Markov Random Fields in Image Segmentation

Markov Random Fields in Image Segmentation

Zoltan Kato

Image Processing and Computer Graphics Dept.
University of Szeged
Szeged 6720
Hungary
kato@inf.u-szeged.hu

Josiane Zerubia

INRIA Sophia Antipolis-Mediterranee
Sophia Antipolis
06902 Cedex
France
Josiane.Zerubia@inria.fr

the essence of knowledge

Boston – Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is Z. Kato and J. Zerubia, Markov Random Fields in Image Segmentation, Foundations and Trends® in Signal Processing, vol 5, nos 1–2, pp 1–155, 2011

ISBN: 978-1-60198-588-0

Foundations and Trends® in Signal Processing
Volume 5 Issues 1–2, 2011

Editorial Board

Editorial Scope

Foundations and Trends® in Signal Processing will publish survey and tutorial articles on the foundations, algorithms, methods, and applications of signal processing including the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital and multirate signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
  - classification and detection
  - estimation and regression
  - tree-structured methods

Information for Librarians

Foundations and Trends® in Signal Processing, 2011, Volume 5, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Foundations and Trends® in
Signal Processing
Vol. 5, Nos. 1–2 (2011) 1–155

DOI: 10.1561/2000000035

Markov Random Fields in Image Segmentation

Zoltan Kato[1] and Josiane Zerubia[2]

[1] *Image Processing and Computer Graphics Dept., University of Szeged, Arpad ter 2, Szeged, 6720, Hungary, kato@inf.u-szeged.hu*

[2] *INRIA Sophia Antipolis-Mediterranee, 2004 Route des Lucioles, Sophia Antipolis, 06902 Cedex, France, Josiane.Zerubia@inria.fr*

Abstract

This monograph gives an introduction to the fundamentals of Markovian modeling in image segmentation as well as a brief overview of recent advances in the field. Segmentation is considered in a common framework, called image labeling, where the problem is reduced to assigning labels to pixels. In a probabilistic approach, label dependencies are modeled by Markov random fields (MRF) and an optimal labeling is determined by Bayesian estimation, in particular maximum a posteriori (MAP) estimation. The main advantage of MRF models is that prior information can be imposed locally through clique potentials. The primary goal is to demonstrate the basic steps to construct an easily applicable MRF segmentation model and further develop its multiscale and hierarchical implementations as well as their combination in a multilayer model. MRF models usually yield a non-convex energy function. The minimization of this function is crucial in order to find the most likely segmentation according to the MRF model. Besides classical optimization algorithms, like simulated annealing or

deterministic relaxation, we also present recently introduced graph cut-based algorithms. We briefly discuss the possible parallelization techniques of simulated annealing, which allows efficient implementation on, e.g., GPU hardware without compromising convergence properties of the algorithms. While the main focus of this monograph is on generic model construction and related energy minimization methods, many sample applications are also presented to demonstrate the applicability of these models in real life problems such as remote sensing, biomedical imaging, change detection, and color- and motion-based segmentation. In real-life applications, parameter estimation is an important issue when implementing completely data-driven algorithms. Therefore some basic procedures, such as expectation-maximization, are also presented in the context of color image segmentation.

Note: A sample implementation of the most important segmentation algorithms is available in grey scale at http://dx.doi.org/10.1561/2000000035_demogray and in color at http://dx.doi.org/10.1561/2000000035_democolor.

Contents

Dedication

"To the memory of my mother" Zoltan Kato

"To the memory of my beloved sister Elise who passed away in August 2012" Josiane Zerubia

1

Introduction

An image processing system involves a sensing device (usually a camera) and computer algorithms to interpret the picture. The term *image* (more precisely, *monochrome image*) refers to a two-dimensional light intensity function whose value at any point is proportional to the brightness (*gray-level*) of the image at that point [70]. A *digital image* is a discretized image both in spatial coordinates and in brightness. It is usually represented as a two-dimensional matrix, the elements of such a digital array are called pixels. The digitized image is the starting point of any kind of computer analysis. In some applications, the sensing device may be more specific responding to other forms of light: infrared imaging, photon emission tomography, radar imaging [182], ultrasonic imaging, etc.

Many image processing tasks deal directly with raw pixel data involving image compression [2], restoration [35, 64, 91, 219, 220, 223], edge detection [65, 200, 219, 220, 223], segmentation [51, 52, 60, 61, 74, 83, 98, 115, 195, 196, 221], texture analysis [43, 66, 122], motion detection [90, 213], optical flow and motion analysis [87, 90, 167], etc. Most of these problems can be formulated in a general framework, called *image labeling*, where we associate a label to each pixel from a finite set. The meaning of this label depends on the problem that we are trying to solve. For image restoration, it means a gray-level; for edge detection, it means the presence or the direction of an edge; for image segmentation, it means a region; etc. The problem here is how to choose a label for a pixel, which is *optimal* in a certain sense. Herein, we deal with a statistical approach of *labeling*. In real scenes, neighboring pixels usually have similar features (intensity, color, texture, etc). In a probabilistic framework, such regularities are well expressed mathematically by Markov random fields. In this survey, we will focus on the fundamental problem of image segmentation using Markovian models.

1.1 Image Segmentation

The primary goal of any segmentation algorithm is to divide the domain R of the input image into the disjoint parts R_i such that they belong to distinct objects in the scene. The solution of this problem sometimes requires high level knowledge about the shape and appearance of the objects under investigation [46, 123, 183, 202]. In many applications, however, such information is not available or impractical to use. Hence low-level features of the surface patches are used for the segmentation process [9, 141, 224]. Herein, we are interested in the latter approach. In either case, we have to summarize all relevant information in a model which is then adjusted to fit the image data.

One broadly used class of models is the so called *cartoon model*, which has been extensively studied from both probabilistic [64] and variational [19, 163, 169] viewpoints. The model assumes that the real world scene consists of a set of regions whose observed low-level features change slowly, but across the boundary between them, these features change abruptly. What we want to infer is a *cartoon* ω consisting of a simplified, abstract version of the input image $\mathcal{I}$: regions R_i have a constant value (called a *label* in our context) and the discontinuities between them form a curve Γ — the contour. The pair (ω, Γ) specifies a *segmentation*. Region based methods are mainly focused on ω while edge based methods try to determine Γ directly.

Taking the probabilistic approach, one usually wants to come up with a *probability measure* on the set Ω of all possible segmentations of $\mathcal{I}$ and then select the one with the highest probability. Note that

Ω is finite, although huge. A widely accepted standard, also motivated by the human visual system [121, 162], is to construct this probability measure in a Bayesian framework [37, 161, 214]: We shall assume that we have a set of observed (Y) and hidden (X) random variables. In our context, any observed value $y \in Y$ represents the low-level features used for partitioning the image, and the hidden entity $x \in X$ represents the segmentation itself. First, we have to quantify how well any occurrence of x fits y. This is expressed by the probability distribution $P(y|x)$ — the *imaging model*. Second, we define a set of properties that any segmentation x must possess regardless the image data. These are described by $P(x)$, the *prior*, which tells us how well any occurrence x satisfies these properties. Factoring these distributions and applying the Bayes theorem gives us the *posterior* distribution $P(x|y) \propto P(y|x)P(x)$. Note that the constant factor $1/P(y)$ has been dropped as we are only interested in $\widehat{x}$ which *maximizes* the posterior, that is, the maximum a posteriori (MAP) estimate of the hidden field X.

The models of the above distributions also depend on certain parameters that we denote by Θ. Supervised segmentation assumes that these parameters are either known or a set of joint realizations of the hidden field X and observations Y (called a *training set*) is available [64, 205]. This is known in statistics as the *complete data* problem which is generally easier to solve than the *incomplete case* [37]. Although the prior knowledge of the parameters is a strong assumption, supervised methods are still useful alternatives when working in a controlled environment. Many industrial applications, like quality inspection of agricultural products [166], fall into this category. In the unsupervised case, however, we know neither Θ nor X. This is called the *incomplete data* problem where both Θ and X have to be inferred from the only observable entity Y. Hence our MAP estimation problem becomes $(\widehat{x}, \widehat{\Theta}) = \arg\max_{x,\Theta} P(x, \Theta|y)$. *Expectation Maximization* (EM) [48] and its variants (Stochastic EM [33, 149], Gibbsian EM [36]), as well as *Iterated Conditional Expectation* (ICE) [30, 108] are widely used to solve such problems. It is important to note, however, that these methods calculate a local maximum [37].

