Restriction of a Markov random field on a graph and multiresolution image analysis

Patrick Pérez, Fabrice Heitz

Unité de recherche INRIA Lorraine, Technôpole de Nancy-Brabois, Campus scientifique,
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399

[29] Z. KATO, M. BERTHOD, and J. ZERUBIA. – Multiscale Markov random field models for parallel image classification. – In *Proc. Int. Conf. Computer Vision*, pages 253–257, Berlin, May 1993.

[30] R. KINDERMANN and J.L. SNELL. – *Markov random fields and their applications, Vol. 1.* – Amer. Math. Soc., 1980.

[31] J. KONRAD and E. DUBOIS. – Bayesian estimation of motion vector fields. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 14, No. 9 : pages 910–927, September 1992.

[32] S. LAKSHMANAN and H. DERIN. – Gaussian Markov random fields at multiple resolutions. – In R. Chellappa and A.K. Jain, editors, *Markov Random Fields : Theory and Applications*, pages 131–157, Boston, 1993. Academic Press.

[33] B.C. LEVY. – Multiscale models and estimation of discrete Gauss-Markov random fields. – In *2nd SIAM Conf. Linear Algebra in Systems Control and Signal Processing*, San Francisco, CA, November 1990.

[34] M. LUETTGEN. – *Image Processing with Multiscale Stochastic Models.* – PhD thesis, MIT Dept. of EECS, May 1993.

[35] F. MARQUES, J. CUNILLERA, and A. GASULL. – Hierarchical segmentation using compound Gauss-Markov random fields. – In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages III.53–III.56, San Francisco, 1992.

[36] J.L. MARROQUIN, S. MITTER, and T. POGGIO. – Probabilistic solution of ill-posed problems in computational vision. – *J. American Statis. Assoc.*, Vol. 82 : pages 76–89, 1987.

[37] J.W. MODESTINO and J. ZHANG. – A Markov Random Field model-based approach to image interpretation. – In *IEEE Conf. Comp. Vision Pattern Rec.*, pages 458–465, June 1989.

[38] G. K. NICHOLLS. – An order preserving hierarchy of markov random fields. – In *Proc. 1st New Zealand Conf. Image and Vision Computing*, pages 205–212, Awckland, New Zealand, 1993.

[39] P. PÉREZ. – *Markov Random Fields and multiresolution image analysis. Application to image sequence analysis (in French).* – PhD thesis, Rennes I University, France, July 1993.

[40] P. PÉREZ and F. HEITZ. – Global Bayesian estimation, constrained multiscale Markov random fields and the analysis of visual motion. – In A. Mohammad-Djafari and G. Demoments, editors, *Maximum Entropy and Bayesian Methods*, pages 383–388. Academic Publishers, 1993.

[41] R.H. SWENDSEN and J.-S. WANG. – Non-universal critical dynamics in Monte Carlo simulations. – *Phys. Rev. Lett.*, Vol. 58 : pages 86–88, 1987.

[42] D. TERZOPOULOS. – Image analysis using multigrid relaxation methods. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 8, No. 2 : pages 129–139, 1986.

[15] X. DESCOMBES. – Markov Random Fields and Image Analysis. – PHD Thesis (in French), École Nationale Supérieure des Télécommunications, 46, rue Barrault, 75013 Paris, December 1993.

[16] X. DESCOMBES and F. PRÊTEUX. – Potentiels canoniques et modèles markoviens. (in French) – In *Proc. 13ème Colloque Trait. Sign. et Image, GRETSI*, pages 812–824, Juan-Les-Pins, September 1991.

[17] J.M. DINTEN. – Tomographic reconstruction of axially symmetric objects : Regularization by a Markovian modelization. – In *Proc. 10th Int. Conf. Pattern Recognition*, volume 2, pages 153–158, Atlantic City, Juin 1990.

[18] D. GEMAN, S. GEMAN, C. GRAFFIGNE, and D. PONG. – Boundary detection by constrained optimization. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, No. 7 : pages 609–628, July 1990.

[19] S. GEMAN and D. GEMAN. – Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 6, No. 6 : pages 721–741, Nov. 1984.

[20] H.-O. GEORGII. – *Gibbs measures and phase transition.* – Walter de Gruyter, 1988.

[21] B. GIDAS. – A renormalization group approach to image processing problems. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 11, No. 2 : pages 164–180, February 1989.

[22] M. GONDRAN and M. MINOUX. – *Graphs et algorithms.* – John Wiley & Sons Ltd., Chichester, 1984.

[23] F. HEITZ and P. BOUTHEMY. – Multimodal estimation of discontinuous optical flow using Markov random fields. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 15, No. 12 , December 1993.

[24] F. HEITZ, H. MAITRE, and C. de COUESSIN. – Event detection in multisource imaging using contextual estimation. – In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages 1647–1650, Glasgow, May 1989.

[25] F. HEITZ, P. PÉREZ, and P. BOUTHEMY. – Multiscale minimization of global energy functions in some visual recovery problems. – *CVGIP : Image Understanding*, Vol. 59, No. 1, 1994.

[26] M. HURN and C. JENNISON. – A study of simulated annealing and a revised cascade algorithm for image reconstruction. – Technical Report 93:04, Univ. of Bath, UK, June 1992.

[27] F.C. JENG. – Subsampling of Markov random fields. – *J. of Visual Com. and Image Representation*, Vol. 3, No. 3 : pages 225–229, September 1992.

[28] M. JUBB and C. JENNISON. – Aggregation and refinement in binary image restoration. – In Possolo, editor, *Spatial Statistics and Imaging*, pages 150–162. Institute of Mathematical Sciences Lecture Notes, Hayward, CA, 1991.

References

[1] R. AZENCOTT, editor. – *Simulated annealing. Parallelization techniques.* – John Wiley & Sons, Inc., New-York, 1992.

[2] S.T. BARNARD. – Stochastic stereo matching over scale. – *Int. J. Comp. Vis.*, Vol. 3 : pages 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. Information Theory*, Vol. 38, No. 2 : pages 766–784, March 1992.

[4] J. BESAG. – Spatial interaction and the statistical analysis of lattice systems. – *J. Royal Statist. Soc.*, Serie B, Vol. 36 : pages 192–236, 1974.

[5] J. BESAG. – On the statistical analysis of dirty pictures. – *J. Royal Statist. Soc.*, Vol. 48, Serie B, No. 3 : pages 259–302, 1986.

[6] L. BLANC-FÉRAUD, M. BARLAUD, and T. GAIDON. – Motion estimation involving discontinuities in a multiresolution scheme. – Technical Report 92–75, Université de Nice, Laboratoire I3S, December 1992.

[7] C. BOUMAN. – Multispectral image segmentation using a multiscale model. – In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, pages III.565–III.568, San Francisco, April 1992.

[8] C. BOUMAN and B. LIU. – Multiple resolution segmentation of textured images. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, No. 2 : pages 99–113, February 1991.

[9] P. BOUTHEMY and E. FRANÇOIS. – Motion segmentation and qualitative dynamic scene analysis from an image sequence. – *Int. J. Computer Vision*, Vol 10, No 2: pages 157–182, 1993.

[10] P. BOUTHEMY and P. LALANDE. – Recovery of moving object masks in an image sequence using local spatio-temporal contextual information. – *Optical Engineering*, Vol.32, No 6: pages 1205–1212, June 1993.

[11] B. CHALMOND. – Image restoration using an estimated Markov model. – *Signal Processing*, Vol. 15, No. 2 : pages 115–129, September 1988.

[12] R. CHELLAPPA and A.K. JAIN, editors. – *Markov Random Fields : Theory and Applications.* – Academic Press, 1993.

[13] P.B. CHOU and C.M. BROWN. – The theory and practice of Bayesian image modeling. – *Int. J. Comp. Vis.*, Vol. 4 : pages 185–210, 1990.

[14] F.S. COHEN and D.B. COOPER. – Simple parallel hierarchical and relaxation algorithms for segmenting noncausal Markovian random fields. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 9, No. 2 : pages 195–219, March 1987.

$x_r,\ r \in S - \{s\}\}$ defined on $\mathcal{E} \times \Omega_{S-\{s\}}$ is not independent of x_t. As in the second part of the previous proof, one can show that there exists a collection $\{V_A, A \in \mathcal{P}(S), A \neq \emptyset\}$ of a-canonical potential functions (with some reference configuration a) such that:

$$Pr\{X_s \in T_s \mid X_r = x_r,\ r \in S - \{s\}\} \propto \exp\left\{ - \sum_{A \in \mathcal{P}(S): A \neq \emptyset} V_A(x) \right\} .$$

Therefore:

$$t \in \mathcal{G}_s \ \Leftrightarrow \sum_{A \in \mathcal{P}(S): \{s,t\} \subset A} V_A(x) \text{ depends on } x_t \ .$$

Since all V_A's are a-irreducible, if $\sum_{A \in \mathcal{P}(S): \{s,t\} \subset A} V_A(x)$ does not depend on x_s, then the special case $x_s = a_s$ shows that it is the zero function. Furthermore, it does not depend on x_t. Therefore, $s \in \mathcal{G}_t$ and $\mathcal{G} = \{\mathcal{G}_s, s \in S\}$ fulfills all the properties of a neighborhood system. ■

for any configuration x and any site s. Therefore, the a-canonicity of potential functions V_A leads to:

$$Pr\{X_s = a_s \mid X_r = x_r, r \in S - \{s\}\} = \frac{1}{\sum_{\lambda \in \Lambda} \exp\left\{-\sum_{A \in \mathcal{P}(S): s \in A} V_A(x^{s,\lambda})\right\}} .$$

Finally:

$$\frac{Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\}}{Pr\{X_s = a_s \mid X_r = x_r, r \in S - \{s\}\}} = \exp\left\{-\sum_{A \in \mathcal{P}(S): s \in A} V_A(x)\right\} . \quad (87)$$

Let t be an arbitrary site (not belonging to $\mathcal{G}_s$). Applied to the configuration x which agrees with a on $S - \{s,t\}$, (87) becomes:

$$\frac{Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\}}{Pr\{X_s = a_s \mid X_r = x_r, r \in S - \{s\}\}} = \exp\left\{-V_{\{s\}}(x) - V_{\{s,t\}}(x)\right\} ,$$

which must not depend on x_t due to the markovianity of X on G. That is to say, the a-canonical potential function $V_{\{s,t\}}$ does not depend on x_t. This is possible only if $V_{\{s,t\}} = 0$. Let us assume that $V_A = 0$ for any subset A such that $A \notin \mathcal{C}$ and $|A| \leq n$, for some integer $n \geq 2$. Let A be a subset of $n+1$ sites which does not belong to $\mathcal{C}$. Then A contains at least two sites s and t which are not neighbors. Applied to the configuration x which agrees with a on $\bar{A}$, (87) yields:

$$\begin{aligned} \frac{Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\}}{Pr\{X_s = a_s \mid X_r = x_r, r \in S - \{s\}\}} &= \exp\left\{-\sum_{B \in \mathcal{P}(A): s \in B} V_B(x)\right\} \\ &= \exp\left\{-\sum_{c \in \mathcal{P}(A) \cap \mathcal{C}: s \in c} V_c(x) - V_A(x)\right\} , \end{aligned}$$

which must not depend on x_t. Since any clique containing s does not contain t, this means that the a-canonical potential function V_A does not depend on x_t. This is possible only if $V_A = 0$. Thus, we have recursively proven that for any subset A which is not a clique of graph G, the potential function V_A is zero. The distribution (86) becomes:

$$Pr\{X = x\} = Pr\{X = a\} \exp\left\{-\sum_{c \in \mathcal{C}} V_c(x)\right\} . \quad (88)$$

Therefore, X is a Gibbs random field on $G = [S, \mathcal{G}]$.

Let us now complete the proof of Proposition 3. Since $\mathcal{V} = \{V_c, c \in \mathcal{C}\}$ is a collection of zero or irreducible potential functions (being a-canonical), we have only to show that the minimality of G implies (85). As in the first part of this proof, one can find the local conditional distribution at any site s:

$$\forall x \in \Omega, \quad Pr\{X_s = x_s \mid X_{\mathcal{G}_s} = x_{\mathcal{G}_s}\} = \frac{1}{Z_s(x_{\mathcal{G}_s})} \exp\left\{-\sum_{c \in \mathcal{C}: s \in c} V_c(x)\right\} \quad (89)$$

If G is minimal, then this distribution is not independent of x_t, for any neighbor t of s. That is to say, $\sum_{c \in \mathcal{C}: \{s,t\} \subset c} V_c(x)$ is not independent of x_t. It comes that the potential functions associated with cliques containing $\{s,t\}$ cannot be all zero simultaneously: there exists $c \in \mathcal{C}$ such that $\{s,t\} \in c$ and $V_c \neq 0$. ■

Proof of Proposition 1 (existence of the minimal graph). Let $X = (\Omega, \mathcal{T}, P)$ be a Markov Random field with finite site set S. For any site s, let $\mathcal{G}_s$ be the set of sites $t \neq s$ such that $Pr\{X_s \in T_s \mid X_r =$

It comes that X is a Markov random field with respect to the neighborhood system $\mathcal{G}$. Furthermore, this provides the *local conditional distribution* at any site s:

$$\forall\, x \in \Omega, \quad Pr\{X_s = x_s \mid X_{\mathcal{G}_s} = x_{\mathcal{G}_s}\} = \frac{1}{Z_s(x_{\mathcal{G}_s})} \exp\left\{ - \sum_{c \in \mathcal{C} : s \in c} V_c(x) \right\} \tag{83}$$

with:

$$Z_s(x_{\mathcal{G}_s}) \stackrel{\Delta}{=} \sum_{\lambda \in \Lambda} \exp\left\{ - \sum_{c \in \mathcal{C} : s \in c} V_c(x^{s,\lambda}) \right\}. \tag{84}$$

Let us show that the graph G is minimal for X if $\mathcal{V}$ is a collection of zero and irreducible potential functions such that:

$$\forall\, \{s,t\} \in \mathcal{C},\ \exists c \in \mathcal{C}\ :\ \{s,t\} \subset c \text{ and } V_c \neq 0\ . \tag{85}$$

Let $s \in S$ be some site. Let us assume that there exists $t \in \mathcal{G}_s$ such that:

$$Pr\{X_s = x_s \mid X_r = x_r, r \in \mathcal{G}_s\} \ = \ Pr\{X_s = x_s \mid X_r = x_r, r \in \mathcal{G}_s - \{t\}\}$$

for any configuration x. This means that $\sum_{c \in \mathcal{C} : \{s,t\} \subset c} V_c(x)$ does not depend on x_t. Since $\{c \in \mathcal{C} : \{s,t\} \subset c \text{ and } V_c \neq 0\}$ is nonempty according to (85), it containts at least one maximal element for the inclusion, denoted by c_o. For any $x \in \Omega$:

$$V_{c_o}(x) = V_{c_o}(x^{c_0}) = \sum_{c \in \mathcal{C} : \{s,t\} \subset c} V_c(x^{c_0}) - \sum_{c \neq c_0 : \{s,t\} \subset c} V_c(x^{c_0})\ ,$$

where the mapping $\left[x \in \Omega \mapsto \sum_{c \in \mathcal{C} : \{s,t\} \subset c} V_c(x^{c_0}) \right]$ is a potential function on $c_0 - \{t\}$ according to the independency assumption. Moreover, for any clique $c \neq c_0$ such that $\{s,t\} \subset c$, the mapping $[x \in \Omega \mapsto V_c(x^{c_0})]$ is a potential function on $(c \cap c_0) \subset c_0$ which is different from c_0 since c cannot contain the maximal clique c_0. Finally, V_{c_0} would be a sum of potential functions on subsets of c_0 different from c_0. This contradicts the irreducibility of V_{c_0}. Therefore, the Gibbs random field X fulfills the minimality property with respect to neighborhood system $\mathcal{G}$. This completes the proof of the sufficient condition in Proposition 3.

