Analytic approximation of Gibbs potentials to model stochastic textures

Georgy Gimel'farb[*]

Abstract

Gibbs random fields with multiple pairwise pixel interactionshave good potentialities in modeling natural image texturesbecause allow for learning both the structure and strengths ofpixel interactions from a given training sample. The learningscheme is based on the maximum likelihood estimate (MLE) ofGibbs potentials that specify the interaction strenghts. Thisscheme is amplified here by deducing an explicit, to scalingfactors, analytic form of the potentials from an additionalfeasible top rank principle. It suggests that the training samplemay possess a feasible top rank in its total Gibbs energy withinthe parent population. Under this condition, only the scaling factorshave to be learnt using their MLE. As a result, the introducedconditional MLE of the potentials extends capabilities of the Gibbsimage models under consideration.

[*]   The University of Auckland, Tamaki Campus, Computing and Information
     Technology Research, Computer Vision Unit, Auckland, New Zealand

Analytic approximation of Gibbs potentials to model stochastic textures

Georgy Gimel'farb

Computing and Information Technology Research
Computer Science Department, Tamaki Campus
The University of Auckland, Auckland
New Zealand
E-mail: georgy@cs.auckland.ac.nz

Abstract

Gibbs random fields with multiple pairwise pixel interactions have good potentialities in modeling natural image textures because allow for learning both the structure and strengths of pixel interactions from a given training sample. The learning scheme is based on the maximum likelihood estimate (MLE) of Gibbs potentials that specify the interaction strenghts. This scheme is amplified here by deducing an explicit, to scaling factors, analytic form of the potentials from an additional feasible top rank principle. It suggests that the training sample may possess a feasible top rank in its total Gibbs energy within the parent population. Under this condition, only the scaling factors have to be learnt using their MLE. As a result, the introduced conditional MLE of the potentials extends capabilities of the Gibbs image models under consideration.

Keywords: *Gibbs random field, multiple pairwise interactions, potential, MLE, image texture*

1: Models with multiple pairwise pixel interactions

Gibbs random field models describe images in terms of an explicit geometric structure and quantitative strengths of pixel interactions (see, for instance, [2, 4, 9]). The *pixel interaction* depends on how probabilities of the signals (that is, gray levels) relate to those of the independent random field (IRF): the more probable a particular spatial signal configuration, the stronger the interaction between these signals.

We restrict the consideration only to translation invariant *pairwise pixel interactions* and, therefore, to a specific class of spatially homogeneous image textures called *stochastic textures* in [6, 7]. The stochastic texture has pixels and pixel pairs as (primitive) elements and its interaction structure is given by a first-order family of the pixels themselves and by several second-order families of the translation invariant pixel pairs. In the Markov/Gibbs models these pairs are *cliques*, or complete subgraphs of the neighbourhood graph [2]. For brevity, the terms "clique" and "clique family" are preserved in [7] also for the non-Markov Gibbs models taking account of admissible gray range changes.

The interaction strength for a given clique family is specified by a (Gibbs) *potential* being a scalar function of the signals in the clique: the stronger the interaction, the higher the potential value.

Traditional Gibbs models, in particular, auto-binomial and auto-normal ones [2, 4, 9], possess mostly pre-defined interaction structures and potentials. Thus they have rather restricted potentialities in modeling different

natural and artificial image textures. Models with multiple pairwise pixel interactions introduced in [6, 7] are adapted more easily to various textures because both the interaction structure and the potentials are learnt from a given training sample using the MLE of the potentials. The learning scheme involves three steps: (*i*) an analytic first approximation of the potentials, (*ii*) a search for most characteristic clique families, and (*iii*) a stochastic approximation refinement of the potentials for the chosen families.

This learning scheme is amplified here by using an additional *feasible top rank principle*. It suggests that the training sample may possess a feasible top rank in its total Gibbs energy within the parent population. This condition results in an explicit, to scaling factors (one per a clique family), analytic form of the potentials so that only the MLE of the factors has to be approximated. Such a learning scheme with the conditional MLE of the potentials extends capabilities of the models with multiple pairwise pixel interactions.