Due to the difficulty of estimating the number of pixel classes (or clusters), unsupervised algorithms often suppose that this parameter is *known a priori* [68, 77, 141, 145, 149]. When the number of pixel classes is also being estimated, the unsupervised segmentation problem may be treated as a *model selection* problem over a combined model space [102, 202, 203].

1.2 Markov Random Fields

In the early 20th century, mostly inspired by the Ising model [170], a new type of stochastic process appeared in the theory of probability, called *Markov random field* (MRF). MRFs rapidly became a broadly used tool in a variety of problems, not only in statistical mechanics. The use of MRFs in image processing became popular with the seminal paper of S. Geman and D. Geman [64] in 1984, but its first use in the domain dates to the early 70s [16, 215]. Here, we give a brief introduction to the theory of MRFs [39, 54, 57, 79, 125, 144, 160, 184, 214].

1.2.1 The Ising Model

Following Ising [10, 69, 170], we consider a sequence, $0,1,2,\ldots,n$ on the line. At each point, there is a small spin which is either *up* or *down* at any given moment (see Figure 1.1). Now, we define a probability measure on the set Ω of all possible configurations $\omega = (\omega_0,\omega_1,\ldots,\omega_n)$. In this context, each spin is a function

$$\delta_i(\omega) = \begin{cases} 1 & \text{if } \omega_i \text{ is } up \\ -1 & \text{if } \omega_i \text{ is } down \end{cases} \tag{1.1}$$

An *energy* $U(\omega)$ is assigned to each configuration:

$$U(\omega) = -J\sum_{i,j}\delta_i(\omega)\delta_j(\omega) - mH\sum_i \delta_i(\omega). \tag{1.2}$$

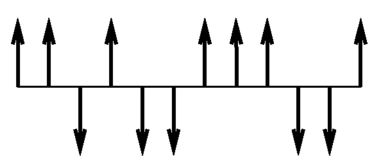

Fig. 1.1 One dimensional Ising model.

In the first sum, Ising made a simplifying assumption that only interactions of points with one unit apart need to be taken into account. This term represents the energy caused by the spin-interactions. The constant J is a property of the material. If $J > 0$, the interactions tend to keep neighboring spins in the same directions (*attractive case*). If $J < 0$, neighboring spins with opposite orientation are favored (*repulsive case*). The second term represents the influence of an external magnetic field of intensity H and $m > 0$ is a property of the material. The probability on Ω is then given by

$$P(\omega) = \frac{\exp\left(-\frac{1}{kT}U(\omega)\right)}{Z}, \tag{1.3}$$

where T is the temperature and k is a universal constant. The normalizing constant (also called *partition function*) Z is defined by

$$Z = \sum_{\omega \in \Omega} \exp\left(-\frac{1}{kT}U(\omega)\right). \tag{1.4}$$

The probability defined in Equation (1.3) is called a *Gibbs distribution.* One could extend the model to two dimensions in a natural way. The spins are arranged on a lattice, they are represented by two coordinates and a point have 4 neighbors unless it is on the boundary. In the two-dimensional case, the limiting measure P is unstable, there is a *phase transition.* As it is pointed out in [125], considering the *attractive* case and an external field h, the measure P_h converges to P^- if h goes to zero through negative values but it converges to $P^+ \neq P^-$ if h goes to zero through positive values. It has been shown, that there exists a *critical temperature* T_C and below this temperature phase transition always occurs. The temperature depends on the vertical (J_1) and horizontal (J_2) interaction parameters.

As a special example, we mention the *Cayley tree model* [125], originally proposed by Bethe [10] as an approximation to the Ising model. In this case, the points sit on a tree (see Figure 1.2). The root is called the 0th level. From the root, we have q branches ($q = 2$ in Figure 1.2). The $q = 1$ case simply gives a one-dimensional Markov chain. A configuration on a tree of n levels is an assignment of a label *up* or *down* to each point. We can define a similar energy function as for the Ising model.

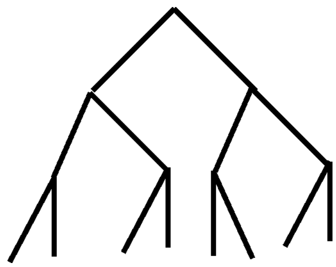

Fig. 1.2 Cayley tree model.

1.2.2 The Potts Model

Another important extension of the Ising model to more than two states per points is the Potts model [10, 195, 216]. The problem is to regard the Ising model as a system of interacting spins that can be either parallel or antiparallel. More generally, we consider a system of spins, each spin pointing one of the q equally spaced directions. These vectors are the linear combinations of q unit vectors pointing in the q symmetric directions of a hypertetrahedron in $q-1$ dimensions. For $q = 2,3,4$, examples are shown in Figure 1.3. The energy function of the Potts model can be written as

$$U(\omega) = \sum_{i,j} J(\Theta_{ij}), \tag{1.5}$$

where $J(\Theta)$ is 2π periodic and Θ_{ij} is the angle between two neighboring spins in i and j. The $q = 2$ case is equivalent to the Ising model.

1.2.3 Gibbs Distribution and MRFs

The most natural way to define MRFs [2, 64, 184] related to image models is to define them on a lattice. However, here we will define

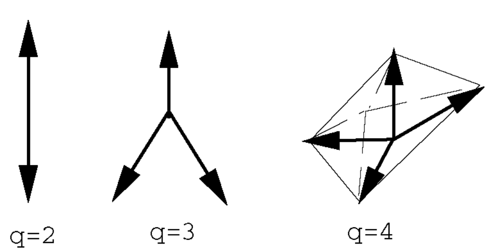

Fig. 1.3 The Potts model.

MRFs more generally on graphs. It will be useful in Section 2 for the study of hierarchical models. Let $\mathcal{G} = (\mathcal{S}, \mathcal{E})$ be a graph where $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$ is a set of vertices (or sites) and $\mathcal{E}$ is the set of edges.

Definition 1.1 (Neighbors). Two points s_i and s_j are neighbors if there is an edge $e_{ij} \in \mathcal{E}$ connecting them. The set of points which are neighbors of a site s (that is, the neighborhood of s) is denoted by $\mathcal{G}_s$.

Definition 1.2 (Neighborhood system). $\mathcal{G} = \{\mathcal{G}_s \mid s \in \mathcal{S}\}$ is a neighborhood system for $\mathcal{G}$ if

(1) $s \notin \mathcal{G}_s$
(2) $s \in \mathcal{G}_r \Leftrightarrow r \in \mathcal{G}_s$

To each site of the graph, we assign a label λ from a finite set of labels Λ. Such an assignment is called a configuration ω having some probability $P(\omega)$. The restriction to a subset $\mathcal{T} \subset \mathcal{S}$ is denoted by $\omega_{\mathcal{T}}$ and $\omega_s \in \Lambda$ denotes the label given to the site s. In the following, we are interested in the probabilities assigned to the set Ω of all possible configurations. First, let us define the *local characteristics* as the conditional probabilities $P(\omega_s \mid \omega_r, r \neq s)$.

Definition 1.3 (Markov random field). $\mathcal{X}$ is a Markov random field (MRF) with respect to $\mathcal{G}$ if

(1) for all $\omega \in \Omega$: $P(\mathcal{X} = \omega) > 0$,
(2) for every $s \in \mathcal{S}$ and $\omega \in \Omega$:
$P(X_s = \omega_s \mid X_r = \omega_r, r \neq s) = P(X_s = \omega_s \mid X_r = \omega_r, r \in \mathcal{G}_s)$.

To continue our discussion about probabilities on Ω, the notion of *cliques* will be very useful.

Definition 1.4 (Clique). A subset $C \subseteq \mathcal{S}$ is a clique if every pair of distinct sites in C are neighbors. $\mathcal{C}$ denotes the set of cliques and $\deg(\mathcal{C}) = \max_{C \in \mathcal{C}} | C |$.

Using the above definition, we can define a *Gibbs measure* on Ω. Let V be a *potential* which assign a number $V_{\mathcal{T}}(\omega)$ to each subconfiguration $\omega_{\mathcal{T}}$. V defines an *energy* $U(\omega)$ on Ω by

$$U(\omega) = -\sum_{\mathcal{T}} V_{\mathcal{T}}(\omega). \tag{1.6}$$

Definition 1.5 (Gibbs distribution). A Gibbs distribution is a probability measure π on Ω with the following representation:

$$\pi(\omega) = \frac{1}{Z}\exp\left(-U(\omega)\right), \tag{1.7}$$

where Z is the normalizing constant (also called *partition function*):

$$Z = \sum_{\omega} \exp\left(-U(\omega)\right),$$

If $V_{\mathcal{T}}(\omega) = 0$ whenever $\mathcal{T}$ is not a clique then V is called a *nearest neighbor Gibbs potential*. In the following, we will focus on such potentials. The next famous theorem establish the equivalence between Gibbs measures and MRFs [16, 160].