□□ Let us assume, now, that X is a Markov random field with respect to $\mathcal{G}$, and that it fulfills the positiveness property. Let $a \in \Omega$ be some "reference" configuration. We can define the following real-valued mapping on Ω:

$$\forall\, x \in \Omega, \quad U(x) \stackrel{\Delta}{=} -\ln(Pr\{X = x\})\ .$$

According to Proposition 6, there exists a unique collection $\{V_A, A \in \mathcal{P}(S), A \neq \emptyset\}$ of a-canonical potential functions such that:

$$U(x) - U(a) = \sum_{A \in \mathcal{P}(S) : A \neq \emptyset} V_A(x).$$

Then:

$$Pr\{X = x\} = Pr\{X = a\} \exp\left\{ - \sum_{A \in \mathcal{P}(S) : A \neq \emptyset} V_A(x) \right\}\ . \tag{86}$$

Let us show that each potential function V_A is zero if A is not a clique of G. As previously, one can easily show that:

$$Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\} \ = \ \frac{\exp\left\{ - \sum_{A \in \mathcal{P}(S) : s \in A} V_A(x) \right\}}{\sum_{\lambda \in \Lambda} \exp\left\{ - \sum_{A \in \mathcal{P}(S) : s \in A} V_A(x^{s,\lambda}) \right\}}\ ,$$

B Hammersley-Clifford's Theorem and Corollary Properties

Let $G = [S, \mathcal{G}]$ be a finite simple nondirected graph, and $X = (\Omega, \mathcal{T}, P)$ be a random field on S with finite state space Λ. We show that X is a Markov random field on G such that $P(T) > 0,\ \forall T \in \mathcal{T}$, if and only if X is a Gibbs random field on G. At first we show that if X is a Gibbs random field, then the local conditional distribution at any site depends only on the neighboring sites, which provides the Markovian property. In a second step, we show that a Markov random field fulfilling the positiveness property has a Gibbs distribution, by exhibiting a collection of potential functions (we have chosen to be a-canonical), which provides the right distribution. At the same time, we give the proof of Proposition 3 which caracterises the minimality of the neighborhood using irreducible potential functions.

Proof of Theorem 1 and Proposition 3.
□ Let us assume that X is a Gibbs random fields on graph G: there exists a collection $\mathcal{V} = \{V_c,\ c \in \mathcal{C}\}$ of potential functions such that:

$$\forall\, x \in \Omega, \quad Pr\{X = x\} = \frac{1}{Z} \exp\left\{ - \sum_{c \in \mathcal{C}} V_c(x) \right\} \tag{79}$$

where $Z = \sum_{x \in \Omega} \exp\{-\sum_{c \in \mathcal{C}} V_c(x)\}$. The positiveness property is obviously verified:

$$\forall\, x \in \Omega, \quad Pr\{X = x\} > 0 \ .$$

Let $x \in \Omega$ be an arbitrary configuration and s be some site of S.

$$\begin{aligned} Pr\{X_s = x_s \quad | \quad & X_r = x_r,\ r \in S - \{s\}\} \\ &= \frac{Pr\{X = x\}}{Pr\{X_r = x_r,\ r \in S - \{s\}\}} && (80) \\ &= \frac{Z^{-1} \exp\{-U(x)\}}{\sum_{\lambda \in \Lambda} Pr\{X_s = \lambda, X_r = x_r,\ r \in S - \{s\}\}} && (81) \\ &= \frac{\exp\left\{-\sum_{c \in \mathcal{C}} V_c(x)\right\}}{\sum_{\lambda \in \Lambda} \exp\left\{-\sum_{c \in \mathcal{C}} V_c(x^{s,\lambda})\right\}} && (82) \end{aligned}$$

where $x^{s,\lambda}$ is the configuration such that $x_s^{s,\lambda} = \lambda$, and $x_r^{s,\lambda} = x_r, \forall\, r \neq s$. Since, for any clique c not containing s, $V_c(x) = V_c(x^{s,\lambda})$, the expression (82) of the conditional distribution becomes:

$$Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\} \ = \ \frac{\exp\left\{-\sum_{c \in \mathcal{C}: s \in c} V_c(x)\right\}}{\sum_{\lambda \in \Lambda} \exp\left\{-\sum_{c \in \mathcal{C}: s \in c} V_c(x^{s,\lambda})\right\}}.$$

This distribution does not depend on variables $x_r, r \notin \mathcal{G}_s \cup \{s\}$. On the other hand:

$$\begin{aligned} & Pr\{X_s = x_s \mid X_r = x_r, r \in \mathcal{G}_s\} \\ = & \sum_{x_{\overline{\mathcal{G}_s \cup \{s\}}} \in \Omega_{\overline{\mathcal{G}_s \cup \{s\}}}} \underbrace{Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\}}_{\text{independent of } x_{\overline{\mathcal{G}_s \cup \{s\}}}} \times Pr\{X_r = x_r, r \in \overline{\mathcal{G}_s \cup \{s\}}\} \\ = & \ Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\} \times \underbrace{\sum_{x_{\overline{\mathcal{G}_s \cup \{s\}}} \in \Omega_{\overline{\mathcal{G}_s \cup \{s\}}}} Pr\{X_r = x_r, r \in \overline{\mathcal{G}_s \cup \{s\}}\}}_{=1} \ . \end{aligned}$$

Therefore:

$$Pr\{X_s = x_s \mid X_r = x_r, r \in S - \{s\}\} \ = \ Pr\{X_s = x_s \mid X_r = x_r, r \in \mathcal{G}_s\} \ .$$

Hence:

$$\begin{aligned}\sum_{B\in\mathcal{P}(A):B\neq\emptyset} V_B(x) &= V_A(x^A) - V_A(x^\emptyset) \\ &= V_A(x) - V_A(a)\ ,\end{aligned}$$

since $x^A_A = x_A \Rightarrow V_A(x^A) = V_A(x^A_A)$ (V_A being a potential function on A), and $x^\emptyset = a$.

Let us show that the potential functions V_B are a-canonical. Let B be a nonempty subset of A and x be a configuration of Ω such that $x_s = a_s$ for some site $s \in B$:

$$\begin{aligned}V_B(x) &= \sum_{C\in\mathcal{P}(B)} (-1)^{|B-C|} V_A(x^C) \\ &= \sum_{C\in\mathcal{P}(B):s\notin C} (-1)^{|B-C|} V_A(x^C) + \sum_{C\in\mathcal{P}(B):s\in C} (-1)^{|B-C|} V_A(x^C) \\ &= \sum_{C\in\mathcal{P}(B):s\notin C} (-1)^{|B-C|} V_A(x^C) + \sum_{C\in\mathcal{P}(B):s\notin C} (-1)^{|B-C|-1} V_A(x^{C\cup\{s\}}) \\ &= \sum_{C\in\mathcal{P}(B):s\notin C} (-1)^{|B-C|} V_A(x^C) - \sum_{C\in\mathcal{P}(B):s\notin C} (-1)^{|B-C|} V_A(x^{C\cup\{s\}}) \\ &= 0\end{aligned}$$

since $x_s = a_s \implies x^C = x^{C\cup\{s\}}$, potential functions V_B are a-canonical.

Let us assume now, that there exists another collection $\{V'_B, B \in \mathcal{P}(A), B \neq \emptyset\}$ of a-canonical potential functions such that:

$$V_A = \sum_{B\in\mathcal{P}(A):B\neq\emptyset} V_B = \sum_{B\in\mathcal{P}(A):B\neq\emptyset} V'_B.$$

For any $x \in \Omega$ and for any $s \in S$, the previous equality leads to $V_{\{s\}}(x^{\{s\}}) = V'_{\{s\}}(x^{\{s\}})$ for the configuration $x^{\{s\}}$ which agrees with a on $S - \{s\}$ and takes state x_s at site s. Since $V_{\{s\}}(x^{\{s\}}) = V_{\{s\}}(x)$, $V'_{\{s\}}(x^{\{s\}}) = V'_{\{s\}}(x)$, then $V_{\{s\}} = V'_{\{s\}}$: potential functions associated with singletons are the same. Proceeding the same way for increasing subset cardinals, one can show the two collections are the same. This proves unicity. ■

The a-canonicity criterion we have presented implies the irreducibility for nonzero potential functions:

Proposition 7 *Let S be a finite set site, $\Omega = \Lambda^S$ be a configuration space on S, A be a subset of S and a be some configuration in Ω. If V_A is a a-canonical potential function on A, then V_A is irreducible or zero.*

Proof. This is straightforward in case A is a singleton. In case $|A| \geq 2$, let us assume that the a-canonical potential function V_A is nonzero and non irreducible. Then, there exists a finite collection $\{A_i, i \in I\}$ of nonempty subsets of A, different from A, and an associated collection $\{V_{A_i}, i \in I\}$ of nonzero potential functions such that $V_A = \sum_{i\in I} V_{A_i}$. According to Proposition 6, each potential function V_A is, apart from some constant, a sum of nonzero a-canonical potential functions on subsets of A_i. Then, it is easy to see that the a-canonical potential function V_A is a sum of nonzero a-canonical potential functions on nonempty subsets of A, different from A. This is contradicts the unicity property given by Proposition 6. ■

Proposition 2 follows easily from Propositions 6 and 7.

A Irreducible and Canonical Potential Functions

The canonical potential functions used in Appendices A and B have been introduced by Descombes [15]. They generalize the canonical potential functions used by Besag [4] and Snell and Kindermann [16, 30]. Their definition is related to the specification of some "reference" configuration $a \in \Omega$:

Definition 8 *Let S be a finite site set, $\Omega = \Lambda^S$ be a configuration space on S, A be a subset of S and a be some configuration of Ω. A potential function V_A on A is <u>a-canonical</u> if:*

$$\forall\, x \in \Omega,\ (\exists s \in A\ :\ x_s = a_s) \implies V_A(x) = 0\ . \tag{76}$$

The λ-canonicity ($\lambda \in \Lambda$ being some reference state) introduced by Snell and Kindermann [30] is a particular case of a-canonicity with $a_s = \lambda,\ \forall\, s \in S$. Besag [4] has used 0-canonical potential functions (assuming that $\Lambda \subset \mathbb{R}$ and $0 \in \Lambda$).

Proposition 6 *For any potential function V_A on the nonempty subset $A \subset S$, and for any configuration $a \in \Omega$, there exists a unique collection $\{V_B, B \in \mathcal{P}(A), B \neq \emptyset\}$ of a-canonical potential functions associated with nonempty subsets of A, such that:*

$$\forall\, x \in \Omega,\quad V_A(x) - V_A(a) \ = \sum_{B \in \mathcal{P}(A) : B \neq \emptyset} V_B(x)\ . \tag{77}$$

Proof. For any nonempty subset $B \subset A$, let us define the following potential function:

$$V_B(x) \stackrel{\triangle}{=} \sum_{C \in \mathcal{P}(B)} (-1)^{|B-C|} V_A(x^C),\quad \forall\, x \in \Omega\ , \tag{78}$$

where x^C is the configuration which agrees with x on C and agrees with a on $S - C$. Let us compute the sum of all these potential functions:

$$\begin{aligned}
\sum_{B \in \mathcal{P}(A) : B \neq \emptyset} V_B(x) \ &= \sum_{B \in \mathcal{P}(A) : B \neq \emptyset}\ \sum_{C \in \mathcal{P}(B)} (-1)^{|B-C|} V_A(x^C) \\
&= \sum_{C \in \mathcal{P}(A)}\ \sum_{\substack{B \in \mathcal{P}(A)\ : \\ B \neq \emptyset, C \subset B}} (-1)^{|B-C|} V_A(x^C) \\
&= \sum_{C \in \mathcal{P}(A)} \underbrace{\left(\sum_{\substack{B \in \mathcal{P}(A)\ : \\ B \neq \emptyset, C \subset B}} (-1)^{|B-C|} \right)}_{①} V_A(x^C)\ .
\end{aligned}$$

If $C \neq \emptyset$ and $C \neq A$, then:

$$① = \sum_{k=0}^{|A-C|} \left(\begin{array}{c} |A-C| \\ k \end{array} \right) (-1)^k = [1 + (-1)]^{|A-C|} = 0\ .$$

If $C = A$, then, $① = 1$. Finally, if $C = \emptyset$,

$$① = \sum_{k=1}^{|A-C|} \left(\begin{array}{c} |A-C| \\ k \end{array} \right) (-1)^k = [1 + (-1)]^{|A-C|} - 1 = -1\ .$$

(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 [19]) or to the computation of Bayesian estimates in MRF-based statistical estimation.

Acknowledgments

This work has been partly supported by MRT (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. We thank F. Prêteux, B. Levy and D. Petritis for fruitful discussions.

If matrix $B = \Sigma_{\underline{X}}^{-1}$ is decomposed in a similar way, i.e.:

$$B = \begin{array}{c} \overbrace{}^{s\in\bar{A}_i} \quad \overbrace{}^{s\in S-\bar{A}_i} \\ \left[\begin{array}{c|c} B^{(i)} & C^{(i)} \\ \hline \bar{C}^{(i)} & \bar{B}^{(i)} \end{array}\right] \end{array} \tag{73}$$

it comes:

$$B\,\Sigma_{\underline{X}} = I \implies \left[B^{(i)}\right]^{-1} = \Sigma^{(i)} - \Delta^{(i)} \left[\bar{\Sigma}^{(i)}\right]^{-1} \bar{\Delta}^{(i)} \ .$$

Let us finally notice that X_A is a Gaussian random field with covariance matrix $\Sigma_{\underline{X}_A} \stackrel{\triangle}{=} \tilde{B}^{-1}$ where the coefficient $\tilde{b}_{st}$ of matrix $\tilde{B}$ associated with the doubleton $(s,t) \in A \times A$ is given by:

$$\tilde{b}_{st} \stackrel{\triangle}{=} b_{st} \ , \quad \text{si } (s,t) \notin \mathcal{G}_{\bar{A}_i} \times \mathcal{G}_{\bar{A}_i} , \quad \forall\, i \in I \ ; \tag{74}$$

$$\tilde{b}_{st} \stackrel{\triangle}{=} b_{st} - \sum_{\substack{s' \in \mathcal{G}_s \cap \bar{A}_i \\ t' \in \mathcal{G}_t \cap \bar{A}_i}} b_{ss'}\, c^{(i)}_{s't'}\, b_{t't} \ , \quad \text{si } (s,t) \in \mathcal{G}_{\bar{A}_i} \times \mathcal{G}_{\bar{A}_i} \ . \tag{75}$$

5 Conclusion

A key problem that arises in many recent approaches to multiresolution MRF modeling (renormalization group approach, subsampling of stochastic processes, MRFs 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 non-directed 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 MRFs. 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 MRFs. 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 a MRF defined on such a hierarchical structure show 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. Finally, the special case of Gauss-Markov random fields has been considered and the precise conditions under which the restriction of these models may be specified locally have been derived.

