Mean field approximation for PDE-Markov random field models in image analysis

S. ZIMERAS
University of the Aegean
Department of Statistics and Actuarial-Financial Mathematics
G.R.832 00, Karlovassi, Samos,
GREECE

*Abstract***:** Markov random fields (M.R.F.) on a lattice system and Gibbs distribution provide a wide area of models for interacting particle systems in image analysis, mechanical physics and statistical mechanics. Physical properties of the neighbors could be explained by partial differential equation (PDE) inside the potential function introducing PDE-MRF models. In image analysis, they have been used to describe the local characteristics of the spatial interaction between pixels. Although, in model image reconstruction, a number of fundamental issues remain unexplored, such as the specification of M.R.F. models, performance evaluation of the neighborhood structure of these models, and the phase transition phenomenon. In this work, spatial behavior of the auto-exponential model in a rectangular lattice would be investigated, concentrating on the first-order neighborhood structures. A simple deterministic model based on a univariate iterative scheme is studied which predicts the properties of these models and realizations have been generating using the Gibbs sampler to illustrate the properties. For well defined regions in the parameter space this iterative scheme is unstable leading to catastrophic and 2-cycle behavior.

Keywords: Markov random fields, Image reconstruction, Auto-exponential model, Gibbs sampler, texture analysis, phase transition, partial differential equations, mean field.

1 Introduction

Stochastic models and statistical methods have been successfully applied in image processing and pattern recognition. Particularly, interesting are Bayesian methods based on local characteristics. Key components of any statistical analysis are the choice of an appropriate model as the prior and the estimation of the prior model parameter. The Bayesian approach to reconstruction in spatial processes involves the modeling of prior information in addition to a component describing the degradation process. This prior components, describe local characteristics of the spatial process, which is often modeled using a Markov random field [5]. However, this prior component usually involves unknown prior parameters which control the influence of the prior distribution. In many applications, appropriate values for these parameters will be found by *trial-and-error*, in other cases a fully Bayesian approach will be adopted and the prior parameters estimated in the same way as other model parameters. However, in all these cases it is expected that the procedures depend *smoothly* on the prior parameters and that there is a unique relationship between the parameters and different types of behavior of the process introducing influence of partial differential operators (pde-MRF models) The foundations of Markov random fields lie in the physics literature on fermomagnetism originating in the work of [14]. Models based on Markov random fields can be used to extract statistical information about the structure of real word phenomena which could be represents in some sense as textures. The main goal of texture analysis is to extract useful textural information from an image using statistical tools like Markov random fields; [6, 18, 19] used these models to generate texture images from the auto-binomial model, [22] investigated the Markov random field models defining general properties using the definitions of random fields and Gibbs random fields. A particular subclass of Markov random fields is the auto-models, introduced in [5] and further studied in [4, 6, 1, 26, 27]. In practice estimates using Bayesian methods cannot be computed analytically. For this reason Monte Carlo algorithms can be used to

generate samples from the posterior distribution and parameter estimates calculated from this sample. The fundamental idea is to use an algorithm, which generates a discrete time Markov chain converging to the desired distribution. The most commonly algorithms include the Gibbs sampler [9] and the Metropolis-Hasting algorithm [10, 14]. Applications of these methods cover a wide range of areas including: detection of lesions in medical imaging [2]); astronomy [20], medical biology [8, 27], medical imaging in SPECT [3, 11, 24, 26, 27]. In this work, it would be shown that for one particular group of models the behavior is far from being straight forward and smoothly dependent on prior parameters. A simple deterministic iterative scheme would be considered which can be used to predict the behavior of the spatial process. We shall look in detail at the corresponding equations for the auto-exponential model, which have a various applications in biology and medicine. A relationship between the prior parameters would be obtained, which shed light on the choice of parameters used by [6] to simulate texture images using auto-exponential model. Finally, the Gibbs sampler would be used to simulate example images for various parameter combinations which demonstrate the behavior predicted by the simple iterative scheme.

