Restriction of a Markov Random Field on a Graph and Multiresolution Statistical Image Modeling

Patrick Pérez and Fabrice Heitz

Abstract—**The association of statistical models and multiresolution data analysis in a consistent and tractable mathematical framework remains an intricate theoretical and practical issue. Several consistent approaches have been proposed recently to combine Markov Random Field (MRF) models and multiresolution algorithms in image analysis: renormalization group, subsampling of stochastic processes, MRF's defined on trees or pyramids, etc. For the simulation or a practical use of these models in statistical estimation, an important issue is the preservation of the *local* Markovian property of the representation at the different resolution levels. It is shown in this paper that this key problem may be studied by considering the restriction of a Markov random field (defined on some simple finite nondirected graph) to a part of its original site set. Several general properties of the restricted field are derived. The general form of the distribution of the restriction is given. "Locality" of the field is studied by exhibiting a neighborhood structure with respect to which the restricted field is an MRF. Sufficient conditions for the new neighborhood structure to be "minimal" are derived. Several consequences of these general results related to various "multiresolution" MRF-based modeling approaches in image analysis are presented.**

Index Terms—**Markov Random Fields, multiresolution statistical image modeling, minimal neighborhood, irreducible potential function, restriction of an MRF, subsampling.**

I. Introduction

MARKOV Random Field (MRF) models have been successfully introduced in many fundamental issues of image analysis and computer vision such as image restoration [5], [14], edge detection [13], image segmentation [9], [13], computed tomography [11], surface reconstruction [9], [30], stereovision [2], motion analysis [18], [24], [37], or scene interpretation [33]. The mathematical framework is a statistical one: entities of interest in a given task are described by statistical models (Markov random fields) and Bayesian estimation theory is used to extract the relevant information from the observed images. By defining comprehensive global statistical models, the MRF theory leads to significant improvement over local methods. Markov random fields define an efficient and powerful framework for specifying nonlinear interactions between features of the same nature or of a different one. They help to combine and organize spatial and temporal information by introducing strong generic knowledge about the features to be estimated. When they are associated with the Maximum A Posteriori (MAP) criterion, they lead to the minimization of a global energy function which may exhibit local minima [14]. This minimization is generally performed using deterministic [5] or stochastic [14] relaxation algorithms. Stochastic algorithms may be drastically time-consuming while deterministic schemes often get "stuck" in local minima of the energy function. In addition, it is known that multigrid methods can improve significantly the convergence rate of iterative relaxation schemes [40]. They are useful when the energy to be minimized presents many local minima. It has indeed been conjectured that multiresolution analysis may, to a certain extent, *smooth* the energy landscape. Deterministic relaxation schemes can then be used at coarse scales to get a good initial guess, which may be refined over increasing resolution.

The combination of Markovian models and multigrid methods in a consistent and tractable mathematical framework is an intricate theoretical and practical issue. The first approaches proposed in the literature were essentially based on heuristic associations of MRF models with multiresolution decompositions of the images to be processed. Gaussian pyramids, wavelet decompositions have for instance been used, but generally the same model was considered at each resolution (same parameters, same neighborhood structure, and same potential functions) [2], [24]. Yet, in multigrid implementations of statistical models such as MRF's, the key problem remains the derivation of the model parameters at different scales. When global mathematical consistency is not guaranteed, the parameters and the neighborhood structure associated with the model can only be adjusted over scale in an *ad hoc* way. Gidas [16] has described a consistent mathematical framework for multiscale Markov modeling, based on the *renormalization group* approach. Unfortunately, this standard technique of statistical physics leads in general to hardly tractable computational schemes (apart from particular models and scale transformations [16]) because of the loss of locality of the model at the coarse scales. In [26], Lakshmanan has shown that simple resolution transformations such as subsampling or block averaging also result in a loss of locality

Manuscript received May 15, 1994; revised June 19, 1995. This work was partly supported by MRT (the French Ministry of Research and Technology) and CNRS in the context of the PRC Program "Man Machine Interface" under Contract PMFE 88F1 548, by the GDR TdSI 134, and by Brittany County Council through contribution to student grant. The material in this paper was presented in part at the IEEE Workshop on Multidimensional Signal Processing, Cannes, France, September 1993.

P. Pérez is with IRISA/INRIA–Rennes, Campus Universitaire de Beaulieu, 35042 Rennes Cédex, France.

F. Heitz is with LSIIT/ENSPS, Bd. Sébastien Brant, 67400 Illkirch, France.

Publisher Item Identifier S 0018-9448(96)00563-9.

for Gaussian MRF models. Jeng [21] has studied the loss of locality in a periodic subsampling of MRF models with infinite support. Other approaches introduced recently aim at defining MRF models on hierarchical structures such as trees [3], [6], [28], or other pyramidal graph structures [6], [22], [28].[1] See [17] for a comprehensive overview of hierarchical MRF-based approaches to image modeling.

For a practical use of these different models, an important issue is the preservation of the local Markovian property of the model at the different resolution levels. It is shown here that this key problem may be studied by considering the restriction of a Markov random field to a part of its original site set. We consider MRF's with either discrete or continuous state spaces, defined on simple finite nondirected graphs. The general form of the distribution of the restriction is given. A neighborhood structure with respect to which the restricted field is an MRF is exhibited and sufficient conditions for this structure to be "minimal" are derived. These general results may be applied to the different multiresolution approaches previously mentioned. Several special cases related to various multiresolution MRF-based image analysis approaches are studied.

This study generalizes and unifies in some sense the works of Jeng [21] and Lakshmanan [26], both devoted to the subsampling of MRF's. Lakshmanan [26] studies the subsampling of Gauss–Markov random fields. The mathematical framework described here applies to general classes of nonlinear MRF's and handles several other approaches to multiresolution MRF modeling. In [21], Jeng considers the periodic subsampling of MRF's defined on infinite lattices. MRF's defined on infinite lattices raise several difficult theoretical issues as far as their definition is concerned. One has to introduce the concept of *Gibbs specification* associated with an interaction potential [15]. The existence and uniqueness of a Gibbs distribution, associated with a Gibbs specification, is an intricate problem which cannot be solved with the standard mathematical tools used in [21]. In this paper we focus on *finite support* MRF models defined on arbitrary graphs which are the models usually considered for modeling images. The conditions given here for the preservation of locality in the particular case of subsampling are thus *not equivalent* to the conditions given by Jeng in [21].

The paper is organized as follows. In Section II, we briefly introduce the terminology and notation related to graphs and random fields used in the paper. Markov random fields are defined in association with simple finite nondirected graphs. The general properties of the restriction of a Markov random field to a part of its original site set are described in Section III. The markovianity of the resulting random field is studied and the distribution of the restricted field is derived. General conditions for preservation of the locality of the field are given for finite-support MRF's. Section IV is devoted to various consequences of this general framework in special cases related to statistical multiresolution image modeling. The first consequence concerns the subsampling of MRF's. The general properties of several subsampling schemes are examined. It is shown that most standard subsampling schemes lead to a loss of locality. Examples of subsampling schemes which preserve a local Markov property are also presented. The statistical properties of MRF models defined on trees [3], [6], [28], or other pyramidal graph structures [6], [22], [28], are also studied. The restriction, at a given scale, of these hierarchical models show long-range interactions over the whole field and, as a consequence, cannot be specified locally. The well-known loss of locality induced by renormalization group transformations is also examined within the proposed framework. Section V contains concluding remarks and some open questions. For the sake of conciseness, several proofs have been omitted, that can be found in [38].[2]

[1] Alternate approaches, involving various kinds of coarsening operators on a *single-resolution model*, have also been developed recently [7], [19], [20], [27], [34], [36], [37]. These approaches are not subject to a loss of locality and hence will not be considered in this paper.