A tight control of the situations in which the Markovian property is preserved should enable the definition of consistent and tractable "multiresolution" relaxation algorithms associated with specific classes of field restrictions. A systematic recording of these "good" situations

This derivation yields the following expression for potential function V_i:

$$V_i(x_A) \;=\; -\ln\left\{\frac{(2\pi)^{\frac{1}{2}|\bar{A}_i|}}{(\det B^{(i)})^{\frac{1}{2}}}\right\} - \frac{1}{2}\,\underline{u}(x_{\mathcal{G}_{\bar{A}_i}})^T\,[B^{(i)}]^{-1}\,\underline{u}(x_{\mathcal{G}_{\bar{A}_i}})\;. \tag{64}$$

As a consequence, the restricted random field X_A is associated with a collection of potential functions $\tilde{\mathcal{V}}_A = \{\tilde{V}_c, c \in \tilde{\mathcal{C}}_A\}$ which are zero for $|c| > 2$, and are defined by:

$$\forall\, x \in \Omega,\quad \tilde{V}_{\{s,t\}}(x_A) \stackrel{\triangle}{=} \tilde{b}_{st}\,(x_s - a_s)(x_t - a_t),$$

on two site cliques $\{s,t\} \in \tilde{C}_A$, with:

$$\tilde{b}_{st} \;\stackrel{\triangle}{=}\; b_{st}\;,\;\text{ if } \{s,t\} \in \mathcal{C} \text{ et } \{s,t\} \not\subset \mathcal{G}_{\bar{A}}\;; \tag{65}$$

$$\tilde{b}_{st} \;\stackrel{\triangle}{=}\; b_{st} - \sum_{\substack{s' \in \mathcal{G}_s \cap \bar{A}_i \\ t' \in \mathcal{G}_t \cap \bar{A}_i}} b_{ss'}\,c^{(i)}_{s't'}\,b_{t't}\;,\;\text{ if } \{s,t\} \subset \mathcal{G}_{\bar{A}_i}\;, \tag{66}$$

where $c^{(i)}_{s't'}$ is the coefficient of matrix $[B^{(i)}]^{-1}$ corresponding to sites s' and t' (multiplied by $\frac{1}{2}$ if $s' = t'$). Since $\tilde{\mathcal{V}}_A$ is composed of zero and irreducible potential functions, one can easily show that the sufficient conditions (H1) and (H2) for the graph to be minimal become in the Gaussian case:

$$\text{(H1) and (H2)} \;\stackrel{X \text{ Gaussian}}{\Longleftrightarrow}\; \left(\forall i \in I,\ \forall\,\{s,t\} \subset \mathcal{G}_{A_i},\quad b_{st} - \sum_{\substack{s' \in \mathcal{G}_s \cap \bar{A}_i \\ t' \in \mathcal{G}_t \cap \bar{A}_i}} b_{ss'}\,c^{(i)}_{s't'}\,b_{t't} \neq 0\right)\;. \tag{67}$$

Coefficients $c^{(i)}_{st}$ may be interpreted as following: since $[B^{(i)}]^{-1}$ is the covariance matrix of the Gaussian random vector $\underline{X}_{\bar{A}_i}$ conditioned by some realization $\underline{x}_{S-\bar{A}_i}$ of random vector $\underline{X}_{S-\bar{A}_i}$, it comes:

$$\begin{aligned} c^{(i)}_{st} &= \operatorname{cov}(X_s, X_t \mid X_{S-\bar{A}_i} = x_{S-\bar{A}_i})\;,\quad \forall\,\{s,t\} \subset \bar{A}_i\;; \\ c^{(i)}_{ss} &= \frac{1}{2}\operatorname{var}(X_s \mid X_{S-\bar{A}_i} = x_{S-\bar{A}_i})\;,\quad \forall\, s \in \bar{A}_i\;. \end{aligned}$$

Matrix $[B^{(i)}]^{-1}$ can be computed using a block decomposition of $\Sigma_{\underline{X}}$. Let us order the sites of S such that $\bar{A}_i = \{s_k, k = 1, \ldots, |\bar{A}_i|\}$ and let us decompose $\Sigma_{\underline{X}}$ in the following way:

$$\Sigma_{\underline{X}} = \left[\begin{array}{c|c} \overbrace{\Sigma^{(i)}}^{s\in\bar{A}_i} & \overbrace{\Delta^{(i)}}^{s\in S-\bar{A}_i} \\ \hline \bar{\Delta}^{(i)} & \bar{\Sigma}^{(i)} \end{array}\right] \tag{68}$$

where the four blocks are:

$$\Sigma^{(i)} \;\stackrel{\triangle}{=}\; \mathbb{E}[(\underline{X}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})\,(\underline{X}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})^T] \tag{69}$$

$$\bar{\Sigma}^{(i)} \;\stackrel{\triangle}{=}\; \mathbb{E}[(\underline{X}_{S-\bar{A}_i} - \underline{a}_{S-\bar{A}_i})\,(\underline{X}_{S-\bar{A}_i} - \underline{a}_{S-\bar{A}_i})^T] \tag{70}$$

$$\Delta^{(i)} \;\stackrel{\triangle}{=}\; \mathbb{E}[(\underline{X}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})\,(\underline{X}_{S-\bar{A}_i} - \underline{a}_{S-\bar{A}_i})^T] \tag{71}$$

$$\bar{\Delta}^{(i)} \;\stackrel{\triangle}{=}\; \mathbb{E}[(\underline{X}_{S-\bar{A}_i} - \underline{a}_{S-\bar{A}_i})\,(\underline{X}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})^T] = \left[\Delta^{(i)}\right]^T \tag{72}$$

Now, let us study the restriction of random field X to a subset $A \subset S$. The potential functions V_i of the restricted field are defined by (26). They will be decomposed as a sum of irreducible potential functions. This decomposition will give a better insight on conditions (H1) and (H2) in this particular case. Let $\bar{A}_i$ be an arbitrary connected component of $\bar{A}$ and let x be a configuration in Ω. According to (26) :

$$\exp\{-V_i(x_A)\} = \int_{x_{\bar{A}_i} \in \mathbb{R}^{\bar{A}_i}} \exp\left\{ -\frac{1}{2} \sum_{s \in \bar{A}_i} b_{ss}(x_s - a_s)^2 - \sum_{\{s,t\} \subset \bar{A}_i} b_{st}(x_s - a_s)(x_t - a_t) \right.$$
$$\left. - \sum_{s \in \bar{A}_i} \sum_{t \in \mathcal{G}_s \cap A} b_{st}(x_s - a_s)(x_t - a_t) \right\} dx_{\bar{A}_i} \ . \tag{62}$$

To each restricted configuration $x_{\bar{A}_i} \in \Omega_{\bar{A}_i}$, we associate the column vector $\underline{x}_{\bar{A}_i}$ of $\mathbb{R}^{|\bar{A}_i|}$ (sites have been ordered from 1 to $|\bar{A}_i|$). The previous expression becomes:

$$\exp\{-V_i(x_A)\} = \int_{\underline{x}_{\bar{A}_i} \in \mathbb{R}^{|\bar{A}_i|}} \exp\left\{ -\frac{1}{2} (\underline{x}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})^T B^{(i)} (\underline{x}_{\bar{A}_i} - \underline{a}_{\bar{A}_i}) \right.$$
$$\left. - (\underline{x}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})^T \underline{u}(x_{\mathcal{G}_{\bar{A}_i}}) \right\} d\underline{x}_{\bar{A}_i} \ , \tag{63}$$

where $B^{(i)}$ is the symmetric matrix of order $|\bar{A}_i|$ whose coefficient $b_{st}^{(i)}$ associated with couple $(s,t) \in \bar{A}_i \times \bar{A}_i$ is the same as the corresponding one in matrix B ([9]):

$$\forall\, (s,t) \in \bar{A}_i \times \bar{A}_i, \quad b_{st}^{(i)} \triangleq b_{st} \ ,$$

and $\underline{u}(x_{\mathcal{G}_{\bar{A}_i}})$ is the vector of $\mathbb{R}^{|\bar{A}_i|}$ whose component $u_s(x_{\mathcal{G}_{\bar{A}_i}})$ associated with site s of $\bar{A}_i$ is defined as:

$$u_s(x_{\mathcal{G}_{\bar{A}_i}}) \triangleq \sum_{t \in \mathcal{G}_s \cap A} b_{st}\,(x_t - a_t) \ .$$

It comes:

$$-\frac{1}{2} (\underline{x}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})^T B^{(i)} (\underline{x}_{\bar{A}_i} - \underline{a}_{\bar{A}_i}) - (\underline{x}_{\bar{A}_i} - \underline{a}_{\bar{A}_i})^T \underline{u}(x_{\mathcal{G}_{\bar{A}_i}})$$
$$= -\frac{1}{2} (\underline{x}_{\bar{A}_i} - \underline{\mu}_{\bar{A}_i})^T B^{(i)} (\underline{x}_{\bar{A}_i} - \underline{\mu}_{\bar{A}_i}) + \frac{1}{2} \underline{u}(x_{\mathcal{G}_{\bar{A}_i}})^T [B^{(i)}]^{-1} \underline{u}(x_{\mathcal{G}_{\bar{A}_i}}) \ ,$$

with:

$$\underline{\mu}_{\bar{A}_i} \triangleq \underline{a}_{\bar{A}_i} - [B^{(i)}]^{-1} \underline{u}(x_{\mathcal{G}_{\bar{A}_i}}) \ .$$

Finally:

$$\exp\{-V_i(x_A)\} = \exp\left\{ \frac{1}{2} \underline{u}(x_{\mathcal{G}_{\bar{A}_i}})^T [B^{(i)}]^{-1} \underline{u}(x_{\mathcal{G}_{\bar{A}_i}}) \right\}$$
$$\times \underbrace{\int_{\underline{x}_{\bar{A}_i} \in \mathbb{R}^{|\bar{A}_i|}} \exp\left\{ -\frac{1}{2} (\underline{x}_{\bar{A}_i} - \underline{\mu}_{\bar{A}_i})^T B^{(i)} (\underline{x}_{\bar{A}_i} - \underline{\mu}_{\bar{A}_i}) \right\} d\underline{x}_{\bar{A}_i}}_{\frac{(2\pi)^{\frac{1}{2}|\bar{A}_i|}}{\left(\det B^{(i)}\right)^{\frac{1}{2}}}} \ .$$

[9] $B^{(i)}$ is positive definite. As a matter of fact, B defines on $\mathbb{R}^N$ a positive definite quadratic form on $\mathbb{R}^N$ and $B^{(i)}$ is the matrix of the restriction of this quadratic form to a sub-space of dimension $|\bar{A}_i|$. Therefore, it is positive definite.

We consider continuous random fields with state space $\Lambda = \mathbb{R}$. Sites are numbered from 1 to N :

$$S = \{s_1, s_2, \ldots, s_N\} \ .$$

An arbitrary configuration $x = \{x_s, s \in S\}$ in $\Omega = \mathbb{R}^S$, will be represented by vector $\underline{x} = [\underline{x}_i]_{i \in \{1,\ldots,N\}} \stackrel{\triangle}{=} [x_{s_i}]_{i \in \{1,\ldots,N\}}$ in $\mathbb{R}^N$. The real-valued random field X on S with probability density p_X will also be represented by the corresponding N-component random vector $\underline{X}$ whose probability density is:

$$\forall\, \underline{x} \in \mathbb{R}^N, \quad p_{\underline{X}}(\underline{x}) = p_X(x) \ .$$

In the remainder of this Section, if $\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_N \end{bmatrix}$ is a column vector with N components, we will denote by $\underline{x}^T = [\underline{x}_1 \ \ldots \ \underline{x}_N]$ its transpose row vector.

Definition 7 *The real-valued random field X with site set S and with probability density p_X is a Gaussian random field if and only if its associated random vector $\underline{X}$ is Gaussian, i.e.:*

$$\forall\, x \in \Omega = \mathbb{R}^S, \ p_X(x) = \frac{1}{(2\pi)^{\frac{N}{2}} (\det \Sigma_{\underline{X}})^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} (\underline{x} - \underline{a})^T \, \Sigma_{\underline{X}}^{-1} \, (\underline{x} - \underline{a}) \right\} \ , \qquad (56)$$

where $\underline{a} \stackrel{\triangle}{=} \mathbb{E}[\underline{X}]$ is the expectation of $\underline{X}$ and $\Sigma_{\underline{X}} \stackrel{\triangle}{=} \mathbb{E}[(\underline{X} - \underline{a})\,(\underline{X} - \underline{a})^T]$ is its symmetric positive definite covariance matrix.

Let $Z \stackrel{\triangle}{=} (2\pi)^{\frac{N}{2}} (\det \Sigma_{\underline{X}})^{\frac{1}{2}}$ and $B = [B_{ij}] \stackrel{\triangle}{=} \Sigma_{\underline{X}}^{-1}$. $\Sigma_{\underline{X}}$ and B are symmetric, positive definite matrices. For two arbitrary sites $s = s_i$ and $t = s_j$ of S, b_{st} will denote the coefficient B_{ij} of matrix B. Let a be the configuration associated with the expectation vector $\underline{a}$. With these conventions, the probability density of the Gaussian random field X becomes:

$$\forall\, x \in \Omega, \quad p_X(x) = \frac{1}{Z} \exp\left\{ -\frac{1}{2} \sum_{s \in S} b_{ss} (x_s - a_s)^2 - \sum_{\{s,t\} \subset S} b_{st} (x_s - a_s)(x_t - a_t) \right\} \ . \qquad (57)$$

X is a Markov random field on the graph $G = [S, \mathcal{G}]$ whose neighborhood system is defined in the following way:

$$\forall\, \{s,t\} \subset S, \quad t \in \mathcal{G}_s \iff s \in \mathcal{G}_t \iff b_{st} = b_{ts} \neq 0 \ . \qquad (58)$$

The potential functions of the Gauss-Markov random field $\mathcal{V} = \{V_c, c \in \mathcal{C}\}$ are defined by:

$$\forall\, s \in S, \ \forall\, x \in \Omega, \qquad V_{\{s\}}(x) \stackrel{\triangle}{=} \frac{1}{2} b_{ss} (x_s - a_s)^2 \ ; \qquad (59)$$

$$\forall\, \{s,t\} \in \mathcal{C}, \ \forall\, x \in \Omega, \qquad V_{\{s,t\}}(x) \stackrel{\triangle}{=} b_{st} (x_s - a_s)(x_t - a_t) \ ; \qquad (60)$$

$$\forall\, c \in \mathcal{C} \ : \ |c| \geq 3, \qquad V_c \stackrel{\triangle}{=} 0 \ . \qquad (61)$$

Since these potential functions are a-canonical, they are irreducible or zero (*cf* definitions and properties in Appendix A). Furthermore, they verify:

$$\forall\, \{s,t\} \in \mathcal{C}, \ V_{\{s,t\}} \neq 0 \ ,$$

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

$$l_s(x_{\mathcal{D}(s)}, y_s) = - \sum_{t \in \mathcal{D}(s)} \rho \, y_s \, x_t, \tag{51}$$

which is a sum of irreducible potential functions on two site cliques $\{s,t\}, t \in \mathcal{D}(s)$. Subsets $\mathcal{D}(s)$ which are 2×2 square blocks are then, at the same time, cliques of the 8-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 an unique father $f(t)$ in S':

$$\forall\, t \in S,\ \mathcal{D}_t^{-1} = \{f(t)\}\ . \tag{52}$$

Therefore, neighborhood system $\mathcal{G}^*$ of graph G^* is given by:

$$\left\{ \begin{array}{ll} \forall\, s \in S, & \mathcal{G}_s^* = \mathcal{G}_s \cup \{f(t)\} \\ \forall\, s \in S', & \mathcal{G}_s^* = \mathcal{D}(s) \end{array} \right. \tag{53}$$

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

$$\forall\, (x,y) \in \Omega^*,\ V_c^*(x,y) \stackrel{\triangle}{=} \left\{ \begin{array}{ll} V_c(x) & \text{if } c \in \mathcal{C}\ ; \\ -\rho\, y_s\, x_t & \text{if } c = \{s,t\} \text{ with } s \in S' \text{ and } t \in \mathcal{D}(s)\ ; \\ 0 & \text{otherwise,} \end{array} \right. \tag{54}$$

is composed with zero and irreducible potential functions which verify (49) : X^* is a Markov random field 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 ([8]).