1.1: Assumptions and notation

Let $\mathbf{R} = \{(m, n) : m = 0, \ldots, M - 1; n = 0, \ldots, N - 1\}$ be a 2D finite rectangular lattice with $M \cdot N$ pixels $i = (m, n)$ supporting digital grayscale images $g : \mathbf{R} \to \mathbf{Q}$. Here, $\mathbf{Q} = \{0, \ldots, q_{\max}\}$ is a set of gray values.

Let $\mathbf{A}$ be an index set for the pairwise clique families and $\mathbf{C}_a = \{(i, j) : i, j \in \mathbf{R}; i - j = const_a\}$ denote a particular family with $a \in \mathbf{A}$. This family has a specific pixel arrangement in the pairs specified by a fixed 2D interpixel shift $const_a \equiv (\mu_a, \nu_a)$: if $i = (m, n)$ and $j = (m', n')$ then $m' - m = \mu_a$ and $n' - n = \nu_a$.

Let the following assumptions hold for the stochastic textures.

- Grayscale images $g \in \{g_1, g_2, \ldots\}$ that differ only by gray ranges should have the same Gibbs probability as their reference image g^{rf} obtained by normalisation. This latter maps initial gray ranges $[\min_{i \in \mathbf{R}} g(i), \max_{i \in \mathbf{R}} g(i)]$ onto the maximum range $[0, q_{\max}]$.

- The interaction strength over the first-order clique family $\mathbf{R}$ is given by a Gibbs

potential function $V : \mathbf{R} \to \mathcal{R}$ of a gray level (GL) in the pixel. Here, $\mathcal{R}$ denotes the set of real numbers.

- The interaction strength over each second-order family $\mathbf{C}_a$ is a function $V_a : \mathbf{D} \to \mathcal{R}$ of a gray level difference (GLD) $d = g(i) - g(j) \in \mathbf{D} = \{-q_{\max}, \ldots, 0, \ldots, q_{\max}\}$ in the clique $(i, j) \in \mathbf{C}_a$.

1.2: Non-Markov Gibbs model

This model proposed in [7] embeds the gray range normalisation $g \to g^{\mathrm{rf}}$ directly into the Gibbs potentials. For brevity, indices "rf" will be omitted below. The Gibbs probability distribution (GPD) for the model is as follows:

$$\Pr(g|\mathbf{V}) = \frac{1}{Z_{\mathbf{V}}} \cdot \exp\left(E(g|\mathbf{V})\right). \tag{1}$$

where $E(g|\mathbf{V}) = e(g|V) + \sum_{a \in \mathbf{A}} e_a(g|V_a)$ is a *total Gibbs energy* of pixel interactions in the image g under the potentials $\mathbf{V} = (V, V_a : a \in \mathbf{A})$ for all the clique families, $e(g|V) = \sum_{i \in \mathbf{R}} V(g(i))$ denotes a *partial energy* of the pixelwise interactions, $e_a(g|V_a) = \sum_{(i,j) \in \mathbf{C}_a} V_a(g(i) - g(j))$ is a partial energy of pairwise pixel interactions for the clique family $\mathbf{C}_a$, and $Z_{\mathbf{V}} = \sum_{g \in \mathbf{G}} \exp(E(g|\mathbf{V}))$ denotes a scaling factor. Here, $\mathbf{G}$ is the parent population of all the grayscale images supported by the lattice $\mathbf{R}$.