II. Markov Random Field on Simple Finite Nondirected Graphs

A. Graph Terminology

A *simple finite nondirected graph* $G = [S, \mathcal{U}]$ is defined as a finite set S of *vertices* or *sites* and a set $\mathcal{U}$ of *edges* (pairs of distinct vertices). If $u = \{s, t\}$, $s \neq t$, is an edge of G, then sites s and t are said to be *adjacent* or *neighboring*, and they are the *endpoints* of u. The *neighborhood system* of G is the mapping $\mathcal{G}$ which associates with each site s its *neighborhood*, i.e., the set of its neighbors.[3] The graph $G = [S, \mathcal{U}]$ will equivalently be denoted by $G = [S, \mathcal{G}]$. This second notation will be used throughout this paper, unless otherwise stated.

The *degree* of a site s, $d_G(s) \triangleq |\mathcal{G}_s|$, is the number of its neighbors. The *degree of the graph* G, $d_G \triangleq \max_{s \in S} d_G(s)$, is the maximal degree of a site of the graph. If all sites have the same degree d_G, the graph (or the neighborhood system) is said to be *regular*. If all sites are mutually neighboring, then the graph is regular of degree $|S| - 1$: it is *complete*.

A *clique* is a subset c of S containing only one site or such that any two sites of c are neighboring. The set of all cliques of G will be denoted by $\mathcal{C}$.

The *subgraph* generated by a subset A of S is the graph G_A whose site set is A, and whose edges are those of G with their two endpoints in A. Its clique set is $\mathcal{C}_A \triangleq \mathcal{C} \cap \mathcal{P}(A)$, where $\mathcal{P}(A)$ is the set of subsets of A. In addition, we will call the *neighborhood of* A the set $\mathcal{G}_A$ of sites in $\overline{A}$[4] having at least one neighbor in A.

A *chain* of length q is a sequence of edges $L = (u_1, \cdots, u_q)$ such that edge u_i $(2 \leq i \leq q-1)$ has one common endpoint with u_{i-1} and a second common endpoint with u_{i+1}, and such that two consecutive edges of the sequence are different. If the two endpoints u_1 and u_q coincide, the chain is called a *cycle*. The *connected components* of graph G are the equivalence classes induced by the equivalence relation $s \mathcal{R} t \iff (s = t$ or there exists a chain joining s and $t)$. The

[2] Reference [38] is available by anonymous ftp at ftp.irisa.fr as file techreports/1994/PI-784.ps.Z.

[3] Due to the definition of edge set $\mathcal{U}$, a neighborhood system has the following properties: $\forall s \in S$, $s \notin \mathcal{G}_s$ and $\forall \{s, r\} \subset S, r \in \mathcal{G}_s \iff s \in \mathcal{G}_r$.

[4] $\overline{A}$ is the complement of A in S: $\overline{A} \triangleq S - A$.

graph is said to be connected if, for any pair $\{s, t\}$ of sites, there exists a chain joining s and t. In this paper we will only consider finite simple nondirected connected graphs.

A *d-dimensional lattice* ($d \in \mathbb{N}^*$) is a subset of $\mathbb{Z}^d$. Such kinds of site sets are widely used in image analysis: bidimensional lattices are often used to represent images (each site corresponds to a *pixel*); tridimensional lattices are used to represent 3D data (in medical imaging or in image sequences for instance). Given a distance D on S, a sequence of regular *neighborhoods of order n with respect to distance D* can be defined on the lattice by the following recurrence [14]:

$$\mathcal{G}_s^1 \triangleq \underset{t \neq s}{\operatorname{argmin}}\ D(s, t);$$
$$\forall n \geq 2, \quad \mathcal{G}_s^{n+1} \triangleq \underset{t \notin \mathcal{G}_s^n \cup \{s\}}{\operatorname{argmin}}\ D(s, t) \cup \mathcal{G}_s^n. \tag{1}$$

Among these different neighborhood systems, the most often used are the first- and second-order neighborhoods (with respect to the Euclidean distance). In the two-dimensional case, they are also called four-neighborhood or eight-neighborhood systems, since they associate four or eight neighbors to any site which is not located on the border of the lattice.

B. Markov Random Fields

Given a finite site set S, we will consider collections $\{x_s, s \in S\}$ of variables indexed by elements of S and belonging to the discrete or continuous state space $\Lambda \subset \mathbb{R}^q$ ($q \in \mathbb{N}^*$). The definition of Λ is of course problem-dependent. In image analysis, various state spaces have been defined, depending on the application field. A recent review on (Markov) random fields and applications may be found in [8]. A measure being defined on the state space Λ, we consider mappings which associate an element of Λ with each site of S:

Definition 1: Let S be a finite site set, and $(\Lambda, \mathcal{E})$ be a measurable space with positive measure κ on the σ-algebra $\mathcal{E}$. A *configuration on S with state space* Λ is a mapping x from S into Λ.

We denote $x_s \triangleq x(s), \forall s \in S$. The state space Λ will generally be a subset of $\mathbb{R}^q$, equipped with its usual σ-algebra and the corresponding measure. The product σ-algebra $\mathcal{T} \triangleq \mathcal{E}^S$ is defined on the set $\Omega \triangleq \Lambda^S$ of all possible configurations, and a probability measure on $\mathcal{T}$ defines a *random field*:

Definition 2: Let S be a finite site set and $(\Lambda, \mathcal{E}, \kappa)$ be a state space. A *random field with site set S and with state space* Λ is any triple $X = (\Omega, \mathcal{T}, \mathrm{P})$ such that $(\Omega, \mathcal{T}) = (\Lambda, \mathcal{E})^S$, and P is a probability measure on the product σ-algebra $\mathcal{T}$.

Restriction to a site subset: Let A be a nonempty subset of S. Consider the canonical surjective mapping from Ω into $(\Omega_A, \mathcal{T}_A) \triangleq (\Lambda, \mathcal{E})^A$, denoted by f_A and which associates with any configuration $x = \{x_s, s \in S\}$ in Ω its *restriction* $x_A \triangleq \{x_s, s \in A\}$. The (marginal) probability measure generated on $\mathcal{T}_A$ by the mapping f_A defines a random field $X_A = (\Omega_A, \mathcal{T}_A, \mathrm{P}_A)$ with site set A and with state space Λ.[5] This field will be called the *restriction of random field X to the subset A*. A standard example of restriction is the *subsampling* of a random field. This case will be considered thoroughly in the following sections. A configuration $x \in \Omega$ will be sometimes denoted by $(x_A, x_{\overline{A}})$ in order to distinguish the "components" corresponding to the sites of A from the others. Finally, for any $T \in \mathcal{T}$, T_A will denote its image under f_A: $T_A = \{x_A, x \in T\}$.

Definition 3: Let $\mathcal{G}$ be a neighborhood system on S. X is *Markovian with respect to* $\mathcal{G}$ (or Markovian on the graph $G = [S, \mathcal{G}]$), if and only if

$$\Pr\{X_s \in T_s | X_r = x_r, r \in S - \{s\}\} = \Pr\{X_s \in T_s | X_r = x_r, r \in \mathcal{G}_s\} \tag{2}$$

for any site s of S, and for any couple (T_s, x) in $\mathcal{E} \times \Omega$.

An intuitive interpretation of the Markovian property (2) is that information on s provided by $S - \{s\}$ is the same as the contextual information provided by the neighbors of s [9], [14]. Markov random fields define global statistical interactions through the combination of "local" interactions.

Let us, however, point out that, according to the previous definition, any random field defined on S is Markovian with respect to the neighborhood system $\mathcal{G}_s = S - \{s\}, \forall s \in S$, which corresponds to the complete graph on S. But, as far as practical applications of these models are concerned, "locality" (i.e., a small degree for the graph) is required for an efficient implementation. The stochastic or deterministic relaxation algorithms generally used in the simulation of these models, may indeed be decomposed in elementary computations only involving a site and its neighbors. The updating of a site thus remains local if a small degree is maintained for the graph. These models are also well-suited to massively parallel implementations provided that the neighborhood structure remains simple and local [32] (the efficiency of parallel schemes decreases rapidly when the chromatic number of the graph becomes large [1]). In the case of a complete graph structure, the computation burden usually becomes prohibitive.