The same result applies for the following stochastic transformation (also proposed by Gidas in the renormalization group approach [21]) which corresponds to a Gaussian interaction model between resolution levels:

$$Pr\{Y = y \mid X = x\} = \prod_{s \in S'} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\,\sigma^2} \left(y_s - \frac{1}{|\mathcal{D}(s)|} \sum_{t \in \mathcal{D}(s)} x_t \right)^2 \right\}\ , \tag{55}$$

where σ^2 is the variance of a Gaussian white noise.
More generally, we have shown that stochastic transformations verifying the positivity condition (42) 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 used in the renormalization group approach.

4.5 Restrictions of Gauss-Markov Random Fields

In this section we study the restriction of Gauss-Markov random field X to a subset $A \subset S$. The expression of the potential functions associated with the restriction of X are given and the conditions (H1) and (H2) under which the graph of the restriction is minimal are developed.

[8] This conclusion does not concern the "majority rule" transform also proposed by Gidas which is a deterministic transform that does not verify condition (42).

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

$$\begin{aligned} P^*(T,T') &\triangleq Pr\{X \in T,\ Y \in T'\} \\ &= \int_{x \in T} Pr\{Y \in T' \mid X = x\}.\frac{1}{Z}\exp\left\{-\sum_{c \in \mathcal{C}} V_c(x)\right\} \kappa^S(dx)\ . \end{aligned}$$

for any couple $(T, T') \in \mathcal{T} \times \mathcal{T}'$. Using (45), one obtains:

$$P^*(T,T') = \frac{1}{Z}\int_{(x,y) \in T \times T'} \exp\left\{-\sum_{c \in \mathcal{C}} V_c(x) - \sum_{s \in S'}\sum_{k \in K_s} V_{s,k}(x,y)\right\} \kappa^S(dx)\, \kappa'^{S'}(dy), \qquad (46)$$

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

$$\left\{ \begin{array}{ll} \forall\, s \in S, & \mathcal{G}^*_s \triangleq \mathcal{G}_s \cup \mathcal{D}_s^{-1} \cup \{t \in S - \{s\} \mid \mathcal{D}_s^{-1} \cap \mathcal{D}_t^{-1} \neq \emptyset\} \\ \forall\, s \in S', & \mathcal{G}^*_s \triangleq \mathcal{D}(s) \end{array} \right. \qquad (47)$$

Let us associate with clique set $\mathcal{C}^*$ the potential function collection $\mathcal{V}^* = \{V^*_c, c \in \mathcal{C}^*\}$ defined as:

$$\forall\,(x,y) \in \Omega^*,\ V^*_c(x,y) \triangleq \left\{ \begin{array}{ll} V_c(x) & \text{if } c \in \mathcal{C} \text{ and } c \notin \{A_{s,k}, s \in S', k \in K_s\}\ ; \\ V_{s,k}(x,y) & \text{if } c = A_{s,k} \text{ and } A_{s,k} \notin \mathcal{C}\ ; \\ V_{s,k}(x,y) + V_c(x) & \text{if } c = A_{s,k} \text{ and } A_{s,k} \in \mathcal{C}\ ; \\ 0 & \text{otherwise,} \end{array} \right. \qquad (48)$$

for any clique $c \in \mathcal{C}^*$. For any $A_{s,k}$ belonging to $\mathcal{C}$, let us assume potential function $[(x,y) \mapsto V_{s,k}(x,y) + V_c(x)]$ to be irreducible or zero. Since all potential functions $V_c, c \in \mathcal{C}$ and $V_{s,k}, s \in S', k \in K_s$ are irreducible or zero, $\mathcal{V}^*$ is a collection of zero and irreducible potential functions. In addition, let us assume that this collection verifies:

$$\forall\,\{s,t\} \in \mathcal{C}^*,\ \exists c \in \mathcal{C}^*\ :\ \{s,t\} \subset c \text{ and } V^*_c \neq 0\ . \qquad (49)$$

Therefore, according to (46), X^* is a Markov random field on the minimal graph G^*. 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.

In the renormalization group approach [21], Gidas considers a Markov random field on a regular bidimensional lattice associated with an 8-neighborhood system, which is minimal and he proposes the following stochastic local transformation:

$$Pr\{Y = y \mid X = x\} = \prod_{s \in S'} \frac{\exp\{\rho\, y_s(\sum_{t \in \mathcal{D}(s)} x_t)\}}{\sum_{\lambda \in \Lambda'} \exp\{\rho\, \lambda(\sum_{t \in \mathcal{D}(s)} x_t)\}}\ , \qquad (50)$$

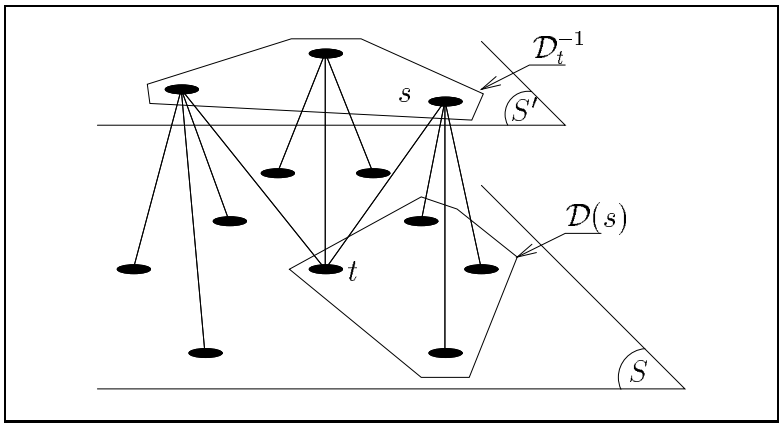

Figure 14: *Example of "link" between S and S' according to a local stochastic transformation of a random field with set site S into a random field with site set S'*

Let us consider a Markov random field $Y = (\Omega', \mathcal{T}', P')$ with site set S', with state space Λ' and such that:

$$Pr\{Y \in T' \mid X = x\} = \prod_{s \in S'} Pr\{Y_s \in T'_s \mid X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\} \tag{41}$$

and

$$Pr\{Y_s \in T'_s \mid X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\} > 0, \quad \forall\, s \in S'. \tag{42}$$

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

$$Pr\{Y_s \in T'_s \mid X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\} = \int_{y_s \in T'_s} \exp\left\{-l_s(x_{\mathcal{D}(s)}, y_s)\right\} \kappa'(dy_s) \tag{43}$$

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:

$$\exp\left\{-l_s(x_{\mathcal{D}(s)}, y_s)\right\} = Pr\{Y_s = y_s \mid X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\} \ .$$

In the continuous case, this assumption means that the local conditional distributions have densities. According to Proposition 2, for any site $s \in S'$, the potential function $\left[(x, y) \mapsto l_s(x_{\mathcal{D}(s)}, y_s)\right]$ on $\mathcal{D}(s) \cup \{s\}$ is a sum of irreducible potential functions $V_{s,k}, k \in K_s$, associated with a collection $\{A_{s,k}, k \in K_s\}$ of nonempty subsets of $\mathcal{D}(s) \cup \{s\}$:

$$\forall\, (x, y) \in \Omega^*,\ l_s(x_{\mathcal{D}(s)}, y_s) = \sum_{k \in K_s} V_{s,k}(x, y) \ . \tag{44}$$

Combining (41), (43) and (44) provides:

$$\begin{aligned} Pr\{Y \in T' \mid X = x\} &= \prod_{s \in S'} \int_{y_s \in T'_s} \exp\left\{-l_s(x_{\mathcal{D}(s)}, y_s)\right\} \kappa'(dy_s) \\ &= \int_{y \in T'} \exp\left\{-\sum_{s \in S'} \sum_{k \in K_s} V_{s,k}(x, y)\right\} \kappa'^{S'}(dy) \ . \end{aligned} \tag{45}$$

for any $\{s,t\} \in \mathcal{C}$, and

$$v_c \stackrel{\triangle}{=} v_c^* \tag{39}$$

for any clique $c \in \mathcal{C}$ such that $|c| \neq 2$.

These results can be used in order to build a Markov random field on G^* whose restriction to S coincide with a given Markov random field on G: let Y be a Markov random field on graph G with state space Λ, associated with the collection of potential function $\mathcal{V} = \{V_c = v_c \circ f_c,\ c \in \mathcal{C}\}$. Let us consider the Markov random field (X, X') on G^* with state space Λ on S and $\Lambda' = \{q, 1-q\}$ (with $q \in]0,1[$) on $\mathcal{U}$, and associated with the potential function collection $\mathcal{V}^* = \{V_c^*,\ c \in \mathcal{C}^*\}$ defined as:

$$\forall\, \{s,t\} \in \mathcal{C},\ \left\{ \begin{array}{l} V^*_{\{s,t\}} \stackrel{\triangle}{=} 0,\ V^*_{\{s,u\}} \stackrel{\triangle}{=} 0,\ V^*_{\{t,u\}} \stackrel{\triangle}{=} 0,\ V^*_{\{u\}} \stackrel{\triangle}{=} 0 \\ V^*_{\{s,t,u\}}(x,x') \stackrel{\triangle}{=} V_{\{s,t\}}(x) - \ln(x'_u),\ \ \forall\,(x,x') \in \Omega^* \end{array} \right.$$

where $u = <s,t>$, and:

$$V_c^* \stackrel{\triangle}{=} V_c\ ,$$

for any clique c such that $|c| \neq 2$. Then, it comes:

$$\begin{aligned} -\ln \left[\int_{x'_u \in \Lambda'} \exp\left\{ -v^*_{\{u\}}(x'_u) - v^*_{\{s,u\}}(x_s, x'_u) - v^*_{\{t,u\}}(x_t, x'_u) - v^*_{\{s,t,u\}}(x_s, x_t, x'_u) \right\} \kappa'(dx'_u) \right] & \\ = \ -\ln\left[\exp\left\{ -V_{\{s,t\}}(x) + \ln(q) \right\} + \exp\left\{ -V_{\{s,t\}}(x) + \ln(1-q) \right\} \right] & \\ = \ V_{\{s,t\}}(x)\ . & \end{aligned}$$

Finally, according to the previous results, the restriction of this field to S is the same Markov random field as Y.

Thus, given a Markov random field on a graph, by "attaching" auxiliary random variables to the edges, it is possible to build an "extented" Markov random field whose restriction on S corresponds to the original random field. This coupling may be very appealing when the new field is easier to handle than the original one. This is for instance the case of the very efficient Swendsen-Wang algorithm which is used to sample the Potts model [41].

4.4 Stochastic Transformations and the Renormalization Group Approach

Let $X = (\Omega, \mathcal{T}, P)$ be a Markov random field on the minimal graph $G = [S, \mathcal{G}]$ with the following probability measure P:

$$\forall T \in \mathcal{T},\ \ P(T) = \frac{1}{Z} \int_{x \in T} \exp\left\{ -\sum_{c \in \mathcal{C}} V_c(x) \right\} \kappa^S(dx)\ , \tag{40}$$

where $\mathcal{V} = \{V_c, c \in \mathcal{C}\}$ is a collection of irreducible or zero potential functions verifying (22). Let S' be another site set and $(\Lambda', \mathcal{E}', \kappa')$ be another state space. Let us associate with each site $s \in S'$ a nonempty subset $\mathcal{D}(s) \subset S$. For any site $t \in S$, we will denote by $\mathcal{D}_t^{-1}$ the subset $\{s \in S' \mid t \in \mathcal{D}(s)\}$ of S' (*fig.* 14).

4.3 Restriction and Coupling

Let $G^* = [S^*, \mathcal{G}^*]$ be a coupling of two graphs $G = [S, \mathcal{G}]$ and $G' = [S', \mathcal{G}']$, and $X^* = (\Omega^*, \mathcal{T}^*, P^*)$ be a Markov random field on G^* with state space $(\Lambda, \mathcal{E}, \kappa)$ for the sites of S, and $(\Lambda', \mathcal{E}', \kappa')$ for the sites of S'. Configurations of this random field may be identified with couples (x, x') such that $x \in \Omega = \Lambda^S$ and $x' \in \Omega = \Lambda'^{S'}$. That is to say:

$$\Omega^* = \Lambda^S \times \Lambda'^{S'} = \Omega \times \Omega' \ .$$

The following notation $X^* = (X, X')$ is adopted, where X is the restriction of X^* to S, and X' is its restriction to S'. The general framework developed in §3 will enable us to learn more about these restrictions in two particular cases.

First consider the case where G^* is the coupling of a connected graph G with its edge graph, according to:

$$s \in \mathcal{G}^*_u \Longleftrightarrow u \in \mathcal{G}^*_s \Longleftrightarrow (\ \exists t \in S \ : \ u = <s,t> \ ) \tag{37}$$

for any (s, u) of $S \times \mathcal{U}$. Since G is connected, any site s of S has at least one neighbor t in G, and the edge-site u corresponding to the edge $\{s, t\}$ is a neighbor of s in G^*. So, S is the neighborhood of $\mathcal{U}$ in the coupling G^*:

$$S = \mathcal{G}^*_{\mathcal{U}} \ .$$

On the other hand, since G is connected, it is easy to see that its edge graph is also connected, and then $\mathcal{U}$ is a connected subset of graph G^*. We can apply Corollary 1: if the Markov random field X^* on G^* fulfills conditions (H1) and (H2) with $A = S$, its restriction X to S has the complete graph on S as minimal graph. In the restoration model developed by Geman and Geman [19] for example, the "intensity-process" (random field of luminance on pixel graph) and the "line-process" (binary random field defined on the edge graph) are generally nonlocal, as far as their marginal distributions are concerned.

We now turn our attention to the coupling – defined by (37)–, between the connected graph $G = [S, \mathcal{U}]$ and the trivial graph $G' = [\mathcal{U}, \emptyset]$ in which the neighborhood of each site is empty. In G^*, the neighborhood of any edge-site $<s,t>$ contains only the two sites s and t:

$$\mathcal{G}^*_{<s,t>} = \{s, t\} \ .$$

Therefore, given a Markov random field (X, X') on G^* associated with the collection of potential functions $\mathcal{V}^* = \{V^*_c = v^*_c \circ f^*_c, \ c \in \mathcal{C}^*\}$ ([6]), its restriction to S is a Markov Random Field on graph G whose Gibbs distribution is given by the collection of potential functions $\mathcal{V} = \{V_c = v_c \circ f_c, \ c \in \mathcal{C}\}$ ([7]) defined as:

$$\begin{aligned} &\forall\, x \in \Omega, \quad V_{\{s,t\}}(x) = v_{\{s,t\}}(x_s, x_t) \stackrel{\triangle}{=} \\ &- \ln \left[\int_{x'_u \in \Lambda'} \exp \left\{ -v^*_{\{u\}}(x'_u) - v^*_{\{s,u\}}(x_s, x'_u) - v^*_{\{t,u\}}(x_t, x'_u) - v^*_{\{s,t,u\}}(x_s, x_t, x'_u) \right\} \kappa'(dx'_u) \right] \\ &+ \ v^*_{\{s,t\}}(x_s, x_t) \end{aligned} \tag{38}$$

[6] f^*_c denotes the canonical surjective map from Ω^* into Ω^*_c and v^*_c denotes some measurable real-valued mapping from Ω^*_c.