2 Markov random fields modeling

Markov random fields are often useful models for a sampled image with local dependencies. The most important characteristic of these models is that the global representation of the image can be formed as a stochastic propagation of interactions defining a local neighborhood structure. In this section basic definitions and notations are given which were introduced by [1, 4, 5, 6]. The reader is directed to these papers for further details.

Suppose that we have a two-dimensional space, S, which has been partitioned into n pixels, labeled by the integers Λ_i={1, 2,..., n} defined as state space. Each pixel variable can be discrete (finite or infinite) or continuous. We refer to the possible values of the pixel variable as intensities. Let X_i be a random variable assigned to a pixel x_i and **X** = (X_1, X_2,...,X_n), be a random variable vector for an image, then a intensity **x** = (x_1, x_2,..., x_n), with $x_i \in \Lambda$ can be assumed as a realization of **X**. We use the function $p(x_i|...)$ to denote the conditional probability distribution (or density function) of X_i given the other arguments. We shall be considering a 2-D rectangular lattice, each site i of which has a random variable x_i. The configuration space for the variable **x** is denoted by Ω with: $\Omega = \prod_{i \in S} \Lambda_i$. For simplicity each pixel of the finite lattice can takes colors from the set Λ= {0, 1, 2, ..., c-1}. To define a Markov random field, a neighborhood structure $\mathcal{N}$ is needed, which defines the range of interaction from one pixel to another. A neighborhood system $\mathcal{N}$={N_i, $\forall$ i$\in$ S} is a collection of subsets of S for which: (i) i$\notin$ N_i (a site is not part of its neighborhood) and (ii) j $\in$ $N_i \Leftrightarrow$ i $\in$ N_j (i is in the neighborhood of j if and only if j is in the neighborhood of i). In general, $\forall$ s$\in$ S: s=(i,j), an n^{th} order homogeneous neighborhood system could be defined as $N^n = \{N_{(i,j)} : (i,j) \in S\}$ and $N^n_{(i,j)} = \{(k,l) \in S : (k-i)^2 + (l-j)^2 \le n\}$ .Obviously, sites near the boundary have fewer neighbors than interior ones (free boundary condition). Furthermore $N^0 \equiv S$ for all n$\geq$0: $N^n \subset N^{n+1}$.Considering conditional probabilities on lattices, site j is said to be a neighbor of site i ($\neq$j), if and only if, the functional form of the conditional probability distribution of X_i given all other pixel values depends upon x_j [4]. The form of neighbors for each site can be defined as $p(x_i \mid x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) = p(x_i \mid x_{\partial i})$,where ∂i is the set of pixels which are neighbours of site i, and $x_{\partial i}$ the set of values of pixels which are neighbours of pixel x_i. The usual neighbourhood system in image analysis defines the first-order neighbours of a pixel as the four pixels sharing a side with the given pixel (Fig. 1). Higher order neighbors are defined in an analogous manner.

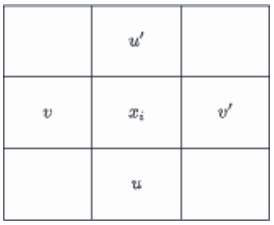

Fig.1. First order neighbours

A clique C is a subset of S for which every pair of sites is neighbors. Single pixels are also considering cliques. The set of all cliques on a lattice is called *C*. Cliques are important when considering the equivalence between MRFs and the Gibbs distribution. Fig. 2 illustrates the set of cliques for the first order or nearest-neighbors system with form: {(i, j)}, {(i-1,j), (i, j)}, {(i,j-1), (i, j)} (Fig. 2).