Definition 4: An *interaction* on Ω is a collection of real-valued measurable mappings $V = \{V_A, A \in \mathcal{A}\}$ ($\mathcal{A} \subset \mathcal{P}(S)$ being a covering of S) such that each V_A is $\mathcal{T}_A$-measurable (i.e., it only depends on variables defined on sites of A: $\forall (x, y) \in \Omega \times \Omega, x_A = y_A \Longrightarrow V_A(x) = V_A(y)$).

The real-valued mapping V_A[6] attached to site subset A is, therefore, the composition of the canonical surjective mapping f_A from Ω into Ω_A with some measurable real-valued mapping v_A defined on Ω_A: $V_A(x) = [v_A \circ f_A](x) = v_A(x_A), \forall x \in \Omega$.

Definition 5: Let $\mathcal{G}$ be a neighborhood system on S. X is a *Gibbs random field with respect to* $\mathcal{G}$ (or on the graph $G = [S, \mathcal{G}]$) if and only if there exists an interaction $V =$

[5] If $A = \{s\}$, the restricted field is denoted X_s. It is a Λ-valued random variable on Ω.

[6] In the remainder, V with some subscript will always denote a real-valued measurable function.

$\{V_c, c \in \mathcal{C}\}$ such that:

$$\text{i)} \qquad Z \triangleq \int_\Omega \exp\left\{-\sum_{c\in\mathcal{C}} V_c(x)\right\} \kappa^S(\mathrm{d}x) < +\infty \quad (3)$$

$$\text{ii)} \quad \forall T \in \mathcal{T}$$

$$\mathrm{P}(T) = \frac{1}{Z}\int_T \exp\left\{-\sum_{c\in\mathcal{C}} V_c(x)\right\} \kappa^S(\mathrm{d}x) \quad (4)$$

where κ^S is the product measure generated by κ on $\mathcal{E}^S$. Z is a normalization constant called the *partition function*. The real-valued function

$$U(x) \triangleq \sum_{c\in\mathcal{C}} V_c(x)$$

defined on Ω is called the *energy function* or the *Hamiltonian* of the Gibbs random field X, and V_c is *the potential* associated with clique c. Note that the interaction associated with a Gibbs random field is not unique. Nevertheless, it can be unique under certain *canonicity criteria* [4], [10], [23].

The Hammersley–Clifford theorem drastically simplifies the specification of MRF's [4]:

Theorem 1 [Hammersley–Clifford]: X is a Markov random field with respect to the neighborhood system $\mathcal{G}$, and verifies $\mathrm{P}(T) > 0$, $\forall T \in \mathcal{T}$ (positiveness property) if and only if X is a Gibbs random field with respect to $\mathcal{G}$.

III. Restriction of a MRF to a Subset of Sites

Different approaches have been proposed recently to combine MRF models and multigrid algorithms in image analysis: renormalization group [16], subsampling of stochastic processes [21], [26], MRF's defined on trees or pyramids [3], [6], [22], [28], etc. For the simulation or a practical use of these models in statistical estimation, an important issue is the preservation of a *local* Markovian property of the representation at the different resolution levels. It is shown in the next section that this key problem may be studied by considering the restriction of a Markov random field defined on a simple finite nondirected graph to a part of its original site set. Several general properties of the restricted field are derived. The general form of the distribution of the restriction is given. The local Markovian properties of the field are studied by exhibiting a neighborhood structure with respect to which the restricted field is an MRF. Sufficient conditions for the new neighborhood structure to be "minimal" are derived.

A. Minimal Graph and Irreducible Potential Functions

The markovianity of random field X with respect to a neighborhood system $\mathcal{G}$ on S does not necessarily account for the "real" dependencies between the variables of the field: as a matter of fact, according to Definition 3, X will obviously be Markovian with respect to any neighborhood system $\mathcal{G}'$ such that $\mathcal{G}_s \subset \mathcal{G}'_s$, $\forall s \in S$. In order to have a better characterization of the "locality" of a random field, and following Besag [4], we will consider the "smallest" neighborhood system with respect to which the random field X is Markovian:

Proposition 1: If for any site s, $\mathcal{G}_s$ stands for the set of sites $t \neq s$ such that the function $\Pr\{X_s \in T_s | X_r = x_r, r \in S - \{s\}\}$ defined on $\mathcal{E} \times \Omega_{S-\{s\}}$ depends on x_t,[7] then the collection $\mathcal{G} = \{\mathcal{G}_s, s \in S\}$ is a neighborhood system on S with respect to which X is Markovian. This neighborhood system will be called the *minimal neighborhood system of field* X.

We will also say that X is Markovian on the *minimal graph* $G = [S, \mathcal{G}]$. The simple proof of this proposition is given in [38]. The minimal graph of a Markov random field is useful to capture its "local" character: if the degree of this minimal graph is small with respect to the total number of sites, it means that the random field is highly "local;" on the other hand, if the minimal graph is complete, the random field is not "local" and its simulation may become intractable.

The "minimality" property, like the Markovian property may be expressed using the potential functions of the Gibbs distribution of the field. To this end we will consider functions V_A which really depend on variables x_s, $s \in A$, and which cannot be decomposed in sums of functions depending on less variables. The second point includes in fact the first one. As a matter of fact, if V_A does not depend on variables x_s, $s \in B$ with $B \subset A$, $B \neq A$, then V_A is the sum of a $\mathcal{T}_{A-B}$-measurable function and the zero function which is $\mathcal{T}_B$-measurable.

Definition 6: Let A be a nonempty subset of S. A $\mathcal{T}_A$-measurable function V_A is *irreducible* if it is nonzero and there exists no collection $\{A_i, i \in I\}$ of subsets of A distinct from A, such that

$$V_A = \sum_{i\in I} V_{A_i}$$

with V_{A_i} being some $\mathcal{T}_{A_i}$-measurable function. The collection $V = \{V_A, A \in \mathcal{A}\}$ is an *irreducible interaction* if any V_A is zero or irreducible and $\forall \{s, t\}$ in some $A \in \mathcal{A}$, there exists $A' \in \mathcal{A}$ containing $\{s, t\}$ and such that $V_{A'}$ is nonzero.

The class of irreducible potentials contains the different canonical nonzero potential functions defined by Besag [4], Snell and Kindermann [23], and Descombes [10] (see [38] for the proof). Using this fact, it is easy to prove the following property (see [38] for the detail):

Proposition 2: For any $\mathcal{T}_A$-measurable function V_A, there exists a (nonunique) collection $\{A_i, i \in I\}$ of nonempty subsets of A (distinct from A if V_A is not irreducible) and a collection $\{V_{A_i}, i \in I\}$ of irreducible $\mathcal{T}_{A_i}$-measurable functions, such that

$$V_A = \sum_{i\in I} V_{A_i}.$$

For polynomial models (including Ising model, Potts model, and Gaussian model), such decompositions of the energy

[7] i.e., there exists some $T_s \in \mathcal{E}$, $x_{S-\{s,t\}} \in \Omega_{S-\{s,t\}}$, and $(\lambda, \lambda') \in \Lambda \times \Lambda$ such that

$$\Pr\{X_s \in T_s | X_{S-\{s,t\}} = x_{S-\{s,t\}}, x_t = \lambda\} \neq \Pr\{X_s \in T_s | X_{S-\{s,t\}} = x_{S-\{s,t\}}, x_t = \lambda'\}.$$

functions under concern are provided right away by considering the monomials involved, which are obviously irreducible. Nevertheless, it is not possible to give a practical way of getting these decompositions *in general.*

Using irreducible functions, it is possible to express the minimality of a neighborhood system for a given Markov random field:

Proposition 3: $\mathcal{G}$ is the minimal neighborhood system of X if and only if there exists an irreducible interaction $V = \{V_c, c \in \mathcal{C}\}$ such that X is the Gibbs random field associated with V.