[7] f_c denotes the canonical surjective map from Ω into Ω_c and v_c denotes some measurable real-valued mapping from Ω_c.

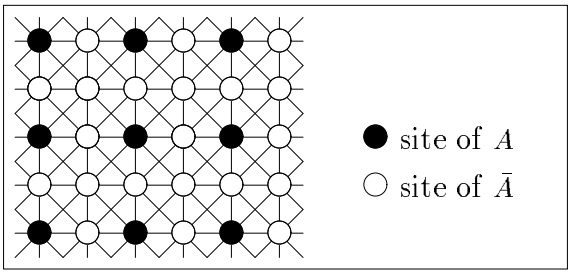

Figure 12: *Decimation by a factor of 2 in each direction of a bidimensional regular lattice associated with an 8-neighborhood system*

level k as minimal graph: for this field, all sites at level k become neighbors. Figure 13 gives an illustration of this phenomenon of neighborhood through upper levels, in the restriction to the deepest level of a pyramid.

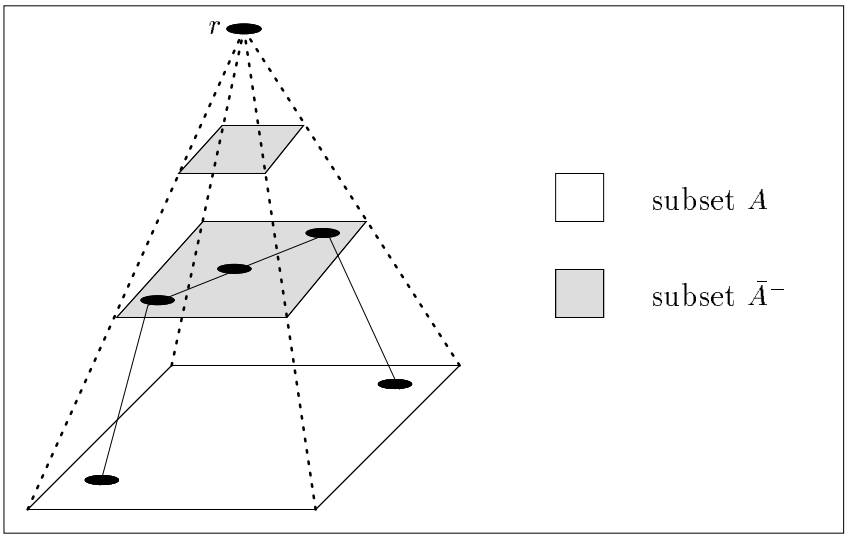

Figure 13: *Restriction to the deepest level of a pyramid*

Thus, when handling a "global" Markovian model defined on a hierarchical structure such as a tree or a pyramid, 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 [7, 34] make use of the whole hierarchical structure to derive efficient algorithms.

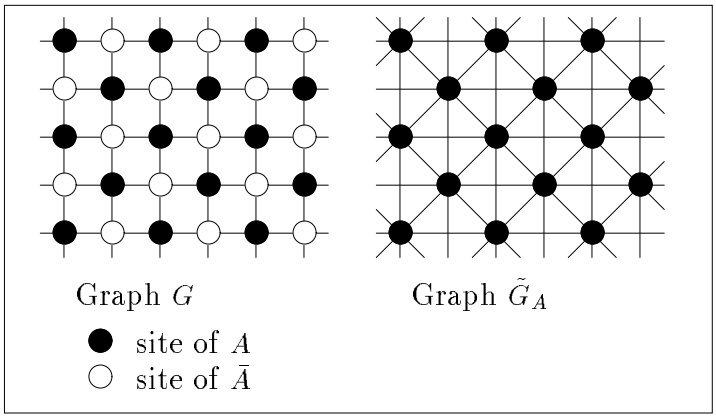

Figure 11: *Quincunx decimation of a regular bidimensional lattice with the 4-neighborhood system, and graph $\tilde{G}_A$ associated with the restricted random field X_A*

$(1 < j < q)$ 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 $\bar{A}$, l is also included in $\bar{A}$. 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 $\bar{A}$ and whose endpoints are s and t: $\bar{A}$ 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. If, for example, 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 *Let X be a Markov random field on a regular d-dimensional lattice with a neighborhood system of order greater than 1, and A be some monochromatic component of the corresponding graph. Assuming (H1) and (H2), the restricted random field 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 8-neighborhood system (*fig.* 12) corresponds to this case: it generally transforms a Markov random field on this graph to a field which is no more local.

4.2 Restrictions on Trees and Pyramids

Let us consider a Markov random field on a tree or a pyramid with root r, and let us consider its restriction to a given depth level S_k $(k > 0)$: $A = S_k$ (see Section 2.1.2). The reunion $\bar{A}^- \stackrel{\triangle}{=} \cup_{l=0,\ldots,k-1} S_l$ of the upper levels (i.e levels toward the root) is a connected component of $\bar{A}$. Moreover any site of A has a neighbor (its father) in $\bar{A}^-$: $A = \mathcal{G}_{\bar{A}^-}$. If (H1) and (H2) are verified, according to Corollary 1, the restricted field X_A has the complete graph on the

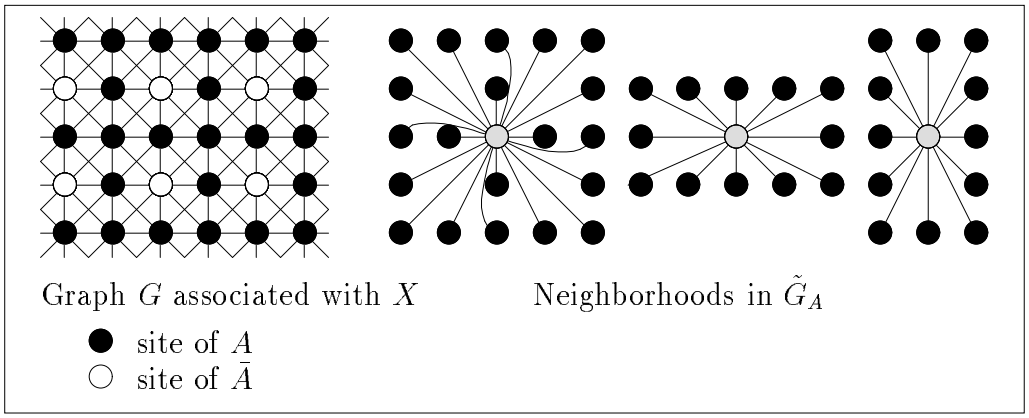

Figure 10: *Restriction to the complementary of a monochromatic component of a bidimensional regular lattice with an 8-neighborhood system. Different classes of neighborhoods in graph* $\tilde{G}_A$

Let us return to the general case. If, in addition, the graph G is 2-chromatic and 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 $\bar{A}$. 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 (excepted 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 with respect to the euclidian distance. If $d = 2$, the graph $\tilde{G}_A$ associated with the restricted random field is regular of degree 8 (*fig.* 11).

4.1.2 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 $\bar{A} = \cup_{k \neq k_0} C_k$. Assuming the neighborhood of any site of G is a connected subset, one can establish the following result:

Proposition 5 *Let $G = [S, \mathcal{G}]$ be a p-chromatic finite simple nondirected graph such that each element of $\mathcal{G}$ is connected. Let A be a monochromatic component of G. For any Markov random field with G as minimal graph and such that (H1) and (H2) hold, its restriction to 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 $\bar{A}$ and due to the assumption of connectedness of all neighborhoods, $\bar{A}$ is connected. As a matter of fact, for any two sites s and t of $\bar{A}$, since G is connected, there exists a chain $L = (\{s_1 = s, s_2\}, \{s_2, s_3\}, \ldots, \{s_{q-1}, s_q = t\})$ joining s and t. Let us assume that some s_j

4.1 Coloring and Subsampling

$G = [S, \mathcal{G}]$ being a p-chromatic ($p \geq 2$) simple nondirected graph, let us consider a decomposition of its site set S in p monochromatic components C_k, $k = 1, \ldots, p$. Let us recall that two arbitrary sites of the same monochromatic component are never neighbors: the connected components of any C_k are singletons.

4.1.1 Restriction to the Complementary of a Monochromatic Component

X being a Markov random field on G, we are going to focus on its restriction to the complementary of a monochromatic component C_{k_0}: $A = \cup_{k \neq k_0} C_k$ and $\bar{A} = C_{k_0}$.

Proposition 4 *Let X be a Markov random field on the minimal graph G and A be the complementary 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 $\bar{A}$: s has $q \geq 1$ neighbors in $\bar{A}$, denoted by $t_1, \ldots, t_q$. q is given by:

$$q = |\mathcal{G}_s \cap \bar{A}| = d_G(s) - |\mathcal{G}_s \cap A| \ .$$

Each singleton $\{t_i\}$ is a connected component of $\bar{A}$: $\mathcal{G}_{t_i} \subset A$. According to (29), 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:

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

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 $\tilde{G}_A$:

$$d_{\tilde{G}_A} \leq d_G(d_G - 1) - (d_G - 2) \min_{s \in A} |\mathcal{G}_s \cap A| \quad (35)$$

$$\leq d_G(d_G - 1) \ . \quad (36)$$

■

To illustrate this kind of restriction, let us consider the coloring with four colors of a regular bidimensional lattice associated with an 8-neighborhood system ($d_G = 8$). Given a Markov random field on such a graph, its restriction to the complementary of one of the four monochromatic components is a Markov random field on a graph of degree less or equal to 56. In this particular case, it is easy to see that $\min_{s \in A} |\mathcal{G}_s \cap A| = 4$. A more accurate bound is given by (35) which provides an upper bound of 32. $\tilde{G}_A$ is in fact a nonregular graph of degree 20 (*fig.* 10).

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

Corollary 1 *If A has no interior sites and if $\bar{A}$ 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 analysis with Markovian models: G being a bidimensional regular lattice associated with a 4-neighborhood system, the decimation by a factor of two in each direction (*fig.* 9) 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 to a nonlocal random field in which every two sites are functionally dependent, assuming (H1) and (H2).

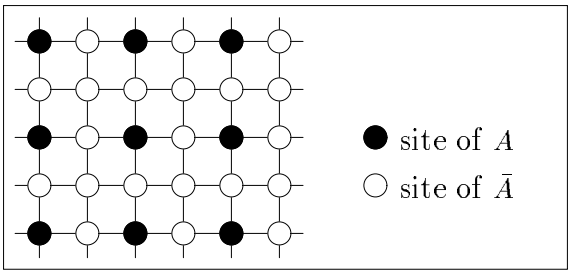

Figure 9: *Subsampling by a factor of two in each direction applied to a bidimensional regular lattice associated with a 4-neighborhood system*

4 Consequences and Applications

In this section, we present various consequences and applications of the general framework developed in Section 3, as far as multiresolution statistical image analysis is concerned. The first consequence concerns the subsampling of MRFs. 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, 7, 34] or other pyramidal graph structures [7, 29, 34] are studied. The restriction, at a given scale, of a model defined on such a hierarchical structure show 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. Finally, the special case of Gauss-Markov random fields is considered and the general conditions derived in Section 3 are developed for these linear models.

3.2.2 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 potential functions V_i associated with the site subsets $\mathcal{G}_{\bar{A}_i}$: for any $i \in I$ there exists a collection $\{A_{i,k}, k \in K_i\}$ of $|K_i|$ nonempty subsets of $\mathcal{G}_{\bar{A}_i}$ and there exists a collection $\{V_{i,k}, k \in K_i\}$ of irreducible potential functions on these subsets such that:

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

Using these decompositions, one can rewrite (27) :

$$P_A(T_A) = \frac{1}{Z} \int_{x_A \in T_A} \exp\{- \sum_{c \in \mathcal{C}_A} V_c(x_A, a_{\bar{A}}) - \sum_{i \in I} \sum_{k \in K_i} V_{i,k}(x_A)\}\, \kappa^A(dx_A)\ . \tag{32}$$

Let us define on the clique set $\tilde{\mathcal{C}}_A$ the following new collection of potential functions denoted by $\tilde{\mathcal{V}}'_A = \{\tilde{V}'_c,\ c \in \tilde{\mathcal{C}}_A\}$:

$$\tilde{V}'_c(x_A) \triangleq \begin{cases} V_c(x_A, a_{\bar{A}}) & \text{if } c \in \mathcal{C}_A \text{ and } c \notin \{A_{i,k}, i \in I, k \in K_i\}\ ; \\ V_{i,k}(x_A) & \text{if } c = A_{i,k} \text{ and } A_{i,k} \notin \mathcal{C}_A\ ; \\ V_{i,k}(x_A) + V_c(x_A, a_{\bar{A}}) & \text{if } c = A_{i,k} \text{ and } A_{i,k} \in \mathcal{C}_A\ ; \\ 0 & \text{otherwise,} \end{cases} \tag{33}$$

for any clique c of $\tilde{\mathcal{C}}_A$. Then (32) may be rewritten as:

$$P_A(T_A) = \frac{1}{Z} \int_{x_A \in T_A} \exp\{- \sum_{c \in \tilde{\mathcal{C}}_A} \tilde{V}'_c(x_A)\}\, \kappa^A(dx_A)\ ,$$

for any $T_A \in \mathcal{T}_A$, with $Z = \int_{x_A \in \mathcal{T}_A} \exp\{- \sum_{c \in \tilde{\mathcal{C}}_A} \tilde{V}'_c(x_A)\}\, \kappa^A(dx_A) < +\infty$.

