M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 200
Appeared: Proc. 26th Annual Asilomar Conf. on Signals, Systems,and Computers, Pacific Grove, CA Oct. 92, pp 1011-1015.

Structured Patterns from Random Fields

Rosalind W. Picard
M.I.T. Media Laboratory
20 Ames Street
Cambridge, MA 02139

Abstract

We discuss the problem of modeling structured textures and image data with random field models. Random field models have recently been applied to a variety of problems such as image enhancement, segmentation, and coding. The models have been shown to successfully reproduce a variety of stochastic patterns, but have not been successful for capturing structure in patterns. Two possible solutions to the structured pattern problem are examined in this paper. The first solution is based on the structure present in the initial state of the random process, and the second is based on a relationship between "internal" random field interactions and coupling to a set of "external" constraints.

1 Introduction

Structure from randomness is a tantalizing process. The ability of a random field model to organize itself into any pattern, structured or random, would be beneficial to build into a practical model. Such a model could theoretically represent all kinds of image data, assisting a huge variety of applications in image understanding, coding, reconstruction, and enhancement.

To identify a good image model it helps to define the criteria for "good". Two of the most common criteria include: 1. the ability to faithfully reconstruct the data, and 2. parameter efficiency, i.e., the size of the model parameters compared to the size of the data. Note that in many applications it is not necessary that the model converge pointwise to the data to satisfy the first criterion. "Visual similarity" is often a sufficient criterion. Also, within the second criterion it is important to consider not just the number of parameters, but also the accuracy (bits of precision) to which they must be determined. Other criteria for a good model may include: 3. low computational complexity and 4. "semantic" efficiency. A semantically efficient model is one where pattern parameters take on attributes such as "orientation," "periodicity," or "randomness," making it easy for the layperson to interact with the signal through its parameters. Most physically-based models can be considered semantic, e.g., the temperature parameter to control regular structure formation in crystals.

2 Structured pattern modeling

To date, the most popular image "models" are orthogonal sets of sinusoidal basis functions, e.g., the discrete cosine transform (DCT), perhaps combined with a subband decomposition. These linear models satisfy the first three criteria nicely in smooth regions of an image. The fourth criterion is rarely satisfied, however. The second criteria also tends to break down where there is texture or inhomogeneous structure. In these regions, the number of parameters may grow significantly.

For modeling textured image data more efficiently, a variety of autoregressive or linear predictive models have been considered. However, the basic assumption of the autoregressive model, i.e., a continuous spectral density, is almost always violated when structure appears in the data, introducing singularities in the periodogram. The autoregressive model is theoretically inappropriate for textures which are nonhomogeneous or nonstochastic, i.e., where the spectral density does not exist (strictly speaking) due to non-stationarity or due to the presence of singular components.

By contrast, the Markov random field model can theoretically synthesize any pattern, structured or stochastic, with nonzero probability. Cross and Jain [1] showed that this model is capable of synthesizing a variety of stochastic-looking textures. However, in studies applying this model to structured patterns, the resulting synthesized patterns were not visually similar to the originals [1, 2]. The structural components were not reconstructed. The use of only pairwise statistics in the model is one limitation; better results are to be expected with higher order statistics [3]. However, it is generally impractical to get an accurate estimate of the statistics as their order is increased. In general, it is hard to find a model which represent both structure and randomness efficiently.

A brief discussion on the definition of "structure" is apropos before proceeding. In 1D, it is easy to imagine three cases as illustrated by the following character strings:

sh ysSTehtignaSio m

CellCellCellCelllCell

This Says Something

The first string represents a sample of a stochastic process. It has relatively low information content and is analogous in 2D to a random texture pattern. The second string also has low information content; it is specified by a primitive element, a replication rule, and perhaps a noisy perturbation. Its structure is the harmonic components of a deterministic random field. In 2D its analogue is a periodic texture. Combinations of these first two strings, in higher dimensions, e.g., periodic in one direction and random in the other like a plowed field, periodic in both like a checkerboard, or random in both like a pile of leaves, span the basic categories typically used for texture classification [4].

The third string is quite different from the first two strings. Although an anagram of the first string, the third string has a *specific* structure. Consequently, it has the most information content of the three cases. Its 2D analogue is typically an object with carefully arranged components, e.g. a building or a chair.