Proof: (See [38] for the details.) Using Moebius inversion formula, the interaction $V = \{V_c, c \in \mathcal{C}\}$ associated with the Gibbs distribution of X can be chosen to be composed of canonical (thus zero or irreducible) clique potentials. Then, we just have to show that $\mathcal{G}$ is minimal if and only if (iff)

$$\forall \{s, t\} \in \mathcal{C}, \quad \exists c \in \mathcal{C}: \{s, t\} \subset c \quad \text{and} \quad V_c \neq 0. \tag{5}$$

Assuming (5), suppose $\mathcal{G}$ is not minimal for X. Due to the local conditional distribution expression [4], this means that there exist two sites s and $t \in \mathcal{G}_s$ such that

$$\sum_{c \in \mathcal{C}: \{s,t\} \subset c} V_c(x)$$

does not depend on x_t. Since set $\{c \in \mathcal{C}: \{s, t\} \subset c$ and $V_c \neq 0\}$ is nonempty according to (5), it contains at least one maximal element for the inclusion, denoted by c_o. For any $x \in \Omega$

$$\begin{aligned} V_{c_o}(x) &= V_{c_o}(x_{c_0}, a) \\ &= \sum_{c \in \mathcal{C}: \{s,t\} \subset c} V_c(x_{c_0}, a) - \sum_{c \neq c_0: \{s,t\} \subset c} V_c(x_{c_0}, a) \end{aligned}$$

with a being some fixed configuration of Ω_{S-c_0}. The first sum on the right-hand side is $\mathcal{T}_{c_0-\{t\}}$-measurable according to the independency assumption, and for any clique $c \neq c_0$ such that $\{s, t\} \subset c$, the mapping $[x \in \Omega \mapsto V_c(x_{c_0}, a)]$ depends only on $x_{c \cap c_0}$ where $c \cap c_0 \subset c_0$ is different from c_0 since c cannot contain the maximal clique c_0. The previous decomposition of V_{c_0} then contradicts its irreducibility.

Assume G is minimal. Due to the expression of the local conditional distribution, this implies that for any $s \in S$ and $t \in \mathcal{G}_s$

$$\sum_{c \in \mathcal{C}: \{s,t\} \subset c} V_c(x)$$

is not independent of x_t. This implies that the potentials associated with cliques containing $\{s, t\}$ cannot be all zero simultaneously: there exists $c \in \mathcal{C}$ such that $\{s, t\} \in c$ and $V_c \neq 0$. ∎

B. Properties of the Restricted Field

Let $X = (\Omega, \mathcal{T}, \mathrm{P})$ be a Markov random field on a finite connected *minimal* graph $G = [S, \mathcal{G}]$, associated with the irreducible interaction $V = \{V_c, c \in \mathcal{C}\}$, i.e., a collection of zero or irreducible potential functions verifying

$$\forall \{s, t\} \in \mathcal{C}, \quad \exists c \in \mathcal{C}: \{s, t\} \subset c \quad \text{and} \quad V_c \neq 0. \tag{6}$$

Let A be a nonempty subset of S, $A \neq S$. We are interested in the statistical properties of the restricted field $X_A = (\Omega_A, \mathcal{T}_A, \mathrm{P}_A)$. A graph structure on A with respect to which this random field is Markovian will first be exhibited. Then, sufficient conditions for this graph to be the minimal graph of X_A will be given.

1) Distribution of the Restriction and Markovianity:

Theorem 2: The restricted field X_A is Markovian on graph $\tilde{G}_A = [A, \tilde{\mathcal{G}}]$ in which two different sites of A are neighbors if and only if they are neighbors in G or if they belong to the neighborhood of a same connected component of $\overline{A} = S - A$.

Proof: Let us denote by $\overline{A}_i$, $i \in I$ the connected components of $\overline{A}$. Let c be an arbitrary clique of $\mathcal{C}$. It cannot contain simultaneously two sites belonging to two different connected components, hence its intersection with $\overline{A}$ is either empty or included in only one connected component of $\overline{A}$ (i.e., there exists an unique $i \in I$ such that $(c \cap \overline{A}) \subset \overline{A}_i$). The clique set $\mathcal{C}$ can thus be partitioned in the following way:

$$\mathcal{C} = \bigcup_{i \in I} \underbrace{\{c \in \mathcal{C} | c \cap \overline{A}_i \neq \emptyset\}}_{\triangleq \mathcal{C}_i} \bigcup \underbrace{\{c \in \mathcal{C} | c \subset A\}}_{\triangleq \mathcal{C}_A}. \tag{7}$$

Using this decomposition, the marginal probability measure P_A defining X_A is rewritten. Let T_A be some element of $\mathcal{T}_A$

$$\begin{aligned} \mathrm{P}_A(T_A) &= \mathrm{P}[f_A^{-1}(T_A)] \\ &= \frac{1}{Z} \int_{T_A} \int_{\Omega_{\overline{A}}} \exp\left\{-\sum_{c \in \mathcal{C}} V_c(x_A, x_{\overline{A}})\right\} \\ &\quad \cdot \kappa^A(\mathrm{d}x_A)\, \kappa^{\overline{A}}(\mathrm{d}x_{\overline{A}}) \\ &= \frac{1}{Z} \int_{T_A} \left[\exp\left\{-\sum_{c \in \mathcal{C}_A} V_c(x_A, a_{\overline{A}})\right\}\right. \\ &\quad \cdot \int_{\Omega_{\overline{A}}} \prod_{i \in I} \exp\left\{-\sum_{c \in \mathcal{C}_i} V_c(x_A, x_{\overline{A}})\right\} \\ &\quad \left. \cdot \kappa^{\overline{A}}(\mathrm{d}x_{\overline{A}})\right] \kappa^A(\mathrm{d}x_A) \end{aligned}$$

where $a \in \Omega$ is an arbitrary configuration, since

$$\exp\left\{-\sum_{c \in \mathcal{C}_A} V_c(x)\right\}$$

is independent of $x_{\overline{A}}$. Then

$$\begin{aligned} \mathrm{P}_A(T_A) = \frac{1}{Z} \int_{T_A} &\left[\exp\left\{-\sum_{c \in \mathcal{C}_A} V_c(x_A, a_{\overline{A}})\right\}\right. \\ &\cdot \prod_{i \in I} \int_{\Omega_{\overline{A}_i}} \exp\left\{-\sum_{c \in \mathcal{C}_i} V_c(x_A, x_{\overline{A}_i}, a_{\overline{A}-\overline{A}_i})\right\} \\ &\left. \cdot \kappa^{\overline{A}_i}(\mathrm{d}x_{\overline{A}_i})\right] \kappa^A(\mathrm{d}x_A) \end{aligned}$$

since

$$\exp\left\{-\sum_{c \in \mathcal{C}_i} V_c(x)\right\}$$

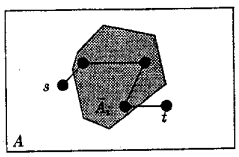


Fig. 1. Sites s and t of A which become neighbors in the restriction of X to A.

is independent of $x_{\overline{A}-\overline{A}_i}$.

$$\int_{\Omega_{\overline{A}_i}} \exp\left\{-\sum_{c\in\mathcal{C}_i} V_c(x_A, x_{\overline{A}_i}, a_{\overline{A}-\overline{A}_i})\right\} \kappa^{\overline{A}_i}(\mathrm{d}x_{\overline{A}_i})$$

being a positive function of variables $x_{\mathcal{G}_{\overline{A}_i}}$, one can define for any $i \in I$ the following $\mathcal{T}_{\mathcal{G}_{\overline{A}_i}}$-measurable function:

$$\begin{aligned} V_i: \Omega_A &\longrightarrow \mathbb{R} \\ x_A &\longmapsto V_i(x_A) \\ &\triangleq -\ln\left[\int_{\Omega_{\overline{A}_i}} \exp\left\{-\sum_{c\in\mathcal{C}_i} V_c(x_A, x_{\overline{A}_i}, a_{\overline{A}-\overline{A}_i})\right\}\right. \\ &\qquad \left. \cdot\, \kappa^{\overline{A}_i}(\mathrm{d}x_{\overline{A}_i})\right] \end{aligned} \tag{8}$$

where a is an arbitrary configuration of Ω. The expression of the image measure becomes

$$\begin{aligned} \mathrm{P}_A(T_A) = \frac{1}{Z}\int_{T_A} \exp\left\{-\sum_{c\in\mathcal{C}_A} V_c(x_A, a_{\overline{A}}) - \sum_{i\in I} V_i(x_A)\right\} \\ \cdot\, \kappa^A(\mathrm{d}x_A). \end{aligned} \tag{9}$$

Let us now introduce the graph $\tilde{G}_A = [A, \tilde{\mathcal{G}}]$ in which two different sites s and t of A are neighbors if and only if they are neighbors in G (i.e., $t \in \mathcal{G}_s$), or they belong to the neighborhood of a same connected component of $\overline{A}$ (i.e., $\exists i \in I: \{s, t\} \subset \mathcal{G}_{\overline{A}_i}$), that is, there exists a chain joining s and t all sites of which, with the exception of its endpoints s and t, belong to $\overline{A}$ (Fig. 1).