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

- *(H1) for any $A_{i,k} = c$ belonging to $\mathcal{C}_A$, $[x_A \mapsto V_{i,k}(x_A) + V_c(x_A, a_{\bar{A}})]$ is irreducible;*
- *(H2) for any two different sites s and t of $\mathcal{G}_{\bar{A}_i}$, 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 potential functions V_c and the irreducibility of potential functions $V_{i,k}$, (H1) implies that $\tilde{\mathcal{V}}'_A$ is a collection of irreducible or zero potential functions. Let $\{s,t\}$ be some element of $\tilde{\mathcal{C}}_A$:

- if $\{s,t\} \not\subset \mathcal{G}_{\bar{A}_i}$, 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 $\{s,t\} \not\subset \mathcal{G}_{\bar{A}}$. Finally, $c \in \tilde{\mathcal{C}}_A$ and $\tilde{V}'_c(.) = V_c(., a_{\bar{A}}) \neq 0$.
- If $\{s,t\} \subset \mathcal{G}_{\bar{A}_i}$ (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 $\tilde{\mathcal{V}}'_A$ defined in (33) contains only irreducible or zero potential functions such that

$$\forall\, \{s,t\} \in \tilde{\mathcal{C}}_A,\ \ \exists c \in \tilde{\mathcal{C}}_A\ :\ \{s,t\} \subset c \text{ and } \tilde{V}'_c \neq 0\ . \tag{34}$$

As a consequence, according to Proposition 3, $\tilde{G}_A$ is the minimal graph of X_A. ■

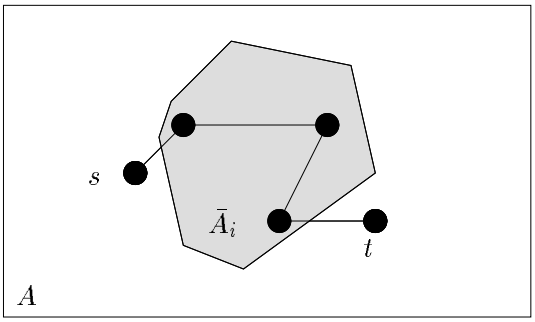

Figure 8: *Sites s and t of A which become neighbors in the restriction of X to A*

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

$$\tilde{\mathcal{G}}_s \stackrel{\Delta}{=} \mathcal{G}_s \ ; \tag{28}$$

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

$$\tilde{\mathcal{G}}_s \stackrel{\Delta}{=} (\mathcal{G}_s \cap A) \cup (\cup_{i \in K} \mathcal{G}_{\bar{A}_i} - \{s\}) \tag{29}$$

  where K is the set of indexes of connected components of $\bar{A}$ which have s in their neighborhood:

$$s \in \cap_{i \in K} \mathcal{G}_{\bar{A}_i} \text{ and } s \notin \cup_{i \in I-K} \mathcal{G}_{\bar{A}_i}.$$

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 $\bar{A}$:

$$\tilde{\mathcal{C}}_A \ \supset \ \mathcal{C}_A \cup \{\mathcal{G}_{\bar{A}_i}, \ i \in I\} \ . \tag{30}$$

Let us associate with this clique set the potential function collection $\tilde{\mathcal{V}}_A = \{\tilde{V}_c, \ c \in \tilde{\mathcal{C}}_A\}$ defined as:

$$\tilde{V}_c(x_A) \stackrel{\Delta}{=} \begin{cases} V_c(x_A, a_{\bar{A}}) & \text{if } c \in \mathcal{C}_A \text{ and } c \notin \{\mathcal{G}_{\bar{A}_i}, i \in I\} \ ; \\ V_i(x_A) & \text{if } c = \mathcal{G}_{\bar{A}_i} \text{ and } \mathcal{G}_{\bar{A}_i} \notin \mathcal{C}_A \ ; \\ V_i(x_A) + V_c(x_A, a_{\bar{A}}) & \text{if } c = \mathcal{G}_{\bar{A}_i} \text{ and } \mathcal{G}_{\bar{A}_i} \in \mathcal{C}_A \ ; \\ 0 & \text{otherwise,} \end{cases} \tag{31}$$

for any clique c in $\tilde{\mathcal{C}}_A$. Expression (27) becomes:

$$P_A(T_A) = \frac{1}{Z} \int_{x_A \in T_A} \exp\{-\sum_{c \in \tilde{\mathcal{C}}_A} \tilde{V}_c(x_A)\} \, \kappa^A(dx_A) \ ,$$

for any $T_A \in \mathcal{T}_A$, with $Z = \int_{x_A \in \mathcal{T}_A} \exp\{-\sum_{c \in \tilde{\mathcal{C}}_A} \tilde{V}_c(x_A)\} \, \kappa^A(dx_A) < +\infty$. This proves that X_A is Markovian on the graph $\tilde{G}_A$. ■

Let us now study the probability measure P_A generated on the measurable space $(\Omega_A, \mathcal{T}_A)$ by f_A, the canonical mapping from Ω into Ω_A. Let T_A be some element of $\mathcal{T}_A$. Its pre-image under f_A is:

$$\begin{aligned} f_A^{-1}(T_A) &= \{x \in \Omega \mid x_A \in T_A\} \\ &= \{(x_A, x_{\bar{A}}),\ x_A \in T_A,\ x_{\bar{A}} \in \Omega_{\bar{A}}\} \ . \end{aligned}$$

The expression of the probability measure P_A may be derived as following:

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

where a is an arbitrary configuration, since $\exp\{-\sum_{c \in \mathcal{C}_A} V_c(x)\}$ is independent of $x_{\bar{A}}$,

$$= \frac{1}{Z} \int_{x_A \in T_A} \left[\exp\{-\sum_{c \in \mathcal{C}_A} V_c(x_A, a_{\bar{A}})\} \prod_{i \in I} \int_{x_{\bar{A}_i} \in \Omega_{\bar{A}_i}} \exp\{-\sum_{c \in \mathcal{C}_i} V_c(x_A, x_{\bar{A}_i}, a_{\bar{A}-\bar{A}_i})\} \kappa^{\bar{A}_i}(dx_{\bar{A}_i}) \right] \kappa^A(dx_A)$$

since $\exp\{-\sum_{c \in \mathcal{C}_i} V_c(x)\}$ is independent of $x_{\bar{A}-\bar{A}_i}$.

$\int_{x_{\bar{A}_i} \in \Omega_{\bar{A}_i}} \exp\{-\sum_{c \in \mathcal{C}_i} V_c(x_A, x_{\bar{A}_i}, a_{\bar{A}-\bar{A}_i})\}\, \kappa^{\bar{A}_i}(dx_{\bar{A}_i})$ being a positive function of variables $x_{\mathcal{G}_{\bar{A}_i}}$, one can define for any $i \in I$ the following potential function on $\mathcal{G}_{\bar{A}_i}$ (neighborhood of the connected component $\bar{A}_i$):

$$\begin{aligned} V_i \ : \ \Omega_A &\longrightarrow \mathbb{R} \\ x_A &\longmapsto V_i(x_A) \triangleq -\ln \left[\int_{x_{\bar{A}_i} \in \Omega_{\bar{A}_i}} \exp\{-\textstyle\sum_{c \in \mathcal{C}_i} V_c(x_A, x_{\bar{A}_i}, a_{\bar{A}-\bar{A}_i})\}\, \kappa^{\bar{A}_i}(dx_{\bar{A}_i}) \right] \end{aligned} \tag{26}$$

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

$$P_A(T_A) = \frac{1}{Z} \int_{x_A \in T_A} \exp\{-\sum_{c \in \mathcal{C}_A} V_c(x_A, a_{\bar{A}}) - \sum_{i \in I} V_i(x_A)\}\, \kappa^A(dx_A) \ . \tag{27}$$

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 $\bar{A}$ (i.e. $\exists i \in I \ : \{s,t\} \subset \mathcal{G}_{\bar{A}_i}$) , that is to say, there exists a chain joining s and t and whose all sites, excepted its endpoints s and t, belong to $\bar{A}$ (*fig.* 8).
The neighborhood system $\tilde{\mathcal{G}}$ of graph $\tilde{G}_A$ is thus defined in the following way: for any site s of A,

Let us notice that this decomposition of any potential function as a sum of irreducible potential functions is not unique.

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

Proposition 3 *Let $X = (\Omega, \mathcal{T}, P)$ be a Markov random field on graph $G = [S, \mathcal{G}]$. $\mathcal{G}$ is the minimal neighborhood system of X if and only if there exists a collection $\mathcal{V} = \{V_c, c \in \mathcal{C}\}$ of zero or irreducible potential functions indexed by the clique set $\mathcal{C}$ of graph G such that X is the Gibbs random field associated with $\mathcal{V}$ and:*

$$\forall \{s,t\} \in \mathcal{C},\ \exists c \in \mathcal{C}\ :\ \{s,t\} \subset c \text{ and } V_c \neq 0\ . \tag{22}$$

The proof is given in Appendix B, together with the proof of Hammersley-Clifford's Theorem.

3.2 Properties of the Restricted Field

Let $X = (\Omega, \mathcal{T}, P)$ be a Markov random field on a finite connected minimal graph $G = [S, \mathcal{G}]$, associated with the collection $\mathcal{V} = \{V_c, c \in \mathcal{C}\}$ of zero or irreducible potential functions verifying (22). Let A be a nonempty subset of S, $A \neq S$. We are interested in the statistical properties of the restricted field $X_A = (\Omega_A, \mathcal{T}_A, P_A)$. 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.

3.2.1 Joint Distribution of the Restriction and Markovianity

Theorem 2 *Let X be a Markov random field on the finite graph $G = [S, \mathcal{G}]$ and A be a nonempty subset of S, $A \neq S$. Restriction X_A of X to the site subset A is a Markov random field on 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 or if they belong to the neighborhood of a same connected component of $\bar{A}$.*

Proof. Let us denote by $\bar{A}_i, i \in I$ the connected components of $\bar{A} = S - A$. 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 $\bar{A}$ is either empty or included in only one connected component of $\bar{A}$ (i.e. there exists an unique $i \in I$ such that $(c \cap \bar{A}) \subset \bar{A}_i$). The clique set $\mathcal{C}$ can thus be partitioned in the following way:

$$\mathcal{C} = \left(\bigcup_{i \in I} \{c \in \mathcal{C} \mid c \cap \bar{A}_i \neq \emptyset\} \right) \bigcup \{c \in \mathcal{C} \mid c \subset A\} \tag{23}$$

that is to say:

$$\mathcal{C} = (\cup_{i \in I} \mathcal{C}_i) \cup \mathcal{C}_A \tag{24}$$

where $\mathcal{C}_i$ denotes the set $\{c \in \mathcal{C} \mid c \cap \bar{A}_i \neq \emptyset\}$ of $\mathcal{C}$ and $\mathcal{C}_A$ designates the set $\{c \in \mathcal{C} \mid c \subset A\}$. Using this decomposition, we can write for any configuration $x \in \Omega$:

$$\exp\{-\sum_{c \in \mathcal{C}} V_c(x)\} \ = \ \exp\{-\sum_{c \in \mathcal{C}_A} V_c(x)\}.\prod_{i \in I} \exp\{-\sum_{c \in \mathcal{C}_i} V_c(x)\}\ . \tag{25}$$

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 *Let $X = (\Omega, \mathcal{T}, P)$ be a Markov random field with site set S and with state space $(\Lambda, \mathcal{E}, \kappa)$. If for any site s, $\mathcal{G}_s$ denotes the set of sites $t \neq s$ such that the function $Pr\{X_s \in T_s \mid X_r = x_r,\ r \in S - \{s\}\}$ defined on $\mathcal{E} \times \Omega_{S-\{s\}}$ depends on x_t ([5]), 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 proof of that proposition is given in Appendix B. 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 potential functions V_A which really depend on variables $x_s, s \in A$, and which cannot be decomposed in sums of potential functions on subsets of A different from A. 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 potential function on $A - B$ and the zero potential function on B.

Definition 6 *Let S be a finite site set, Ω be a configuration space on S, A be a nonempty subset of S and V_A be a potential function on A. V_A is called an irreducible potential function on A if and only if:*

- *$|A| = 1$ and V_A is nonzero,*
- *or $|A| \geq 2$ and there exists no collection $\{V_{A_i}, i \in I\}$ of potential functions indexed by a finite collection $\{A_i, i \in I\}$ of subsets of A, distinct from A, such that $V_A = \sum_{i \in I} V_{A_i}$.*

The class of irreducible potential functions contains the different canonical nonzero potential functions defined by Besag [4], Snell & Kindermann [16, 30] and Descombes [15] (see Appendix A for the proof). Using this fact, it is easy to prove the following property (see Appendix A for the detail):

Proposition 2 *For any potential function V_A on a subset A of S, there exists a finite 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 potential functions such that:*

$$V_A \ = \ \sum_{i \in I} V_{A_i} \ . \tag{21}$$

[5] 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

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

.

Theorem 1 (Hammersley–Clifford) *Let $X = (\Omega, \mathcal{T}, P)$ be a random field with site set S and $\mathcal{G}$ be a neighborhood system on S. X is a Markov random field with respect to $\mathcal{G}$ and verifies $P(T) > 0$, $\forall T \in \mathcal{T}$ (positiveness property) if and only if X is a Gibbs random field with respect to $\mathcal{G}$.*

In Appendix B, we give in the case of a finite configuration space a proof of this theorem using the generalized canonical potential functions introduced by Descombes *et al.* [15] and presented in Appendix A.

The *local conditional distribution* at any site can be expressed with potential functions indexed by all the cliques containing the site under concern (for the proof see Appendix B or [4]):

$$Pr\{X_s \in T_s \mid X_{\mathcal{G}_s} = x_{\mathcal{G}_s}\} = \frac{1}{Z_s(x_{\mathcal{G}_s})} \int_{x_s \in T_s} \exp\{- \sum_{c \in \mathcal{C} \,:\, s \in c} V_c(x)\} \, \kappa(dx_s) \tag{19}$$

where the local normalization constant is function of $x_{\mathcal{G}_s}$:

$$Z_s(x_{\mathcal{G}_s}) \stackrel{\triangle}{=} \int_{\lambda \in \Lambda} \exp\{- \sum_{c \in \mathcal{C} \,:\, s \in c} V_c(x^{s,\lambda})\} \, \kappa(d\lambda) \; , \tag{20}$$

$x^{s,\lambda}$ being the configuration which agrees with x on $S - \{s\}$ and takes the value λ at site s.

3 Restriction of a MRF to a Subset of Sites

Different approaches have been proposed recently to combine Markov Random Field (MRF) models and multigrid algorithms in image analysis: renormalization group [21], subsampling of stochastic processes [27, 32], MRFs defined on trees or pyramids [3, 7, 29, 34] 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 finite arbitrary nondirected graph) to a part of its original site set. This section addresses this problem by considering the restriction of MRFs (defined on finite arbitrary nondirected graphs) to a subset of nodes of the original graph. 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 a MRF. Sufficient conditions for the new neighborhood structure to be "minimal" are derived.

3.1 Minimal Graph and Irreducible Potential Functions

Let us consider a random field X defined on site set S. The markovianity of X with respect to a neighborhood system $\mathcal{G}$ on S does not necessary account for the "real" functional 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:

$$\forall \, s \in S, \;\; \mathcal{G}_s \subset \mathcal{G}'_s \; .$$

Definition 4 *Let Ω be a configuration space on site set S and A be a nonempty subset of S. A potential function on A is a measurable map from Ω into $\mathbb{R}$ which is $\mathcal{T}_A$-measurable, that is to say which depends only on variables x_s, $s \in A$.*

A real-valued map V_A defined on Ω is a potential function on A if it is the composition of the canonical surjective map f_A from Ω into Ω_A with some measurable real-valued map v_A defined on Ω_A:

$$\begin{array}{rccl} V_A & : & \Omega & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & V_A(x) = [v_A \circ f_A](x) = v_A(x_A) \end{array}$$

Particularly, any two configurations of Ω which agree together on A have the same image under a potential function on A:

$$\forall\, (x, y) \in \Omega \times \Omega, \quad x_A = y_A \implies V_A(x) = V_A(y) \ .$$

Definition 5 *Let $X = (\Omega, \mathcal{T}, P)$ be a random field with site set S and $\mathcal{G}$ 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 a collection $\mathcal{V} = \{V_c,\ c \in \mathcal{C}\}$ of potential functions indexed by the clique set $\mathcal{C}$ of graph G such that:*

$$(i) \quad Z \triangleq \int_{x \in \Omega} \exp\{-\sum_{c \in \mathcal{C}} V_c(x)\}\, \kappa^S(dx) \ < \ +\infty \ , \tag{13}$$

$$(ii) \quad \forall\, T \in \mathcal{T}, \quad P(T) = \frac{1}{Z} \int_{x \in T} \exp\{-\sum_{c \in \mathcal{C}} V_c(x)\}\, \kappa^S(dx) \ . \tag{14}$$

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. Let us notice that the collection of potential functions associated with a Gibbs random field is not unique. Nevertheless, it can be unique under certain *canonicity criteria* [4, 16, 30] (*cf* Appendix A). In the case of a continuous field, properties (13) and (14) become:

$$(ii) \quad Z \triangleq \textstyle\int_{x \in \Omega} \exp\{-\sum_{c \in \mathcal{C}} V_c(x)\}\, dx \ < \ +\infty \ , \tag{15}$$

$$(iii) \quad \forall\, x \in \Omega, \quad p_X(x) = \tfrac{1}{Z} \exp\{-\textstyle\sum_{c \in \mathcal{C}} V_c(x)\}, \tag{16}$$

where p_X is the *probability density* of measure P. In the case of discrete random fields, properties (13) and (14) become:

$$(ii) \quad Z \triangleq \textstyle\sum_{x \in \Omega} \exp\{-\sum_{c \in \mathcal{C}} V_c(x)\} \ < \ +\infty \ , \tag{17}$$

$$(iii) \quad \forall\, x \in \Omega, \quad p_X(x) = \tfrac{1}{Z} \exp\{-\textstyle\sum_{c \in \mathcal{C}} V_c(x)\}, \tag{18}$$

where $p_X(x)$ denotes the probability $Pr\{X = x\}$.