These three cases and everything in between them may occur in images. The first two are generally considered "textures", characterized by collective properties, and hence full of redundancy. The third tends to be classified with the "objects" or items where specificity of arrangement is important. The latter two can both be considered "structured". It is both of these kinds of structure that are addressed in this paper.

3 Gibbs random fields

The random field model considered here is the discrete Gibbs random field (GRF). There exists a now well-known equivalence between certain Gibbs and Markov models. We focus on the Gibbs which is technically more accurate when the histogram of the image is constrained.

The following assumptions are made in the definition of the GRF. Let an image be represented by a finite rectangular $M \times N$ lattice $\mathcal{S}$ with a neighborhood structure $\mathbf{N} = \{\mathcal{N}_s, s \in \mathcal{S}\}$ where $\mathcal{N}_s \subseteq \mathcal{S}$ is the set of sites which are neighbors of the site $s \in \mathcal{S}$. Every site has a nonnegative graylevel value, $x_s \in \Lambda = \{0, 1, \ldots, n-1\}$. Let $\mathbf{x}$ be the vector $(x_s, 1 \leq s \leq |\mathcal{S}|)$ of site graylevel values and Ω be the set of all configurations taken by $\mathbf{x}$. A neighborhood structure is *symmetric* if $\forall s, r \in \mathcal{S}$, $s \in \mathcal{N}_r$ if and only if $r \in \mathcal{N}_s$.

For the finite periodic lattice $\mathcal{S}$ with a symmetric neighborhood structure one can define a Gibbs energy. There are many ways to define the energy; the one used in this paper is the *autobinomial* energy originally defined by Besag [5], and used for texture by Cross and Jain [1] and Garand and Weinman [6]. The autobinomial energy is

$$E(\mathbf{x}) = -\sum_{s \in \mathcal{S}} \left(\alpha_s x_s + \sum_{r \in \mathcal{N}_s} \beta_{sr} x_s x_r \right), \qquad (1)$$

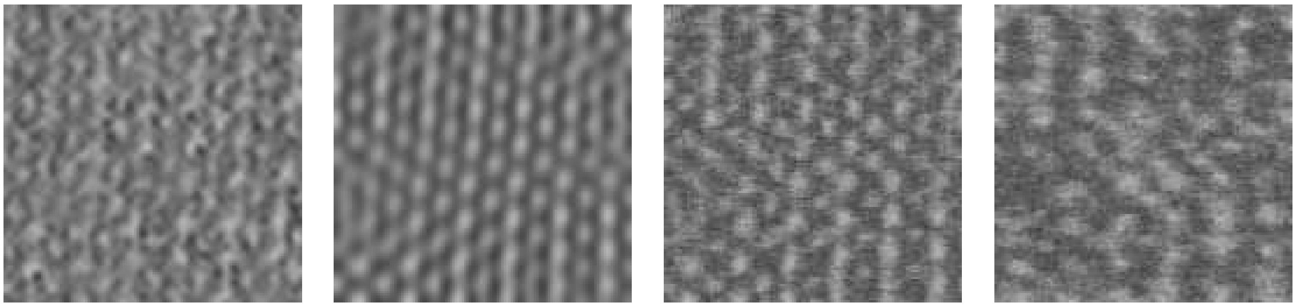
where the model parameters are α_s, the possibly spatial varying *external field*, and β_{sr}, the possibly spatial varying bonding parameters for the *internal field*. In this paper the internal field is always homogeneous so that $\beta_{sr} = \beta_r$. The homogeneous internal field may also be isotropic, in which case $\beta_r = \beta$. A joint probability distribution is assigned to the Gibbs energy yielding the *Gibbs random field*,

$$P(\mathbf{x}) = \frac{1}{Z} \exp\left(-\frac{1}{T} E(\mathbf{x}) \right), \qquad (2)$$

where Z is a positive normalizing constant known in the physics literature as a *partition function* and T is the "temperature" of the field. A GRF image pattern is a sample from this probability distribution.

The Gibbs distribution is easily incorporated into a Bayesian framework, and frequently used as an image model in the context of maximum a posteriori (MAP) estimation [7]. In this context, one may be interested in finding samples from this probability distribution which maximize the probability $P(\mathbf{x})$, consequently minimizing the Gibbs energy.

