the class of PRFs with additive potentials is sufficiently large to model a wide variety of situations. For example, an additive potential can include terms which affect the length of a coloring's discontinuities, the number of its edges, and its coloring at individual sites.

For our specific example, we will choose the prior distribution induced by the Arak process. That is, let

$$F(\chi) = 2\Gamma$$

This potential is isotropic and tends to limit the total length of an image's discontinuities. Recall that we assume that  $F(\chi) = \infty$  for any coloring which has T-shaped, X-shaped, or more complex intersections of edges. Also, because this is the potential induced by the Arak process, the PRF exists.

### 3.2.2 Stochastic Measurement Model

There are obviously many ways that a stochastic measurement model could be formulated to represent observations of a given coloring  $\chi$ . For example, observations could be modeled by measurements of  $\chi$  at specific points in  $T$ , where the measurements are blurred according to a point-spread process. However, the resulting posterior distribution would no longer have the Markov property, and the Markov property is important for generating sample colorings of PRFs. We will require a measurement model which yields posterior distributions with additive potentials.

For our model, we shall consider measurements of a given image  $\chi$  at a fixed set of points  $t_i \in T$ , where  $i = 1, 2, \dots, n$ . The measurement at each point  $t_i$  shall be a function of the color of the image at  $t_i$  plus an independent noise term. Let  $\chi(t_i)$  be the color of the image at  $t_i$ , let  $g : J \rightarrow \mathfrak{R}$ , and let  $\{v_i\}$  be a set of independent identically distributed Gaussian random variables with mean 0 and variance  $\sigma^2$ . For each point  $t_i$ , we will obtain a measurement  $y_i$ , where

$$y_i = g(\chi(t_i)) + v_i. \quad (3.1)$$

Thus, the measurements  $y_i$  are normally distributed with mean  $g(\chi(t_i))$  and variance  $\sigma^2$ . Here,  $g$  can be interpreted as a function which assigns a real valued intensity to a color.

Let the measurement  $y$  be  $y = (y_1, y_2, \dots, y_n)$ . The measurement likelihood function,  $L(y|\chi)$ , is the density of  $y$  given the image  $\chi$ . Clearly, given  $\chi$ , the measurement  $y$  has a Gaussian distribution. Therefore, the measurement likelihood function is

$$L(y|\chi) = c \cdot \exp \left[ -\frac{1}{2\sigma^2} \sum_i [y_i - g(\chi(t_i))]^2 \right] \quad (3.2)$$

where  $c$  is a constant.

For the specific example of this section,  $J = \{-1, 1\}$ . Thus, we can simply let  $g$  be the identity function. (If we had chosen  $J = \{\text{white}, \text{black}\}$ , we could have assigned



$g(\text{white}) = -1$  and  $g(\text{black}) = 1$ .) Let the sites of  $t_i$  be a 21 by 21 grid, spaced at 0.05 intervals. Thus, the likelihood function for the example is:

$$L(y|\chi) \propto \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^{441} (y_i - \chi(t_i))^2 \right]. \quad (3.3)$$

In the next subsection, we will see that this measurement yields a posterior distribution with an additive potential. However, it should be noted that other measurement models would also have this property. Clifford and Middleton[4] suggest an observation process in which the measurement  $y$  is an inhomogeneous Poisson point process with intensity  $f(\chi(t))$ . Alternatively, the measurements could be modeled as a binary symmetric channel with error rate  $\epsilon$  at the sites  $t_i$ . Both of these methods yield additive posterior densities. However, for our purposes, we will be concerned only with the observation process formulated above. This fixed grid, real valued formulation is used here because it is compatible with the sampling we will do in chapter 4 when we reconstruct piecewise smooth functions.

### 3.2.3 Posterior Distribution

Given a measurement  $y = (y_1, y_2, \dots, y_n)$ , we wish to compute the posterior distribution. Let the prior density be  $e^{-F(x)}$ , and again let  $A \subset \Omega_T$ . Let  $e^{-F^*(x)}$  denote the probability density of the posterior distribution.

From Bayes' rule,

$$e^{-F^*(x)} = \frac{L(y|\chi) e^{-F(x)}}{f_y(y)} \quad (3.4)$$

where  $f_y(y)$  is the unconditioned probability density for the measurements. For a given set  $y$  of measurements,  $f_y(y)$  is simply a constant. Combining this with equation 3.4 yields

$$e^{-F^*(x)} = c \cdot \exp \left[ -\frac{1}{2\sigma^2} \sum_i [y_i - g(\chi(t_i))]^2 \right] \cdot \exp [-F(\chi)]$$

where  $c$  is a constant. The proportionality constant in the above density function will cancel out in the probability measure. The important feature of the above density is

that

$$e^{-F^*(\chi)} \propto \exp \left[ -F(\chi) - \frac{1}{2\sigma^2} \sum_i [y_i - g(\chi(t_i))]^2 \right]$$

and so the posterior distribution has potential:

$$F^*(\chi) = F(\chi) + \frac{1}{2\sigma^2} \sum_i [y_i - g(\chi(t_i))]^2. \quad (3.5)$$

The summation above is a sum over a fixed grid of points in the region  $T$ , and it is thus an additive function. Furthermore, the prior potential  $F(\chi)$  is also an additive function. Therefore, the posterior potential is an additive function. Thus, this problem formulation will always produce posterior distributions with the spatial Markov property.

Finally, note that each term  $[y_i - g(\chi(t_i))]^2$  in the sum of equation 3.5 is non-negative. Therefore  $F^*(\chi) \geq F(\chi)$ , and so the PRF with potential  $F^*$  exists.

For the specific example of this section, the prior potential is  $F(\chi) = 2\Gamma$ , and the measurement likelihood function is given in equation 3.3. Combining these terms, the posterior distribution for this example is therefore

$$F^*(\chi) = 2\Gamma + \frac{1}{2\sigma^2} \sum_{i=1}^{441} (y_i - \chi(t_i))^2. \quad (3.6)$$

This potential has a simple explanation. In the prior distribution, the likelihood of a coloring was less if it had extensive discontinuities. A coloring was penalized by a term proportional to  $e^{-2\Gamma}$ , and this penalty is also present in the posterior distribution. However, the posterior distribution also includes a penalty based on the difference between each measurement  $y_i$  and the coloring at each measured point  $\chi(t_i)$ . Thus, colorings which disagree with the measurements will be penalized. Also, when  $\sigma^2$  is small, meaning that the measurements are accurate, the penalty for disagreeing with the measurements will be large.

The full model for image segmentation problem of this chapter has now been presented: the prior distribution, the fixed grid stochastic measurement model, and

the posterior distribution. We have seen that an additive prior potential gives rise to an additive posterior potential, so that the posterior PRFs will have the Markov property. However, we do not yet have a method to generate sample colorings from the posterior distribution. This is the problem which we now address.

### 3.3 Obtaining Realizations of PRFs with Additive Potentials

So far, the only procedure to generate sample realizations of PRFs which we have discussed is the particle model of the Arak process. Yet although the Arak process is simple and intuitively appealing, it is also very limited: it can only produce realizations from a PRF with density  $F(\chi) = 2\Gamma$ . The reconstruction problem formulated in this chapter has a much more general posterior distribution:

$$F^*(\chi) = F(\chi) + \frac{1}{2\sigma^2} \sum_i [y_i - g(\chi(t_i))]^2$$

where  $F(\chi)$  is an additive function. To sample from the posterior distribution, we therefore need to generate sample configurations from PRFs with arbitrary additive potentials.

In this section, a Monte Carlo procedure which generates sample colorings of PRFs with additive potentials is discussed. The procedure is similar to Monte Carlo procedures used to generate realizations of Markov random fields. The key to the algorithm is to create a Markov chain on  $\Omega_T$ , the set of polygonal colorings of  $T$ , whose equilibrium is the density from which we wish to sample. The algorithm is a variation of an algorithm proposed by Clifford[5], which in turn is based on a class of algorithms discussed by Hastings[8].

We begin this section with a simple Monte Carlo algorithm which is similar to the algorithm for generating sample colorings. Next, we present the algorithm which generates sample colorings, and then we show that it produces colorings sampled from

the correct distribution.

For reconstruction problems based on the use of MRFs, one generally seeks an optimal estimate: for example, the sample configuration which maximizes posterior probability. However, as we shall discuss, such estimates for the PRFs of this chapter will not be useful. Therefore, we are now simply concerned with creating sample functions from PRFs with additive potentials. The question of the choice of appropriate statistics as estimates remains to be investigated.

### 3.3.1 Monte Carlo Method Preliminaries

In this subsection, the principles underlying the PRF algorithm are discussed, including the “detailed balance” criterion for Markov chains and a simple Monte Carlo procedure for sampling from a discrete probability density  $p(x)$ . This simple algorithm is presented because it closely parallels the algorithm for PRFs presented in the next subsection.

First, let us review the detailed balance criterion. Let a regular Markov chain have transition probabilities  $P(X \rightarrow Y)$ , and let  $p(X)$  be a probability distribution for its states. Recall that if

$$p(X) \cdot P(X \rightarrow Y) = p(Y) \cdot P(Y \rightarrow X)$$

for all states  $X$  and  $Y$ , then  $p(X)$  is the steady-state probability distribution for the Markov chain (Gallager [6], p. 5.21).

The detailed balance criterion is fundamental to the algorithm which produces sample colorings of PRFs. However, the inherent complexities of PRFs make that algorithm somewhat complicated. Therefore, we will begin by demonstrating how a similar algorithm can be used to sample from a probability distribution with probability mass function  $p(X)$  on a set  $\Omega$ . The steps of the PRF algorithm will closely follow the steps of this algorithm. The algorithm is as follows:

1. Starting from a state  $X \in \Omega$ , choose a new candidate state  $Y$  according to a

probability mass function  $q_X(Y)$ , where  $q_X(Y) > 0$  for all  $Y \in \Omega$ .

2. Accept new coloring  $Y$  with probability

$$\left[ 1 + \frac{p(X)q_X(Y)}{p(Y)q_Y(X)} \right]^{-1}.$$

If the new state is not accepted, retain the old state.

3. Return to step 1.

The algorithm specifies a regular Markov chain for which the elements of  $\Omega$  are the states. Furthermore, note that the transition probability  $P(X \rightarrow Y)$  equals  $q_X(Y)$  times the acceptance probability for  $Y$ , assuming  $X \neq Y$ . Therefore, for  $X \neq Y$ ,

$$\begin{aligned} p(X) \cdot P(X \rightarrow Y) &= p(X)q_X(Y) \cdot \left[ 1 + \frac{p(X)q_X(Y)}{p(Y)q_Y(X)} \right]^{-1} \\ &= \frac{p(X)q_X(Y)p(Y)q_Y(X)}{p(Y)q_Y(X) + p(X)q_X(Y)} \\ &= p(Y)q_Y(X) \cdot \left[ 1 + \frac{p(Y)q_Y(X)}{p(X)q_X(Y)} \right]^{-1} \\ &= p(Y) \cdot P(Y \rightarrow X) \end{aligned}$$

Thus, the algorithm meets the detailed balance criterion, and its equilibrium distribution will therefore be  $p(X)$ .

Finally, note that this algorithm has several characteristics which make it an appropriate basis for developing an algorithm for generating sample colorings of PRFs. This algorithm does not require that we sample directly from the distribution  $p(X)$ , but it does require both that we sample from some distribution  $q_X(Y)$  and that we can compute the ratio  $p(X)/p(Y)$ . These features take advantage of several characteristics of PRFs with additive potentials. It is not possible to sample from an arbitrary PRF with an additive potential in a direct way. However, it is easy to compute the ratio of the value of a PRF's density for two sample colorings,  $F^*(\chi_1)/F^*(\chi_2)$ .

Further, it is possible to sample from a PRF's local conditional distribution to choose a new candidate coloring for the PRF. These features will form the essential part of the algorithm for generating sample configurations of PRFs presented in the next subsection.

### 3.3.2 Algorithm for the Generation of PRFs with Additive Potentials

In this subsection, we present an algorithm to generate sample colorings from PRFs with additive potentials  $F^*(\chi)$  and with underlying Poisson line process intensity  $\lambda$ . For simplicity, let the region  $T$  be a rectangle. At each step in the algorithm, we will choose a rectangular subregion  $S$  of  $T$ . We will then make a transition along a Markov chain with states in  $\Omega_T$  whose equilibrium is the conditional probability distribution  $P_S(A_S|\chi^{T\setminus S})$  with density  $F^*(\chi)$ . To make matters more concrete, we will begin by presenting the algorithm. The algorithm will then be analyzed in the next subsection.

Algorithm to generate a sample coloring of a PRF with potential  $F^*(\chi)$ :

1. Randomly select a rectangular subregion  $S$  in  $T$ , where the current coloring of  $T$  is  $\chi$ . Let  $\ell_{cur}$  be the set of lines which contribute segments of discontinuity to  $\chi^S$ , the current coloring of  $S$ .
2. Run a Poisson line process with intensity  $\lambda$  on  $S$ . Let  $\ell^{S\setminus\partial S}$  be the set of lines generated by the Poisson line process.
3. Let  $\ell^{\partial S}$  be the set of lines which form the existing discontinuities of the coloring  $\chi$  on  $\partial S$ . Combine this set of lines with the lines generated in step 2: let  $\ell_{new} = \ell^{S\setminus\partial S} \cup \ell^{\partial S}$ .
4. The set  $\ell_{new}$  and the coloring  $\chi^{T\setminus S}$  determine  $\Omega_S^{\ell_{new}}(\chi^{T\setminus S})$ , the set of colorings

$\chi^S$  consistent with  $\chi^{T \setminus S}$  and having segments of the lines of  $\ell_{new}$  for their edges. Similarly,  $\ell_{cur}$  and  $\chi^{T \setminus S}$  determine the set  $\Omega_S^{\ell_{cur}}(\chi^{T \setminus S})$ . (See section 2.4 for the definition of  $\Omega_S^{\ell}(\chi^{T \setminus S})$ .)