1.3: Sufficient statistics

The partial interaction energy can be represented as a dot product of the *centered* potential vector and the vector of relative sample GL or GLD frequencies, that is, of normalised GL or GLD histogram (H) collected over the reference image g (see [6, 7]):

$$e(g|V) = |\mathbf{R}| \cdot \sum_{q \in \mathbf{Q}} V(q) \cdot F(q|g);$$

$$e_a(g|V_a) = |\mathbf{R}| \cdot \rho_a \cdot \sum_{d \in \mathbf{D}} V_a(d) \cdot F_a(d|g). \tag{2}$$

Here, $F(q|g) = \frac{1}{|\mathbf{R}|} \cdot \sum_{i \in \mathbf{R}} \delta(q - g(i))$ and $F_a(d|g) = \frac{1}{|\mathbf{C}_a|} \cdot \sum_{(i,j) \in \mathbf{C}_a} \delta(d - (g(i) - g(j)))$ are

the normalised GLH and GLDH, respectively, $|\dots|$ denotes the set cardinality, $\rho_a = \frac{|\mathbf{C}_a|}{|\mathbf{R}|}$, and $\delta()$ is the Kronecker function.

The potential centering

$$\sum_{q \in \mathbf{Q}} V(q) = 0; \quad \forall a \in \mathbf{A} \sum_{d \in \mathbf{D}} V_a(d) = 0 \qquad (3)$$

can be deduced from the unique representation of the Gibbs probability distribution (GPD) by relative Hamiltonian [5]. This centering implies the similar centering of the histograms in (2). Below, both the potentials and histograms are assumed to be centered.

The resulting exponential family representation of the GPD [1] shows that the centered GLH and GLDHs for all the clique families form *sufficient statistics* for the model. It can be proven that conditions imposed in [1] to ensure strict log-concavity of the GPD (or unimodality of the likelihood function) hold for the model.

1.4: Learning the parameters

Both the characteristic clique families and the potentials are learnt from a given training sample g° using analytic and stochastic approximation of the MLE of the potentials. This learning scheme introduced in [6, 7] is as follows:

(*i*) Analytic first approximation:

$$\forall q \in \mathbf{Q} \quad V_{[0]}(q) = \lambda_{[0]} \cdot F(q|g^\circ);$$
$$\forall a \in \mathbf{A}; \; \forall d \in \mathbf{D} \; V_{a,[0]}(d) = \qquad (4)$$
$$\lambda_{[0]} \cdot (F_a(d|g^\circ) - M_{\mathrm{dif}}(d))$$

where $M_{\mathrm{dif}}(d)$ denotes the centered marginal probabilities of the GLD for the IRF (it is easily shown that $M_{\mathrm{dif}}(d) = \frac{1+q_{\max}-|d|}{(1+q_{\max})^2} - \frac{1}{1+2 \cdot q_{\max}}$) and the factor $\lambda_{[0]}$ is computed from the same centered normalised GLH and GLDHs, too (see [6] for more details).

(*ii*) Search for most characteristic interaction structure using approximate partial Gibbs energies (2) with the potentials (4) for comparing a big many possible clique families.

(*iii*) Refinement of the potential estimates for the chosen families by stochastic approximation techniques (see [10]).

This approach gives good results in simulating natural textures that may be considered as the stochastic ones [6, 7]. Figures 1 and 2 present samples 128×128 of natural stochastic textures from [3] simulated by the proposed approach. But, such a learning involves rather big number of parameters to be refined by stochastic approximation, namely, all the centered potential values $\mathbf{V}$, that is, in total $q_{\max} \cdot (2 \cdot |\mathbf{A}| + 1)$ scalar values.