The neighborhood system of $\tilde{G}_A$ is thus defined as follows:

- if $s \in A$ is not in the neighborhood of $\overline{A}$, i.e., if s is an "interior" site of A, then $\mathcal{G}_s \subset A$ and

$$\tilde{\mathcal{G}}_s \triangleq \mathcal{G}_s \tag{10}$$

- if $s \in A$ is in the neighborhood of $\overline{A}$, i.e., if s is on the "border" of A, then

$$\tilde{\mathcal{G}}_s \triangleq (\mathcal{G}_s \cap A) \cup (\cup_{i\in K} \mathcal{G}_{\overline{A}_i} - \{s\}) \tag{11}$$

where K is the set of indices of connected components of $\overline{A}$ which have s in their neighborhood: $s \in \cap_{i\in K} \mathcal{G}_{\overline{A}_i}$ and $s \notin \cup_{i\in I-K} \mathcal{G}_{\overline{A}_i}$.

The clique set of this graph, denoted by $\tilde{\mathcal{C}}_A$, contains all cliques of $\mathcal{C}$ included in A and the neighborhoods of all connected components of $\overline{A}$: $\tilde{\mathcal{C}}_A \supset \mathcal{C}_A \cup \{\mathcal{G}_{\overline{A}_i}, i \in I\}$. For any clique c in $\tilde{\mathcal{C}}_A$, let us introduce the measurable function $\tilde{V}_c$ defined on Ω_A as

$$\tilde{V}_c(x_A) \triangleq \begin{cases} V_c(x_A, a_{\overline{A}}), & \text{if } c\in\mathcal{C}_A \text{ and } c\notin\{\mathcal{G}_{\overline{A}_i}, i\in I\} \\ V_i(x_A), & \text{if } c=\mathcal{G}_{\overline{A}_i} \text{ and } \mathcal{G}_{\overline{A}_i}\notin\mathcal{C}_A \\ V_i(x_A) + V_c(x_A, a_{\overline{A}}), & \text{if } c=\mathcal{G}_{\overline{A}_i} \text{ and } \mathcal{G}_{\overline{A}_i}\in\mathcal{C}_A \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

The collection $\tilde{V} \triangleq \{\tilde{V}_c, c \in \tilde{\mathcal{C}}_A\}$ obviously constitutes an interaction on Ω_A, and (9) becomes

$$\mathrm{P}_A(T_A) = \frac{1}{Z}\int_{T_A} \exp\left\{-\sum_{c\in\tilde{\mathcal{C}}_A} \tilde{V}_c(x_A)\right\} \kappa^A(\mathrm{d}x_A)$$

for any $T_A \in \mathcal{T}_A$, with

$$Z = \int_{\mathcal{T}_A} \exp\left\{-\sum_{c\in\tilde{\mathcal{C}}_A} \tilde{V}_c(x_A)\right\} \kappa^A(\mathrm{d}x_A) < +\infty.$$

This proves that X_A is Markovian on the graph $\tilde{G}_A$. ■

2) Sufficient Conditions of Minimality of $\tilde{G}_A$: In order to find sufficient conditions for graph $\tilde{G}_A$ to be the *minimal* graph of X_A, let us apply the decomposition property (Proposition 2) to the new $\mathcal{T}_{\mathcal{G}_{\overline{A}_i}}$-measurable functions V_i: for any $i \in I$ there exists a collection $\{A_{i,k}, k \in K_i\}$ of $|K_i|$ nonempty subsets of $\mathcal{G}_{\overline{A}_i}$ and there exists a collection $\{V_{i,k}, k \in K_i\}$ of irreducible $\mathcal{T}_{A_{i,k}}$-measurable functions such that

$$V_i = \sum_{k\in K_i} V_{i,k}, \forall i \in I.$$

Using these decompositions, one can rewrite (9)

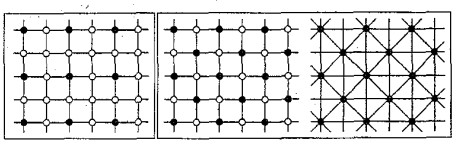
$$\begin{aligned} \mathrm{P}_A(T_A) = \frac{1}{Z}\int_{T_A} \exp\left\{-\sum_{c\in\mathcal{C}_A} V_c(x_A, a_{\overline{A}}) - \sum_{i\in I}\sum_{k\in K_i} V_{i,k}(x_A)\right\} \\ \cdot\, \kappa^A(\mathrm{d}x_A). \end{aligned} \tag{13}$$

For any clique $c \in \tilde{\mathcal{C}}_A$, let us introduce the measurable function $\tilde{V}'_c$ defined on Ω_A as

$$\tilde{V}'_c(x_A) \triangleq \begin{cases} V_c(x_A, a_{\overline{A}}), & \text{if } c \in \mathcal{C}_A \text{ and } c \notin \{A_{i,k}, i \in I, k \in K_i\} \\ V_{i,k}(x_A), & \text{if } c = A_{i,k} \text{ and } A_{i,k} \notin \mathcal{C}_A \\ V_{i,k}(x_A) + V_c(x_A, a_{\overline{A}}), & \text{if } c = A_{i,k} \text{ and } A_{i,k} \in \mathcal{C}_A \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

The collection $\tilde{V}' \triangleq \{\tilde{V}'_c, c \in \tilde{\mathcal{C}}_A\}$ constitutes an interaction on Ω_A, and (13) may be rewritten as

$$\mathrm{P}_A(T_A) = \frac{1}{Z}\int_{T_A} \exp\left\{-\sum_{c\in\tilde{\mathcal{C}}_A} \tilde{V}'_c(x_A)\right\} \kappa^A(\mathrm{d}x_A)$$



(a) (b)

Fig. 2. Decimation of a regular bidimensional lattice with the four-neighborhood system (•: site of A, ○: site of $\overline{A}$): (a) by a factor two in each direction and (b) quincunx decimation and graph associated with the restricted random field.

for any $T_A \in \mathcal{T}_A$, with

$$Z = \int_{\mathcal{T}_A} \exp\left\{ - \sum_{c \in \tilde{\mathcal{C}}_A} \tilde{V}'_c(x_A) \right\} \kappa^A(\mathrm{d}x_A) < +\infty.$$

Theorem 3: $\tilde{G}_A$ is the minimal graph of X_A if:

1) (H1) for any $A_{i,k} = c$ belonging to $\mathcal{C}_A$,

$$[x_A \mapsto V_{i,k}(x_A) + V_c(x_A, a_{\overline{A}})]$$

is irreducible;

2) (H2) for any two different sites s and t of $\mathcal{G}_{\overline{A}_i}$, there exists at least one subset $A_{i,k}$ which contains them.