Sample uniformly from the finite set  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  and combine this coloring with the current coloring of  $\chi^{T \setminus S}$  to choose a new candidate coloring  $\chi_{new}$  of  $T$ .

5. Accept the new coloring  $\chi_{new}$  with probability

$$\left[ 1 + \frac{\exp[-F^*(\chi)] |\Omega_S^{\ell_{new}}(\chi^{T \setminus S})|^{-1}}{\exp[-F^*(\chi_{new})] |\Omega_S^{\ell_{cur}}(\chi^{T \setminus S})|^{-1}} \right]^{-1} \quad (3.7)$$

where  $|\cdot|$  denotes the cardinality of the set. If the new coloring is not accepted, retain the old coloring  $\chi$ .

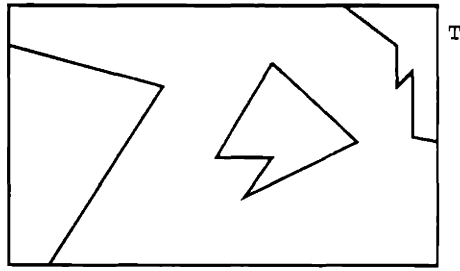
6. Return to step 1.

The first four steps of this algorithm correspond to the first step of the simple algorithm of the previous subsection. In these steps, we choose a new candidate coloring for  $T$ . Furthermore, the fifth step here corresponds to the second step of the simple algorithm, when we decide whether or not to accept the new coloring. The additional complexity of the PRF algorithm is largely due to the complexity of the process of choosing a new candidate coloring and the fact that this is a generalization of the previous algorithm to the continuous case. Figure 3.1 on the next page illustrates the steps of the algorithm for a particular coloring.

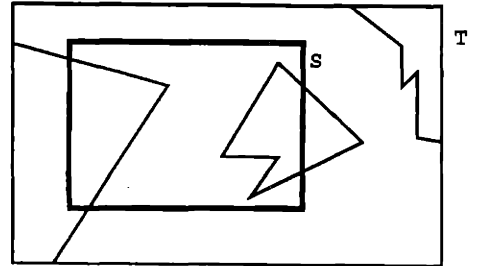
### 3.3.3 Analysis of the PRF Sample Coloring Algorithm

In this subsection, we show that the algorithm of the previous subsection generates samples from the PRF with density  $F^*(\chi)$ . This argument has two parts. First, that an iteration of the algorithm is a step along a Markov chain whose equilibrium is the conditional distribution  $P_S(A_S | \chi^{T \setminus S})$ ; that is, the equilibrium distribution is the conditional distribution with density  $F^*$ . Second, that iterating on successive subregions  $S$  of  $T$  leads to the proper distribution on the whole of  $T$ .

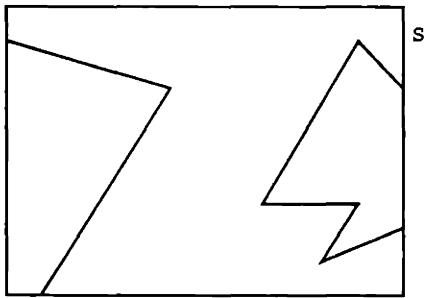
Figure 3.1: The steps of the algorithm to generate sample colorings of PRFs.



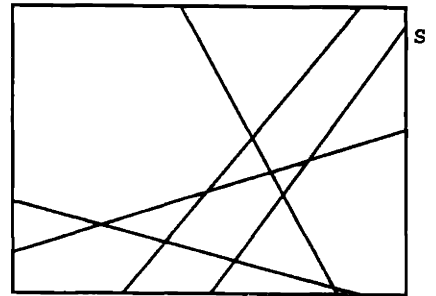
Begin with a coloring of the region T.



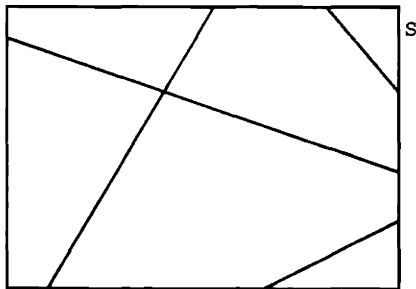
Step 1: Choose a subregion S in T



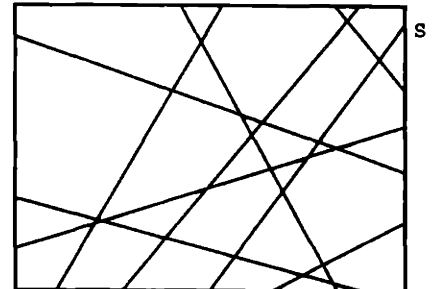
Expanded view of subregion S.



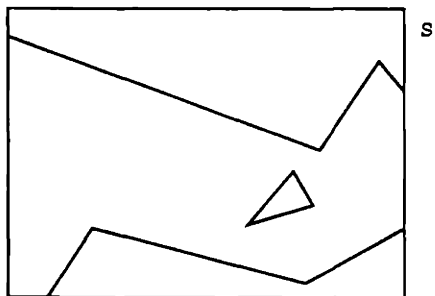
Step 2: Lines from a Poisson line process on S.



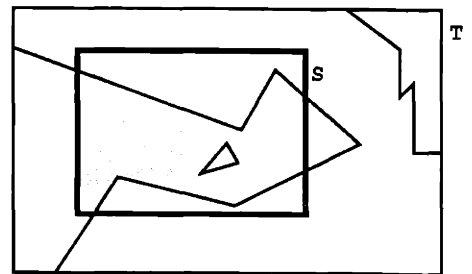
Step 3: Lines from the edges which cause discontinuity points on the boundary of S.



Step 3: The combined set of lines.



Step 4: A possible new candidate coloring. It uses one segment from each line and is consistent with the boundary conditions.



Step 5: The new coloring of region T, assuming the old coloring is replaced on S.



There is one difficulty in the argument here. The argument relies on use of the detailed balance criterion to show that the algorithm reaches the proper equilibrium distribution. But the detailed balance criterion holds for a Markov chain with a countable number of states, while a PRF has a continuous distribution. More work would be necessary to address the question of whether the detailed balance criterion holds for the continuous distribution of PRFs. Though we assume that we can use the detailed balance criterion here, we will not actually have fully proven the convergence for the algorithm.

Assume that the subregion  $S$  has been chosen, and consider a single iteration of the algorithm. We will show that the detailed balance criterion is satisfied for the conditional distribution on  $S$  with potential  $F^*$ .

Given the current coloring  $\chi_1$ , the algorithm induces a probability distribution on the candidates  $\chi_2$  to be the new coloring  $\chi_{new}$ . (Colorings consistent with  $\chi_1$  on  $\chi_1^{T \setminus S}$  may become the new coloring.) The algorithm runs a Poisson line process with intensity  $\lambda$  on  $S$ . Those lines are augmented with the lines which form discontinuities on the boundary of  $S$  to produce the set of lines  $\ell_{new}$ . The algorithm then chooses at random from the acceptable colorings of  $S$  which can be obtained from those lines – that is, it samples with uniform distribution from the set  $\Omega_S^{\ell_{new}}(\chi_1^{T \setminus S})$ . Thus, each element of  $\Omega_S^{\ell_{new}}(\chi_1^{T \setminus S})$  is chosen with probability  $|\Omega_S^{\ell_{new}}(\chi_1^{T \setminus S})|^{-1}$ .

Recall now that the conditional polygonal coloring measure  $\gamma_S$  is simply the expected number of colorings  $|\Omega_S^{\ell_{new}}(\chi_1^{T \setminus S})|$  which can be obtained from this procedure. Let  $q_{\chi_1}(\chi_2)$  be the density of the new candidate coloring distribution induced by the algorithm. We will denote  $\ell_{new}$  by  $\ell_2$  to emphasize the association between  $\ell_{new}$  and  $\chi_2$ . From this argument, we see that the density  $q_{\chi_1}(\chi_2)$  is proportional to  $|\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1}$  with respect to the measure  $\gamma_S$ .

The new candidate coloring  $\chi_{new}$  is influenced by the old coloring  $\chi_1$  only through the discontinuities along the boundary of  $S$ , which remain unchanged in an iteration.

Therefore, for any colorings  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  which are compatible on the boundary of  $S$ ,

$$q_{\chi_1}(\chi_3) = q_{\chi_2}(\chi_3).$$

The density for the new candidate colorings  $\chi_2$  will hence be written as  $q(\chi_2)$ , keeping in mind that it is dependent on the current coloring outside of  $S$ .

Let  $P(\chi_1 \rightarrow \chi_2)$  be the probability density with respect to  $\gamma_S$  that we move from coloring  $\chi_1$  to coloring  $\chi_2$ . Since this is simply  $q(\chi_2)$  times the acceptance probability, we have

$$P(\chi_1 \rightarrow \chi_2) = q(\chi_2) \cdot \left[ 1 + \frac{\exp[-F^*(\chi_1)] |\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1}}{\exp[-F^*(\chi_2)] |\Omega_S^{\ell_1}(\chi^{T \setminus S})|^{-1}} \right]^{-1}.$$

Now we are finally ready to verify that the detailed balance criterion holds for the local conditional distribution. Since the conditional probability density of a coloring  $\chi$  is  $\exp[-F^*(\chi)]$ , we have

$$\exp[-F^*(\chi_1)] P(\chi_1 \rightarrow \chi_2) = \exp[-F^*(\chi_1)] q(\chi_2) \left[ 1 + \frac{\exp[-F^*(\chi_1)] |\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1}}{\exp[-F^*(\chi_2)] |\Omega_S^{\ell_1}(\chi^{T \setminus S})|^{-1}} \right]^{-1}.$$

Substituting  $|\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1}$  for the density  $q(\chi_2)$  and a little algebra yields:

$$\begin{aligned} & \exp[-F^*(\chi_1)] P(\chi_1 \rightarrow \chi_2) \\ &= \exp[-F^*(\chi_1)] |\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1} \left[ 1 + \frac{\exp[-F^*(\chi_1)] |\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1}}{\exp[-F^*(\chi_2)] |\Omega_S^{\ell_1}(\chi^{T \setminus S})|^{-1}} \right]^{-1} \\ &= \frac{\exp[-F^*(\chi_1)] |\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1} \cdot \exp[-F^*(\chi_2)] |\Omega_S^{\ell_1}(\chi^{T \setminus S})|^{-1}}{\exp[-F^*(\chi_1)] |\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1} + \exp[-F^*(\chi_2)] |\Omega_S^{\ell_1}(\chi^{T \setminus S})|^{-1}} \\ &= \exp[-F^*(\chi_2)] |\Omega_S^{\ell_1}(\chi^{T \setminus S})|^{-1} \left[ 1 + \frac{\exp[-F^*(\chi_2)] |\Omega_S^{\ell_1}(\chi^{T \setminus S})|^{-1}}{\exp[-F^*(\chi_1)] |\Omega_S^{\ell_2}(\chi^{T \setminus S})|^{-1}} \right]^{-1} \\ &= \exp[-F^*(\chi_2)] P(\chi_2 \rightarrow \chi_1) \end{aligned}$$

Thus, an iteration of the algorithm satisfies the detailed balance criterion. Therefore, under the assumption that the detailed balance criterion is sufficient to show that the algorithm converges, the equilibrium distribution for the Markov chain on  $S$  is the conditional distribution with density  $F^*$ .

The fact that each iteration of the algorithm is a step along a Markov chain whose equilibrium is the local conditional distribution is sufficient to guarantee that the algorithm generates samples on all of  $T$  with the proper distribution. This result is analogous to the result for Markov random fields. Again, let  $\chi_1$  and  $\chi_2$  be two colorings of  $T$ . Note that it is possible to go between any two colorings in a finite number of steps, so the Markov chain on  $\Omega_T$  has only one class. As usual, to prove our result we need to use the detailed balance criterion and argue that  $\exp[-F^*(\chi_1)] P(\chi_1 \rightarrow \chi_2) = \exp[-F^*(\chi_2)] P(\chi_2 \rightarrow \chi_1)$  on the whole of  $T$ . But this is simply a repeat of the argument above, because a transition can take place from  $\chi_1$  to  $\chi_2$  only on an appropriate subregion  $S$  which supports the transition. This completes the proof that the equilibrium distribution of the given algorithm is a sample coloring of a PRF with additive potential  $F^*$ .

Before turning to the practical considerations involved in implementing the algorithm, it is worthwhile to mention two possible variations of the algorithm. First, the algorithm is amenable to parallel implementation. Because PRFs have the Markov property, an iteration of the algorithm on a subregion  $S$  depends only on the boundary of  $S$ . Therefore, we could choose multiple subregions  $S_i$  and make transitions on all of them concurrently. The only restriction on this procedure is the obvious one: the subregions  $S_i$  cannot intersect each other.