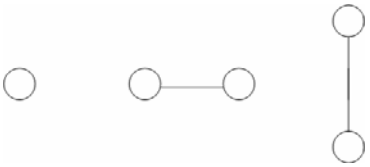

Fig. 2. Clique type for first order neighbors system

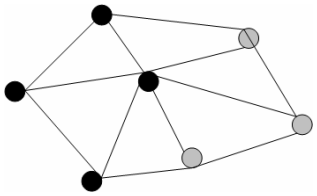

Fig.3. Independent graph associated to an image partition

A random field, with respect to a neighborhood structure is a Markov random field if the joint probability density on the set of all possible intensities values **x** satisfies the following properties: (i) p(**X**)>0 for all **X**, (ii) p(all pixels in the lattice except x_i)=p(X_i| neighbors of x_i). According to the Hammersley-Clifford theorem [13, 4, 5], an MRF can equivalently be characterized by a Gibbs distribution. Thus the joint probability is given by $p(x)=\frac{1}{Z}\exp\{-U(\mathbf{x})\}$, where Z is the normalised constant or partition function and U(x) is the energy function with form $U(\mathbf{x})=\sum_{c\in C}V_c(\mathbf{x})$ with the summation is over the local clique potentials set $V_c(\mathbf{x})$ over all possible cliques C [9]. We make the additional assumption that the MRF is homogeneous; i. e., the potential function is the same for all cliques. This property gives rise to translation-invariance of an MRF model. Without loss of generality we assume the maximal cliques in the MRF are square pixel patches of a fixed size; other, non-square, neighborhoods could be used. Configurations of low energies are the more likely, whereas high energies correspond to low probabilities. The interaction structure induced by the factorized form ($p(x)=\prod_{c\in C}f_c(x_c)$, where $f_c(x_c)$ depends only on the variable subset $x_c=\{x_i, i\in c\}$ and $\prod_{c\in C}f_c(x_c)$ is summable over Ω) presented by an independent graph: the independence graph associated with the factorization is the undirected graph G=[E,S] with vertex set S={1,2,…,n} and edge set E defined as: $\{i,j\}\in E\Leftrightarrow\exists c\in C:\{i,j\}\subset c$. As a consequence of the definition, any subset c is either a singleton or composed of mutually neighbouring sites: *C* is a set of cliques for G. Fig. 3 represent an independent graph associated with partitions.

Energy function U(x) can be written in terms of clique potentials which accounts for the contribution to the energy from the cliques.

$$U(\mathbf{x})=\sum_{c\in C}V_C(x_c)=$$
$$=\sum_{1\le i\le S}x_iG_i(x_i)+\sum_{i\le i<j\le S}\sum x_ix_jG_{i,j}(x_i,x_j)$$

G_i (.) is arbitrary functions. The above models are called ***auto-models*** [5]. The family of auto-models can now be generated using the following assumption [5]: the probability structure only depends on contributions from sites taken either singularly or in pairs and the conditional probability distribution is a member of the regular exponential family of distributions,

$p(x_i\mid x_{\partial i})=\exp\{A_i(\omega_i)B_i(x_i)+C_i(x_i)+D_i(\omega_i)\}$ where ω_i is a model parameter associated with site i and is a function of the values at sites neighbouring site i. As a direct consequence of the above assumptions, A_i must satisfy, $A_i(\omega_i)=\alpha_i+\sum\beta_{ij}B_j(x_j)$, where $\beta_{ij}=\beta_{ji}$ and $\beta_{ij}=0$ unless sites i and j are neighbours. The functions B_j are linear in x_j with the exception for the auto-exponential model where B_j is equal to $-x_j$. With B_j is equal to x_j the equation becomes, $A_i(\omega_i)=\alpha_i+\sum\beta_{ij}x_j$ .In this paper only homogeneous schemes are considered, so that the dependency of α_i and β_{ij} on i is removed. Also toroidal boundary conditions is assumed. For the first-order neighbourhood system introduced above, the equation is given by $A_i(\omega_i)=\alpha+\beta_1(u+u^*)+\beta_2(v+v^*)$, where α and β_1 and β_2 are parameters of the model. If $\beta_1=\beta_2=\beta$ we have isotropic model, otherewise we have anisotropic model.