Theorem 1.6 (Hammersley–Clifford). $\mathcal{X}$ is a MRF with respect to the neighborhood system $\mathcal{G}$ if and only if $\pi(\omega) = P(\mathcal{X} = \omega)$ is a Gibbs distribution with a nearest neighbor Gibbs potential V, that is

$$\pi(\omega) = \frac{1}{Z}\exp\left(-\sum_{C\in\mathcal{C}} V_C(\omega)\right) \tag{1.8}$$

The main benefit of this equivalence is that it provides us a simple way to specify MRFs, namely specifying potentials instead of local characteristics (see Definition 1.3), which is usually very difficult.

1.2.4 Spatial Lattice Schemes

In this section, we deal with a particular subclass of MRFs which are the most commonly used schemes in image processing. In this case,

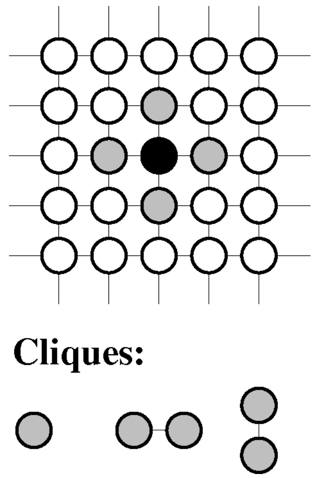

Fig. 1.4 First order neighborhood system.

we consider $\mathcal{S}$ as a lattice $\mathcal{L}$ so that $\forall s \in \mathcal{S} : s = (i,j)$ and define the so-called *nth order homogeneous neighborhood systems* as

$$\mathcal{G}^n = \{\mathcal{G}^n_{(i,j)} : (i,j) \in \mathcal{L}\}, \tag{1.9}$$

$$\mathcal{G}^n_{(i,j)} = \{(k,l) \in \mathcal{L} : (k-i)^2 + (l-j)^2 \leq n\}. \tag{1.10}$$

Obviously, sites near the boundary have fewer neighbors than interior ones (free boundary condition). Furthermore, $\mathcal{G}^0 \equiv \mathcal{S}$ and for all $n \geq 0 : \mathcal{G}^n \subset \mathcal{G}^{n+1}$. Figure 1.4 shows a first-order neighborhood corresponding to $n = 1$. The cliques are $\{(i,j)\}, \{(i,j),(i,j+1)\}, \{(i,j),(i+1,j)\}$. In practice, more than two order systems (cf. Figure 1.5) are rarely used since the energy function would be too complicated requiring a lot of computation. Although not as widespread as orthogonal lattice schemes, hexagonal lattices [45, 193] as well as MRFs on graphs [204] have also been studied in the literature.

1.3 Related Approaches

1.3.1 Weak Membrane Model

The *weak membrane model* was introduced in image reconstruction by A. Blake and A. Zisserman [19]. The problem is to reconstruct surfaces

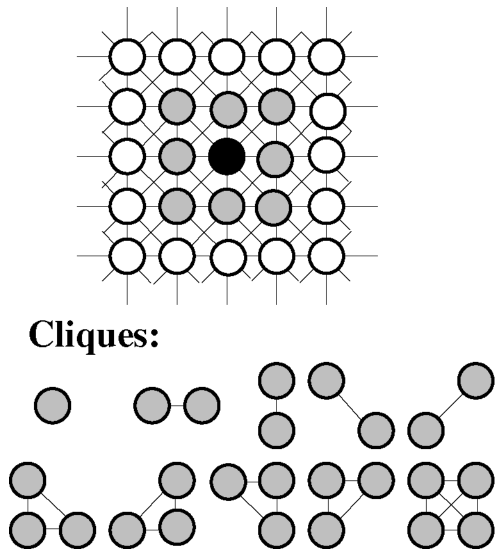

Fig. 1.5 Second order neighborhood system.

which are *continuous almost everywhere* or, in other words, continuous in patches. To reach a satisfactory formalization of this principle, they have used a membrane model: Imagine an elastic membrane which we are trying to fit to a surface. The edges will appear as tears in the membrane. Depending on how elastic is the membrane, there may be more or fewer edges. The membrane is described by an energy function (the elastic energy of the membrane) which has to be minimized in order to find an equilibrium state. The energy has three components:

D: A measure of faithfulness to the data:

$$D = \int (u - d)^2 dA, \tag{1.11}$$

where $u(x,y)$ represents the membrane and $d(x,y)$ represents the data.

S: A measure of how the function $u(x,y)$ is deformed:

$$S = \lambda^2 \int (\nabla u)^2 dA. \tag{1.12}$$

λ^2 is a measure of elasticity of the membrane.

P: The sum of penalties α levied for each break in the membrane:

$$P = \alpha Z, \tag{1.13}$$

where Z is a measure of the set of contours along which $u(x,y)$ is discontinuous (see [19] for more details).

The elastic energy of the membrane is then given by

$$E = D + S + P = \int (u - d)^2 dA + \lambda^2 \int (\nabla u)^2 dA + \alpha Z. \tag{1.14}$$

There is a strong relation between the *weak membrane model* and MRF models: An elastic system can also be considered from a probabilistic view-point. The link between the elastic energy E and probability P is

$$P \propto \exp\left(\frac{-E}{T}\right), \tag{1.15}$$

that is the Gibbs distribution. However, the *weak membrane model* operates with mechanical analogies, representing *a priori* knowledge from a mechanical point of view while MRF modelization is purely probabilistic.[1]

1.3.2 Snakes, Variational and Level Set Methods

Active Contours (snakes) are closed curves evolving toward the boundary of the object of interest. The curve evolution is governed by a boundary functional [101] which takes its minimum on the object contour. The main drawback of the parametric snake model is that it cannot handle topological changes easily. Nevertheless, they became quite popular because they make it relatively easy to enforce contour-smoothness; and starting from an appropriate initialization a local minimum of the associated energy function will give good results. One extension of the original model is gradient vector flow [217] snakes

[1] We notice that the weak membrane model has also been used in a Markovian context but originally, as proposed by Blake and Zisserman [19], it was a non-Markovian model.

which make the snake less sensitive to initialization and allow the contour to segment concave objects. Another extension is the so-called balloon force [44] which basically introduces an area minimizing term [31] into the snake energy.

Geodesic active contours [31] are curves of minimum length in the metric defined by a function u. The criterion to minimize is usually of the form $\int_\Gamma u(s)ds$. Most of the time, u is simply a function of the image gradient like $u = 1/(1 + |\nabla\mathcal{I}|)$. The contour evolution equation is as follows [31]:

$$\frac{\partial\Gamma}{\partial t} = (\kappa u \nabla u \cdot N)N, \tag{1.16}$$

where κ is the curvature and N is the inward normal of Γ.

Region based active contours are another class of boundary based methods where region descriptors (usually some kind of statistical features) are introduced into the energy in order to better characterize an object [169, 188, 224].

Variational approaches consider the segmentation as an optimal approximation of the original image $\mathcal{I}$ by a piecewise smooth function f having discontinuities across Γ. The classical Mumford–Shah energy functional [163] is then defined as

$$E(f,\Gamma) = \mu^2 \int\int_R (f - \mathcal{I})^2 dxdy + \int\int_{R-\Gamma} ||\nabla f||^2 dxdy + \nu|\Gamma|. \tag{1.17}$$

Clearly, the minimum is achieved when f approximates $\mathcal{I}$ (first term), f is smooth over each R_i (second term), and the boundaries Γ are as short as possible. Note that dropping any of the above three terms would result in $\inf E = 0$ with some trivial and (from a practical point of view) useless settings for f and Γ. The minimization of the above functional is far from trivial. Note also that in our context, $f = \omega$ is constant over each region R_i, hence the problem can be further simplified to a piecewise constant functional. A closely related model, proposed by Blake and Zisserman, is the so-called *weak membrane* model (see Section 1.3.1) which can be minimized via graduated non-convexity (GNC) [19].