T-D4   G-D4   T-D5   G-D5   T-D9   G-D9

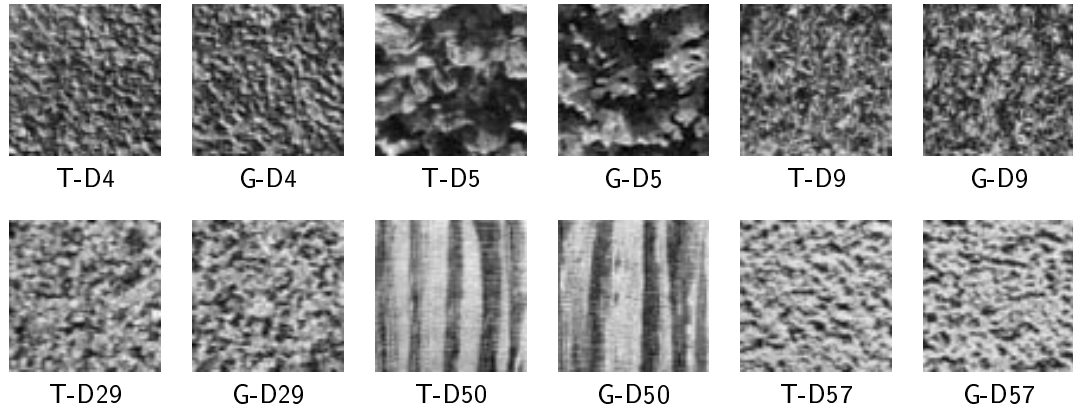T-D29   G-D29   T-D50   G-D50   T-D57   G-D57

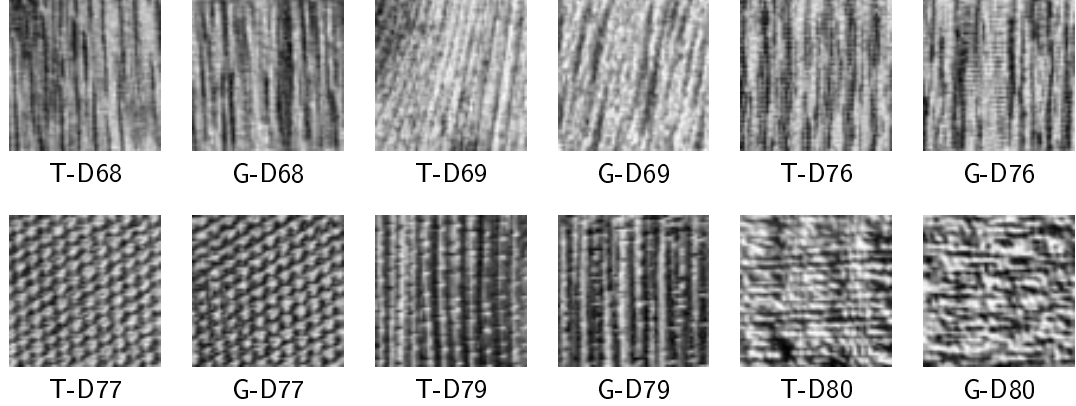Figure 1: Training (T) and simulated (G) samples of natural textures.

Figure 2: Training (T) and simulated (G) samples of natural textures.

2: Explicit form of the potentials

Here, we consider a somewhat different learning approach based on an *explicit*, to scaling factors, form of the potentials. In this case the number of parameters to be computed is reduced to only the number $|\mathbf{A}|$ of the clique families. This explicit form of the potentials is deduced under a specific ranking of the training sample g° within $\mathbf{G}$. The desired potential estimates are obtained using the MLE of the scaling factors or, what is the same, the *conditional* MLE of the potentials (CMLE) provided that the training sample g° may occupy a feasible for it top rank within $\mathbf{G}$ in the total Gibbs energy.

2.1: Ranking in Gibbs energies

Let us rank the images $g \in \mathbf{G}$ in ascending order of the partial Gibbs energy (in particular, $e(g|V)$) for the pixelwise family $\mathbf{R}$ or $e_a(g|V_a$ for the pairwise family $\mathbf{C}_a$; $a \in \mathbf{A}$). It is easily seen from (1) and (2) that this ranking is *invariant* to potential (and energy) normalisation that reduces the corresponding potential vector $\mathbf{V} = (V(q) : q \in \mathbf{Q})$ or $\mathbf{V}_a = (V_a(d) : d \in \mathbf{D})$ to the unit vector $\mathbf{v} = \frac{\mathbf{V}}{|\mathbf{V}|}$ or $\mathbf{v}_a = \frac{\mathbf{V}_a}{|\mathbf{V}_a|}$, respectively.