Finally, we could adapt the algorithm to produce sample colorings with maximum posterior probability. The algorithm currently produces samples from the posterior distribution with density  $F^*(\chi)$ . By adopting a procedure similar to simulated annealing, it would be possible to instead produce realizations with maximum posterior probability. To do this, we would add a time-dependent temperature term to the potentials. That is, we would let our potential be  $F^*(\chi)/T(t)$ , where  $T(t) \rightarrow 0$ . However, the cooling schedule and the convergence properties of this procedure are not clear. Furthermore, the value of maximum posterior probability estimates for recon-

struction problems using PRFs is dubious. For the example of the previous section, with the posterior density given by equation 3.6, we could draw arbitrarily small polygons around each of the measurement sites. Then, the length of the discontinuities would be arbitrarily small, the sum of equation 3.6 would be minimized, the posterior density would be maximized, and our partition would be trivial and useless.

### 3.4 Implementing the PRF Sample Coloring Algorithm

Implementing the algorithm for generating sample colorings of PRFs is significantly different than implementing an algorithm to generate sample realizations of a Markov random field. For a Markov random field, a sample realization can be stored in a computer simply as a matrix of the field's values at each of its sites. Such a simple method cannot work for the continuous PRF. In this section, we discuss the practical considerations necessary to implement the PRF algorithm on a computer. We will begin by mentioning the general structure for storing PRFs, and then we will discuss what is required to implement each step of the algorithm. Except for step 4 of the algorithm, there are no great theoretical difficulties here, and we will not spend much time on these matters. However, in the case of step 4, we will see that the PRF algorithm presents formidable computational difficulties.

First, we mention a possible computer storage structure for binary PRFs. Obviously, we cannot store the color at each point. Note, however, that a binary polygonal coloring is fully defined by its edges and by its coloring at a single point. Furthermore, it has a finite number of edges. Thus, a binary PRF can be stored as an array of the coordinates of the endpoints of its edges along with its coloring at a particular point. A non-binary PRF could be stored as an array of its edges along with an array of its colors on each component of the partition.

We now consider the computations involved in implementing the algorithm to generate sample colorings. Step 4 will be considered last, as it is the only step which will take any time to discuss.

Step 1, choosing a random subregion  $S$  in  $T$ , can be carried out in many ways. The only requirement is that each point in  $T$  must repeatedly be included in the chosen subregions. An algorithm for step 2, which requires us to run a Poisson line process on  $S$ , has been discussed in Section 2.2. For step 3, we must find the lines which cause discontinuities along the boundary of  $S$ . This can be done by checking to see if the edges of the coloring  $\chi$  intersect the segments which form the border of  $S$ , a simple and straightforward computation. And once we have done step 4, step 5 simply requires that we evaluate the function of equation 3.7 and compare it to a random number between 0 and 1.

Now we are ready to examine step 4. There is a significant computational difficulty involved in implementing this step of the PRF algorithm. The difficulty is the following: given the set of lines  $\ell_{new}$  and the coloring  $\chi^{T \setminus S}$ , we must determine the set  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$ . This is, of course, the set of colorings  $\chi^S$  which satisfy:

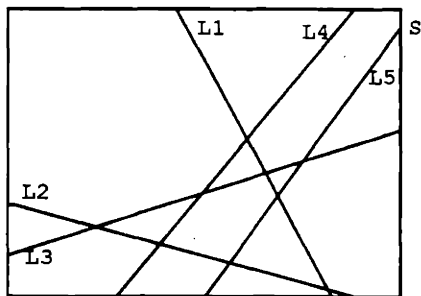
1.  $\chi^S$  is consistent with  $\chi^{T \setminus S}$  on the boundary of  $S$ .
2. For each line  $l_i$  in  $\ell_{new}$ , there is a segment  $[l_i] \subset l_i$  such that  $[l_i]$  is an edge of  $\chi^S$ .
3.  $\cup_i [l_i]$  is the set of discontinuities of  $\chi^S$ .

In figure 3.1, we chose a coloring from the nine lines of  $\ell_{new}$  which fit these criteria. However, figure 3.1 did not address the question of how many other such colorings could have been chosen. To sample uniformly from  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  and to compute  $|\Omega_S^{\ell_{new}}(\chi^{T \setminus S})|^{-1}$ , this question must be answered.

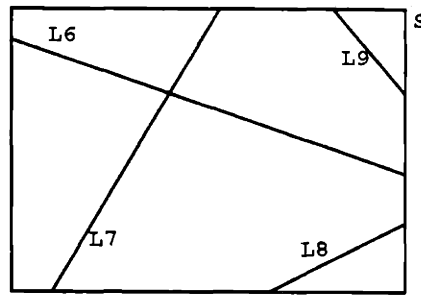
We will show that it is possible to enumerate  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  through an association between the lines of  $\ell_{new}$  and the nodes of a graph. The key observation is the

following. The lines of  $\ell_{new}$  can be divided into three sets: those generated by the Poisson line process on  $S$ , those which contribute a single discontinuity to  $\partial S$ , and those which contribute two discontinuities to  $\partial S$ . To form a valid coloring of the subregion  $S$ , an edge from a line that was generated by the Poisson line process must have as its endpoints its points of intersection with exactly two of the other lines in  $\ell_{new}$ . Further, the edge from each line in  $\ell_{new}$  that contributed a single discontinuity point to  $\partial S$  must have that discontinuity point as one of its endpoints, and it must have its intersection with one of the other lines in  $\ell_{new}$  as its other endpoint. Finally, lines which contribute two discontinuity points to  $\partial S$  must have both of those points as its endpoints.

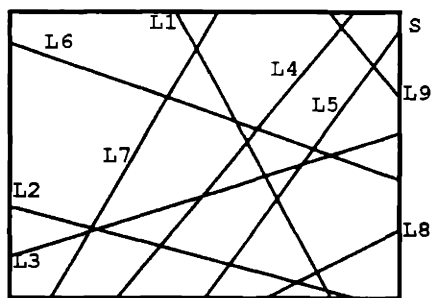
Figure 3.2: Choosing line segments to form a new coloring of  $S$ .



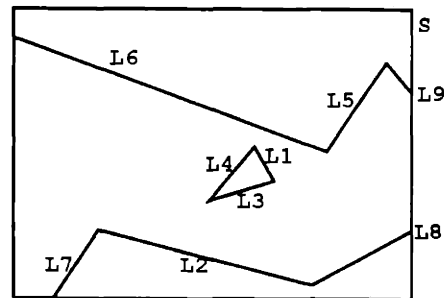
The lines from the Poisson line process on  $S$ .



Lines which form a single point of discontinuity on the boundary of  $S$ .



The combined set of lines



A possible new candidate coloring.

To illustrate this principle, consider figure 3.2. In it, we return to the examples from figure 3.1. Note that there are five lines from the Poisson line process on  $S$ , labeled L1 through L5. In the new candidate coloring, each contributes an edge

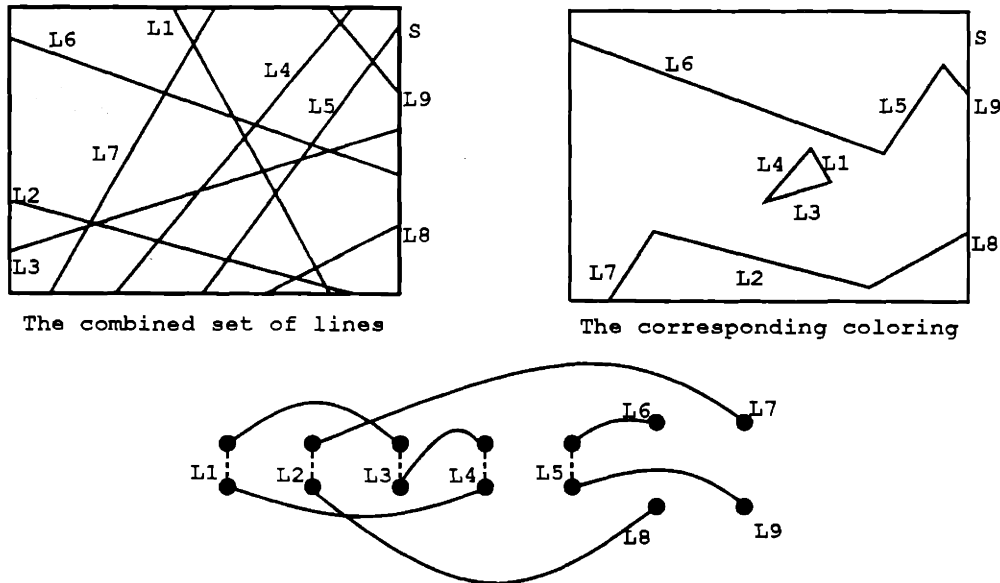
which is entirely within the interior of  $S$ . Thus, the endpoints of each of these edges come from the edge's intersection with one of the other lines. For example, the edge from L4 intersects L6 and L9. Furthermore, there are four lines which contributed a single discontinuity to the  $\partial S$ . These lines are labeled L6 through L9. Each of these edges intersects with only one edge in the interior of  $S$ . For example, L7 intersects L2. Note that there are no lines which contribute two discontinuity points to the  $\partial S$ . However, edges from these lines would have to remain unchanged in  $S$ , and they would not intersect any other edges.

Using this observation, we are ready to turn the enumeration of  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  into a combinatoric problem by associating the elements of  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  with a graph. Let each line from  $\ell_{new}$  which formed a single point of discontinuity on  $\partial S$  be associated with a single node of a graph. Also, let each line from  $\ell_{new}$  which was generated by the Poisson line process be associated with a pair of nodes of the graph. Finally, if two edges of a coloring from  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  intersect in  $S$ , let this intersection be associated with an arc connecting the nodes corresponding to the two edges. Under this association, every element of  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  will be associated with a graph in which each node is connected by an arc to exactly one other node. Here, a graph in which each node is connected to exactly one other node is called a node-pairing graph. Figure 3.3 shows the lines of figure 3.1, a node-pairing graph corresponding to those lines, and the element of  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  which corresponds to the graph shown.

Of course, given a set of lines  $\ell_{new}$ , not all node-pairing graphs corresponding to these lines will correspond to elements of  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$ . But the converse is true: every element of  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  will correspond to a node-pairing graph. Thus, to enumerate  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$ , we have the following algorithm: for a set of lines  $\ell_{new}$ , determine the corresponding set of nodes. Then, for each node-pairing graph of the set of nodes, check to see if the corresponding coloring is valid. These checks include ensuring that the boundary conditions are not violated, ensuring that no node is connected to a node corresponding to the same line, and ensuring that no edges cross each other.

There is, however, a problem with this algorithm. Computationally, it is extremely expensive. If there are  $n$  nodes, where  $n$  is even, there are  $(n-1) \cdot (n-3) \cdot (n-5) \cdot \dots \cdot (3)$  different node-pairing graphs. For the example of figure 3.3, where there are 14 nodes, the algorithm requires checking  $13 \cdot 11 \cdot 9 \cdot 7 \cdot 5 \cdot 3 = 135135$  possible colorings. Even for this fairly simple case, this would take a considerable amount of time. There are some techniques which could be used to shorten the algorithm. For example, in figure 3.3, the line L9 can only be connected to L4 or L5, so node-pairing graphs which do not have one of these connections would not have to be checked. However, such techniques cannot avoid the basic difficulty:  $|\Omega_S^{\ell_{new}}(\chi^{T \setminus S})|$  apparently grows exponentially with the number of lines in  $\ell_{new}$ . Thus, in order for the algorithm to run in a reasonable amount of time, it is necessary to choose the region  $S$  to be small, so that the number of lines in  $\ell_{new}$  remains small.

Figure 3.3: Associations between a set of lines, a graph, and a coloring



Each pair of nodes connected by a dashed line represents a line from the Poisson process on  $S$ . Each single node represents a line which formed a point of discontinuity on the boundary of  $S$ . Each node is connected to one other node by an arc. Such an arc means that the line segments from the two corresponding lines have endpoints at the two lines' point of intersection.

It may also be possible to make a reasonable approximation which avoids the dif-



difficulty in enumerating  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$ . The algorithm does not require that we calculate  $|\Omega_S^{\ell_{new}}(\chi^{T \setminus S})|$ ; it just requires that we calculate

$$\frac{|\Omega_S^{\ell_{new}}(\chi^{T \setminus S})|^{-1}}{|\Omega_S^{\ell_{cur}}(\chi^{T \setminus S})|^{-1}}.$$

If we assume that  $|\Omega_S^{\ell}(\chi^{T \setminus S})|$  is proportional to the number of node-pairing graphs which can be formed from the nodes corresponding to the lines of  $\ell$ , or that it is proportional to some other easily calculated approximation, the ratio above would be easy to calculate.

## 3.5 Experimental Results

We conclude this chapter with a few results from the algorithm to generate sample colorings from PRFs. We will begin with a pair of examples designed to illustrate the versatility of PRFs as prior potentials, and then we will present the results of a pair of simulations based on the image segmentation model of section 3.2. We will use a binary PRF on the unit square for our model. In each case, the sample colorings shown are taken from the PRF sample coloring algorithm once it achieves equilibrium.