To minimize the energy, Monte Carlo stochastic relaxation algorithms such as the Metropolis exchange [1, 6] are often applied. These algorithms are iterative, making typically small changes at each iteration in an effort to minimize the global energy function. An iteration of the Metropolis exchange algorithm swaps pairs of elements in the current state (image) according to a probabilistic rule. Unlike gradient descent methods, stochastic relaxation permits small increases in the energy at any given iteration. In the Metropolis exchange, since the new state is formed by swapping elements in the previous state, the histogram of the initial image is also preserved. Samples taken after



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

Figure 1: All these images are 128×128 with 120 graylevels. Image (a) is the original fabric texture. Image (b) is a reconstruction of (a) keeping the largest 16 components from its DCT and rescaling linearly to the original dynamic range. Images (c) and (d) were synthesized using a GRF with internal field parameters estimated from (a), and initial state image of (b). Image (c) ran for 50 iterations, and image (d) for 225 iterations, both at temperature $T = 17$.

only a small number of iterations usually correspond to patterns from non-minimum energy configurations of the distribution.

4 Structure from initial state

The first method we consider for incorporating structure into the random field is based on putting structure into the initial state of the model. Garand and Weinman [6] showed examples where this approach worked well for images of marine stratocumulus clouds.

Let $F(\mathbf{y})$ be the discrete Fourier transform (DFT) of the image to be modeled, $\mathbf{y}$. Garand and Weinman found that the $N_c = 6$ largest components of $|F(\mathbf{y})|^2$ corresponded to the frequencies which provided the predominant structure in their cloud images. Inverse transforming the magnitude and phase components at these six locations and scaling to approximate the original image histogram resulted in a new image, $\hat{\mathbf{y}}$. This image was then used as the initial state for the GRF texture synthesis.

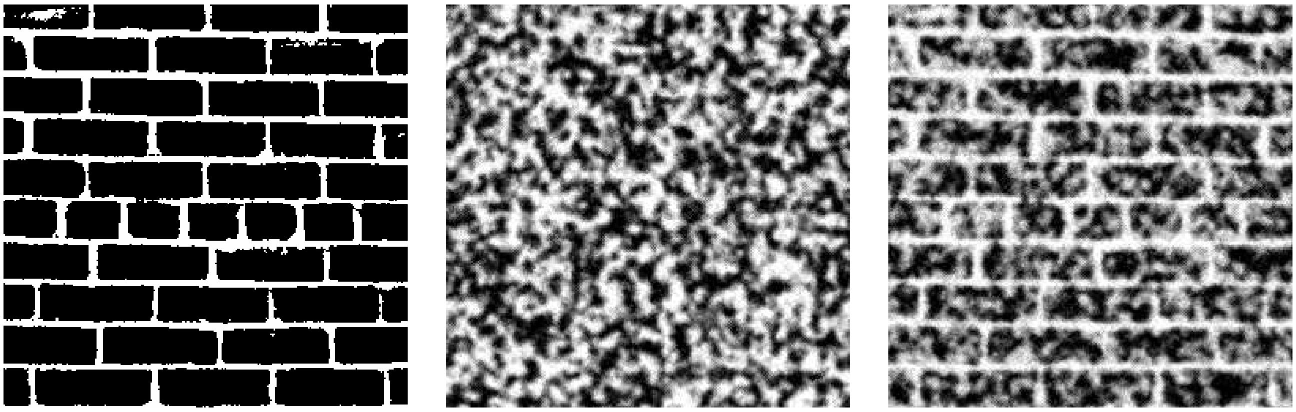
From a signal processing standpoint, this algorithm corresponds to abrupt filtering which one might expect to cause ringing. It also provides no framework to estimate the (data-dependent) choice of N_c, the number of coefficients to keep, nor does it separate out this component before estimating the GRF parameters. In practice, however, their results were visually quite similar to the original textured cloud images. Also, the method was efficient with respect to the number of parameters, and somewhat semantic for identifying properties of clouds.