More recently, the level set formulation [192] of the piecewise constant Mumford–Shah energy functional proposed by Chan and Vese [38] have become a popular framework for image segmentation. The contour Γ is represented as the zero level set of an embedding function (the level set function) $\phi : R \to \Re$ on the image domain R: $\phi(\Gamma) = 0$. The main advantage of this formulation is that it handles topological changes of the evolving contour. This makes the level set formalism well suited to the segmentation of multiple objects. The region based level set scheme for foreground–background segmentation consists in minimizing the following functional:

$$E_{CV}(c_1, c_2, \phi) = \int_R (\mathcal{I} - c_1)^2 H(\phi) dx + \int_R (\mathcal{I} - c_2)^2 (1 - H(\phi)) dx \\ + \nu \int_R |\nabla H(\phi)| dx, \tag{1.18}$$

c_1 and c_2 are the means of the regions, where $\phi > 0$ (outside or background) and $\phi < 0$ (inside or foreground), and $H(.)$ is the Heaviside function. The last term measures the length of the zero crossing of ϕ (i.e., the contour). The Euler–Lagrange equation for this model is implemented by the following gradient descent:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \mathrm{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - (\mathcal{I} - c_1)^2 + (\mathcal{I} - c_2)^2 \right]. \tag{1.19}$$

Unfortunately, even with the narrow band implementation [1], the level set approach has a rather high computational complexity. The fast marching method [192] has a lower complexity but it requires that the speed function doesn't change sign during evolution.

1.3.3 Conditional Random Fields

Conditional Random Fields (CRF) directly model the posterior distribution of $P(X|Y)$ as a Gibbs field [86, 135, 136, 206]. Unlike the generative image models commonly used in MRFs, CRFs can depend on arbitrary non-independent characteristics of the observation $Y = y$. Originally, CRFs were proposed for segmenting 1D text sequences [140, 212], but it is straightforward to extend these consepts to 2D images.

Basically, a CRF is a random field globally conditioned on the observation Y. Following [140], we can formally define CRFs on graph:

Definition 1.7 (Conditional Random Field). Let $G = (V, E)$ be a graph such that the label field X is indexed by the vertices: $X = \{X_v\}_{v \in V}$ and neighboring elements $v \sim w$ of the field are connected by edges in G, i.e., $(v,w) \in E$. Then (Y, X) is a *conditional random field (CRF)* if the random variables X_v, when conditioned on Y, obey the Markov property with respect to the graph: $P(X_v|Y, X_w, w \neq v) = P(X_v|Y, X_w, w\tilde{v})$.

The simplest example of such a graph structure is a lattice where vertices correspond to pixels and neighboring lattice sites are connected by edges (see Section 1.2.4 for various neighborhood structures on lattices). Considering a first order neighborhood, the posterior distribution can be easily expressed using the Hammersley–Clifford theorem (see Theorem 1.6):

$$P(x|y) = \exp\left(\sum_{e \in E, k} \lambda_k f_k(e, x|_e, y) + \sum_{v \in V, k} \mu_k g_k(v, x|_v, y))\right), \quad (1.20)$$

where x is a labeling of a given input image y and $x|_S$ is the set of components of x associated to the vertices in the subgraph S. Furthermore, the features f_k and g_k are assumed to be known and fixed, and the parameter values λ_k and μ_k are to be learned from training data [140]. As we can see from Equation (1.20), standard CRFs use two forms of feature functions, which can be interpreted in 2D as follows [86]:

- state feature function $g_k(s, x_s, y)$ of the label x_s at a site s and the observed image y,
- transition feature function $f_k(s, r, x_s, x_r, y)$ of the labels x_s and x_r at neighboring sites $s \sim r$ and the observed image y.

In image processing applications, state feature functions are usually defined as unary (also known as singleton) clique potentials based on classifier responses (such as Ada-boost [194] or kernel SVMs [191]), while transition feature functions are defined as pairwise (also known as doubleton) potentials modeling the correlation between pairs of random variables. Recently, CRFs became popular in image segmentation [86, 206], especially CRFs coupled with graph cut energy minimization [128, 138, 208].

References

[1] D. Adalsteinsson and J. A. Sethian, "A fast level set method for propagating interfaces," *Journal of Computational Physics*, vol. 118, pp. 269–277, 1995.

[2] R. Azencott, "Markov fields and image analysis," in *Proceedings of Association Francaise pour la Cybernetique Economique et Technique*, Antibes, France, 1987.

[3] R. Azencott, "Parallel simulated annealing: An overview of basic techniques," in *Simulated Annealing: Parallelization Techniques*, (R. Azencott, ed.), pp. 37–46, New York, NY: John Wiley & Sons, 1992.

[4] R. Azencott, ed., *Parallel Simulated Annealing. Parallelization Techniques.* New York, NY: John Wiley & Sons, 1992.

[5] R. Azencott and C. Graffigne, "Parallel annealing by periodically interacting multiple searches: Acceleration rates," in *Simulated Annealing: Parallelization Techniques*, (R. Azencott, ed.), pp. 81–90, John Wiley & Sons: New York, NY, 1992.

[6] L. Baratchart, M. Berthod, and L. Pottier, "Optimization of positive generalized polynomials under l^p constraints," Technical Report RR-2750, INRIA, December 1995.

[7] A. Barbu and S.-C. Zhu, "Generalizing Swendsen-Wang to sampling arbitrary posterior probabilities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1239–1253, 2005.

[8] Y. Bard, *Nonlinear Parameter Estimation.* New York, NY: Academic Press, 1974.

[9] S. A. Barker and P. J. W. Rayner, "Unsupervised image segmentation using Markov random field models," *Pattern Recognition*, vol. 33, pp. 587–602, April 2000.

[10] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*. London: Academic Press, 1990.

[11] C. Benedek and T. Sziranyi, "Change detection in optical aerial images by a multi-layer conditional mixed markov model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 3416–3430, October 2009.

[12] C. Benedek, T. Sziranyi, Z. Kato, and J. Zerubia, "A multi-layer MRF model for object-motion detection in unregistered airborne image-pairs," in *Proceedings of International Conference on Image Processing*, pp. 141–144, San Antonio, Texas, USA, September 2007.

[13] C. Benedek, T. Sziranyi, Z. Kato, and J. Zerubia, "A three-layer MRF model for object motion detection in airborne images," Research Report 6208, INRIA, France, June 2007.

[14] C. Benedek, T. Sziranyi, Z. Kato, and J. Zerubia, "Detection of object motion regions in aerial image pairs with a multilayer Markovian model," *IEEE Transactions on Image Processing*, vol. 18, pp. 2303–2315, October 2009.

[15] M. Berthod, Z. Kato, and J. Zerubia, "DPA: A deterministic approach to the MAP," *IEEE Transactions on Image Processing*, vol. 4, pp. 1312–1314, September 1995.

[16] J. Besag, "Spatial interaction and the statistical analysis of lattice systems (with discussion)," *Journal of the Royal Statistical Society, Series B*, vol. 36, no. 2, pp. 192–236, 1974.

[17] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society, Series B*, vol. 48, no. 3, pp. 259–302, 1986.

[18] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr, "Interactive image segmentation using an adaptive GMMRF model," in *Proceedings of European Conference on Computer Vision*, pp. 428–441, 2004.

[19] A. Blake and A. Zisserman, *Visual Reconstruction*. USA: MIT Press, 1987.

[20] T. Blaskovics, Z. Kato, and I. Jermyn, "A Markov random field model for extracting near-circular shapes," in *Proceedings of International Conference on Image Processing*, pp. 1073–1076, Cairo, Egypt, November 2009.

[21] E. Boros, P. L. Hammer, and X. Sun, "Network flows and minimization of quadratic pseudo-boolean functions," Research Report RRR 17-1991, RUTCOR, May 1991.

[22] C. Bouman, "A multiscale image model for Bayesian image segmentation," Technical Report TR-EE 91-53, Purdue University, 1991.

[23] C. Bouman and B. Liu, "Multiple resolution segmentation of texture images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 99–113, 1991.

[24] C. Bouman and M. Shapiro, "Multispectral image segmentation using a multiscale model," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pp. 565–568, San Francisco, California, USA, March 1992.

[25] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient n-D image segmentation," *International Journal of Computer Vision*, vol. 70, pp. 109–131, February 2006.

[26] Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in n-D images," in *Proceedings of International Conference on Computer Vision*, pp. 105–112, July 2001.

[27] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1124–1137, September 2004.

[28] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1222–1239, November 2001.

[29] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1222–1239, November 2001.