Let $\mathbf{F}(g^\circ) = (F(q|g^\circ) : q \in \mathbf{Q})$ and $\mathbf{F}_a(g^\circ) = (F_a(d|g^\circ) : q \in \mathbf{Q})$ denote, respectively, the *centered* vectors of marginal GL sample frequencies and of marginal GLD

sample frequencies for the clique family $\mathbf{C}_a$. Then it is readily shown that unit vectors $\mathbf{v}^\circ = \frac{\mathbf{F}(g^\circ)}{|\mathbf{F}(g^\circ)|}$ and $\mathbf{v}_a^\circ = \frac{\mathbf{F}_a(g^\circ)}{|\mathbf{F}_a(g^\circ)|}$ maximise the normalised partial energies $e(g^\circ|\mathbf{v})$ and $e_a(g^\circ|\mathbf{v}_a)$, respectively.

A particular potential vector obtained by arbitrary scaling of such a unit vector ranks the training sample g° in a corresponding partial energy to the same top place that may be feasible among the samples $g \in \mathbf{G}$ as compared to any other potential vector. Let this *feasible top rank principle* be applied to the ranking in the partial energy for each the clique family of the model (1). Then the potentials $\mathbf{V} = (V, V_a : a \in \mathbf{A})$ ranking the training sample to a feasible top place within $\mathbf{G}$ in the *total* Gibbs energy have to possess the following explicit, to scaling factors, form:

$$\mathbf{V}^\circ(\mathbf{\Lambda}) = (\lambda \cdot \mathbf{F}(g^\circ), \lambda_a \cdot \mathbf{F}_a(g^\circ) : a \in \mathbf{A}). \quad (5)$$

Here, $\mathbf{\Lambda} = (\lambda, \lambda_a : a \in \mathbf{A})$ is a vector of arbitrary positive scaling factors.

2.2: CMLE of the potentials

Therefore, the CMLE of the Gibbs potentials $\mathbf{V}^\star \equiv \mathbf{V}^\circ(\mathbf{\Lambda}^\star)$ for the image model (1) is as follows:

$$\mathbf{V}^\star = (\lambda^\star \cdot \mathbf{F}(g^\circ), \lambda_a^\star \cdot \mathbf{F}_a(g^\circ) : a \in \mathbf{A}) \quad (6)$$

where the desired scaling factors are computed by maximising the likelihood function $L(\mathbf{\Lambda}|g^\circ) = \ln \Pr(g^\circ|\mathbf{V}^\circ(\mathbf{\Lambda}))$:

$$\mathbf{\Lambda}^\star = \arg\max_{\mathbf{\Lambda}} L(\mathbf{\Lambda}|g^\circ). \quad (7)$$

Generally, this CMLE may differ from the unconditional MLE of the potentials. But, for the GPD (1) both the estimates are supposedly fairy close if not equivalent. This conjecture that needs further theoretical investigations has some supporting considerations: in particular, (*i*) the CMLE (6) and the analytic first approximation of the unconditional MLE (4) have very similar forms, (*ii*) the refined potentials are usually close to their first approximations, and (*iii*) the samples possess fixed ranks not only under the uniform scaling of the potentials but also, by symmetry, under the like scaling of the centered sample histograms.

2.3:   Learning the factors

The desired factors (7) are learnt in a similar way as the potentials themselves in [6, 7]: first, by analytic first approximation and searching for a characteristic interaction structure and then by refining the factors for chosen clique families using stochastic approximation.