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

Stochastic Transformations Let $X = (\Omega, \mathcal{T}, P)$ be a random field with site set S and with state space Λ. Let S' be another site set and let Λ' be another state space. A random field $Y = (\Omega', \mathcal{T}', P')$ on S' can be defined through the conditional distribution:

$$\begin{array}{rcl} \mathcal{T}' \times \Omega & \longrightarrow & [0,1] \\ (T', x) & \longmapsto & Pr\{Y \in T' \mid X = x\} \end{array} \tag{9}$$

We will say that Y is generated under a *stochastic transformation* of random field X. Under the additional assumption that, for any site s of S', conditional distribution of Y_s depends only on $X_{\mathcal{D}(s)}$, with $\mathcal{D}(s) \subset S$, that is to say:

$$Pr\{Y_s \in T'_s \mid X = x,\ Y_r \in T'_r,\ r \neq s\} = Pr\{Y_s \in T'_s \mid X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\}\ , \tag{10}$$

the conditional distribution of Y given X takes the following form:

$$\forall\, T' \in \mathcal{T}',\ \forall\, x \in \Omega,\quad Pr\{Y \in T' \mid X = x\} = \prod_{s \in S'} Pr\{Y_s \in T'_s \mid X_{\mathcal{D}(s)} = x_{\mathcal{D}(s)}\}\ . \tag{11}$$

This kind of "local" stochastic transformations are used to model the relationship between an "observable" random field Y and a "hidden" random field X in the Bayesian estimation framework [13]. Such transformations are also used in the Renormalization Group approach [21].

2.2.2 Markov Random Fields and Gibbs Random Fields

Definition 3 *Let $X = (\Omega, \mathcal{T}, P)$ be a random field with finite site set S and with state space $(\Lambda, \mathcal{E}, \kappa)$. 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:*

$$Pr\{X_s \in T_s \mid X_r = x_r,\ r \in S - \{s\}\} \ = \ Pr\{X_s \in T_s \mid X_r = x_r,\ r \in \mathcal{G}_s\} \tag{12}$$

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 (12) is that the contextual information provided by $S - \{s\}$ is the same as the statistical information provided by the neighbors of s [19, 13] ([4]). 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 of these models. 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 (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.

[4] This property of course does not mean that two arbitrary random variables X_s and X_t are independent if s and t are not neighboring. But we will say that these two variables are *functionally independent.*

- $\mathcal{E}$ is the Borel σ-algebra and κ is the Borel measure, if Λ is not of measure zero;
- $\mathcal{E}$ is the subset σ-algebra and κ the countage measure, if Λ is a finite or countable subset of $\mathbb{R}^q$.

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. One calls random field with site set S and with state space Λ any triple $X = (\Omega, \mathcal{T}, P)$ such that:*

- $(\Omega, \mathcal{T}) = (\Lambda, \mathcal{E})^S$,
- *P is a probability measure on the product σ-algebra $\mathcal{T}$.*

If Λ is finite or countable, it's a *discrete random field*. If Λ is not of measure zero for the Borel measure on $\mathbb{R}^q$, the random field is said *continuous*. Given a random field X, we consider for any site s the random variable X_s with values in Λ and defined as:

$$\begin{array}{rccc} X_s \ : & (\Omega, \mathcal{T}, P) & \longrightarrow & (\Lambda, \mathcal{E}) \\ & x & \longmapsto & X_s(x) = x_s \end{array}$$

To simplify, one often writes $X = \{X_s, s \in S\}$.

Deterministic Transformation of a Random Field Let $X = (\Omega, \mathcal{T}, P)$ be a random field on S with state space Λ. Let S' be another site set and $(\Lambda', \mathcal{E}')$ be another state space. If f is a measurable map from $(\Omega, \mathcal{T}, P)$ on $(\Omega', \mathcal{T}') \triangleq (\Lambda', \mathcal{E}')^{S'}$, then it defines a probability measure P_f on $(\Omega', \mathcal{T}')$:

$$\forall T \in \mathcal{T}', \ P_f(T) \triangleq P(f^{-1}(T))$$

where $f^{-1}(T)$ is the pre-image of the set T under the map f. The random field $X' \triangleq (\Omega', \mathcal{T}', P_f)$, is the image of the random field X under the map f. In the remainder, we will consider particular cases of deterministic transformations called *restrictions*.

Restriction of a Random Field Let A be a nonempty subset of S. Let us consider the canonical surjective map 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:

$$\begin{array}{rccc} f_A \ : & (\Omega, \mathcal{T}, P) & \longrightarrow & (\Omega_A, \mathcal{T}_A) \\ & x = \{x_s, \ s \in S\} & \longmapsto & x_A = \{x_s, \ s \in A\} \end{array} \tag{8}$$

The probability measure generated on $\mathcal{T}_A$ by the map f_A defines a random field $X_A = (\Omega_A, \mathcal{T}_A, P_A)$ with site set A and with state space Λ. This field will be called *restriction of random field X to 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 $(x_A, x_{\bar{A}})$ 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: $T_A = \{x_A, \ x \in T\}$.

Trees provide a useful hierarchical structure for representing stochastic processes. They have been introduced recently in signal processing [3] as well as in image processing 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 [7] and optical flow computation [34].
In trees, two different sites at a same depth level are not neighbors (otherwise, a cycle would result). A different hierarchical structure (called "pyramid") with connected depth levels has been proposed [29]. Let $\mathcal{A} = [S, \mathcal{U}]$ be a tree of root r. The *pyramid associated with* $\mathcal{A}$ is any graph $G = [S, (\cup_i \mathcal{U}_i) \cup \mathcal{U}]$ such that the subgraph generated by each level S_i is $G_{S_i} = [S_i, \mathcal{U}_i]$, and is connected (*fig.* 7).

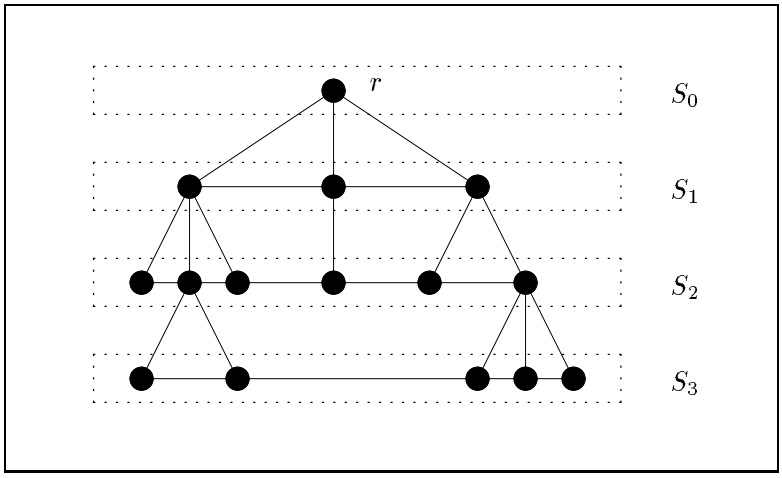

Figure 7: *Pyramid associated with a tree of root r*

2.2 Markov Random Fields on Finite Simple Nondirected Graphs

2.2.1 Random Fields: Definitions

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 state space Λ. Λ will generally be a subset of $\mathbb{R}^q$ ($q \in \mathbb{N}^*$). 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 [12]. A measure being defined on the state space Λ, we consider maps 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 map from S into Λ:*

$$\begin{array}{rccl} x & : & S & \longrightarrow & (\Lambda, \mathcal{E}, \kappa) \\ & & s & \longmapsto & x(s) = x_s \end{array}$$

As mentioned previously, the state space Λ will generally be a subset of $\mathbb{R}^q$, with its usual σ-algebra and the corresponding measure:

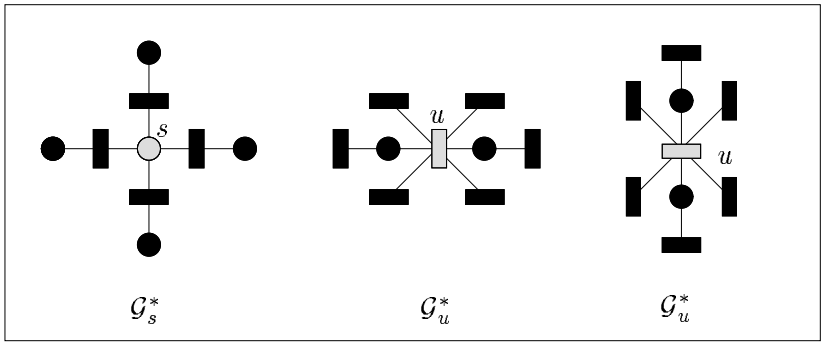

Figure 5: *Example of neighborhoods associated with the coupling between a two-dimensional lattice (with the 4-neighborhood system) and its edge graph*

Trees and Pyramids 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. Let us denote by $d(s,t)$ the length of the unique chain joining sites s and t. d is a distance on S. It is interesting –for multiresolution image analysis purposes for instance– to define a *depth* in the tree. To this end let us consider some reference site called *root* and denoted by r. The *level of depth* k, denoted by S_k, will be the set of sites at distance k from the root (*fig.* 6) :

$$s \in S_k \iff d(r,s) = k \ . \tag{7}$$

Any site s in S_k, with $k > 0$ (i.e. s is not the root) has an unique neighbor on the level S_{k-1}, called its *father*. The other neighbors of s are on the level S_{k+1} (if this level exists). They form the set of its *children*.

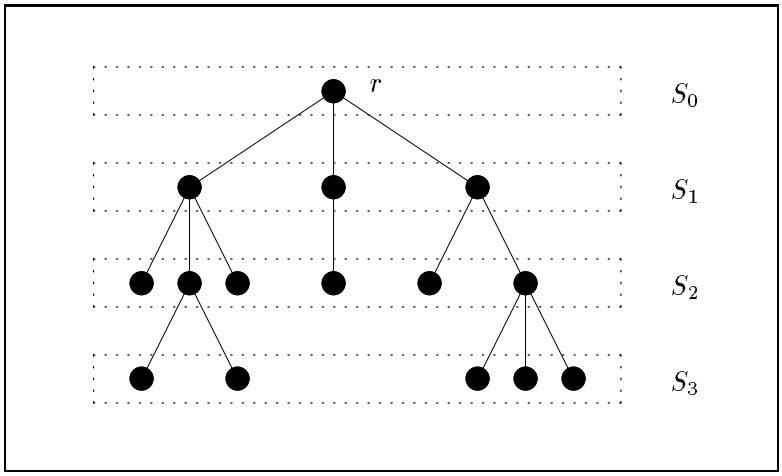

Figure 6: *Tree with root r*

Let us notice that a d-dimensional lattice with a neighborhood system of order $n \in \{1,2\}$ with respect to the euclidian distance is $2n$-chromatic. Thus a bidimensional lattice with the 4-neighborhood system (resp. 8-neighborhood system) can be colored with 2 (resp. 4) colors.

Coupling of a Graph and its Edge Graph Let $G = [S,\mathcal{U}] = [S,\mathcal{G}]$ be a graph on a lattice S. Its edge graph is denoted by $L(G)$. S and $\mathcal{U}$ are combined in a single site set S^*. An element of S^* will be called a "site" if it corresponds to an element of S, and an "edge-site" if it corresponds to an element of $\mathcal{U}$: the edge-site associated with the edge $\{s,t\} \in \mathcal{U}$ will be denoted by $<s,t>$. In image analysis, a widely used coupling of graphs G and $L(G)$ is the graph $G^* = [S^*, \mathcal{G}^*]$ in which a site s and an edge-site u are neighboring if and only if the edge corresponding to u has s as an endpoint:

$$s \in \mathcal{G}^*_u \iff u \in \mathcal{G}^*_s \iff (\ \exists t \in S\ :\ u = <s,t>\ )\ .$$

Figure 4 depicts, for instance a common case where G is defined on a two-dimensional regular lattice with the 4-neighborhood system. Its edge graph $L(G)$ is regular of degree 6 and they are two kinds of edges: "horizontal" edges and "vertical" ones (*fig.* 4 and 5).

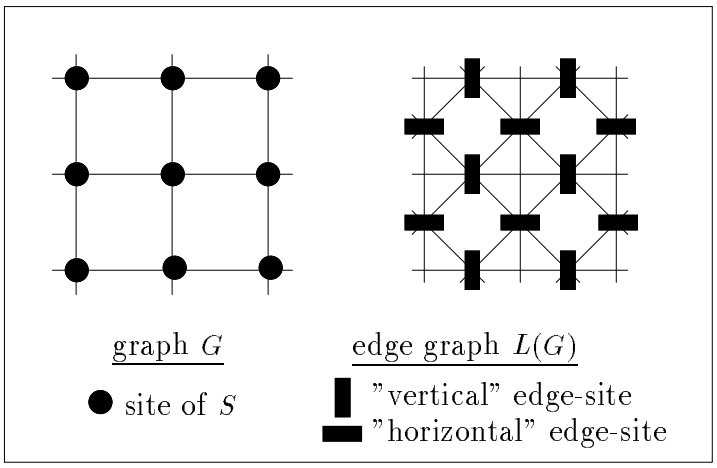

Figure 4: *Example of a bidimensional lattice with the 4-neighborhood system and its associated edge graph*

In the coupling G^*, the neighborhood of a site s is composed with its four neighbors in G and the four edge-sites associated with the four edges containing s (*fig.* 5) :

$$\mathcal{G}^*_s = \mathcal{G}_s \cup \{<s,t>,\ t \in \mathcal{G}_s\}\ .$$

The neighborhood in G^* of an edge-site $u = <s,t>$ is composed with its six neighbors in $L(G)$ and by its endpoints s and t (*fig.* 5) :

$$\mathcal{G}^*_u = \{<s,s'>,\ s' \in \mathcal{G}_s - \{t\}\}\ \cup\ \{<t,t'>,\ t' \in \mathcal{G}_t - \{s\}\}\ \cup\ \{s,t\}.$$

The sequence of *neighborhoods of order n with respect to distance D* (*fig.* 2) is defined by the following recurrence:

$$\mathcal{G}_s^{n+1} \stackrel{\Delta}{=} \arg\min_{t \notin \mathcal{G}_s^n \cup \{s\}} D(s,t) \cup \mathcal{G}_s^n \ . \tag{6}$$

9	8	7	6	7	8	9
8	5	4	3	4	5	8
7	4	2	1	2	4	7
6	3	1	s	1	3	6
7	4	2	1	2	4	7
8	5	4	3	4	5	8
9	8	7	6	7	8	9

Figure 2: *Successive neighborhoods with respect to the Euclidian distance on a two-dimensional lattice*

Among these different neighborhood systems, the most often used are the first and second order neighborhoods (with respect to the euclidian distance). In the two-dimensional case, they are also called 4-neighborhood or 8-neighborhood systems, since they associate 4 or 8 neighbors to any site which is not located on the border of the lattice (*fig.* 3).