We show an example of using an algorithm similar to theirs in Figure 1. The primary difference is we use the DCT instead of the DFT, avoiding the need for a prefiltering window. One can see the effect of the GRF on the initial image is one of slow random modification by the parameters, $\beta_{11} = \beta_{12} = .52$, $\beta_{21} = \beta_{22} = -.27$, corresponding to the west, north, northwest, and northeast neighbors respectively.

There is a basic theoretical problem [7] with this method of putting structure in the random field. In theory, the final state of the GRF should be independent of the initial state. Clearly, this is not the case here, as the structure in the initial state is intentionally preserved. Problems will arise with this method if careful attention is not made to the temperature, T, of the model during the Metropolis exchange. The Metropolis exchange method swaps the graylevels at two sites with probability

$$\min(1, \exp(\frac{-\Delta E}{T})),$$

where $\Delta E = E(\mathbf{y}) - E(\mathbf{x})$, $\mathbf{x}$ is the initial configuration, and $\mathbf{y}$ is the configuration with the elements at two sites swapped. Consider the probability of a swap



(a) (b) (c)

Figure 2: Image (a) is a 256×256 binary image made from a photo of a brick wall. Image (b) is a 256×256 $n = 16$ graylevel GRF with internal field parameters $\beta_{11} = \beta_{12} = 1.0$ and $\beta_{21} = \beta_{22} = 2.0$ and no external field. Image (c) is the result of using (a) as the external field and (b) as the internal field. The result in (c) is a GRF with structure.

when some $T_2 > T_1$:

$$\min(1, \exp(\frac{-\Delta E}{T_1})) \leq \min(1, \exp(\frac{-\Delta E}{T_2})).$$

Thus, the higher T, the higher the probability of swapping. When $T \rightarrow \infty$ then all pixels swap, and since there is no restriction on their range, the pixels in the image become completely "mixed", as if someone had stirred them into a random solution. Hence, the structure in the initial image is lost depending on T. Furthermore, one can show that the behavior of increasing T is approximately the same as the behavior of decreasing the magnitude of the β parameters or decreasing the number of graylevels [8]. For certain combined ranges of these values, the effect will be that of low temperature, which will preserve initial structure for a long time. Garand and Weinman ran the Metropolis Exchange for approximately 12 iterations. Over so few iterations, and over their relatively low temperature conditions, the initial state structure is preserved.

Thus, one practical problem with this model is the careful monitoring required to ensure that the initial structure is preserved. By empirical study of temperature dependent "phase transition" behavior [8], one can determine the temperature ranges (or temperature annealing schedules) that are appropriate. For non-isotropic parameters, however, this can be a lot of work. In general, the success of this method depends on fine-tuning for a particular class of data.

5 Structure from external field

The second method for putting structure into a GRF involves the external field term in the energy function. Recall that the external field is specified by the α_s's in (1). When the Metropolis exchange algorithm is used, and when $\alpha_s = \alpha$, i.e., the external field is homogeneous, then one can show that the external field term cancels and hence has no effect [6]. This is the case in the texture study of Cross and Jain [1], so that their results, which did not achieve any structured patterns, can be attributed to interactions of the internal field β parameters only.

In applications of the GRF used with MAP estimation it is useful to consider a non-homogeneous external field. Consider again the energy function of (1), but now with a homogeneous internal field. Furthermore, let us set the nonhomogeneous external field to the values in some reference image, $\mathbf{y}$. Hence the value of $\alpha_s = y_s$, the graylevel value at site s in the image $\mathbf{y}$. This yields an energy of:

$$E(\mathbf{x}) = -\sum_{s \in \mathcal{S}} \left(y_s x_s + \sum_{r \in \mathcal{N}_s} \beta_r x_s x_r \right), \qquad (3)$$

Although one can synthesize samples from any energy range of the Gibbs distribution, the most probable samples correspond to those with the least energy. Consider the "attractive" case where the parameters β_r are nonnegative. In this case, the energy is minimized when the products $y_s x_s$ and $\sum_{r \in \mathcal{N}_s} x_s x_r$ are maximized. If we let $\mathbf{P}$ be a permutation matrix, then the first product can be rewritten in the form $\mathbf{y}'\mathbf{P}\mathbf{x}$. For some $\mathbf{P}$, i.e., some configuration of the GRF $\mathbf{x}$, this product will be maximized. If the images $\mathbf{y}$ and $\mathbf{x}$ have the same first order statistics, then this product will be maximized when $\mathbf{y} = \mathbf{P}\mathbf{x}$. If not for competition from the internal field product, the synthesized random field would align itself perfectly with the desired external field.