### 3.5.1 Prior Potentials

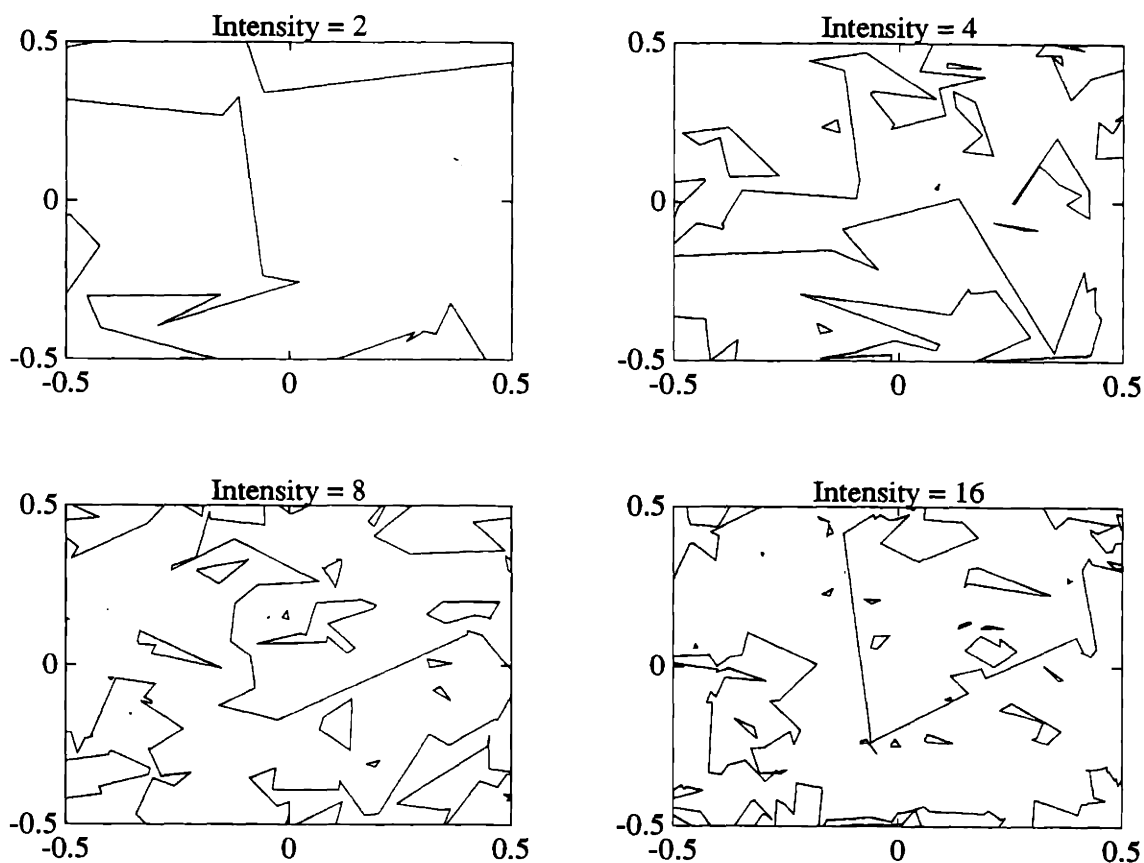
We have previously stated that the class of PRFs with additive potentials is sufficiently broad to model a wide variety of situations. This versatility enables us to choose our prior distribution in a manner which reflects our prior knowledge about the nature of the images with which we are dealing. In this subsection, we present a pair of examples which show how the choice of prior potentials affects sample colorings.

We begin by illustrating how the Poisson line process with intensity  $\lambda$  affects the scale of the images reconstructed. Again, recall that the Poisson line measure  $\mu$  is induced by this Poisson line process, and that for a Poisson line process on  $T$ , doubling

$\lambda$  is equivalent to doubling the side length of  $T$ . Figure 3.4 shows sample colorings from a PRF with  $F(\chi) = 2\Gamma$  when the intensity  $\lambda$  is 2, 4, 8, and 16. The figure shows that as  $\lambda$  increases, sample colorings tend to have more edges, and the edges tend to be shorter. The connection to the notion of scale is clear: sample colorings with  $\lambda = 2$  resemble subregions of sample colorings with  $\lambda = 16$ .

Figure 3.4: Varying the intensity of the PRF's underlying Poisson line process.

Here,  $F(\chi) = 2\Gamma$  and  $\lambda = 2, 4, 6, 8$ .



Although the choice of  $\lambda$  helps determine the scale of the sample colorings, it is the choice of  $F(\chi)$  which best enables us to influence the sample colorings. Clearly, the class of additive potentials is quite large, enabling us to affect the length, orientation, coloring, number of segments, and many other characteristics of our sample PRFs. Here, we give an example to show how  $F(\chi)$  can be chosen to promote the formation

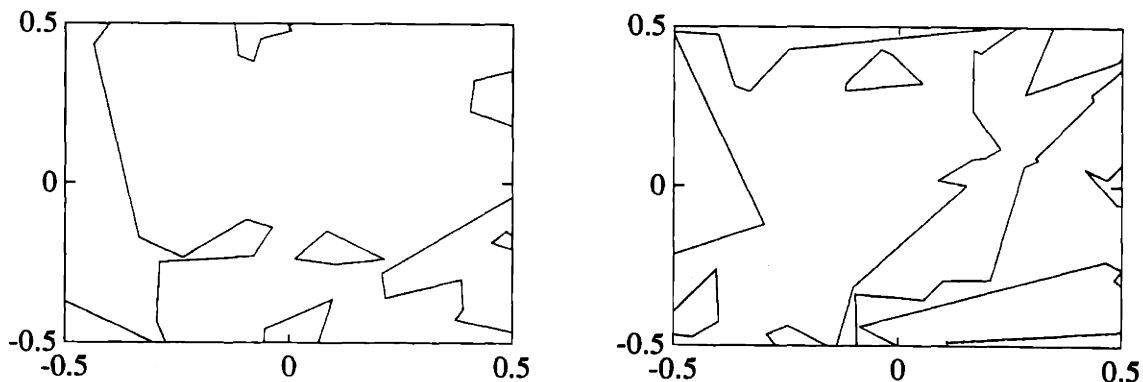
of obtuse angles for the intersections between edges. Let  $\lambda = 2$ , and consider the potential

$$F(\chi) = 0.1\Gamma + 3 \sum_{\theta_i} \left(1 - \frac{\theta_i}{\pi}\right)^3$$

where the sum over  $\theta_i$  means to sum over the angle  $\theta$  formed whenever two segments meet, where  $0 \leq \theta \leq \pi$ . Note that this potential is additive. The term  $0.1\Gamma$  means that there will be only a slight length penalty.  $(1 - \theta_i/\pi)^3$  terms are designed to penalize heavily colorings with small angles between segments, while only mildly penalizing colorings with obtuse angles. For example, if  $\theta_i = \pi/6$ , the penalty is 1.74; if  $\theta_i = \pi/2$ , the penalty is 0.38; and if  $\theta_i = 5\pi/6$ , the penalty is 0.014. Figure 3.5 shows a pair of colorings sampled independently from this potential. Comparing these to the PRFs shown in figure 3.4, it is clear that these sample colorings have a much higher ratio of obtuse angles to acute angles than the previous sample colorings.

Figure 3.5: Two sample colorings generated from a PRF with density

$$F(\chi) = 0.1\Gamma + 3 \sum_{\theta_i} \left(1 - \frac{\theta_i}{\pi}\right)^3$$



### 3.5.2 The Image Segmentation Problem

We are now ready to discuss simulations based on the image segmentation model of section 3.2. In this subsection, we will present two examples. For the first, we

reconstruct a simple artificially created image which includes a triangle and a parallelogram. For the second, we reconstruct a particular image sampled from a PRF with  $F(\chi) = 2\Gamma$  and  $\lambda = 2$ . Although both of these images are fairly simple, the results are indicative of the performance of the PRF image segmentation model for more complex images, since these simple images could simply be subregions from a larger, more complicated image.

We begin by reviewing the general framework for these two reconstruction problems. The images we attempt to reconstruct are binary PRFs; we will let the colors have values -1 and 1. The measurements  $y_i$  are taken of the images  $\chi$  on a fixed grid. For this section, the grid will be an evenly spaced 21 by 21 grid, so the measurements are made at 0.05 intervals. For each image, we will make three separate measurements: one with  $\sigma^2 = 0.1$ , one with  $\sigma^2 = 0.5$ , and one with  $\sigma^2 = 0.7$ . For the prior probability distribution, we will let  $F(\chi) = 2\Gamma$  and  $\lambda = 2$ . Given this prior density, from section 3.2 we know that the posterior distribution will have the potential

$$F^*(\chi) = 2\Gamma + \frac{1}{2\sigma^2} \sum_{i=1}^{441} (y_i - \chi(t_i))^2 \quad (3.8)$$

The first image is a triangle and a parallelogram with “color” 1 on a background with “color” -1. Figure 3.6 shows the edges of the image, as well as a three dimensional plot of a set of perfect measurements of the image. The three dimensional mesh is shown for comparison with the noisy measurements now made of the image.

Figure 3.6: The initial image, a triangle and a parallelogram.

The background color has value -1; the figures' color has value 1.

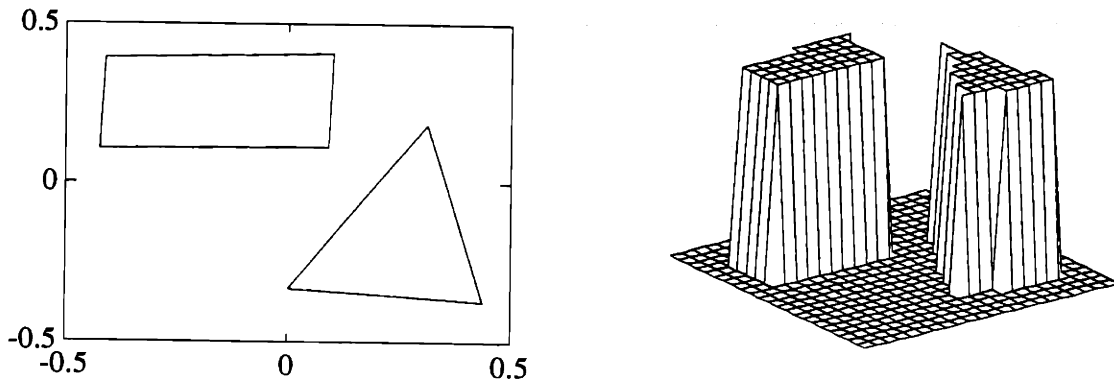


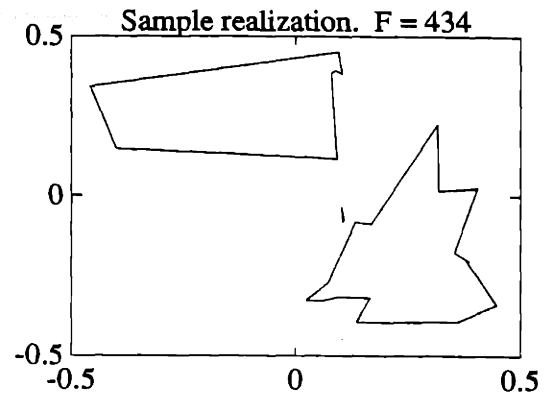
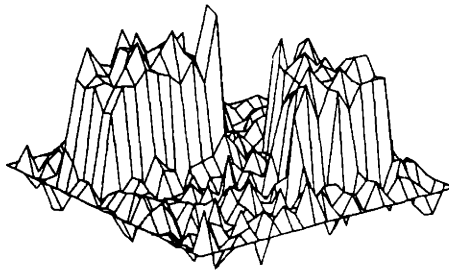
Figure 3.7 on the next page shows the results of the image segmentation algorithm applied to this image. Measurements of the initial coloring of figure 3.6 are taken with error variances  $\sigma^2 = 0.1, 0.5, \text{ and } 0.7$ . On the left side of figure 3.7 is a plot of these measurements. (The base in these plots has value -1.) On the right of figure 3.7 are sample colorings generated by the PRF algorithm using the posterior density given by equation 3.8.

For the three colorings shown in figure 3.7, the posterior density  $F^*(\chi)$  has value 434, 395, and 417. In comparison, the initial image for these three sets of data has value  $F^* = 433, 433, \text{ and } 474$ . In other words, when the measurements are noisy, the sample colorings tend to be better than the initial image with respect to the potential function. When the measurements are very good, the sample colorings generated by the algorithm are forced to agree with the initial image at every single measurement point, and then the total edge length of the sample colorings tends to be a small amount larger than the total edge length of the initial image.

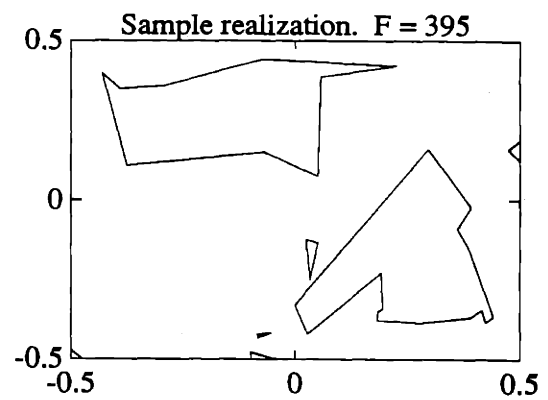
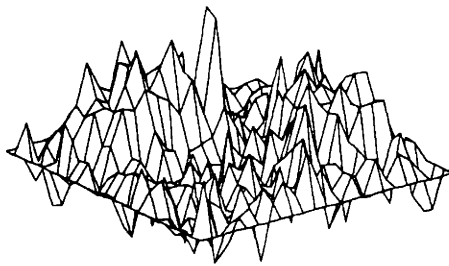
Figure 3.7: Measurements with three different variances and sample colorings from the corresponding posterior distributions.

$F$  is the value of the potential for the realization shown.