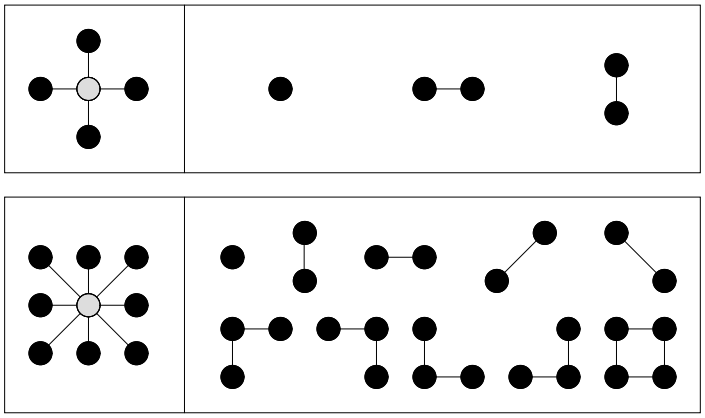

Figure 3: *First and second order neighborhood systems on a two-dimensional regular lattice (euclidian distance); associated cliques*

chain is called a *cycle*. The relation $\mathcal{R}$ defined on S as

$$s\,\mathcal{R}\,t \iff \begin{cases} s = t \\ \text{or} \\ \text{there exists a chain joining } s \text{ and } t \end{cases} \tag{4}$$

is an equivalence relation. The equivalence classes $S_1, \ldots, S_p$ induced on S are the *connected components* of graph G, and p is the *connectedness number* of graph G. The graph is said to be connected if $p = 1$, that is to say, for any pair $\{s, t\}$ of sites, there exists a chain joining s and t. Thus, each connected component S_i induces a connected subgraph G_{S_i}. In this paper we will only consider finite simple nondirected connected graphs.
The *coloring* of a graph is the assignment of a color to each site such that two neighboring sites are not of the same color. In other words it is a partition of S in k subsets $C_1, \ldots, C_k$ such that:

$$\forall j \in \{1, \ldots, k\},\ \forall \{s, t\} \subset C_j,\quad s \notin \mathcal{G}_t\ .$$

Such a partition is also called a *codage* [4]. Subsets C_j will be called *monochromatic components* throughout this paper. The graph is said to be *k-chromatic* if it can be colored with k colors. The *chromatic number* of G is the minimum number of colors necessary to color its sites.
The *edge graph* associated with graph $G = [S, \mathcal{U}]$ is the graph, denoted by $L(G)$, whose vertices represent the edges of G, two of these sites being neighbors if the two corresponding edges in G have a common endpoint. If u is a site of $L(G)$ corresponding to the edge $\{s, t\}$ of G, its degree is:

$$d_{L(G)}(u) = (d_G(s) - 1) + (d_G(t) - 1)\ .$$

If G is a regular graph of degree d_G, then its edge graph is also regular, and its degree is $d_{L(G)} = 2(d_G - 1)$.
Given two graphs $G = [S, \mathcal{G}]$ and $G' = [S', \mathcal{G}']$, it can be interesting to combine them in a single graph. To this end, let us consider a set S^* of sites whose each element corresponds either to an unique site of S, or to an unique site of S'. To simplify, we will denote $S^* = S \cup S'$, and each site of S^* is identified with its corresponding site in S or S'. A *coupling of the two graphs G and G'* is a graph $G^* = [S^*, \mathcal{G}^*]$ on S^* whose neighborhood system $\mathcal{G}^*$ verifies:

$$\begin{cases} \forall \{s, t\} \subset S,\ s \in \mathcal{G}^*_t \iff s \in \mathcal{G}_t \\ \forall \{s, t\} \subset S',\ s \in \mathcal{G}^*_t \iff s \in \mathcal{G}'_t \end{cases} \tag{5}$$

G^* is obtained by introducing edges connecting sites of S and S'.

2.1.2 Special Cases

Finite Lattices A *d-dimensional lattice* ($d \in \mathbb{N}^*$) is a subset of $\mathbf{Z}^d$. 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).
Using metrics on $\mathbb{R}^d$, it is easy to define a *distance* on a lattice. Given such a distance D on S, one can define the following sequence of neighborhood systems: $\mathcal{G}^1_s$ is the set of the nearest neighbors of s,

$$\mathcal{G}^1_s \stackrel{\triangle}{=} \operatorname*{argmin}_{t \neq s} D(s, t)\ ,$$

The *degree* of a site s, denoted by $d_G(s)$, is the number of edges which have s as an endpoint, i.e. the number of its neighbors:

$$d_G(s) \triangleq |\mathcal{G}_s| \ .$$

The *degree of the graph* G, denoted by d_G, is the maximal degree of a site of the graph:

$$d_G \triangleq \max_{s \in S} d_G(s) \ . \tag{3}$$

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 will be denoted by $\mathcal{C}_A \triangleq \mathcal{C} \cap \mathcal{P}(A)$, where $\mathcal{P}(A)$ is the set of subsets of A.
Besides, we will call *neighborhood of* A the set $\mathcal{G}_A$ of sites in $\bar{A}$ ([3]) having at least one neighbor in A (*fig.* 1). The *border of* A is the neighborhood of its complementary, denoted by $\mathcal{B}_A$ (*fig.* 1) and the *interior* $\overset{\circ}{A}$ of A is the set of sites of A which do not belong to the border of A.

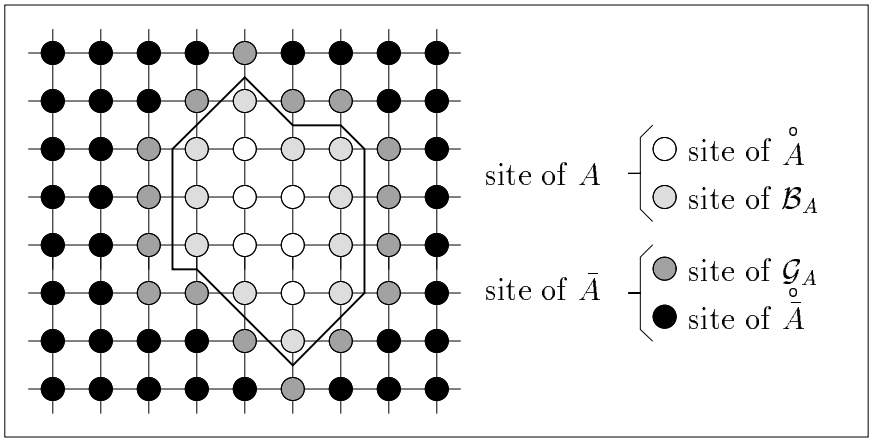

Figure 1: *Neighborhood, interior and border of some site subset A of a regular graph*

A *chain* of length q is a sequence of q edges $L = (u_1, \ldots, u_q)$ 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. The endpoint s of u_1 which does not belong to u_2, and the endpoint t of u_q which does not belong to u_{q-1}, are called *endpoints of the chain* L. The chain L *joins* the sites s and t. If the two endpoints s and t coincide, the

[3] $\bar{A}$ is the complementary of A in S: $\bar{A} \triangleq S - A$.

The different graph structures used in multiresolution Markovian modeling are described: finite lattices, trees and pyramids. The general properties of the restriction of a Markov Random Field to a part of its original site set are described in Section 3. 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 MRFs. Section 4 is devoted to various consequences and applications of this general framework in statistical multiresolution image analysis. The first consequence concerns the subsampling of MRFs. 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, 7, 34] or other pyramidal graph structures [7, 29, 34] are 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 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. Finally, the special case of Gauss-Markov random fields is considered and the general conditions derived in Section 3 are developed for these linear models. Section 5 contains concluding remarks.

2 Markov Random Field on Finite Nondirected Graphs

In this section we present some fundamentals of graph theory that will be used extensively in the remainder of this paper [22]. Markov Random Fields are defined on finite nondirected graphs and the different graph structures used in multiresolution Markovian modeling are presented (finite lattices, trees and pyramids.)

2.1 Nondirected Graphs

2.1.1 Definitions

A *simple finite nondirected graph* $G = [S, \mathcal{U}]$ is defined as a finite set S whose elements are called *vertices* or *nodes* ([2]) and a set $\mathcal{U}$ whose elements are pairs of nodes, called *edges*. If $u = \{s, t\}$ is an edge of G, then sites s and t are said to be *adjacent* or *neighboring*, and they are the *endpoints* of u. Such a graph is called *simple* because there is no loop (i.e. edge whose endpoints coincide) and there is at most one edge between two nodes. A simple nondirected graph can be equivalently defined as its vertice set S and the map $\mathcal{G}$, called *neighborhood system*, which associates with each site s its *neighborhood*, i.e. the set of its neighbors. Therefore, the graph $G = [S, \mathcal{U}]$ will also be denoted by $G = [S, \mathcal{G}]$. This second notation will be used throughout this paper, unless otherwise stated. Due to the definition of edge set $\mathcal{U}$, a neighborhood system has the following properties:

$$(i) \quad \forall s \in S,\ s \notin \mathcal{G}_s\ , \tag{1}$$

$$(ii) \quad \forall \{s, r\} \subset S,\ r \in \mathcal{G}_s \Longleftrightarrow s \in \mathcal{G}_r\ . \tag{2}$$

[2] In image analysis vertices are usually called *sites*

Lakshmanan has shown that simple resolution transformations such as subsampling or block averaging also result in a loss of locality for Gaussian MRF models. Jeng [27] has studied the loss of locality in a periodic subsampling of MRF models with infinite support. Other approaches ([1]) introduced recently aim at defining MRF models on hierarchical structure such as trees [3, 7, 34] or other pyramidal graph structures [7, 29, 34].
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 MRFs with either discrete or continuous state spaces, defined on finite arbitrary nondirected graphs. Several general properties of the restricted field are derived. The general form of the distribution of the restriction is given. A neighborhood structure with respect to which the restricted field is a MRF is exhibited. The local properties of the restriction are studied and sufficient conditions for the new neighborhood structure to be "minimal" are derived. These general results may be applied to the different multiresolution approaches previously mentioned. Several consequences and applications to various multiresolution MRF-based modeling approaches are presented. In particular, the following case studies are examined thoroughly:

- subsampling of a MRF by elimination of a monochromatic component (set of sites associated with the same color in a coloration of the graph), standard decimations by a factor $\sqrt{2}$ or 2 in each direction of a MRF on a regular bidimensional lattice, etc. ;
- analysis of the restriction at a single resolution level of a global Markovian model on a multilevel graph structure (tree or pyramid [3, 7]) ;
- stochastic transform of a Markov random field using the renormalization group approach [21] ;
- expression of the restriction in the particular case of a Gaussian MRF.

This study generalizes and unifies the works of Jeng [27] and Lakshmanan [32], both devoted to the subsampling of MRFs. Lakshmanan [32] studies the subsampling of linear MRFs (namely Gauss-Markov random fields). The mathematical framework described here applies to general classes of nonlinear MRFs and handles several other approaches to multiresolution MRF modeling. In [27], Jeng considers the periodic subsampling of MRFs defined on infinite lattices. MRFs 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 a family of potential functions [20]. 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 [27]. 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 [27].
The paper is organized as follows. In Section 2, we present some fundamentals on graph theory and random fields. Markov Random Fields are defined on finite nondirected graphs.

[1] Alternate approaches, involving various kind of coarsening operators on a *single-resolution model*, have also been developed recently [8, 14, 25, 26, 28, 33, 38, 39, 40]. These approaches are not subject to a loss of locality and hence will not be considered in this paper

1 Introduction

Markov Random Fields (MRF) models have been successfully introduced in many fundamental issues of image analysis and computer vision such as image restoration [5, 11, 19], edge detection [18], image segmentation [13, 18], multisource image analysis [24], computed tomography [17], surface reconstruction [13, 36], stereovision [2], motion analysis [9, 10, 23, 40] or scene interpretation [37]. 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 criterion, they lead to the minimization of a global energy function which may exhibit local minima [19]. This minimization is generally performed using deterministic [5] or stochastic [19] relaxation algorithms.

Stochastic algorithms may be drastically time consuming while deterministic schemes often get "stuck" in local minima of the energy function. Besides, it is known that multigrid methods can improve significantly the convergence rate of iterative relaxation schemes [42]. 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 process. 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, 6, 31, 35]. Yet, in multigrid implementations of statistical models such as MRFs, 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 [21] has described a consistent mathematical framework for multiscale Markov modeling, based on the *renormalization group* approach. Unfortunately, this standard technique of statistical physics does not lead to tractable computational schemes (apart from particular models and scale transformations [21]) because of the loss of locality of the model at the coarse scales. In [32],

Restriction d'un champ markovien sur un graphe et analyse d'image multirésolution

Résumé : La définition de modèles statistiques multirésolutions se pose comme un problème théorique (et pratique) difficile. Différentes approches ont été proposées récemment pour associer champs markoviens et techniques d'analyse multirésolution de l'image : groupe de renormalisation, sous-échantillonnage de champs aléatoires, définition de modèles markoviens sur des arbres ou sur des graphes pyramidaux, etc. Pour une simulation efficace de ces modèles ou leur utilisation dans le contexte de l'estimation bayesienne, il est primordial que des propriétés markoviennes (locales) soient préservées pour le champ multirésolution. Nous montrons dans cet article que ce problème peut être étudié en considérant la restriction d'un champ markovien (défini sur un graphe fini non-orienté) à une partie de son support. La distribution ainsi que les propriétés markoviennes du champ restreint sont établies. Le caractère local de la structure de voisinage associée à ce nouveau modèle markovien est étudié de façon détaillée. Ces résultats généraux sont appliqués à différentes approches récentes de la modélisation markovienne multirésolution en imagerie.

Mots-clé : Graphes non-orientés, champs markoviens, distributions de Gibbs, modèles statistiques multirésolutions, restriction d'un champ aléatoire, potentiels canoniques et irréductibles, voisinage minimal, sous-échantillonnage.

Restriction of a Markov Random Field on a Graph and Multiresolution Image Analysis

Patrick Pérez, Fabrice Heitz

Programme 4 — Robotique, image et vision
Projet TEMIS

Rapport de recherche n° 2170 — Mars 1994 — 45 pages

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, MRFs 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 a finite arbitrary 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 a MRF. Sufficient conditions for the new neighborhood structure to be "minimal" are derived. Several consequences and applications of these general results to various "multiresolution" MRF-based modeling approaches are presented.

Key-words: Nondirected graphs, Markov Random Fields, Gibbs Random Fields, statistical multiresolution image analysis, minimal neighborhood, irreducible potential function, restriction of a MRF, subsampling

(Résumé : tsvp)

Submitted to IEEE Trans. Information Theory

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
Téléphone : (33) 99 84 71 00 – Télécopie : (33) 99 38 38 32

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE

Restriction of a Markov Random Field on a Graph and Multiresolution Image Analysis

Patrick Pérez, Fabrice Heitz

N° 2170

Mars 1994

PROGRAMME 4

Robotique,

image

et vision

ISSN 0249-6399