Proof: Let us assume that (H1) and (H2) are verified. Due to the nullity or irreducibility of functions V_c and the irreducibility of functions $V_{i,k}$, (H1) implies that $\tilde{V}'$ is a collection of irreducible or zero functions. Let $\{s, t\}$ be some element of $\tilde{\mathcal{C}}_A$:

- If $\{s, t\} \not\subset \mathcal{G}_{\overline{A}_i}$, then s and t are neighboring in G and there exists a clique $c \in \mathcal{C}$ including $\{s, t\}$ and such that $V_c \neq 0$ (since G is the minimal graph of X). Furthermore, $c \in \mathcal{C}_A$ since $\{s, t\} \not\subset \mathcal{G}_{\overline{A}}$. Finally, $c \in \tilde{\mathcal{C}}_A$ and $\tilde{V}'_c(\bullet) = V_c(\bullet, a_{\overline{A}}) \neq 0$.
- If $\{s, t\} \subset \mathcal{G}_{\overline{A}_i}$ (with $i \in I$), according to (H2) there exists $k \in K_i$ such that $c = A_{i,k}$ contains s and t. We come back to the previous case and $\tilde{V}'_c \neq 0$.

The collection $\tilde{V}'$ is then an irreducible interaction on Ω_A which implies that $\tilde{G}_A$ is the minimal graph of X_A. ■

In case the set $\overline{A}$ of "discarded" sites is connected and all sites of A are in the neighborhood of $\overline{A}$ (i.e., $A = \mathcal{G}_{\overline{A}}$), it is straightforward that $\tilde{G}_A$ is complete:

Corollary 1: If A has no interior sites and if $\overline{A}$ is connected, then assuming (H1) and (H2), the complete graph on A is the minimal graph of X_A.

In this case, the restricted random field X_A is not local. We can apply this corollary to a widespread situation in the framework of multiresolution image modeling with Markovian models: G being a bidimensional regular lattice associated with a four-neighborhood system, the decimation by a factor of two in each direction (Fig. 2(a)) discards a connected set whose neighborhood is A (it is still true in dimension $d \geq 2$). Such a subsampling turns a Markov random field on G into a nonlocal random field assuming (H1) and (H2).

IV. Consequences on Special Cases

In this section, we present various consequences of the general framework developed in Section III within special cases related to multiresolution statistical image modeling. The first consequence concerns the subsampling of MRF's. The general properties of several subsampling schemes are examined. It is shown that most standard subsampling schemes lead to a loss of locality. Examples of subsampling schemes which preserve a local Markov property are also presented. The statistical properties of MRF models defined on trees [3], [6], [28], or other pyramidal graph structures [6], [22], [28], are studied. The restriction, at a given scale, of a model defined on such a hierarchical structure shows long-range interactions over the whole field and, as a consequence, may not be specified locally. The general properties of the renormalization group approach are also examined. It is shown that the usual scale transforms used in this approach yield a loss of locality for the random fields induced at the coarse resolution levels.

A. Coloring and Subsampling

The *coloring* of $G = [S, \mathcal{G}]$ is the assignment of a color to each site such that two neighboring sites are not of the same color, i.e., it is a partition of S in p subsets $C_1, \cdots, C_p$, that we will call *monochromatic components*, such that: $\forall k \in \{1, \cdots, p\}, \forall \{s, t\} \subset C_k, s \notin \mathcal{G}_t$. In other words, the connected components of any C_k are singletons.

1) Restriction to the Complement of a Monochromatic Component: G being colored with $p \geq 2$ colors, consider first the restriction of X to the complement of a monochromatic component C_{k_0}: $A = \cup_{k \neq k_0} C_k$ and $\overline{A} = C_{k_0}$.

Proposition 4: Let A be the complement of some monochromatic component of G. The restricted random field X_A is Markovian on a graph whose degree is less or equal to $d_G(d_G - 1)$.

Proof: We are going to find an upper bound to the degree of graph $\tilde{G}_A$ with respect to which X_A is Markovian. Let s be an arbitrary site of A, belonging to the neighborhood of $\overline{A}$: s has $q \geq 1$ neighbors in $\overline{A}$, denoted by $t_1, \cdots, t_q$, with $q = |\mathcal{G}_s \cap \overline{A}| = d_G(s) - |\mathcal{G}_s \cap A|$. Each singleton $\{t_i\}$ is a connected component of $\overline{A}$: $\mathcal{G}_{t_i} \subset A$. According to (11), the neighborhood of s in graph $\tilde{G}_A$ is $\tilde{\mathcal{G}}_s = (\mathcal{G}_s \cap A) \cup (\cup_{i=1}^{i=q} \mathcal{G}_{t_i} - \{s\})$. Then

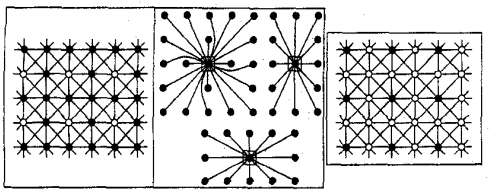
$$\begin{aligned} d_{\tilde{G}_A}(s) &\leq |\mathcal{G}_s \cap A| + q(d_G - 1) \\ &= |\mathcal{G}_s \cap A| + [d_G(s) - |\mathcal{G}_s \cap A|](d_G - 1) \\ &\leq d_G(d_G - 1) - (d_G - 2)|\mathcal{G}_s \cap A|. \end{aligned}$$

This inequality still holds for some interior site of A, since its degree in $\tilde{G}_A$ is the same as its degree in G. Finally, we get an upper bound on the degree of $\tilde{G}_A$

$$\begin{aligned} d_{\tilde{G}_A} &\leq d_G(d_G - 1) - (d_G - 2) \min_{s \in A} |\mathcal{G}_s \cap A| \\ &\leq d_G(d_G - 1). \end{aligned} \tag{15}$$

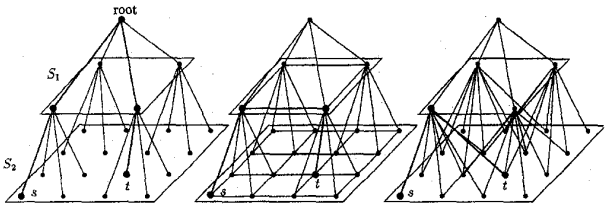
■

To illustrate this kind of restriction, let us consider the coloring with four colors of a regular bidimensional lattice associated with an eight-neighborhood system ($d_G = 8$). Given a Markov random field on such a graph, its restriction to the complement of one of the four monochromatic components



(a) (b)

Fig. 3. Decimation of a bidimensional regular lattice associated with an eight-neighborhood system (•: site of A, ○: site of $\overline{A}$): (a) restriction to the complement of a monochromatic component and different kinds of neighborhood in $\tilde{G}_A$ and (b) decimation by a factor 2 in each direction.



(a) (b) (c)

Fig. 4. Restriction to the deepest level of a pyramidal graph: (a) quad-tree, (b) and (c) pyramidal graphs based on a quad-tree. In each case, sites s and t become neighbors in the restriction to the lower level, through bold chains (among others in (b) and (c)) joining them.

is a Markov random field on a graph of degree less or equal to 56. $\tilde{G}_A$ is in fact a nonregular graph of degree 20 (Fig. 3(a)).

Let us return to the general case. If, in addition, the graph G is only colored with two colors, then A is the second monochromatic component: given some site s of A, its whole neighborhood $\mathcal{G}_s$ is in $\overline{A}$. Therefore, its neighborhood in $\tilde{G}_A$ is

$$\tilde{\mathcal{G}}_s = \bigcup_{t \in \mathcal{G}_s} \mathcal{G}_t - \{s\}.$$

That means that neighbors of s in $\tilde{G}_A$ are the neighbors of its neighbors in G (except s itself). A widespread example of this situation is given by the *quincunx subsampling* of a regular d-dimensional lattice associated with a second-order neighborhood system. If $d = 2$, the graph $\tilde{G}_A$ associated with the restricted random field is regular of degree 8 (Fig. 2(b)).

2) Restriction to a Monochromatic Component: We now focus on the restriction of a Markov random field on G to some monochromatic component C_{k_0}: $A = C_{k_0}$ and $\overline{A} = \cup_{k \neq k_0} C_k$. Assuming the neighborhood of any site of G is a connected subset, one can establish the following result:

Proposition 5: Let A be a monochromatic component of G. If $\mathcal{G}_s$ is connected $\forall\, s \in S$, then, under (H1) and (H2), the restricted field X_A has the complete graph on A as minimal graph.