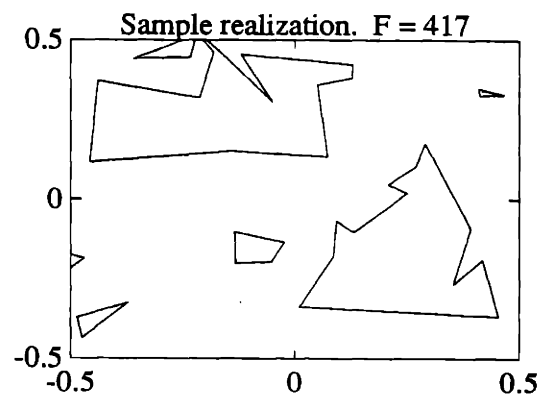
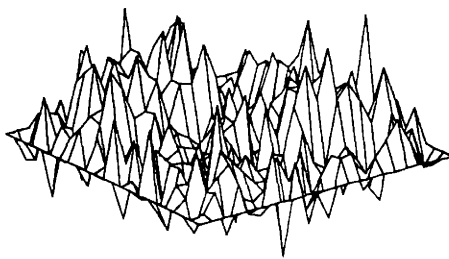
Measurements when variance = 0.1



Measurements when variance = 0.5

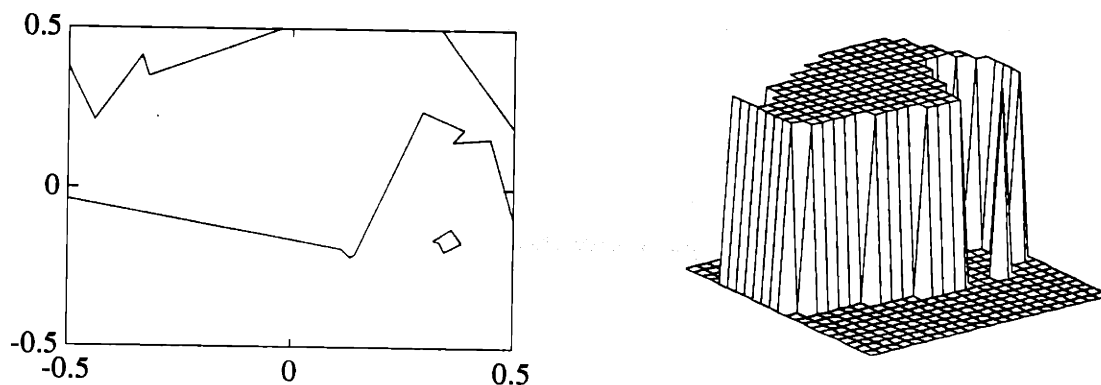


Measurements when variance = 0.7



Our second image is slightly more complicated than our first image. It is a particular binary PRF sampled created by the PRF generation algorithm with  $F(\chi) = 2\Gamma$  and  $\lambda = 2$ . Figure 3.8 shows the image and a three dimensional mesh plot for a set of perfect measurements of the image. Again, the base in the plot has value -1, and the height of the plot is 1.

Figure 3.8: The initial image, generated from Arak density with  $\lambda = 2$ .

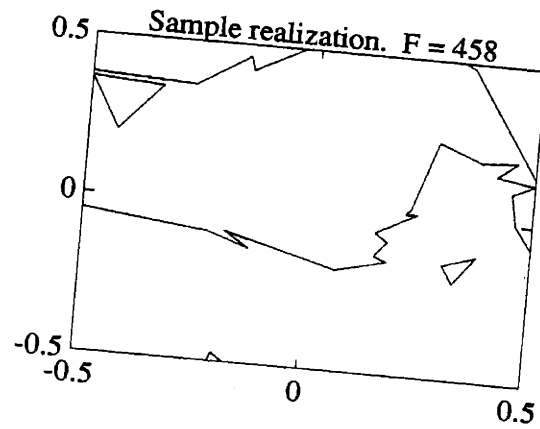
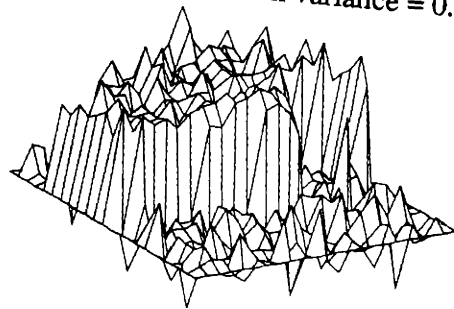


Repeating the same process as before, figure 3.8 shows the results of the image segmentation algorithm applied to this image. Measurements of the initial coloring of figure 3.8 are taken with error variances  $\sigma^2 = 0.1, 0.5, \text{ and } 0.7$ . Figure 3.9 shows the measurements and sample colorings generated by the PRF algorithm using the posterior density given by equation 3.8 and the corresponding measurements.

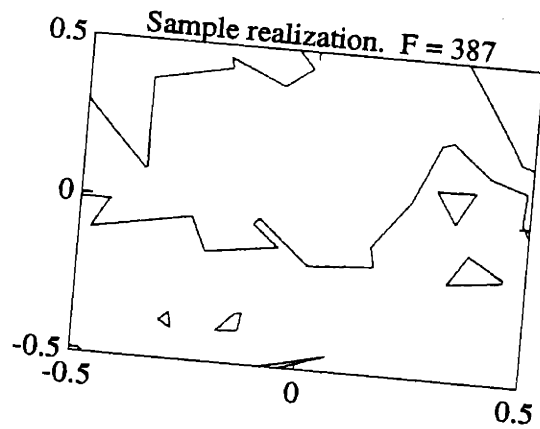
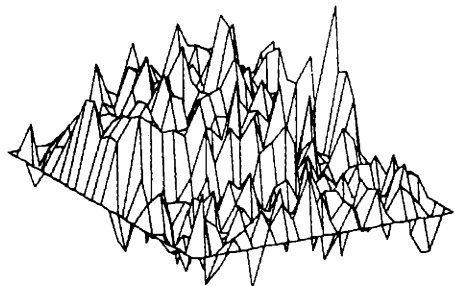
For the three colorings shown in figure 3.9, the posterior density  $F^*(\chi)$  has value 458, 387, and 429. In comparison, the initial image for these three sets of data has value  $F^* = 456, 432, \text{ and } 471$ . As before, when the measurements are noisy, the sample colorings tend to be better than the initial image with respect to the potential function.

Figure 3.9: Measurements with three different variances and sample colorings from the corresponding posterior distributions.  $F$  is the value of the potential for the realization shown.

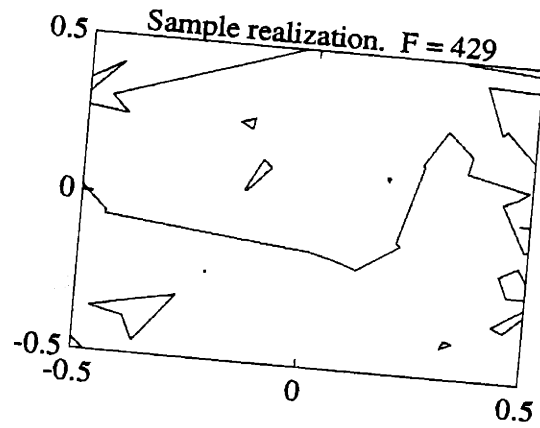
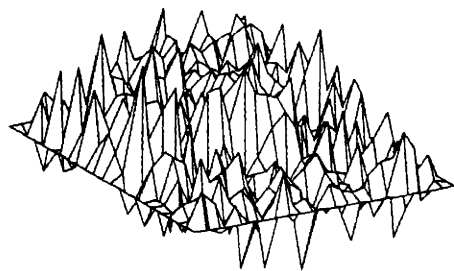
Measurements when variance = 0.1



Measurements when variance = 0.5



Measurements when variance = 0.7





# Chapter 4

## Reconstruction of Piecewise Continuous Functions

### 4.1 Introduction

In this chapter, we discuss an application of PRFs to the problem of restoring and segmenting a piecewise smooth function which has been degraded by noise. In our reconstruction, we will be concerned both with locating the boundaries between smooth regions and interpolating the function over the smooth regions.

Problems of this nature have previously been addressed through both variational formulations and discrete MRF formulations. For example, Mumford and Shaw[12] proposed a variational formulation based on minimizing a cost functional over a set of boundaries with smooth functions within the boundaries. Under this formulation, if the region is  $T$  and the observation is  $g$ , then a reconstructed function  $f$  and its edges  $\Gamma$  are found by minimizing

$$E(f, \Gamma) = c_1 \int_T (f - g)^2 dA + c_2 \int_{T \setminus \Gamma} \|\nabla f\|^2 dA + c_3 L(\Gamma) \quad (4.1)$$

where  $L(\Gamma)$  is the length of  $\Gamma$  and  $c_1, c_2$ , and  $c_3$  are constants.

A MRF approach to the reconstruction of piecewise continuous functions is dis-

cussed in Geman and Geman [7] and in Marroquin [11]. We will describe this approach more fully later, as it is similar to the PRF formulation we develop in this chapter. The MRF formulation is Bayesian. The prior distribution is modeled with two coupled MRFs: one MRF represents the function intensity, and the other represents the presence or absence of edge segments between adjacent sites of the function intensity lattice. The prior potential for this coupled model then includes terms which reflect the cost of edge segments and the cost of discontinuities in the function intensity for adjacent sites not separated by an edge.

One limitation of the coupled MRF model is that it can only model image discontinuities in a few arbitrary directions. For example, in the binary case, all edges are either vertical or horizontal. This limitation is significant: as a result, the coupled MRF model fails to preserve certain intrinsic properties of an image, like the total length of its edges. In contrast, a PRF model for the discontinuities of an image would not have this limitation. Certainly, PRFs seem to be more natural than MRFs for modeling the edges of an image.

These considerations motivate a different coupled model for the reconstruction of piecewise continuous functions: a MRF to represent the function intensity, and a PRF to represent the edges between the smooth regions. The purpose of this chapter is to develop such a model. We will begin by reviewing MRFs and the MRF coupled model, and then we present a PRF–MRF coupled model for reconstruction problems. Next, we discuss a Monte Carlo algorithm for the generation of sample realizations from the coupled model. Finally, we present experimental results from the model for the reconstruction of piecewise continuous functions.

## 4.2 Markov Random Fields

Because MRFs play an important role in the reconstruction problems of this chapter, we begin with a brief discussion of their properties. For a more extensive discussion,

see Marroquin[11] or Kinderman and Snell[10].

Let  $G = (Z, E)$  be an undirected graph, where  $Z$  is a set of  $n$  sites and  $E$  is a set of edges. Two sites connected by an edge are called neighbors. A subset of  $Z$  is a clique if it contains a single site or if each element of the subset is a neighbor of each other element. In this chapter, the graphs we consider are 4-connected lattices, so the cliques are single sites or adjacent pairs of sites.

With each site  $z_i$ , associate a random variable  $F_i$  taking on values in some finite set  $Q$ . Let  $F$  be the family of random variables  $\{F_i\}$ . A sample configuration of  $F$  has the form  $f = (f_1, f_2, \dots, f_n)$ , where  $f_i \in Q$ . Let  $P$  be a probability measure on the set of configurations. Then  $F$  is a MRF if

1.  $P(F = f) > 0$  for all configurations  $f$ .
2.  $P(F_i = f_i | F_j = f_j, j \neq i) = P(F_i = f_i | F_j = f_j, j \text{ neighbors } i)$  for all  $i$ .

Given the definition alone, MRFs would not be particularly useful, as they are difficult to specify or analyze in the above form. There is, however, a result which makes MRFs easy to specify and amenable to analysis: the Hammersley–Clifford theorem. It states that the probability distribution  $P$  for a MRF always has the form

$$P(f) = c \cdot \exp\left[-\sum_C V_C(f)\right]$$

where  $c$  is a constant and the sum is over the cliques  $C$  of the graph. A proof of this theorem can be found in Clifford[5]. Finally, note that  $P$  is a discrete probability mass function on the finite number of configurations  $f$ . Therefore, Monte Carlo algorithms using the detailed balance criterion can be used to generate sample configurations of MRFs.

### 4.3 A MRF Model for the Reconstruction of Piecewise Continuous Functions

Marroquin [11] suggests a formulation for the reconstruction of piecewise continuous functions using a coupled MRF model. In this section, we briefly review this model. This model is discussed here because it is similar to the coupled PRF–MRF model we develop in this chapter, and because part of the algorithm developed by Marroquin is applicable to the coupled PRF–MRF problem.

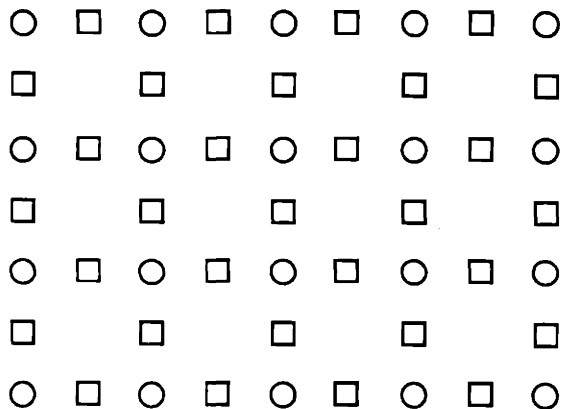
In the coupled MRF model, one MRF represents the function intensity at a lattice of sites. Let this lattice have  $n$  sites, and let the intensity at site  $i$  be  $f_i$ , where  $f_i$  is real valued. A second lattice has sites located between adjacent sites of the intensity lattice. These are the sites for the random variables of the “line process” MRF. At each of these sites, a binary random variable indicates the presence or absence of an “edge” segment between the two adjacent sites of the intensity lattice. If the adjacent function intensity sites are  $i$  and  $j$ , then let the associated “edge” random variable be  $l_{ij}$ , where  $l_{ij} = 1$  if there is an edge segment between the two function intensity sites, and  $l_{ij} = 0$  if there is not an edge segment. Figure 4.1 shows the dual lattice of function intensity and edge segment sites.