[30] B. Braathen, W. Pieczynski, and P. Masson, "Global and Local Methods of Unsupervised Bayesian Segmentation of Images," *Machine Graphics and Vision*, vol. 2, no. 1, pp. 39–52, 1993.

[31] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[32] O. Catoni and A. Trouvé, "Parallel annealing by multiple trials," in *Simulated Annealing: Parallelization Techniques*, (R. Azencott, ed.), pp. 129–144, New York, NY: John Wiley & Sons, 1992.

[33] G. Celeux and J. Diebolt, "The SEM algorithm: A probabilisitic teacher algorithm derived from the EM algorithm for the mixture problem," *Computational Statistics Quaterly*, vol. 2, pp. 73–82, 1985.

[34] V. Černy, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *Journal of Optimization Theory and Applications*, vol. 45, pp. 41–51, January 1985.

[35] B. Chalmond, "Image restoration using an estimated Markov model," *Signal Processing*, vol. 15, pp. 115–129, September 1988.

[36] B. Chalmond, "An iterative Gibbsian technique for reconstruction of M-ary images," *Pattern Recognition*, vol. 22, no. 6, pp. 747–762, 1989.

[37] B. Chalmond, *Modeling and Inverse Problems in Image Analysis*. New York, NY: Springer, 2003.

[38] T. Chan and L. Vese, "An active contour model without edges," in *Proceedings of International Conference on Scale-Space Theories in Computer Vision*, pp. 141–151, 1999.

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

[40] C. Chen, H. Li, X. Zhou, and S. T. C. Wong, "Constraint factor graph cut-based active contour method for automated cellular image segmentation in RNAi screening," *Journal of Microscopy*, vol. 230, pp. 177–191, May 2008.

[41] Y. Chen, H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. Gopinath, R. Briggs, and E. Geiser, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 315–328, 2002.

[42] P. Chou and C. Brown, "The theory and practice of Bayesian image labeling," *International Journal of Computer Vision*, vol. 4, no. 3, pp. 185–210, 1990.

[43] F. S. Cohen and D. B. Cooper, "Simple parallel hierarchical and relaxation algorithms for segmenting noncausal Markov random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 195–219, March 1987.

[44] L. Cohen, "On active contour models and balloons," *Computer Vision, Graphics and Image Processing: Image Understanding*, vol. 53, pp. 211–218, March 1991.

[45] L. Condat, D. V. de Ville, and T. Blu, "Hexagonal versus orthogonal lattices: A new comparison using approximation theory," in *Proceedings of International Conference on Image Processing*, pp. 1116–1119, Genoa, Italy, September 2005.

[46] D. Cremers, F. Tischhauser, J. Weickert, and C. Schnorr, "Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 295–313, 2002.

[47] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, pp. 25–39, January 1983.

[48] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[49] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 800–810, August 2001.

[50] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 39–55, January 1987.

[51] H. Derin, H. Elliott, R. Cristi, and D. Geman, "Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 707–720, November 1984.

[52] H. Derin and C. S. Won, "A parallel segmentation algorithm using relaxation with varying neighborhoods and its mapping to array procesors," *Computer Vision, Graphics and Image Processing*, vol. 40, pp. 54–78, October 1987.

[53] X. Descombes, R. Morris, J. Zerubia, and M. Berthod, "Maximum likelihood estimation of Markov random field paremeters using Markov chain Monte Carlo algorithms," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 133–148, Springer 1997.

[54] P. L. Dobruschin, "The description of a random field by means of conditional probabilities and constructions of its regularity," *Theory of Probability and its Applications*, vol. XIII, no. 2, pp. 197–224, 1968.

[55] O. Faugeras and M. Berthod, "Improving consistency and reducing ambiguity in stochastic labeling: An optimization approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 4, pp. 412–423, 1981.

[56] T. S. Ferguson, *Mathematical Statistics. A Decision Theoretic Approach. Probability and Mathematical Statistics*. New York, NY: Academic Press, 1967.

[57] P. Fieguth, *Statistical Image Processing and Multidimensional Modeling*. New York, NY: Springer, 2011.

[58] K. Fish, "Total internal reflection fluorescence (TIRF) microscopy," *Current Protocols in Cytometry*, vol. Chapter 12, 2009.

[59] I. Gaudron and A. Trouvé, "Massive parallelization of simulated annealing: An experimental and theoretical approach for spin-glass models," in *Simulated Annealing: Parallelization Techniques*, (R. Azencott, ed.), pp. 163–186, New York, NY: John Wiley & Sons, 1992.

[60] D. Geiger and F. Girosi, "Parallel and deterministic algorithms for MRFs: Surface reconstruction and integration," in *Proceedings of European Conference on Computer Vision*, pp. 89–98, Antibes, France, 1990.

[61] D. Geiger and A. Yuille, "A common framework for image segmentation," Technical Report 89-7, Harvard Robotics Lab, 1989.

[62] S. B. Gelfand and S. K. Mitter, "On sampling methods and annealing algorithms," in *Markov Random Fields*, (R. Chellappa, ed.), pp. 499–515, Boston, MA: Academic Press, 1993.

[63] D. Geman, "Bayesian image analysis by adaptive annealing," in *Proceedings of International Geoscience and Remote Sensing Symposium*, pp. 269–277, Amherst, USA, October 1985.

[64] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.

[65] S. Geman, D. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE transactions on pattern analysis and machine intelligence*, vol. 12, pp. 609–628, 1990.

[66] S. Geman and C. Graffigne, "Markov random field image models and their application to computer vision," Research Report, Brown University, 1986.

[67] B. Gidas, "A renormalization group approach to image processing problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 164–180, February 1989.

[68] N. Giordana and W. Pieczynski, "Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 465–475, May 1997.

[69] J.-F. Giovannelli, "Estimation of the ising field parameter thanks to the exact partition function," in *Proceedings of International Conference on Image Processing*, pp. 1441–1444, Hong Kong, China, September 2010.

[70] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ: Prentice Hall, 2008.

[71] C. Graffigne, "A parallel simulated annealing algorithm," Research Report, CNRS, Université Paris-Sud, 1984.

[72] C. Graffigne, "Parallel annealing by periodically interacting multiple searches: An experimental study," in *Simulated Annealing: Parallelization Techniques*, (R. Azencott, ed.), pp. 47–80, New York, NY: John Wiley & Sons, 1992.

[73] C. Graffigne, F. Heitz, P. Pérez, F. Prêteux, M. Sigelle, and J. Zerubia, "Hierarchical Markov random field models applied to image analysis: a review," in *Proceedings of Conference on Neural, Morphological, and Stochastic Methods in Image and Signal Processing*, San Diego, USA, July 1995.

[74] C. Graffigne, J. Zerubia, and B. Chalmond, *Analyse d'images : filtrage et segmentation*, ch. Segmentation région: approches statistiques. Masson, 1995.

[75] U. Grenander, *General Pattern Theory.* New York, NY: Oxford University Press, 1993.

[76] U. Grenander and M. Miller, "Representations of knowledge in complex systems," *Journal of the Royal Statistical Society, Series B*, vol. 56, pp. 549–603, 1994.

[77] L. Gupta and T. Sortrakul, "A Gaussian-mixture-based image segmentation algorithm," *Pattern Recognition*, vol. 31, no. 3, pp. 315–325, 1998.

[78] M. R. Gupta and Y. Chen, "Theory and use of the em algorithm," *Foundations and Trends in Signal Processing*, vol. 4, no. 3, pp. 223–296, 2010.

[79] X. Guyon, *Champs aléatoires sur réseaux: modélisations, statistique et applications.* Masson, 1992.

[80] B. Hajek, "A tutorial survey of theory and applications of simulated annealing," in *Proceedings of International Conference on Decision and Control*, pp. 755–760, Lauderdale, FL, USA, December 1985.

[81] B. Hajek, "Cooling schedules for optimal annealing," *Mathematics of Operations Research*, vol. 13, pp. 311–329, May 1988.

[82] P. L. Hammer, P. Hansen, and B. Simeone, "Roof duality, complementation and persistency in quadratic 0-1 optimization," *Mathematical Programming*, vol. 28, pp. 121–155, 1984.

[83] F. R. Hansen and H. Elliott, "Image segmentation using simple Markov field models," *Computer Vision, Graphics and Image Processing*, vol. 20, pp. 101–132, October 1982.

[84] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems on Man, and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.

[85] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their application," *Biometrika*, vol. 57, pp. 97–109, 1970.