Proof: Since A is a monochromatic component, the neighborhood of any site of A belongs to $\overline{A}$ and due to the assumption of connectedness of all neighborhoods, $\overline{A}$ is connected. As a matter of fact, for any two sites s and t of $\overline{A}$, since G is connected, there exists a chain

$$L = (\{s_1 = s,\, s_2\},\, \{s_2,\, s_3\},\, \cdots,\, \{s_{q-1},\, s_q = t\})$$

joining s and t. Let us assume that some s_j $(1 < j < q)$ belongs to A. Since s_{j-1} and s_{j+1} belong to the neighborhood of s_j which is a connected subset, there exists a chain l included in $\mathcal{G}_{s_j}$ and joining s_{j-1} and s_{j+1}. Since $\mathcal{G}_{s_j}$ is included in $\overline{A}$, l is also included in $\overline{A}$. Therefore, the new chain L' obtained by replacing $(\{s_{j-1},\, s_j\},\, \{s_j,\, s_{j+1}\})$ by l in L, joins s and t, and has one site less than L in A. By iterating this process, we construct a chain included in $\overline{A}$ and whose endpoints are s and t: $\overline{A}$ is a connected site subset. We can then apply Corollary 1: assuming (H1) and (H2), the restricted field X_A has the complete graph on A as minimal graph. ■

The assumption of connectedness of all neighborhoods is often verified. For example, in case S is a regular d-dimensional lattice $(d > 1)$ associated with a neighborhood system of order greater than one, it is easy to see that any neighborhood in $G = [S, \mathcal{G}]$ is connected. This provides the following result:

Corollary 2: If the minimal graph of X is a regular d-dimensional lattice with a neighborhood system of order greater than 1, and A is some monochromatic component of the corresponding graph, then, under (H1) and (H2), the restriction X_A has the complete graph on A as minimal graph.

The standard decimation by a factor of two in each direction of a regular bidimensional lattice with the eight-neighborhood system (Fig. 3(b)) corresponds to this case: it generally transforms a Markov random field on this graph into a field which is no more local.

B. Restrictions on Trees and Pyramidal Graphs

Trees and "pyramidal graphs" provide a useful hierarchical structure for representing stochastic processes. They have been introduced recently in signal processing [3] as well as in image processing [28] to support MRF models. They are the basis for the development of new efficient multigrid image analysis techniques that have been applied to multispectral image classification [6], [25], and optical flow computation [29].

A *tree* is a connected graph without cycles. One can show that a graph is a tree if and only if there exists one and only one chain between any two vertices. Some reference site (the *root*) being chosen, trees naturally support a simple hierarchical structure by associating to each site s the length, denoted $l(s)$, of the chain joining it to the root. The father of s is its unique neighbor t such that $l(t) = l(s) - 1$. The remaining neighbors, called its *children*, are at "distance" $l(s) + 1$ to the root. The very simple *pyramidal graph* that is the most widely used in image analysis, is the *quad-tree*. It is a tree in which each site (apart from those of the last "level") has four children (Fig. 4(a)). More sophisticated pyramidal graphs have been defined by adding in a special way extra edges to a tree. Level k of a tree being the subset $S_k \triangleq \{s \in S : l(s) = k\}$, all sites of a given level are mutually nonneighboring. Then, the extra connections may concern pairs of sites of the same level. For example, each level may be turned into a connected subgraph [22] (see, for example, Fig. 4(b)). The extra edges may also be added between consecutive levels, increasing for each site the number of children, and the number of fathers as well [6] (see, for example, Fig. 4(c)). In any case, the new graph thus obtained is not a tree any more.

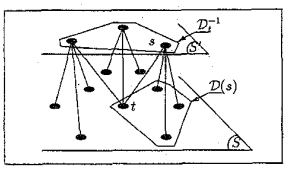


Fig. 5. Example of "link" between S and S' according to a local stochastic transformation of a random field with set site S into a random field with site set S'.

Let us consider a Markov random field on a tree or a pyramidal graph with root r, and let us consider its restriction to a given level S_k $(k > 0)$: $A = S_k$. The reunion $\overline{A}^- \triangleq \cup_{l=0,\cdots,k-1} S_l$ of the upper levels (i.e., levels toward the root) is a connected component of $\overline{A}$. Moreover, any site of A has a neighbor (one of its fathers) in $\overline{A}^-$: $A = \mathcal{G}_{\overline{A}^-}$. If (H1) and (H2) are verified, according to Corollary 1, the restricted field X_A has the complete graph on the level k as minimal graph: for this field, all sites at level k become neighbors. Fig. 4 gives an illustration of this phenomenon of neighborhood through upper levels, in the restriction to the deepest level of a pyramidal graph.

Thus when handling a "global" Markovian model defined on a tree-based hierarchical structure, the simulation of the restriction of the original field at a given level is generally not tractable. For this class of multiresolution models, other techniques, described for instance in [6], [25], [28], [29], make use of the whole hierarchical structure to derive efficient algorithms.

C. Stochastic Transformations and the Renormalization Group Approach

Of major interest in Statistical Physics, the *Renormalization Group Transform* deals with *coarsening* of MRF's into "reduced" versions, according to local, deterministic, or stochastic transformations (see [12], [16], [31], [34], [35], [39] for examples of applications to image analysis problems). The case of (deterministic) decimations falls into the scope of Section IV-A. We consider here stochastic transformations.

Let $X = (\Omega, \mathcal{T}, \mathrm{P})$ be an MRF on the minimal graph $G = [S, \mathcal{G}]$ with irreducible interaction $V = \{V_c, c \in \mathcal{C}\}$. Let S' be another site set and $(\Lambda', \mathcal{E}', \kappa')$ be another state space. With each site $s \in S'$ a ("small") nonempty subset $\mathcal{D}(s)$ of S is associated. For any site $t \in S$, we will denote by $\mathcal{D}_t^{-1}$ the subset $\{s \in S' | t \in \mathcal{D}(s)\}$ of S' (Fig. 5).

Consider a random field $Y = (\Omega', \mathcal{T}', \mathrm{P}')$ with site set S', with state space Λ', conditionally defined as

$$\Pr\{Y \in T' | X = x\} = \prod_{s \in S'} \underbrace{\Pr\{Y_s \in T'_s | X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\}}_{>0, \forall s \in S'} \tag{16}$$

for any $T' \in \mathcal{T}'$ and for any configuration $x \in \Omega$. The positiveness condition enables us to assume local conditional distributions of the following form:

$$\Pr\{Y_s \in T'_s | X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\} = \int_{T'_s} \exp\{-l_s[x_{\mathcal{D}(s)}, y_s]\} \kappa'(\mathrm{d}y_s) \tag{17}$$

where, for any site $s \in S'$, l_s is a measurable real-valued function from $[\Omega_{\mathcal{D}(s)}, \mathcal{T}_{\mathcal{D}(s)}, \kappa^{\mathcal{D}(s)}] \times (\Lambda', \mathcal{E}', \kappa')$. This is always true in the discrete case, with

$$l_s[x_{\mathcal{D}(s)}, y_s] \triangleq -\ln \Pr\{Y_s = y_s | X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\}.$$

In the continuous case, this assumption means that the local conditional distributions have densities. According to Proposition 2, for any site $s \in S'$, there is a collection $\{A_{s,k}, k \in K_s\}$ of nonempty subsets of $\mathcal{D}(s) \cup \{s\}$ and a collection $\{V_{s,k}, k \in K_s\}$ of irreducible $\mathcal{T}_{A_{s,k}}$-measurable functions such that

$$l_s[x_{\mathcal{D}(s)}, y_s] = \sum_{k \in K_s} V_{s,k}(x, y), \forall (x, y) \in \Omega^* \triangleq \Omega \times \Omega'.$$

Combined with (16), (17), it provides

$$\Pr\{Y \in T' | X = x\} = \int_{T'} \exp\left\{-\sum_{s \in S'} \sum_{k \in K_s} V_{s,k}(x, y)\right\} \kappa'^{S'}(\mathrm{d}y). \tag{18}$$