Definition: Suppose that the conditional distribution of x_i has an exponential distribution with mean λ_i dependent on the values of the neighbors. Then the joint probability distribution p(x) is auto-exponential and $A_i(\lambda_i)=1/\lambda_i$. The conditional probability for the auto-exponential model, is given

by: $p(x_i \mid x_{\partial i}) = \frac{1}{\lambda_i}\exp\left(-\frac{1}{\lambda_i}x_i\right) = \exp\left\{-\frac{1}{\lambda_i}x_i - \log(\lambda_i)\right\}$

where $A_i(\lambda_i) = \frac{1}{\lambda_i} \Rightarrow \lambda_i = \frac{1}{(a_i + \sum \beta_{ij} x_i)}$

3 PDE-Markov random fields models

Physical properties of the neighbors, direct us to the assumption of smoothness and homogeneity implementing various Markov random fields models as prior models. The statistical methodology in image analysis allows us to analyze priors that allow moderate discontinuities in the images. Based on the consideration of smoothness-prior assumption for image textures, a suitable form of energy function can be considered explained by partial differential equation (PDE) inside the potential function. Modification of the smoothens assumption could be introduced using total variation of the image. The total variation of a function f, denoted by TV (f), is defined as: $TV(f) = \int |\nabla f(x)| dx$. A function is said to have bounded variation if TV(f)<∞. The TV prior is defined as $p(x) \approx \exp\{-U(x)\} \approx \exp\{-TV(f)\}$. If it is assumed that u-u* (or v-v*) is spatially changing, then interaction between neighboring pixels may be approximated by first-order differential operator with general form: $V_{pde} = \int_{c \in C}\left[f_1 \frac{\partial X}{\partial u} + f_2 \frac{\partial X}{\partial v}\right] du dv$, where f_1 and f_2 are proposed functions, X is the random variable vector for an image and the partial derivatives calculate the interaction between neighboring pixels.

4 Mean field approximation

The mean field theory investigates the influence of the intensity x_i in the calculation of the expectation which can be approximate by the influence of $<x_i>$. By definition the expectation of the intensity can be calculated by: $< x_i > = \sum_{x_i} x_i p(\mathbf{x}) = \frac{1}{Z}\sum_{x_i} x_i \exp\{-U(\mathbf{x})\}$. [25] suggests that the mean field approximation of the above form can be given by: $< x_i > = \frac{1}{Z}\sum_{x_i} x_i \exp\{-U^{MF}(\mathbf{x})\}$, where inflence of the field can be approximated by the mean of the neighbours. In order to find that mean field, an iterative process can take place to solve the complexity of the computations. Recalling the pseudo-likelihood method [4, 26] for every auto-model the conditional expectation of a pixel value given its neighbors is proportional to the model parameter $E[x_i \mid x_{\partial i}] = c\omega_i$. So it is reasonable to replace the value of each neighbor by its conditional expectation (mean field theory, [7, p131]). Since we are generally interested in ω_i, we invert the equation A_i and define Z_i, $\omega_i = Z_i\left(\alpha_i + \sum \beta_{ij} x_j\right)$, where $Z_i = A_i^{-1}$. In this section an isoropic model is assumed, so the equation becomes $\omega_i = Z\{\alpha + \beta(u + u^* + v + v^*)\}$. Then using the mean-field approximation [7], that is assuming a uniform image, except for statistical variations this reduces the above equation to [1, 3, 26, 27] $\omega = Z\{\alpha + b\omega\}$ where $b = m\beta k$ with m denotes the number of neighbours. We could solve equation by setting up the corresponding iterative process $\omega_{r+1} = Z\{\alpha + b\omega_r\}$.

Any fixed point in the process is a solution of equation and a candidate of the parameters in an isotropic steady-state stochastic process. A necessary condition for this is the stability of the process. If we define $F(\omega) = Z\{\alpha + b\omega\}$ so the iterative process may be expressed as $\omega_{r+1} = F(\omega_r)$, then according to standard theory of iterative process, a fixed point ω_F is stable when $|F'(\omega_F)| < 1$ and unstable when $|F'(\omega_F)| > 1$. So for stability,

$$|F'(\omega_F)| = |bZ'(\alpha + b\omega)| < 1 \Rightarrow$$

$$|b| < \frac{1}{|Z'(\alpha + b\omega)|} = \frac{1}{|Z'(A(\omega))|}$$

Assuming isotropy and the mean field approximation equation produces the relationship $\alpha + b\omega - A(\omega) = 0$ which shows that the locus of the points for which ω takes the fixed constant values is a straight line in the $\alpha - b$ plane. With the α-axis horizontal and b-axis vertical the line has gradient $-\omega^{-1}$ and vertical intercept $A(\omega)\omega^{-1}$ [1]. By considering equality and since there is a relation between α, b and ω, the parametric equations for the boudary of the critical regions can be obtained. There are two parts, the upper half in the $\alpha - b$ plane [1, 3, 26, 27]