[86] X. He, R. S. Zemel, and M. A. Carreira-Perpinan, "Multiscale conditional random fields for image labeling," in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 695–702, Washington, DC, USA, July 2004.

[87] F. Heitz and P. Bouthemy, "Multimodal estimation of discontinuous optical flow using Markov random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1217–1232, December 1993.

[88] F. Heitz, E. Memin, P. Perez, and P. Bouthemy, "A parallel multiscale relaxation algorithm for image sequence analysis," in *Proceedings of International Colloquiumon Parallel Image Processing*, Paris, France, June 1991.

[89] F. Heitz, P. Perez, and P. Bouthemy, "Multiscale minimization of global energy functions in some visual recovery problems," *Computer Vision, Graphics and Image Processing: Image Understanding*, vol. 59, no. 1, pp. 125–134, 1994.

[90] F. Heitz, P. Perez, E. Memin, and P. Bouthemy, "Parallel visual motion analysis using multiscale Markov random fields," in *Proceedings of Workshop on Motion*, Princeton, October 1991.

[91] H. P. Hiriyannaiah, G. L. Bilbro, and W. E. Snyder, "Restoration of piecewise-constant images by mean-field annealing," *Journal of the Optical Society of America A*, vol. 6, pp. 1901–1912, December 1989.

[92] R. Huang, V. Pavlovic, and D. N. Metaxas, "A graphical model framework for coupling MRFs and deformable models," in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 739–746, Washington, DC, USA, June 2004.

[93] R. Hummel and S. Zucker, "On the foundations of relaxation labeling processes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, pp. 267–287, May 1983.

[94] H. Ishikawa, "Exact optimization for markov random fields with convex priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1333–1336, October 2003.

[95] H. Ishikawa, "Transformation of general binary mrf minimization to the first order case," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 1234–1249, June 2011.

[96] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

[97] F. C. Jeng and J. M. Woods, "Compound Gauss — Markov random fields for image estimation," *IEEE Transactions on Signal Processing*, vol. 39, pp. 683–697, March 1991.

[98] B. Jeon and D. A. Landgrebe, "Classification with spatio-temporal interpixel class dependency contexts," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 30, pp. 663–672, July 1992.

[99] J. M. Jolion and A. Rosenfeld, *A Pyramid Framework for Early Vision, Engineering and Computer Science*. Dordrecht, Netherlands: Kluwer Academic Publisher, 1994.

[100] T. R. Jones, I. H. Kang, D. B. Wheeler, R. A. Lindquist, A. Papallo, D. M. Sabatini, P. Golland, and A. E. Carpenter, "Cellprofiler analyst: Data exploration and analysis software for complex image-based screens," *BMC Bioinformatics*, vol. 9, p. 482, November 2008.

[101] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[102] Z. Kato, "Segmentation of color images via reversible jump MCMC sampling," *Image and Vision Computing*, vol. 26, pp. 361–371, March 2008.

[103] Z. Kato, M. Berthod, and J. Zerubia, "A hierarchical Markov random field model for image classification," in *Proceedings of International Workshop on Image and Multidimensional Digital Signal Processing*, Cannes, France, September 1993.

[104] Z. Kato, M. Berthod, and J. Zerubia, "Multiscale Markov random field models for parallel image classification," in *Proceedings of International Conference on Computer Vision*, pp. 253–257, Berlin, Germany, May 1993.

[105] Z. Kato, M. Berthod, and J. Zerubia, "Parallel image classification using multiscale Markov random fields," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pp. 137–140, Minneapolis, USA, April 1993.

[106] Z. Kato, M. Berthod, and J. Zerubia, "A hierarchical Markov random field model and multi-temperature annealing for parallel image classification," *Computer Graphics and Image Processing: Graphical Models and Image Processing*, vol. 58, pp. 18–37, January 1996.

[107] Z. Kato, M. Berthod, and J. Zerubia, "Parallel image classification using multiscale Markov random fields," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pp. 137–140, Minneapolis, April 1993.

[108] Z. Kato, M. Berthod, J. Zerubia, and W. Pieczynski, "Unsupervised adaptive image segmentation," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pp. 2399–2402, Detroit, Michigan, USA, May 1995.

[109] Z. Kato and T. C. Pong, "A Markov random field image segmentation model using combined color and texture features," in *Proceedings of International Conference on Computer Analysis of Images and Patterns*, (W. Skarbek, ed.), pp. 547–554, Warsaw, Poland, September 2001.

[110] Z. Kato and T. C. Pong, "Video object segmentation using a multicue Markovian model," in *Joint Hungarian-Austrian Conference on Image Processing and Pattern Recognition*, (D. Chetverikov, L. Czuni, and M. Vincze, eds.), pp. 111–118, Veszprem, Hungary: KEPAF, OAGM/AAPR, Austrian Computer Society, May 2005.

[111] Z. Kato and T. C. Pong, "A Markov random field image segmentation model for color textured images," *Image and Vision Computing*, vol. 24, pp. 1103–1114, October 2006.

[112] Z. Kato and T. C. Pong, "A multi-layer MRF model for video object segmentation," in *Proceedings of Asian Conference on Computer Vision*, (P. J. Narayanan, S. K. Nayar, and H.-Y. Shum, eds.), pp. 953–962, Hyderabad, India: Springer, January 2006.

[113] Z. Kato, T. C. Pong, and G. Q. Song, "Multicue MRF image segmentation: Combining texture and color," in *Proceedings of the International Conference on Pattern Recognition*, pp. 660–663, Quebec, Canada, August 2002.

[114] Z. Kato, T. C. Pong, and G. Q. Song, "Unsupervised segmentation of color textured images using a multi-layer MRF model," in *Proceedings of International Conference on Image Processing*, pp. 961–964, Barcelona, Spain, September 2003.

[115] Z. Kato, J. Zerubia, and M. Berthod, "Image classification using Markov random fields with two new relaxation methods: Deterministic pseudo annealing and modified Metropolis dynamics," Research Report 1606, INRIA, Sophia Antipolis, France, February 1992.

[116] Z. Kato, J. Zerubia, and M. Berthod, "Satellite image classification using a modified Metropolis dynamics," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pp. 573–576, San Francisco, California, USA, March 1992.

[117] Z. Kato, J. Zerubia, and M. Berthod, "Bayesian image classification using Markov random fields," in *Maximum Entropy and Bayesian Methods*, (A. Mohammad-Djafari and G. Demoment, eds.), pp. 375–382, Dordrecht Netherlands: Kluwer Academic Publisher, 1993.

[118] Z. Kato, J. Zerubia, and M. Berthod, "Unsupervised parallel image classification using a hierarchical Markovian model," Research Report 2528, INRIA, Sophia Antipolis, France, April 1995.

[119] Z. Kato, J. Zerubia, and M. Berthod, "Unsupervised parallel image classification using a hierarchical Markovian model," in *Proceedings of International Conference on Computer Vision*, pp. 169–174, Cambridge, MA, USA, June 1995.

[120] Z. Kato, J. Zerubia, and M. Berthod, "Unsupervised parallel image classification using Markovian models," *Pattern Recognition*, vol. 32, pp. 591–604, April 1999.

[121] D. Kersten, P. Mamassian, and A. Yuille, "Object perception as Bayesian inference," *Annual Review of Psychology*, vol. 55, pp. 271–304, 2004.

[122] C. Kervrann and F. Heitz, "A statistical model-based approach to unsupervised texture segmentation," in *Proceedings of Scandinavian Conferences on Image Analysis*, pp. 284–288, Tromso, Norway, May 1993.

[123] C. Kervrann and F. Heitz, "Statistical deformable model-based segmentation of image motion," *IEEE Transactions on Image Processing*, vol. 8, pp. 583–588, 1999.

[124] S. Khan and M. Shah, "Object based segmentation of video using color, motion and spatial information," in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 746–751, Kauai, Hawaii, December 2001.

[125] R. Kindermann and J. L. Snell, *Markov Random Fields and their Applications*. Providence, RI: American Mathematical Society, 1980.

[126] S. Kirkpatrick, C. Gellatt, and M. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, May 1983.

[127] P. Kohli, M. P. Kumar, and P. H. Torr, "$\mathcal{P}^3$ & beyond: Move making algorithms for solving higher order functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, pp. 1645–1656, September 2009.

[128] P. Kohli, L. Ladicky, and P. Torr, "Robust higher order potentials for enforcing label consistency," in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2008.