To reconstruct piecewise continuous functions, Marroquin proposes a prior potential  $U(f, l)$  for the coupled MRF model on this dual lattice, where  $U(f, l)$  is given by:

$$U(f, l) = \sum_i \sum_{N_i} (f_i - f_j)^2 (1 - l_{ij}) + \sum_{C_l} V_{C_l}(l). \quad (4.2)$$

Here, the sum over  $i$  is a sum over the sites of the intensity lattice, and  $N_i$  represents the horizontal and vertical nearest neighbors of those sites. Also,  $C_l$  is the set of cliques for the line process, and the function  $V_{C_l}$  is the potential function for the line process. For our purposes, we will not bother to specify the  $V_{C_l}$  terms explicitly, but they can be chosen to affect both the number and the orientation of edge segments.

Figure 4.1: Dual lattice of function intensity and edge element sites. Each circle represents a site in the function intensity lattice and takes on real values. Each square represents a site in the line process MRF and takes on a binary value indicating the presence or absence of a line segment separating the two adjacent intensity sites.



The terms of this potential have a simple interpretation. The  $(f_i - f_j)^2(1 - l_{ij})$  terms penalize discontinuities over the smooth regions where there are no edges, since  $l_{ij} = 0$  over these regions. However, there is no discontinuity penalty for sites separated by an edge, since  $l_{ij} = 1$  in such cases. Finally, the  $V_{C_i}$  terms discourage excessive discontinuities in a function; otherwise, letting  $l_{ij} = 1$  for all  $i$  and  $j$  would minimize  $U$ . Thus, the potential is chosen to strike a balance between the penalties for discontinuities within smooth regions and the penalties for the boundaries between the smooth regions.

The same observation process can be used for this coupled model as was used for the PRF model in Chapter 3. Specifically, let a piecewise smooth function  $f$  have intensity  $f_i$  at site  $i$  of the intensity lattice. Then let the measurement  $y_i$  at site  $i$  of the intensity lattice be

$$y_i = f_i + v_i$$

where the  $v_i$  are independent identically distributed Gaussian random variables with mean 0 and variance  $\sigma^2$ . This measurement process results in a posterior distribution

with potential

$$U_P(f, l|y) = \sum_i \sum_{N_i} (f_i - f_j)^2 (1 - l_{ij}) + \sum_{C_i} V_{C_i}(l) + \frac{1}{2\sigma^2} \sum_i (f_i - y_i)^2. \quad (4.3)$$

Using this coupled MRF model, Marroquin discusses the computation of optimal Bayesian estimates for the reconstruction of piecewise continuous functions. A computational difficulty of the model prevents the direct application of Monte Carlo methods which produce optimal estimates for MRFs: because the intensity  $f_i$  takes on continuous values, the complexity of the algorithms Marroquin discusses becomes excessively large when the  $f_i$  are discretized sufficiently finely. However, Marroquin points out a way to avoid this problem, and this method will be useful in the PRF-MRF coupled model discussed in this chapter. The key to this method is to note that for any fixed realization of the line process MRF, the posterior potential of equation 4.3 has the quadratic form:

$$U_P(f|l, y) = \sum_i \sum_{N_i: l_{ij}=0} (f_i - f_j)^2 + c + \frac{1}{2\sigma^2} \sum_i (f_i - y_i)^2 \quad (4.4)$$

where  $c$  is a constant. The global minimum for this potential can be found efficiently through simple deterministic minimization procedures. We will discuss a method to find this minimum later, when it is used in the coupled PRF-MRF model. Here, simply note that we are able to use a deterministic procedure to find the minimizer of the conditional potential of equation 4.4, thereby avoiding the computational difficulty caused by discretizing the continuous valued  $f_i$ .

## 4.4 A PRF-MRF Model for the Reconstruction of Piecewise Continuous Functions

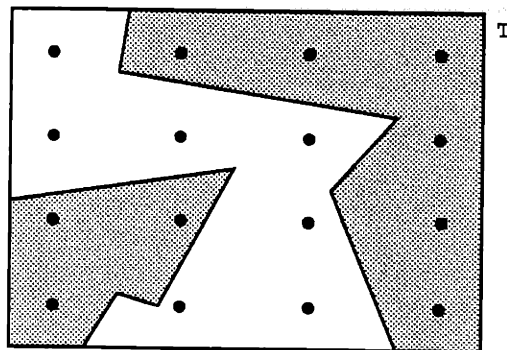
In the coupled MRF model for the reconstruction of piecewise continuous functions, one MRF is used to model image intensity, and the other is used to model the edges of

the image. In many cases, however, a PRF seems to be a more natural model for the edges of an image than a MRF. In this section, we formulate a coupled PRF-MRF model for the reconstruction of piecewise continuous functions. A MRF is again used to model function intensity, but a PRF is now used to model the function's discontinuities. The model formulated in this section has three components: the prior distribution, the stochastic measurement model, and the posterior distribution.

#### 4.4.1 Prior Distribution

We are concerned both with finding the edges of a piecewise smooth function on a bounded convex region  $T$  and with reconstructing the function's intensity at a set of sites in  $T$ . The edges can be modeled with a binary PRF with additive potentials, and the intensity function can be modeled with a continuous valued MRF with nearest neighbor potentials. Let the PRF have colorings  $\chi$  on  $T$ , and let the MRF have sample configurations  $f = (f_1, f_2, \dots, f_n)$  on a lattice of  $n$  sites in  $T$ , where  $f_i$  is the intensity at site  $i$ . Furthermore, let  $\Omega$  denote the set of all pairs of polygonal colorings  $\chi$  and sample configurations  $f$ . That is, an element of  $\Omega$  is a pair  $(\chi, f)$ . We seek a probability distribution for  $\Omega$  which serves as an appropriate prior distribution for the reconstruction of piecewise continuous functions. Figure 4.2 shows an example of a coloring  $\chi$  of a region  $T$  and a set of lattice sites in  $T$ .

Figure 4.2: A polygonal coloring for the region  $T$  and a set of lattice sites in  $T$ . Each lattice site takes on a real value representing function intensity, and the coloring's edges represent the function's discontinuities.



For the PRF, any additive potential  $F(\chi)$  could be chosen to serve as a prior potential. For the sake of simplicity, let  $F(\chi) = c_1 \Gamma$ , where  $c_1 > 0$ , and  $\Gamma$  is again the total length of the edges of  $\chi$ . (We continue to assume that  $F(\chi) = \infty$  for any coloring which has T-shaped, X-shaped, or more complex intersections of edges.) An edge of  $\chi$  will of course be associated with an edge of the piecewise smooth function.

In defining the potential for the MRF, it is necessary to choose a potential which reflects the requirement of piecewise continuity. Thus, the potential must depend both on the difference in intensity between neighboring sites of the lattice and also on whether or not an edge of the coloring  $\chi$  separates the neighboring sites. Given a coloring  $\chi$  and two adjacent lattice sites  $i$  and  $j$ , define

$$\begin{aligned} l_{ij} &= 1 && \text{if } i \text{ and } j \text{ are separated by an edge of } \chi \\ &= 0 && \text{otherwise.} \end{aligned}$$

(An edge of  $\chi$  separates sites  $i$  and  $j$  if the segment from  $i$  to  $j$  intersects an edge of  $\chi$ .) Because  $l_{ij}$  is dependent on the coloring  $\chi$ , the potential for the MRF will also be dependent on  $\chi$ . We can now choose a prior potential for the MRF. For a given  $\chi$ , the requirement of piecewise continuity for the MRF can be modeled by a potential of the form

$$U_\chi(f) = c_2 \sum_i \sum_{N_i} (f_i - f_j)^2 (1 - l_{ij})$$

where  $c_2 > 0$  is a constant, and  $N_i$  represents the nearest neighbors  $j$  to site  $i$ . The motivation for this choice of potential is clear: it penalizes the difference in intensity between adjacent sites in smooth regions, but not between adjacent sites separated by an edge. Also, this potential has the same form as the potential for the intensity lattice of the coupled MRF model given by equation 4.2.

Combining the potential  $F(\chi)$  for the PRF with the potential  $U_\chi(f)$  for the MRF yields the prior potential for the PRF-MRF model:

$$U(\chi, f) = c_1 \Gamma + c_2 \sum_i \sum_{N_i} (f_i - f_j)^2 (1 - l_{ij}). \quad (4.5)$$



This potential determines the probability distribution on  $\Omega$  in the obvious manner. For  $A \subset \Omega$ ,

$$P(A) = \frac{\int_A e^{-U(x,f)} df \gamma(d\chi)}{\int_{\Omega} e^{-U(x,f)} df \gamma(d\chi)}. \quad (4.6)$$

As  $c_1$  and  $c_2$  vary, there is a tradeoff between penalties for total edge length and for the discontinuities over the smooth regions. These constants must be chosen to reflect prior information about the functions to be reconstructed. Also, as  $c_2$  grows large,  $f$  is essentially required to remain constant over smooth regions.

#### 4.4.2 Measurement Model and Posterior Distribution

We will consider the same measurement model for the PRF-MRF model for the reconstruction of piecewise continuous functions as we described in the previous section for the coupled MRF model. Again, other measurement models could be used here; the primary requirement is that the posterior distribution have the Markov property.

Specifically, let  $f$  be a fixed piecewise continuous function on  $T$ . At each of the  $n$  sites of the MRF lattice, let a measurement be

$$y_i = f_i + v_i$$

where the  $v_i$  are independent identically distributed Gaussian random variables with mean 0 and variance  $\sigma^2$ . Writing the measurement as  $y = (y_1, y_2, \dots, y_n)$ , the measurement likelihood function is

$$L(y|f) = c \cdot \exp \left[ -\frac{1}{2\sigma^2} \sum_i (f_i - y_i)^2 \right]. \quad (4.7)$$

Using equations 4.5, 4.7, and Bayes' rule, the posterior distribution for  $\Omega$  has potential

$$U^*(\chi, f) = c_1 \Gamma + c_2 \sum_i \sum_{N_i} (f_i - f_j)^2 (1 - l_{ij}) + \frac{1}{2\sigma^2} \sum_i (f_i - y_i)^2. \quad (4.8)$$

This potential shares with the PRF model of chapter 3 the difficulties which arise from combining a continuous model for the discontinuities with discrete measurements. Note that equation 4.8 can be minimized by drawing arbitrarily small

polygons in  $\chi$  about each of the sites of the MRF. Thus, this model cannot produce meaningful MAP estimates.

Finally, note that the combined PRF–MRF distribution will have the Markov property. This observation follows simply from the fact that both the PRF and the MRF individually have the Markov property. However, because the PRF and the MRF are Markov in somewhat different senses, the PRF–MRF must incorporate what may be called the “boundary information” from both the PRF and the MRF. Consider a region  $S \subset T$  containing a set of lattice sites  $S_l$  for the MRF. For the PRF, the boundary information must include  $\chi^{\partial S}$  and  $\ell^{\partial S}$  (as defined in section 2.5). For the MRF, the boundary information must include  $f_i$  for sites  $i \notin S_l$  but neighboring sites in  $S_l$ , and it must include  $l_{ij}$  for sites  $i \notin S_l$  and sites  $j \in S_l$ .

## 4.5 Monte Carlo Algorithm for the PRF–MRF Model

In this section, we describe a Monte Carlo procedure to produce sample realizations from the coupled PRF–MRFs. For simplicity, we will concentrate on PRF–MRFs with a potential  $U^*(\chi, f)$  of the form given by equation 4.8. However, the algorithm is suitable for any coupled PRF–MRF with an additive PRF potential. The algorithm creates a Markov chain on  $\Omega$ , the set of pairs  $(\chi, f)$ , whose equilibrium is the distribution from which we wish to sample.

Recall that the PRF algorithm of section 3.3 produces samples from a PRF with an additive potential  $F(\chi)$ . Essentially, that algorithm produces a candidate coloring on a region of  $T$ , and it accepts the new candidate coloring with a probability designed to ensure that the chain has an equilibrium distribution with potential  $F(\chi)$ . We could develop a similar algorithm to sample from the distribution on  $\Omega$  with potential  $U^*(\chi, f)$ . Such an algorithm would select a region  $S \subset T$ , generate new candidates for  $\chi$  and  $f$  on  $S$ , and then accept the new  $\chi$  and  $f$  with the necessary

acceptance probability. However, because  $f$  is continuous valued, there are additional convergence questions for such an algorithm, and it would certainly be computationally expensive.

Instead, we will use the fact that for a given  $\chi$ , equation 4.8 has a global minimum  $f_\chi^*$ , and this minimum can be calculated deterministically. The algorithm will run a Markov chain on pairs of the form  $(\chi, f_\chi^*)$ . Let  $\Omega^* \subset \Omega$  be the set of all such pairs  $(\chi, f_\chi^*)$ . Rather than producing samples from the full set  $\Omega$ , the algorithm will produce samples from  $\Omega^*$ . It is apparent that the algorithm will not simply sample at random from the PRF-MRF with potential  $U^*$ . In fact, we shall see that the algorithm produces samples from the distribution that the potential  $U^*(\chi, f)$  induces on  $\Omega^*$ .