Analytic first approximation   of the factors is obtained by a truncated Taylor's series expansion of the likelihood function $L(\mathbf{\Lambda}|g^\circ)$ about the zero point $\mathbf{\Lambda} = \mathbf{0}$. This technique that is quite similar to those proposed in [6, 7] results in the following approximation:

$$\begin{aligned}
\lambda_{[0]} &=& \alpha_{[0]} \cdot \varepsilon_{[0]}; \\
\forall a \in \mathbf{A} \ \lambda_{a,[0]} &=& \alpha_{[0]} \cdot \varepsilon_{a,[0]}, \quad (8)
\end{aligned}$$

where

$$\varepsilon_{[0]} = \sum_{q \in \mathbf{Q}} F^2(q|g^\circ)$$

and

$$\varepsilon_{a,[0]} = \rho_a \cdot \sum_{d \in \mathbf{D}} \left(F_a(d|g^\circ) - M_{\mathrm{dif}}(d) \right) \cdot F_a(d|g^\circ)$$

are relative pixelwise and pairwise total Gibbs energies about the zero point, respectively. The scale factor $\alpha_{[0]}$ is computed from these energies as:

$$\alpha_{[0]} = \frac{\varepsilon_{[0]}^2 + \sum\limits_{a \in \mathbf{A}} \varepsilon_{a,[0]}^2}{\varepsilon_{[0]}^2 \cdot U_{[0]} + \sum\limits_{a \in \mathbf{A}} \varepsilon_{a,[0]}^2 \cdot U_{a,[0]}} \quad (9)$$

where

$$U_{[0]} = \sum_{q \in \mathbf{Q}} \sigma_{\mathrm{irf}} \cdot F^2(q|g^\circ);$$

$$U_{a,[0]} = \rho_a \cdot \sum_{d \in \mathbf{D}} \sigma_{\mathrm{dif}} \cdot F_a^2(d|g^\circ), \quad (10)$$

and σ_{irf} and σ_{dif} denote variances for the IRF of the marginal frequencies of the GL and GLD, respectively.

Search for the interaction structure   exploits in this case the weighted relative partial energies of pairwise pixel interactions: $\mathbf{e}_{[0]} = \left\{ \omega_{a,[0]} \cdot \varepsilon_{a,[0]} : \ a \in \mathbf{A}_{\mathrm{srch}} \right\}$ where the weight $\omega_{a,[0]} = \rho_a \cdot \sum\limits_{d \in \mathbf{D}} F_a^2(d|g^\circ)$ and $\mathbf{A}_{\mathrm{srch}}$ is an index set for the clique families within a large search window. The window is specified by a given range of the intra-clique pixel shifts $|\mu_a| \leq \mu_{\max}, \ |\nu_a| \leq \nu_{\max}$ to be exhausted during the search. The energies $\mathbf{e}_{[0]}$ over the search window form an *interaction map* that allows to choose most characteristic clique families using an appropriate thresholding technique (see [6, 8] for more details).

Stochastic approximation refinement of the factors exploits the similar partial energies that depend on a proximity between the marginal GL and GLD frequencies for each clique family in the training image sample g° and a sample generated by pixelwise stochastic relaxation using the current factors. At each step t of the stochastic approximation, the current factors $\mathbf{\Lambda}_t$ are updated as follows:

$$\lambda_{[t+1]} = \lambda_{[t]} + \alpha_{[t]} \cdot \varepsilon_{[t]}(g_{[t]});$$

$$\forall a \in \mathbf{A} \ \lambda_{a,[t+1]} = \lambda_{a,[t]} + \alpha_{[t]} \cdot \varepsilon_{a,[t]}(g_{[t]}). \ (11)$$

Here, $g_{[t]}$ is the sample generated at this step, $\alpha_{[t]}$ denotes the current scaling factor decreasing from the starting value $\alpha_{[0]} in (9)$ as $\frac{c_0 + 1}{c_1 + c_2 \cdot t}$ (see [10] for theoretical and empirical choices of the control values c_0, c_1, and c_2), and $\varepsilon_{[t]}(g_{[t]})$ and $\varepsilon_{a,[t]}(g_{[t]})$ are the current differential partial energies:

$$\varepsilon_{[t]}(g_{[t]}) = \sum_{q \in \mathbf{Q}} \Delta(q|g_{[t]}) \cdot F(q|g^\circ);$$

$$\varepsilon_{a,[t]}(g_{[t]}) = \rho_a \cdot \sum_{d \in \mathbf{D}} \Delta_a(d|g_{[t]}) \cdot F_a(d|g^\circ). \ (12)$$

where $\Delta(q|g_{[t]}) = F(q|g^\circ) - F(q|g_{[t]})$ and $\Delta_a(d|g_{[t]}) = F_a(d|g^\circ) - F_a(d|g_{[t]})$.