[129] A. Kokaram, *Motion Picture Restoration*. London: Springer, 1998.

[130] V. Kolmogorov, "QPBO algorithm," *software*, 2007.

[131] V. Kolmogorov and C. Rother, "Minimizing nonsubmodular functions with graph cuts-a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1274–1279, 2007.

[132] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 147–159, February 2004.

[133] V. Krylov, G. Moser, S. Serpico, and J. Zerubia, "Enhanced dictionary-based SAR amplitude distribution estimation and its validation with very high-resolution data," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 8, no. 1, pp. 148–152, 2011.

[134] V. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "Supervised high resolution dual polarization SAR image classification by finite mixtures and copulas," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 554–566, 2011.

[135] S. Kumar and M. Hebert, "Discriminative fields for modeling spatial dependencies in natural images," in *Proceedings of Neural Information Processing Systems*, 2003.

[136] S. Kumar and M. Hebert, "A hierarchical field framework for unified context-based classification," in *Proceedings of International Conference on Computer Vision*, pp. 1284–1291, 2005.

[137] P. V. Laarhoven and E. Aarts, *Simulated Annealing: Theory and Applications.* Dordrecht: Kluwer Academic Publisher, 1987.

[138] L. Ladicky, C. Russell, P. Kohli, and P. H. Torr, "Graph cut based inference with co-occurrence statistics," in *Proceedings of European Conference on Computer Vision*, (K. Daniilidis, P. Maragos, and N. Paragios, eds.), pp. 239–253, Crete, Greece, September 2010.

[139] J.-M. Laferte, P. Perez, and F. Heitz, "Discrete Markov modeling and inference on the quad-tree," *IEEE Transactions on Image Processing*, vol. 9, no. 3, pp. 390–404, 2000.

[140] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *International Conference on Machine Learning*, pp. 282–289, 2001.

[141] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 799–813, August 1989.

[142] S. Lakshmanan and H. Derin, "Gaussian Markov Random fields at multiple resolution," in *Markov Random Fields*, pp. 131–157, San Diego: Academic Press, 1993.

[143] M. Leskó, Z. Kato, A. Nagy, I. Gombos, Z. Török, L. V. Jr, and L. Vígh, "Live cell segmentation in fluorescence microscopy via graph cut," in *Proceedings of the International Conference on Pattern Recognition*, pp. 1485–1488, Istambul, Turkey, August 2010.

[144] S. Z. Li, *Markov Random Field Modeling in Image Analysis.* New York NY: Springer, 3rd Edition, 2009.

[145] E. Littmann and H. Ritter, "Adaptive color segmentation — a comparison of neural and statistical methods," *IEEE Transactions on Neural Networks*, vol. 8, pp. 175–185, January 1997.

[146] S. Liu-Yu, "Reconnaissance de formes par vision par ordinateur: application à l'identification de foraminifères planctoniques," PhD thesis, University of Nice, Sophia Antipolis, France, June 1992.

[147] F. Marques, J. Cunillera, and A. Gasull, "Hierarchical segmentation using compound Gauss-Markov random fields," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, San Francisco, California, USA, March 1992.

[148] J. L. Marroquin, "Probabilistic solution of inverse problems," PhD thesis, MIT-Artificial Intelligence Lab., USA, 1985.

[149] P. Masson and W. Pieczynski, "SEM Algorithm and unsupervised statistical segmentation of satellite images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, pp. 618–633, May 1993.

[150] E. Memin, "Algorithmes et architectures parallèles pour les approches markoviennes en analyse d'image," PhD thesis, University of Rennes I, France, 1993.

[151] E. Memin, F. Heitz, and F. Charot, "Efficient parallel non-linear multigrid relaxation algorithms for low-level vision applications," *Journal of Parallel Distributed Computing*, vol. 29, pp. 96–103, August 1995.

[152] D. Metaxas, *Physics-based Deformable Models: Applications to Computer Vision, Graphics and Medical Imaging*. Kluwer Academic Publisher, 1997.

[153] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[154] W. Michiels, E. H. L. Aarts, and J. Korst, *Theoretical Aspects of Local Search*. New York, NY: Springer, 2007.

[155] M. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *International Journal of Computer Vision*, vol. 41, pp. 61–84, February 2001.

[156] M. I. Miller, U. Grenander, O. J. A., and D. L. Snyder, "Automatic target recognition organized via jump-diffusion algorithms," *IEEE Transactions on Image Processing*, vol. 6, pp. 157–174, January 1997.

[157] R. Morris, X. Descombes, and J. Zerubia, "Fully Bayesian image segmentation — an engineering perspective," in *Proceedings of International Conference on Image Processing*, Santa Barbara, USA, October 1997.

[158] G. Moser, V. Krylov, S. Serpico, and J. Zerubia, "High resolution SAR-image classification by Markov random fields and finite mixtures," in *Proceedings of SPIE IS&T/SPIE Electronic Imaging*, pp. 1–8, San Jose, USA, January 2010.

[159] G. Moser, S. B. Serpico, and J. Zerubia, "Dictionary-based Stochastic Expectation Maximization for SAR amplitude probability density function estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 1, pp. 188–199, 2006.

[160] J. Moussouris, "Gibbs and Markov random system with constrainsts," *Journal of Statistical Physics*, vol. 10, pp. 11–33, January 1974.

[161] D. Mumford, "The Bayesian rationale for energy functionals," in *Geometry-Driven Diffusion in Computer Vision*, (B. Romeny, ed.), pp. 141–153, Boston, MA: Kluwer Academic Publisher, 1994.

[162] D. Mumford, "Pattern theory: A unifying perspective," in *Perception as Bayesian Inference*, (D. Knill and W. Richards, eds.), pp. 25–62, Cambridge University Press, 1996.

[163] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[164] E. Nagy, Z. Balogi, I. Gombos, M. Akerfelt, A. Bjorkbom, G. Balogh, Z. Torok, A. Maslyanko, A. Fiszer-Kierzkowska, K. Lisowska, P. Slotte, L. Sistonen, I. Horvath, and L. Vigh, "Hyperfluidization-coupled membrane microdomain reorganization is linked to activation of the heat shock response in a murine melanoma cell line," in *Proceedings of National Academy Science USA*, pp. 7945–7950, 2007.

[165] R. B. Nelsen, *An Introduction to Copulas*. New York, NY: Springer, 2nd Edition, 2006.

[166] J. C. Noordam, G. W. Otten, A. J. M. Timmermans, and B. v. Zwol, "High-speed potato grading and quality inspection based on a color vision system," in *Proceedings of SPIE Machine Vision Applications in Industrial Inspection*, (K. W. T. Jr., ed.), pp. 206–220, 2000.

[167] J.-M. Odobez and P. Bouthemy, "MRF-based motion segmentation exploiting a 2D motion model robust estimation," in *Proceedings of International Conference on Image Processing*, pp. 628–631, Washington, DC, USA, October 1995.

[168] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar images*. New Jersey, NJ: SciTech Publishing, 2004.

[169] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *International Journal of Computer Vision*, vol. 46, pp. 223–247, 2002.

[170] G. Parisi, *Statistical Field Theory*. Westview Press, 1998.

[171] P. Perez, "Champs markoviens et analyse multirésolution de l'image: Application à l'analyse du mouvement," PhD thesis, University of Rennes I, France, 1993.

[172] P. Perez and F. Heitz, "Multiscale Markov random fields and constrained relaxation in low level image analysis," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, pp. 61–64, San Francisco, California, USA, March 1992.

[173] P. Pérez and F. Heitz, "Restriction of Markov random fields on graphs. Application to multiresolution image analysis," Research Report 2170, INRIA, March 1994.

[174] H. Permuter, J. Francos, and I. Jermyn, "A study of Gaussian mixture models of colour and texture features for image classification and segmentation," *Pattern Recognition*, vol. 39, pp. 695–706, April 2006.

[175] W. Pieczynski, "Statistical image segmentation," in *Proceedings of Machine Graphics and Vision*, pp. 261–268, Naleczow, Poland, May 1992.

[176] S. Rajasekaran, "On the convergence time of simulated annealing," Research Report MS-CIS-90-89, University of Pennsylvania, Department of Computer and Information Science, USA, November 1990.

[177] S. Raman, B. Parvin, C. Maxwell, and M. H. Barcellos-Ho, "Geometric approach to segmentation and protein localization in cell cultured assays," in *Advances in Visual Computing*, pp. 427–436, November 2005.