Although the algorithm restricts the Markov chain to  $\Omega^*$  partly for computational reasons, the restriction is also appealing: for a given  $\chi$ , the algorithm considers only the maximum *a posteriori* (MAP) estimate to  $f$ . Although the overall MAP estimate for  $U^*(\chi, f)$  cannot serve as a useful estimate for the reconstruction problem considered here, this limitation does not hold for the MAP estimate to  $f$  given  $\chi$ . In this section, we will first present the algorithm, and then we will show that it samples from the distribution on  $\Omega^*$  induced by the potential  $U^*(\chi, f)$ .

### 4.5.1 Algorithm for the Generation of PRF-MRF Sample Realizations

In this subsection, we present the algorithm to generate sample colorings  $\chi$  and intensity functions  $f$  given the potential  $U^*(\chi, f)$  of equation 4.8. The algorithm will then be analyzed in the next subsection. Let the region  $T$  be a rectangle, and assume we begin with some PRF coloring  $\chi$  and MRF configuration  $f$ . For a region  $S \subset T$ , the algorithm generates a new candidate coloring  $\chi_{new}$  on  $S$ , calculates  $f_{\chi_{new}}^*$  on  $S$ , and then chooses to accept or reject  $\chi_{new}$  and  $f_{\chi_{new}}^*$ . The algorithm is similar to the PRF algorithm of section 3.3, and it uses the same notation for the PRF.

Algorithm to generate a sample realization from  $\Omega^*$  with potential  $U^*(\chi, f)$ :

1. Randomly select a rectangular subregion  $S$  in  $T$ , where the current polygonal coloring of  $T$  is  $\chi$ , and the current MRF configuration is  $f$ . Let  $\ell_{cur}$  be the set of lines which contribute edges to  $\chi^S$ , the current polygonal coloring of  $S$ .
2. Run a Poisson line process with intensity  $\lambda$  on  $S$ . Let  $\ell^{S \setminus \partial S}$  be the set of lines generated by the Poisson line process.
3. Let  $\ell^{\partial S}$  be the set of lines which form the existing discontinuities of the coloring  $\chi$  on  $\partial S$ . Combine this set of lines with the lines generated in step 2: let  $\ell_{new} = \ell^{S \setminus \partial S} \cup \ell^{\partial S}$ .
4. The set  $\ell_{new}$  and the coloring  $\chi^{T \setminus S}$  determine  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$ , the set of colorings  $\chi^S$  consistent with  $\chi^{T \setminus S}$  and having segments of the lines of  $\ell_{new}$  for their discontinuities. Similarly,  $\ell_{cur}$  and  $\chi^{T \setminus S}$  determine the set  $\Omega_S^{\ell_{cur}}(\chi^{T \setminus S})$ . (See section 2.4 for the definition of  $\Omega_S^\ell(\chi^{T \setminus S})$ .)

Sample uniformly from the finite set  $\Omega_S^{\ell_{new}}(\chi^{T \setminus S})$  and combine this coloring with the current coloring of  $\chi^{T \setminus S}$  to choose a new candidate coloring  $\chi_{new}$  of  $T$ .

5. Calculate  $f_{\chi_{new}}^*$ , a minimizer of the potential  $U^*(\chi_{new}, f)$  given  $\chi_{new}$ .
6. Accept the new polygonal coloring  $\chi_{new}$  and the new MRF configuration  $f_{\chi_{new}}^*$  with probability

$$\left[ 1 + \frac{\exp[-U^*(\chi, f)] |\Omega_S^{\ell_{new}}(\chi^{T \setminus S})|^{-1}}{\exp[-U^*(\chi_{new}, f_{\chi_{new}}^*)] |\Omega_S^{\ell_{cur}}(\chi^{T \setminus S})|^{-1}} \right]^{-1} \quad (4.9)$$

where  $|\cdot|$  denotes the cardinality of the set. If the new coloring is not accepted, retain the old  $\chi$  and  $f$ .

7. Return to step 1.

The description of the algorithm will not be complete until we discuss a method to calculate  $f_{\chi_{new}}^*$ . For a given  $\chi_{new}$ , the task is to find the  $f$  which minimizes

$$U^*(f|\chi) = c + c_2 \sum_i \sum_{N_i} (f_i - f_j)^2 (1 - l_{ij}) + \frac{1}{2\sigma^2} \sum_i (f_i - y_i)^2. \quad (4.10)$$

Here, the  $l_{ij}$  are now known constants. This equation is non-negative and quadratic, so it has a global minimum  $f_{\chi_{new}}^*$ . This minimum can be found through a number of deterministic methods. For example, the gradient of equation 4.10 is given by

$$\frac{\partial U^*(f|\chi)}{\partial f_i} = 4c_2 \sum_{N_i} (f_i - f_j)(1 - l_{ij}) + \frac{1}{\sigma^2} (f_i - y_i). \quad (4.11)$$

Since the gradient must be 0 at  $f_{\chi_{new}}^*$ , it follows that the components  $f_i^*$  of  $f_{\chi_{new}}^*$  must satisfy

$$4c_2 \sum_{N_i} f_i^* (1 - l_{ij}) + \frac{1}{\sigma^2} f_i^* = 4c_2 \sum_{N_i} f_j^* (1 - l_{ij}) + \frac{1}{\sigma^2} y_i$$

for all  $i$ . Therefore,  $f_{\chi_{new}}^*$  is a fixed point of the system

$$f_i^{(k+1)} = \frac{4c_2 \sum_{N_i} f_j^{(k)} (1 - l_{ij}) + \frac{1}{\sigma^2} y_i}{4c_2 \sum_{N_i} (1 - l_{ij}) + \frac{1}{\sigma^2}} \quad (4.12)$$

where we let  $f_i^{(k+1)} = f_i^{(k)}$  if the denominator is 0. Furthermore, (Marroquin[11], p 133), the system is stable and converges to  $f_{\chi_{new}}^*$ .

## 4.5.2 Analysis of the PRF–MRF Algorithm

We now show that the algorithm of the previous subsection samples from the distribution on  $\Omega^*$  induced by the potential  $U^*(\chi, f)$ . This is the distribution given by

$$P(A) = \frac{\int_A e^{-U^*(\chi, f_{\chi}^*)} \gamma(d\chi)}{\int_{\Omega^*} e^{-U^*(\chi, f_{\chi}^*)} \gamma(d\chi)} \quad (4.13)$$

for  $A \subset \Omega^*$ . As in chapter 4, we shall show that the algorithm satisfies the detailed balance criterion. Again, because the PRF is continuous valued, this is not fully sufficient to prove the the algorithm's convergence.

Assume that the subregion  $S$  has been chosen, and consider a single iteration of the algorithm. It is clear that the algorithm creates a Markov chain on  $\Omega^*$ . Given the current coloring and intensity  $(\chi_1, f_1)$ , the algorithm induces a probability distribution on the candidates  $(\chi_2, f_2)$  to be the new coloring  $\chi_{new}$  and function intensity  $f_{\chi_{new}}^*$ . Note that the new candidate coloring is chosen in this algorithm in exactly the same way that the new candidate coloring was chosen in the PRF algorithm of section 3.3. Also, the candidate function intensity  $f_2$  is dependent only on  $\chi_2$  and the measurements, since  $f_2$  must be  $f_{\chi_2}^*$ . Thus, both  $\chi_2$  and  $f_2$  are dependent on  $\chi_1$ , but not on  $f_1$ . Furthermore, letting  $q(\chi_2, f_2)$  be the density induced by the algorithm on the new candidate colorings, we see that  $q(\chi_2, f_2)$  has precisely the same value as  $q(\chi_2)$  in section 3.3. Therefore,  $q(\chi_2, f_2)$  is again proportional to  $|\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1}$ .

Let  $P[(\chi_1, f_1) \rightarrow (\chi_2, f_2)]$  be the transition probability density for moving from  $(\chi_1, f_1)$  to  $(\chi_2, f_2)$ . Since this is  $q(\chi_2, f_2)$  times the acceptance probability, we have

$$P[(\chi_1, f_1) \rightarrow (\chi_2, f_2)] = q(\chi_2, f_2) \cdot \left[ 1 + \frac{\exp[-U^*(\chi_1, f_1)] |\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1}}{\exp[-U^*(\chi_2, f_2)] |\Omega_S^{\ell_1}(\chi_1^{T \setminus S})|^{-1}} \right]^{-1}.$$

Now we are now ready to verify that the detailed balance criterion holds when the distribution on the states of  $\Omega^*$  has potential  $U^*(\chi, f)$ . Since the density for a pair  $(\chi, f)$  is then  $\exp[-U^*(\chi, f)]$ , we have

$$\begin{aligned} & \exp[-U^*(\chi_1, f_1)] P[(\chi_1, f_1) \rightarrow (\chi_2, f_2)] \\ &= \exp[-U^*(\chi_1, f_1)] q(\chi_2, f_2) \left[ 1 + \frac{\exp[-U^*(\chi_1, f_1)] |\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1}}{\exp[-U^*(\chi_2, f_2)] |\Omega_S^{\ell_1}(\chi_1^{T \setminus S})|^{-1}} \right]^{-1}. \end{aligned}$$

Substituting  $|\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1}$  for the density  $q(\chi_2, f_2)$  yields:

$$\begin{aligned} & \exp[-U^*(\chi_1, f_1)] P[(\chi_1, f_1) \rightarrow (\chi_2, f_2)] \\ &= \exp[-U^*(\chi_1, f_1)] |\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1} \left[ 1 + \frac{\exp[-U^*(\chi_1, f_1)] |\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1}}{\exp[-U^*(\chi_2, f_2)] |\Omega_S^{\ell_1}(\chi_1^{T \setminus S})|^{-1}} \right]^{-1} \\ &= \frac{\exp[-U^*(\chi_1, f_1)] |\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1} \cdot \exp[-U^*(\chi_2, f_2)] |\Omega_S^{\ell_1}(\chi_1^{T \setminus S})|^{-1}}{\exp[-U^*(\chi_1, f_1)] |\Omega_S^{\ell_2}(\chi_1^{T \setminus S})|^{-1} + \exp[-U^*(\chi_2, f_2)] |\Omega_S^{\ell_1}(\chi_1^{T \setminus S})|^{-1}} \end{aligned}$$

$$\begin{aligned}
&= \exp[-U^*(\chi_2, f_2)] |\Omega_S^{t_1}(\chi^{T \setminus S})|^{-1} \left[ 1 + \frac{\exp[-U^*(\chi_2, f_2)] |\Omega_S^{t_1}(\chi^{T \setminus S})|^{-1}}{\exp[-U^*(\chi_1, f_1)] |\Omega_S^{t_2}(\chi^{T \setminus S})|^{-1}} \right]^{-1} \\
&= \exp[-U^*(\chi_2, f_2)] P[(\chi_2, f_2) \rightarrow (\chi_1, f_1)]
\end{aligned}$$

Thus, an iteration of the algorithm satisfies the detailed balance criterion. Therefore, the equilibrium distribution for the Markov chain on  $\Omega^*$  has density  $U^*(\chi, f_\chi^*)$ , and the algorithm produces samples from the probability distribution given by equation 4.13.

Finally, note the similarity between the probability distribution of equation 4.13 and the probability distribution for a PRF given by equation 2.2. This similarity arises because of the simple one-to-one correspondence between the PRF of the coupled model and the elements of  $\Omega^*$ : for each  $\chi$  in  $\Omega_T$  of the PRF, there is one  $(\chi, f_\chi^*)$  in  $\Omega^*$ . In fact, the distribution of equation 4.13 is the distribution for a PRF. The potential  $U^*(\chi, f_\chi^*)$  in equation 4.13 is actually a function only of  $\chi$ : since  $f_\chi^*$  is determined by  $\chi$ ,  $U^*(\chi, f_\chi^*)$  could be written as  $U^*(\chi)$ , and it is therefore the potential of a PRF on  $\Omega_T$ . Thus, this algorithm reduces the problem of producing samples from the coupled PRF-MRF model to the problem of producing samples from a PRF. However, the PRF potential here is not additive, as the boundary information of the MRF is still required to compute  $U^*(\chi)$ .

## 4.6 Experimental Results

We end with a pair of examples in which we use the algorithm to segment and restore a piecewise smooth function. For each example, we begin by choosing a piecewise smooth function on the unit square. Two sets of noisy measurements are then taken of the function. For each set of measurements, the algorithm of the previous section is used to generate a sample coloring  $\chi$  and intensity function  $f_\chi^*$  from the resulting posterior distribution with potential  $U^*(\chi, f_\chi^*)$  on  $\Omega^*$ .