The internal field product term, $\sum_{r \in \mathcal{N}_s} x_s x_r$ has been shown elsewhere [9] to be maximized when the graylevel sets in the image form configurations which maximize their "separation", analogous to the separation of immiscible fluids. The product $x_s x_r$ is maximized when the same graylevels occur as neighbors. Hence patterns which minimize the number of neighbors that have a different graylevel are most likely to form. Thus, the minimum energy internal field will have minimal length boundaries between pairs of graylevel sets. Moreover, a region of constant graylevel, g, will tend to be surrounded by regions of constant graylevel $g-1$ or $g+1$. The second-order statistics for such patterns can be quite restricted [9].

In Figure 2 an example is shown using an external field to introduce structure. Here a balance is maintained between the external and internal fields so that both of them contribute to the final pattern. This method also tends to be efficient in its use of parameters, but it is still unknown how best to estimate the relative strengths of the two fields.

6 Summary and conclusions

This paper has discussed the problem of incorporating structure into a random field model. Two methods were described for augmenting the ability of the autobinomial GRF for producing structured patterns. The first method uses the dominant components from a deterministic sinusoidal expansion as the initial state. This method exploits the efficiency of a deterministic basis for representing the structured part of the signal, but will fail to preserve the structure if the random field is allowed to run to equilibrium, eventually becoming independent of the initial state. The second method is better motivated theoretically, but it is still unknown how to best balance the parameters so that the competition between the two fields reaches an equilibrium at the desired pattern.

It is important to point out that although in the second case the structure is actually synthesized by the random swapping process, it is not "emerging from the randomness." Rather, it is encoded in a pre-designed external field.

One potentially promising alternate method for incorporating structure is to decompose the random field into its mutually orthogonal *stochastic*, *deterministic*, and *evanescent* components using a Wold-like decomposition as by Francos, et al. [10]. The stochastic component is then represented with an autoregressive model, and the other two components are represented by some kind of structural model. The structural model of Francos, et al. consists of sums of delta functions. Currently, however, the Wold decomposition only works for a homogeneous random field. It also does not provide an efficient way to represent the non-stochastic structural components. We are continuing research along these lines now.

References

[1] G. R. Cross and A. K. Jain. Markov random field texture models. *IEEE T. Patt. Analy. and Mach. Intell.*, PAMI-5(1):25–39, 1983.

[2] R. Chellappa. "Two-Dimensional Discrete Gaussian Markov Random Field Models For Image Processing". In L. N. Kanal and A. Rosenfeld, editors, *Progress In Pattern Recognition 2*. Elsevier Science publishers B. V. (North Holland), 1985.

[3] A. Gagalowicz and S. D. Ma. Sequential synthesis of natural textures. *Comp. Vis., Graph., and Img. Proc.*, 30:289–315, 1985.

[4] R. Haralick. Statistical and structural approaches to texture. *Proc. IEEE*, 67:786–804, May 1979.

[5] J. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *J. Roy. Stat. Soc., Ser. B*, 36:192–236, 1974.

[6] L. Garand and J. A. Weinman. A structural-stochastic model for the analysis and synthesis of cloud images. *J. of Climate and Appl. Meteorology*, 25:1052–1068, 1986.

[7] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE T. Patt. Analy. and Mach. Intell.*, PAMI-6(6):721–741, 1984.

[8] R. W. Picard. Gibbs random fields: Temperature and parameter analysis. In *Proc. ICASSP*, pages III.45–III.48, San Francisco, CA, 1992.

[9] R. W. Picard and I. M. Elfadel. Structure of aura and co-occurrence matrices for the Gibbs texture model. *J. of Mathematical Imaging and Vision*, 2:5–25, 1992.

[10] J. M. Francos, D. P. de Garrido, and J. W. Woods. On the parameter estimation of the harmonic evanescent and purely indeterministic components of homogeneous random fields. In *Proc. ICASSP*, pages III.41–III.44, San Francisco, CA, 1992.