[178] A. Rangarajan and R. Chellappa, "Markov random field models in image processing," in *Handbook of Brain Theory and Neural Networks*, (A. M.A, ed.), pp. 564–567, Cambridge, MA: MIT Press, 1995.

[179] A. Rangarajan, B. Manjunath, and R. Chellappa, "Markov random fields and neural networks with applications in early vision problems," in *Artificial Neural Networks and Statistical Pattern Recognition: Old and New Connections*, (I. Sethi and A. Jain, eds.), Amsterdam: Elsevier Science Publishers, 1991.

[180] B. Reddy and B. Chatterji, "An FFT-based technique for translation, rotation and scale-invariant image registration," *IEEE Transactions on Image Processing*, vol. 5, no. 8, pp. 1266–1271, 1996.

[181] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, vol. 26, pp. 195–239, April 1984.

[182] E. Rignot and R. Chellappa, "Maximum a posteriori classification of multifrequency, multilook, synthetic aperture radar intensity data," *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 10, no. 4, pp. 573–582, 1993.

[183] M. Rochery, I. H. Jermyn, and J. Zerubia, "Higher order active contours and their application to the detection of line networks in satellite imagery," in *Proceedings of Workshop on Variational, Geometric and Level Set Methods in Computer Vision*, New York, NY: ICCV, Nice, France, October 2003.

[184] Y. Rosanov, *Markov Random Fields*. New York, NY: Springer, 1982.

[185] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer, "Optimizing binary MRFs via extended roof duality," in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 1–8, Minneapolis, USA, June 2007.

[186] M. Rousson and N. Paragios, "Shape priors for level set representations," in *Proceedings of European Conference on Computer Vision*, pp. 78–92, Copenhagen, Denmark, 2002.

[187] C. Russell, D. Metaxas, C. Restif, and P. Torr, "Using the P^n Potts model with learning methods to segment live cell images," in *International Conference on Computer Vision*, pp. 1–8, Rio de Janeiro, Brazil, October 2007.

[188] C. Samson, L. Blanc-Feraud, G. Aubert, and J. Zerubia, "A level set model for image classification," *International Journal of Computer Vision*, vol. 40, no. 3, pp. 187–197, 2000.

[189] S. J. Sangwine and R. E. N. Horne, eds., *The Colour Image Processing Handbook*. London: Chapman & Hall, 1998.

[190] M. Schneider, P. Fieguth, W. Karl, and A. Willsky, "Multiscale statistical methods for the segmentation of images," *IEEE Transactions on Image Processing*, vol. 9, pp. 442–455, March 2000.

[191] B. Schölkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge: MIT Press, 2001.

[192] J. Sethian, *Level Set Methods and Fast Marching Methods Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science. Cambridge Monograph on Applied and Computational Mathematics*. Cambridge University Press, 1999.

[193] T. Shima, S. Sugimoto, and M. Okutomi, "Comparison of image alignment on hexagonal and square lattices," in *Proceedings of International Conference on Image Processing*, pp. 141–144, Hong Kong, China, September 2010.

[194] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context," *International Journal of Computer Vision*, vol. 81, pp. 2–23, 2009.

[195] M. Sigelle, C. Bardinet, and R. Ronfard, "Relaxation of classification images by a Markov field technique — application to the geographical classification of Bretagne region," in *Proceedings of European Association of Remote Sensing*, Eger, Hungary, September 1992.

[196] M. Sigelle and R. Ronfard, "Modèles de Potts et relaxation d'images de labels par champs de markov," *Traitement du Signal*, vol. 9, pp. 449–458, March 1993.

[197] T. Simchony, R. Chellappa, and Z. Lichtestein, "Image estimation using 2-D noncausal Gauss Markov random fields," in *Image Restoration*, (A. Katsaggelos, ed.), pp. 109–141, Springer, 1991.

[198] R. H. Swendsen and J.-S. Wang, "Nonuniversal critical dynamics in Monte Carlo simulations," *Physical Review Letters*, vol. 58, pp. 86–88, 1987.

[199] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 1068–1080, June 2008.

[200] H. L. Tan, S. B. Gelfand, and E. J. Delp, "A cost minimization approach to edge detection using simulated annealing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 3–18, January 1991.

[201] A. Trouvé, "Massive parallelization of simulated annealing: A mathematical study," in *Simulated Annealing: Parallelization Techniques*, (R. Azencott, ed.), pp. 145–164, John Wiley & Sons, 1992.

[202] Z. Tu, X. Chen, A. Yuille, and S.-C. Zhu, "Image parsing: Unifying segmentation, detection, and recognition," *International Journal of Computer Vision*, vol. 63, pp. 113–140, July 2005.

[203] Z. Tu and S.-C. Zhu, "Image segmentation by data-driven Markov chain Monte Carlo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 657–673, May 2002.

[204] F. Tupin, H. Maitre, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: Application to road network extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, pp. 434–453, March 1998.

[205] N. Vandenbroucke, L. Macaire, and J. Postaire, "Color image segmentation by supervised pixel classification in a color texture feature space. Application to soccer image segmentation," in *Proceedings of the International Conference on Pattern Recognition*, pp. 621–624, Barcelona, Spain, 2000.

[206] J. Verbeek and B. Triggs, "Scene segmentation with CRFs learned from partially labeled images," in *Proceedings of Advances in Neural Information Processing Systems*, pp. 1553–1560, January 2008.

[207] S. Vicente, V. Kolmogorov, and C. Rother, "Graph cut based image segmentation with connectivity priors," in *Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, June 2008.

[208] Ľubor Ladický, P. Sturgess, K. Alahari, C. Russell, and P. H. Torr, "What, where and how many? combining object detectors and CRFs," in *Proceedings of European Conference on Computer Vision*, (K. Daniilidis, P. Maragos, and N. Paragios, eds.), pp. 424–437, Crete, Greece: Springer, September 2010.

[209] A. Voisin, V. Krylov, G. Moser, S. Serpico, and J. Zerubia, "Classification of very high resolution sar images of urban areas," Technical Report, INRIA Sophia Antipolis Mediterranee, 2011.

[210] A. Voisin, V. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "Multichannel hierarchical image classification using multivariate copulas," in *IS&T/SPIE Electronic Imaging 2012, Proceedings of SPIE, volume 8296, 82960K*, pp. 22–26, San Francisco, USA, January 2012.

[211] A. Voisin, G. Moser, V. Krylov, S. B. Serpico, and J. Zerubia, "Classification of very high resolution SAR images of urban areas by dictionary-based mixture models, copulas and Markov random fields using textural features," in *Proceedings of SPIE*, p. 78300O, 2010.

[212] H. M. Wallach, "Conditional random fields: An introduction," Technical Report MS-CIS-04-21, University of Pennsylvania, USA, February 2004.

[213] Y. Wang and S.-C. Zhu, "Analysis and synthesis of textured motion: Particles and waves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1348–1363, October 2004.

[214] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*. Springer, 2003.

[215] J. H. Woods, "Two-dimensional discrete Markovian fields," *IEEE Transactions on Information Theory*, vol. 18, pp. 232–240, March 1972.

[216] F. Y. Wu, "The Potts model," *Reviews of Modern Physics*, vol. 54, pp. 235–268, January 1982.

[217] C. Xu and J. L. Prince, "Snakes, shapes, gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, pp. 359–369, March 1998.

[218] S. Yu and M. Berthod, "A game strategy approach for image labeling," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 32–37, 1995.

[219] J. Zerubia and R. Chellappa, "Mean field approximation using compound Gauss-Markov random field for edge detection and image restoration," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, Albuquerque, USA, 1990.

[220] J. Zerubia and R. Chellappa, "Mean field annealing using Compound Gauss-Markov Random fields for edge detection and image estimation," *IEEE Transactions on Neural Networks*, vol. 8, pp. 703–709, July 1993.

[221] J. Zerubia and C. Graffigne, *Analyse d'images: filtrage et segmentation*, ch. Segmentation contour: Approches statistiques. Masson, 1995.

[222] J. Zerubia, Z. Kato, and M. Berthod, "Multi-temperature annealing: A new approach for the energy-minimization of hierarchical Markov random field models," in *Proceedings of the International Conference on Pattern Recognition*, pp. 520–522, Jerusalem, Israel, October 1994.

[223] J. Zerubia and F. Ployette, "Detection de contours et restauration d'image par des algorithmes deterministes de relaxation. Mise en oeuvre sur la machine a connexions CM2," Research Report 1291, INRIA, September 1991.
[224] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 884–900, 1996.