3: Concluding remarks

The derived analytic form of the Gibbs potentials, besides its theoretical value, allows not only to reduce the number of the model parameters to be computed during the learning stage but also to extend the number of signal values $|\mathbf{Q}| = q_{\max} + 1$ taken into account in the image model (1).

The learning schemes of [6, 7] based on the unconditional MLE of the potentials presume that the marginal relative sample GL anf GLD frequencies, obtained by normalising the sample GLH and GLDHs, give valid statistical estimates of the corresponding marginal probabilities under the given model (1). But, the bigger the number $|\mathbf{Q}|$, the larger the size of the training sample to obtain such the estimates. To simplify a choice of the training samples, for the texture simulation experiments in [6, 7] this number has been restricted to $|\mathbf{Q}| = 16$).

In practice the training samples have usually relatively small sizes so that do not contain all possible GLs and GLDs if there is sufficiently big number of the gray values, say, $|\mathbf{Q}| = 256$. Of course, in this case one or another known robust approximation to the unobserved marginals from the obtained incomplete sample histograms may be implemented for getting valid potential estimates based on the proposed CMLE. Such way out is possible because errors in the differential partial energies (12) due to approximation errors may influence, during the refinement process (11), only the scaling factors but not the overall form of the potentials.

But, when refining the unconditional MLE of the potentials, each the potential value $V_{[t]}(q)$ for $q \in \mathbf{Q}$ and $V_{a,[t]}(d)$ for $d \in \mathbf{D}$ and $a \in \mathbf{A}$ is updated independently using the distinctions between corresponding marginal frequencies for the training sample and for the current generated one. Therefore, as opposed to the CMLE, errors in the approximations of the marginals at each stochastic approximation step t (that is, for each currently generated image $g_{[t]}$) may result in unpredictable errors in the final potential estimates.

4: References

[1] O. Barndorrf-Nielsen, Information and Exponential Families in Statistical Theory. Wiley, 1978.

[2] J. E. Besag, "Spatial interaction and the statistical analysis of lattice systems", J. Royal Statistical Soc., vol. B36, pp. 192-236, 1974.

[3] P. Brodatz, Textures: a Photographic Album for Artists and Designers. New York: Dover Publications, 1966.

[4] R. Chellappa and A. Jain (Eds), Markov Random Fields: Theory and Application. Academic Press, 1993.

[5] R. L. Dobrushin and S. A. Pigorov, "Theory of random fields", in Proc. 1975 IEEE-USSR Joint Workshop on Information Theory, December 15-19, 1975, Moscow, USSR. IEEE, pp. 39-49, 1976.

[6] G. L. Gimel'farb, "Texture modeling by multiple pairwise pixel interactions", IEEE Trans. Pattern Anal. Machine Intell., vol. 18, pp. 1110-1114, 1996.

[7] G. L. Gimel'farb, "Non-Markov Gibbs texture model with multiple pairwise pixel interactions", in Proc. 13th IAPR Int. Conf. on Pattern Recognition, August 25-29, 1996, Vienna, Austria. Wien: TUWien, vol.II, pp. 591-595, 1996.

[8] G. L. Gimel'farb and A. K. Jain, "On retrieving textured images from an image database", Pattern Recognition, vol. 29, pp. 1461-1483, 1996.

[9] G. Winkler, Image Analysis, Random Fields, and Dynamic Monte Carlo Methods. Springer, 1995.

[10] L. Younes, "Estimation and annealing for Gibbsian fields", Annales de l'Institut Henri Poincare, vol. 24, pp. 269-294, 1988.