Let $X^* = (X, Y)$ be the random field with site set $S^* \triangleq S \cup S'$, whose restriction to S is the random field X, and whose restriction to S' is the random field Y. Its probability measure P^* comes from Bayes' rule and (18)

$$\mathrm{P}^*(T, T') = \frac{1}{Z} \int_{T \times T'} \cdot \exp\left\{-\sum_{c \in \mathcal{C}} V_c(x) - \sum_{s \in S'} \sum_{k \in K_s} V_{s,k}(x, y)\right\} \cdot \kappa^S(\mathrm{d}x)\, \kappa'^{S'}(\mathrm{d}y), \tag{19}$$

for any $(T, T') \in \mathcal{T} \times \mathcal{T}'$. Let $G^* = [S^*, \mathcal{G}^*]$ be the graph with site set S^* and neighborhood system

$$\begin{cases} \forall s \in S, & \mathcal{G}_s^* \triangleq \mathcal{G}_s \cup \mathcal{D}_s^{-1} \\ & \quad \cup \{t \in S - \{s\} | \mathcal{D}_s^{-1} \cap \mathcal{D}_t^{-1} \neq \emptyset\} \\ \forall s \in S', & \mathcal{G}_s^* \triangleq \mathcal{D}(s). \end{cases} \tag{20}$$

For each clique c of its clique set $\mathcal{C}^*$, introduce the measurable function V_c^* defined on Ω^* as

$$\forall (x, y) \in \Omega^*, V_c^*(x, y) \triangleq \begin{cases} V_c(x), & \text{if } c \in \mathcal{C} \text{ and } c \notin \{A_{s,k}, s \in S', k \in K_s\} \\ V_{s,k}(x, y), & \text{if } c = A_{s,k} \text{ and } A_{s,k} \notin \mathcal{C} \\ V_{s,k}(x, y) + V_c(x), & \text{if } c = A_{s,k} \text{ and } A_{s,k} \in \mathcal{C} \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

The collection $V^* \triangleq \{V_c^*, c \in \mathcal{C}^*\}$ constitutes an interaction on Ω^*. If, i) for any $A_{s,k}$ belonging to $\mathcal{C}$, the $\mathcal{T}_{A_{s,k}}$-measurable function $[(x, y) \mapsto V_{s,k}(x, y) + V_c(x)]$ is irreducible or zero, and ii) $\forall \{s, t\} \in \mathcal{C}^*, \exists c \in \mathcal{C}^*: \{s, t\} \subset c$, and $V_c^* \neq 0$, then V^* is an irreducible interaction and X^* is an MRF on the minimal graph G^*, according to distribution (19). Random field Y is the restriction of X^* to S'. G being connected, S is a connected subset of G^* and each site of S' has at least one neighbor in S: $\mathcal{G}_S^* = S'$. Therefore, one can apply Corollary 1: if X^* fulfills conditions (H1) and (H2), then its restriction Y to S' has the complete graph on S' as minimal graph.

For instance, in [16], Gidas considers an MRF on a regular bidimensional lattice associated with an eight-neighborhood system, which is minimal and he proposes the following stochastic local transformation:

$$\Pr\{Y = y | X = x\} = \prod_{s \in S'} \frac{\exp\left\{\rho y_s \sum_{t \in \mathcal{D}(s)} x_t\right\}}{\sum_{\lambda \in \Lambda'} \exp\left\{\rho \lambda \sum_{t \in \mathcal{D}(s)} x_t\right\}} \tag{22}$$

where ρ is a positive parameter. Thus for any site $s \in S'$

$$l_s[x_{\mathcal{D}(s)}, y_s] = -\sum_{t \in \mathcal{D}(s)} \rho y_s x_t$$

which is a sum of irreducible $\mathcal{T}_{\{s,t\}}$-measurable functions, with $t \in \mathcal{D}(s)$. Subsets $\mathcal{D}(s)$ which are 2×2 square blocks are then, at the same time, cliques of the eight-neighborhood system associated with S ($\forall s \in S'$, $\mathcal{D}(s) \in \mathcal{C}$) and they define a partition of S, i.e., each site $t \in S$ has a unique father $f(t)$ in S': $\mathcal{D}_t^{-1} = \{f(t)\}$. Therefore, the neighborhood system $\mathcal{G}^*$ of graph G^* is given by

$$\begin{cases} \forall s \in S, & \mathcal{G}_s^* = \mathcal{G}_s \cup \{f(t)\} \\ \forall s \in S', & \mathcal{G}_s^* = \mathcal{D}(s). \end{cases} \tag{23}$$

The interaction $V^* = \{V_c^*, c \in \mathcal{C}^*\}$ attached to the clique set of G^* and defined as

$$\forall (x, y) \in \Omega^*, \; V_c^*(x, y) \triangleq \begin{cases} V_c(x), & \text{if } c \in \mathcal{C} \\ -\rho y_s x_t, & \text{if } c = \{s, t\} \text{ with } s \in S' \text{ and } t \in \mathcal{D}(s) \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

is composed of zero and irreducible functions which obviously verify ii): X^* is an MRF on the minimal graph G^*. If it fulfills conditions (H1) and (H2) with respect to S', then the random field Y (restriction of X^* to S') is not local and admits the complete graph on S' as minimal graph.

It is easy to show that the same result applies for the Gaussian transformation [16]

$$\Pr\{Y = y | X = x\} = \prod_{s \in S'} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}\left[y_s - \frac{1}{|\mathcal{D}(s)|}\sum_{t \in \mathcal{D}(s)} x_t\right]^2\right\}. \tag{25}$$

More generally, we have shown that stochastic transformations of form (16) yield in general a loss of locality (i.e., the interactions between the variables may not be described locally). In particular, this is the case for the stochastic coarsening transformations usually used in the renormalization group approach.

V. Conclusion

A key problem that arises in many recent approaches to multiresolution MRF modeling (renormalization group approach, subsampling of stochastic processes, MRF's defined on trees or pyramids, ...) is the loss of "locality" of the representation at the different resolution levels, i.e., the impossibility to specify the distribution of the model at different scales using local potential functions. This loss of locality has adverse consequences on the complexity of the algorithms used in the simulation of these models.

In this paper, it has been shown that this problem may be studied (in a unified framework) by considering the restriction of a Markov random field (defined on a finite arbitrary nondirected graph) to a part of its original site set. General conditions for a preservation of the local specification of the restricted field have been given for finite-support MRF's. The consequences of these general results on the recently proposed approaches to multiresolution MRF modeling have been studied. The first consequence concerns the subsampling of MRF's. The general properties of several subsampling schemes have been examined. It has been shown that most standard subsampling schemes lead to a loss of locality. Examples of subsampling schemes which preserve a local Markovian property have been presented. The statistical properties of MRF models defined on trees or other pyramidal graph structures have also been studied. The restriction, at a given scale, of an MRF defined on such a hierarchical structure shows long-range interactions over the whole field and, as a consequence, may not be specified locally. The general properties of the renormalization group approach have also been examined. It has been shown that the usual scale transforms used in this method yield a loss of locality for the coarse scale models.

A tight control of the situations in which the Markovian property is preserved should allow the definition of consistent and tractable "multiresolution" relaxation algorithms associated with specific classes of field restrictions. A systematic recording of these "good" situations (with respect to the original neighborhood structure) would be a profitable extension of this work. These general restriction schemes could in turn be applied to the simulation of particular models (using relaxation algorithms or Monte Carlo techniques [14]) or to the computation of Bayesian estimates in MRF-based statistical estimation.

Acknowledgment

The authors wish to thank F. Prêteux, B. Levy, D. Petritis, and A. Vallat for fruitful discussions.

References

[1] R. Azencott, Ed., *Simulated Annealing. Parallelization Techniques.* New York: Wiley, 1992.
[2] S. T. Barnard, "Stochastic stereo matching over scale," *Int. J. Comp. Vis.*, vol. 3, pp. 17–32, 1989.
[3] M. Basseville, A. Benveniste, K. C. Chou, S. A. Golden, R. Nikoukhah, and A. S. Willsky, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 766–784, 1992.