For each initial function, the measurements  $y_i$  are taken on an evenly spaced 21 by 21 grid. One set of measurements has  $\sigma^2 = 0.1$ , and the other has  $\sigma^2 = 0.7$ . We

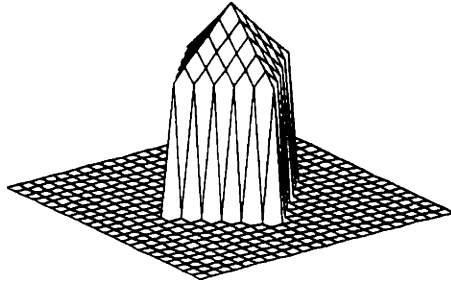
then choose the prior distribution to have  $c_1 = 7$ ,  $c_2 = 2$ , and underlying Poisson process intensity  $\lambda = 2$ . From equation 4.8, the posterior distribution from which the algorithm samples has potential

$$U^*(\chi, f) = 7\Gamma + 2 \sum_i^{441} \sum_{N_i} (f_i - f_j)^2 (1 - l_{ij}) + \frac{1}{2\sigma^2} \sum_i^{441} (f_i - y_i)^2. \quad (4.14)$$

Some care must be taken in choosing  $c_1$  and  $c_2$  in order for the algorithm to obtain reasonable results. If  $c_1$  is too large compared to  $c_2$ , the algorithm tends to find no edges in  $\chi$ . If it is too small, the algorithm finds an extensive set of edges, and the reconstructed function simply mirrors the measurement.

Our first piecewise continuous function is shown in figure 4.3. It is a tower with height (or function intensity) 0 at the base, 3 at the peak, and 2 around the top edges.

Figure 4.3: The initial piecewise continuous function.

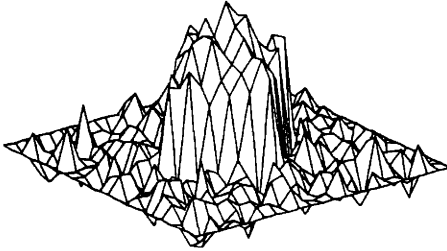


At each site in the grid, measurements of figure 4.3 are taken with error variances  $\sigma^2 = 0.1$  and  $\sigma^2 = 0.7$ . These two sets of measurements are shown in figure 4.4.

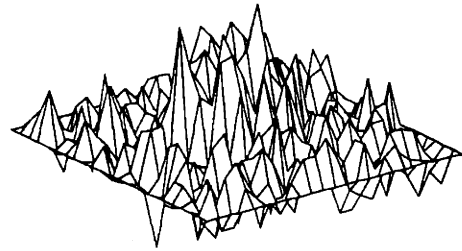


Figure 4.4: Measurements with two different variances of the initial function.

Measurement when variance = 0.1

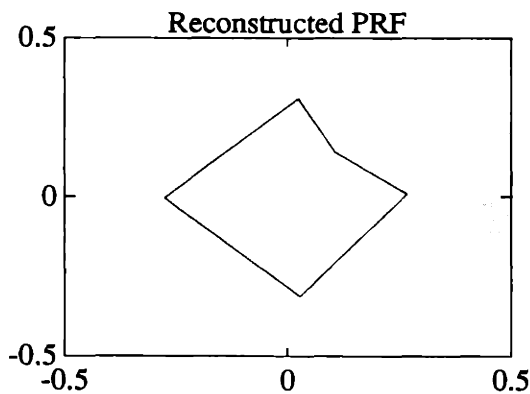


Measurement when variance = 0.7



For each of these sets of measurements, the algorithm generates a sample PRF coloring  $\chi$  and the corresponding optimal intensity function  $f_\chi^*$ . Figure 4.5 shows the reconstructed PRF coloring and intensity function for the measurements with  $\sigma^2 = 0.1$ . For this set of measurements, in which there is relatively little error, the algorithm is effective in locating the edges and restoring the initial function.

Figure 4.5: Reconstructed PRF and intensity function when  $\sigma^2 = 0.1$ .



Reconstructed intensity function

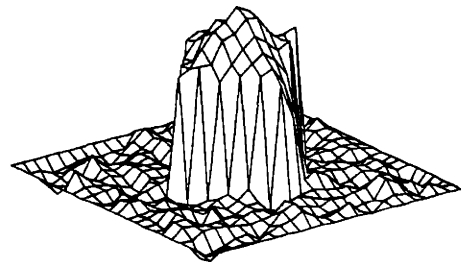
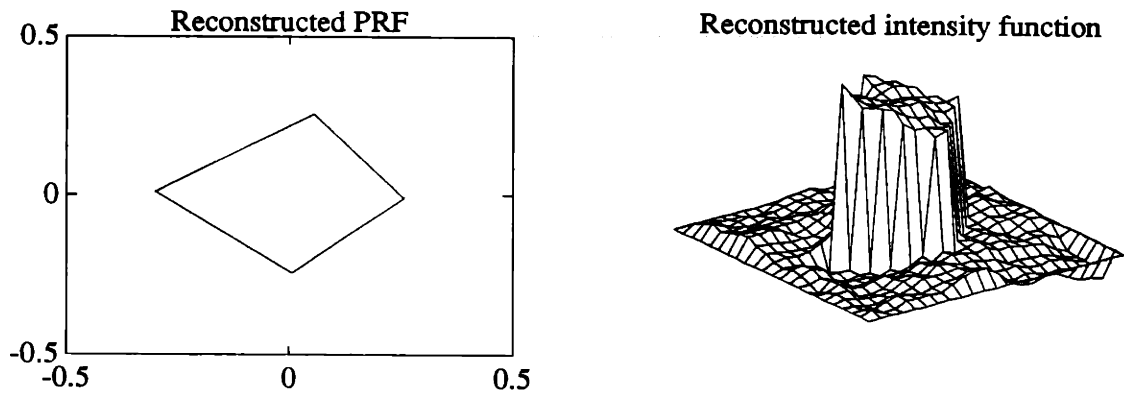


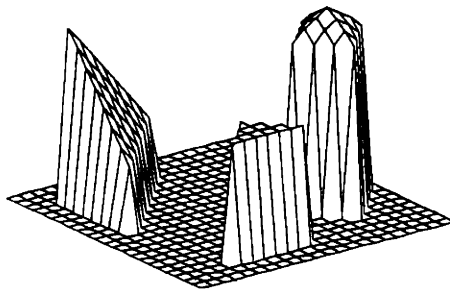
Figure 4.6 shows the reconstructed PRF coloring and intensity function when  $\sigma^2 = 0.7$ . Although the shape of the top of the tower is lost due to the considerable amount of error in the measurements, the algorithm is still able to locate the edges of the function.

Figure 4.6: Reconstructed PRF and intensity function when  $\sigma^2 = 0.7$ .



Our second piecewise continuous function is shown in figure 4.7. It consists of four separate regions: a base with height (or function intensity) 0, a triangle with height 3, a sloping triangle with height ranging from 1.6 to 4.4, and a square tower with height ranging from 4.3 to 5.

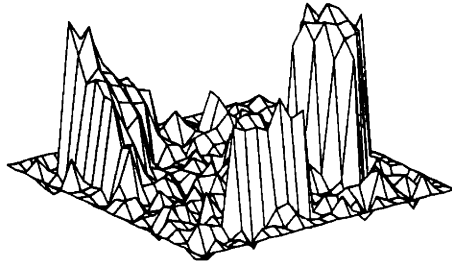
Figure 4.7: The second piecewise continuous function.



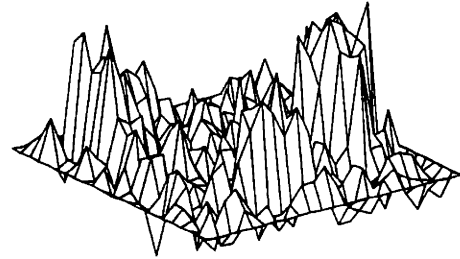
Measurements of figure 4.7 are taken on the 21 by 21 grid with variance  $\sigma^2 = 0.1$  and  $\sigma^2 = 0.7$ . These sets of measurements are shown in figure 4.8.

Figure 4.8: Measurements with two different variances of the second function.

Measurement when variance = 0.1



Measurement when variance = 0.7



Given the measurements of figure 4.8, we again use the algorithm to generate sample PRF colorings  $\chi$  and function intensities  $f_\chi^*$ . Figure 4.9 shows a sample reconstructed PRF coloring and intensity function for the measurements with  $\sigma^2 = 0.1$ . The reconstructed coloring finds the edges of the initial function, and it also includes some additional edges. The reconstructed intensity is similar to the initial intensity function. The addition of extra edge segments is not surprising: in regions where the measurement error happens to be large, additional edge segments will tend to reduce the value of the potential.

Figure 4.9: Reconstructed PRF and intensity function when  $\sigma^2 = 0.1$ .

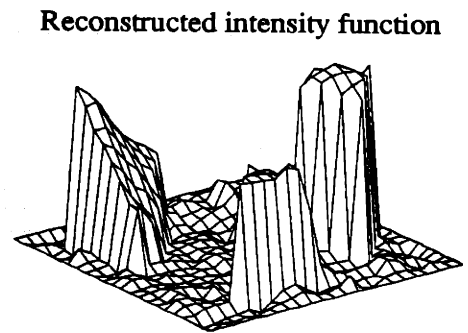
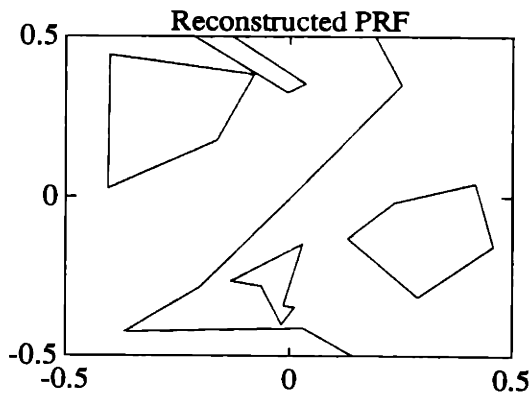
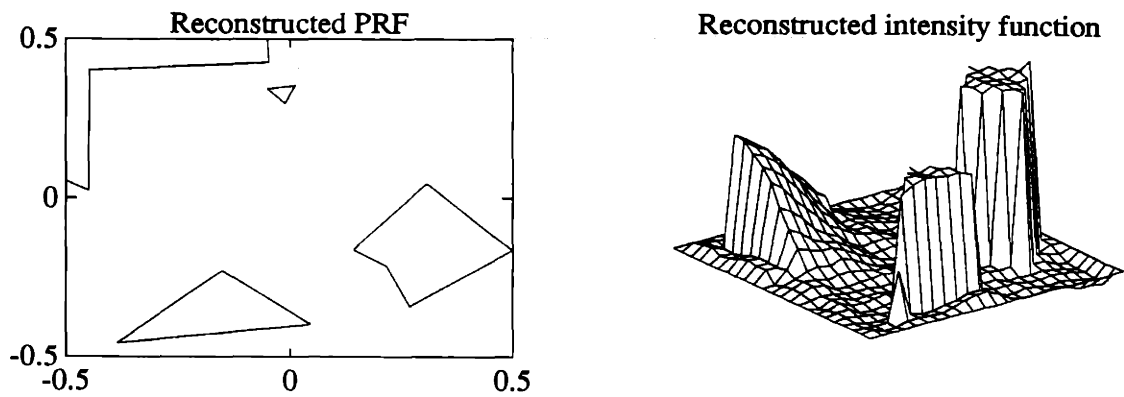


Figure 4.10 shows a sample reconstructed PRF coloring and intensity function when  $\sigma^2 = 0.7$ . In this case, the reconstructed coloring includes the edges of the initial

function, except for the low edge of the sloping triangle. From the measurements, it is clear that the low edge of the sloping triangle cannot be distinguished from the base because the measurement error is large relative to the height of the edge. In the reconstructed intensity function, the upward slope of the triangle is still clear; it simply begins from the region's base.

Figure 4.10: Reconstructed PRF and intensity function when  $\sigma^2 = 0.7$ .



# Bibliography

- [1] T. Arak, "On Markovian Random Fields with a Finite Number of Values," *4th USSR-Japan Symposium on Probability Theory and Mathematical Statistics. Abstracts and Communications*, 1982.
- [2] T. Arak and D. Surgailis, "Markov Fields with Polygonal Realizations," *Probability Theory and Related Fields*, 80, 543-579, 1989.
- [3] T. Arak and D. Surgailis, "Consistent Polygonal Fields," *Preprint to appear in Probability Theory and Related Fields*, 1991.
- [4] P. Clifford and R. D. Middleton, "Reconstruction of Polygonal Images," *Journal of Applied Statistics*, Vol. 16, No. 3, 409-422, 1989.
- [5] P. Clifford, "Markov Random Fields in Statistics," *Disorder in Physical Systems*, G. R. Grimmett and D. J. A. Welsh, Eds, 19-32, (Oxford University Press), 1990.
- [6] R. Gallager, *Discrete Stochastic Processes*, Lecture notes for 6.262, MIT, Fall 1990.
- [7] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741, 1984.
- [8] W. K. Hastings, "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 57, 97-109, 1970.

- [9] M. G. Kendall and P. A. P. Moran, *Geometrical Probability*, Charles Griffin, London, 1963.
- [10] R. Kinderman and J. L. Snell, *Markov Random Fields and Their Application*, Vol 1. American Mathematical Society, 1980.
- [11] J. L. Marroquin, "Probabilistic Solution of Inverse Problems," Ph.D. Thesis, Dept. of EECS, M.I.T., 1985.
- [12] S. K. Mitter and O. Zeitouni, "An SPDE Formulation for Image Segmentation," *Stochastic Partial Differential Equations and Applications*, Eds. G. Da Prato and L. Tubaro, Longman Scientific and Technical, 257-267, 1992.
- [13] D. Mumford and J. Shah, "Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems," *Comm. Pure and App. Math*, XLII, 577-685, 1989.
- [14] P. Switzer, "A Random Set Process in the Plane with a Markov Property," *Annals of Mathematical Statistics*, 36, 1859-1863, 1965.