$$\alpha = A(\omega) - \frac{\omega}{|Z'(A(\omega))|}, \qquad b = \frac{1}{|Z'(A(\omega))|}$$

and in the lower half of the $\alpha - b$ plane [1, 3, 26, 27] $\alpha = A(\omega) + \frac{\omega}{|Z'(A(\omega))|}$, $b = -\frac{1}{|Z'(A(\omega))|}$.

For the auto-exponential model $A_i(\lambda_i) = 1/\lambda$, and making the usual assumptions a-b plane equation becomes $\alpha + b\lambda - 1/\lambda = 0$. The iterative scheme is stable when $|b| < 1/\lambda^2$. The parametric equations for the boundaries of the critical regions are in the upper half of the $\alpha - b$ plane: $\alpha = 0, b = 1/\lambda^2$ and in the lower half of the $\alpha - b$ plane: $\alpha = 2/\lambda, b = -1/\lambda^2$, illustrared in Fig. 4.

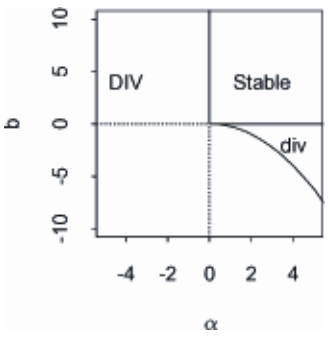

Fig. 4: Regions of district behaviour for the auto-exponential model. Vertical axis: values for a parameter; Horizontial axis: values for b parameter.

For parameter combinations in the region marked DIV in, the iterative scheme diverges for the auto-exponential model. Parameter combinations in the Stable quadrant produce unique stable fixed points in the iterative scheme. The observed images should appear smooth with variation due to exponential noise; the mean should be given by the unique fixed point of the iterative scheme. For parameter combinations in the region marked div, the iterative scheme has two fixed points. For initial values above the upper fixed point the iterative scheme diverges; otherwise it converges to the lower fixed point. However, statistical variation will cause divergence in the spatial process.

5. Image reconstruction algorithms

Equation $\omega = Z(a + b\omega)$ is a representetive of a set of n non-linear simultaneous equations in n-unknows because the values of the neighbours depend on the corresponding conditional distributions. It is not possible to solve thi set of equations analytically, so the ependency of ω, on a and b can be investigated using simulated realisations of the spatial process. Since the configuration space is extremely large (for a binary NxN MRF there is 2^M possible states where $M=N^2$ which for N of the order of 100 is $2^{10000} \approx 10^{3010}$). Markov chain Monte Carlo methods have been studied by various researchers as an alternative to exact probability computation in image analysis. The general method is to construct a Markov chain with the required probability as it equilibrium distribution. [21]. Then realisations of the Markov chain form a pseudo-sample from the required distribution. This pseudo-sample can then be used to estimate various statistical measures of the image. In particular our interest is concentrated on the distribution p(x) using a particular realisation $X^1, X^2 \ldots X^N$ on the Markov chain with transition probability $p(x \to x')$.Typically asymptotic results include:

$X^t \underset{t\to\infty}{\overset{d}{\to}} X \sim p(x)$; $\frac{1}{t}\sum_{i=1}^{t} f(x^i) \xrightarrow[t\to\infty]{} E_p\{f(x)\}$ where the expectation $E_p\{f(x)\}$ is to be estimated. The corresponding empirical average will be used namely: $\bar{f}_N = \frac{1}{N}\sum_{t+1}^{N} f\left(x^{(t)}\right)$. The Gibbs sampler [9], a particular MCMC method, is an iterative process for simulating a sample from a specified probability distribution. At each iteration, the value of only one pixel can change. Although the sequence in which the pixels are visited is arbitrary, it is common to visit them in order 1,2,…,n. The Gibbs sample is a general method for producing samples from a distribution. It is particularly useful when the distribution being sampled is a Gibbs distribution, and the resulting samples form a Markov random field. Let X_s be a finite dimensional random field that takes on values in a discrete and finite set Ω for all s∈S. If we assume that the distribution of X is strictly positive, then without loss of generality we know that the distribution of X can be written in the form $p(x) = \frac{1}{Z}\exp\{-U(\mathbf{x})\}$. The marginal distribution of a pixel can be written as $p(x_s \mid x_i, i \neq s) = \frac{\exp\{-U(x_s \mid x_i, i \neq s)\}}{\sum_{x_s' \in \Omega}\exp\{-U(x_s' \mid x_i, i \neq s)\}}$ We can generate samples from the distribution of p(x) by using the following Gibbs sampler algorithm:

1. Set N = number of pixels

2. Order the N pixels as N=s(0), s(1),…., s(N-1)
3. Repeat for k=0 to ∞
   Form $X^{(k+1)}$ from $X^{(k)}$ via

$$X_r^{(k+1)} = \begin{cases} W, \text{ if } r = s(k) \\ X_r^{(k)}, \text{ if } r \neq s(k) \end{cases}$$

where $W \sim p\left(x_{s(k)} \mid X_i^{(k)}, i \neq s(k)\right)$. This iterative scheme is continued to produce a sequence $\left(X^0, X^1, \ldots, X^t\right)$, which is realisation of a Markov chain with transition probabilities from x^t to x^{t+1} given by: $p\left(x^t \rightarrow x^{t+1}\right) = \prod_{l=1}^{n} p\left(x_l \mid x_j^{t+1}, j < l, x_l^t, j > t\right)$.

After an initial transient period, a realisation from the required probability distribution has occurred. In practice it is important to monitor convergence to the equilibrium distribution. For a discussion of convergence see [12, 23]. For the auto-exponential the valid parameter space is even further restricted, since the spatial scheme is stable only when a > 0 and b > 0. In accordance with the predictions of section 4 the stochastic process with any other parameters diverges. In figure 5(a) a = 0.2 and b = 0.8 and in figure 5(b) a = 2 and b = 10 which is further into the Stable region.

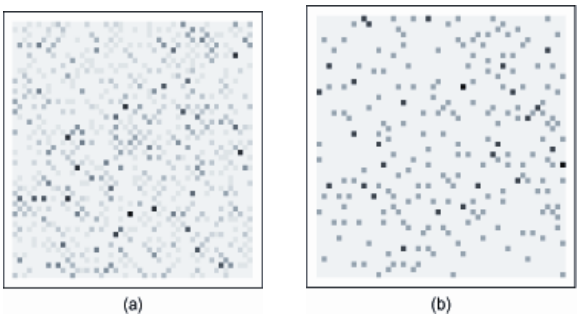

Fig. 5: Examples from the auto-exponential models: (a) α=0.2, b=0.8, (b) α=2, b=10

6. Conclusions

Stochastic models and statistical approaches have been successfully applied in various areas of image analysis. The main task is the choice of an appropriate model as the prior and estimation of model parameters. Because of the complexity of the problems direct estimation can not be used. For that reason estimates can be generated from Monte Carlo Markov chain methods based on samples from the posterior (or the prior) distributions. The most commonly used algorithms include the Gibbs sampler and the Metropolis-Hastings.

In this work we have shown that for the auto-exponential model parameter space can be divided into regions, each with distinct spatial behavior. Spatial properties of the auto-exponential model were studied and realizations from auto-exponential model have been generated using the Gibbs sampler. For appropriate combinations of models parameters different Markov random fields can be introduced and for certain combinations the behavior is not straightforward. Physical properties of the neighbors, smoothness and homogeneity introduce various Markov random fields models which could be explained by partial differential equation (PDE) inside the potential function. As a result, the parameter space can be divided into regions, each with distinct spatial behavior. The iterative procedure classifies each region as either stable or unstable leading to phase transition process. The behavior in the stable region is clear, one solution exists which is the expected value of the particular model. For the unstable region the behavior is more complicated and depends on the model being examined. Clearly this type of behavior is unacceptable if we wish to perform stable estimation.

Reference:

[1] Aykroyd R, Haigh J. G. B., Zimeras S. (1996): Unexpected spatial patterns in exponential family auto-models. *Graphical Models and Image Processing*, 58, 452-463.

[2] Aykroyd, R. G. and Green, P. J. (1991). Global and local priors and the location of lesions using gamma-camera imagery. *Phil. Trans. R. Soc. Lond. A*, 337, 323-342.

[3] Aykroyd R. G., Zimeras S. (1999): Inhomogeneous prior models for image reconstruction, *Journal of American Statistical Association (JASA)*,Vol 94, No 447, 934-946.

[4] Besag, J. (1986): On the statistical analysis of dirty pictures (with discussion). *J.R. Statist. Soc. B*, 48, 259-302.

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

[6] Cross, G. R. and Jain, A. K. (1983): Markov Random Field Texture Models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 5(1), 25-39.

[7] Chandler, D. (1978): *Introduction to Modern Statistical Mechanics*, Oxford University Press: New York.

[8] Diggle P. J. (1983): *Statistical Analysis of Spatial Pattern Point*, Academic Press, London.

[9] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6, 721-741.

[10] Hastings, W. K. (1970). Monte Carlo simulation methods using Markov chains, and their applications. *Biometrika*, 57, 97-109.

[11] Green P. J. (1990): Bayesian reconstructions from emission tomography data using a modified EM algorithm, *IEEE Trans Medical Imaging* 9, 84-93.

[12] Green P. J. and Han X. L. (1992): Metropolis methods, Gaussian proposals and antithetic variables, *Lect. Notes Statistics*, 74, 142-16.

[13] Hamersley J. A. and Clifford P. (1971): *Markov fields on finite graphs and lattices*, unpublished work

[14] Ising E. (1925): Beitrag zur Theorie des Ferromagnetismus *Zeitschrift fr Physik*, 31 253-258.

[15] Kindermann R. and Snell J. L. (1980): *Markov random fields and their applications*, American Mathematical Society.

[16] Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., Teller E. (1953): Equations of state calculations by fast computing machines, *J. Chemical Physics*, 21, 1087-1091.

[17] Poston, T. and Stewart, I. (1978): *Catastrophe theory and its applications*, Pitman: London.

[18] Qian W. and Titterington D. M. (1991): Multidimensional Markov chain model for image texture, *J. R. Stat. Soc. B* 53, 661-674.

[19] Ranngarajan A. and Chellappa Rama (1995): *Markov random fields models in image processing, the handbook of Brain Theory and Neural Networks*, ed. M. Arbib, 564-567, MIT Press.

[20] Ripley, B. D. and Sutherland, A. I. (1990). Finding spiral structures in images of galaxies. *Phil. Trans. R. Soc. Lond. A*, 332, 477-485.

[21] Smith A. F. M. and Robert G. O. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, *J. R. Stat. Soc. B*,55, 3-23.

[22] Spitzer F. (1971): Markov random fields and Gibbs ensembles, *Amer. Math. Monthly*, 78, 142-154.

[23] Sokal A. D.(1989): *Monte Carlo methods in statistical mechanics: foundation and new algorithms*, Cours de Troisieme Cycle de la Physique en Suisse Romande, Lausanne.

[24] Weir I. S. (1993): *Statistical modeling and reconstructions in single photons emission computed tomography*, Ph.D. Thesis, Bristol University.

[25] Zhang Jun (1992): The mean field theory in EM procedures for Markov random fields, *IEEE Trans. On Signal Processing*, Vol 40, No 10, 2570- 2583.

[26] Zimeras S. (1997): *Statistical models in medical image analysis*, Ph. D. Thesis, Leeds University.

[27] Zimeras S. and Georgiakodis F. (2005): Bayesian models for medical image biology using Monte Carlo Markov Chain techniques, *Mathematical and Computer Modeling,* 42, 2005, 759-768.

[28] Zimeras S. (2006): Simulating texture patterns using auto-logistic models, *WSEAS Transactions on Systems, Issue 10, Volume 5, 2269-2276.*