Texture Modelling by Multiple Pairwise Pixel Interactions

G. L. Gimel'farb

Abstract— **A Markov random field model with a Gibbs probability distribution (GPD) is proposed for describing particular classes of grayscale images which can be called spatially uniform stochastic textures. The model takes into account only multiple short- and long-range pairwise interactions between the gray levels in the pixels. An effective learning scheme is introduced to recover a structure and strength of the interactions using maximal likelihood estimates of the potentials in the GPD as desired parameters. The scheme is based on an analytic initial approximation of the estimates and their subsequent refinement by a stochastic approximation. Experiments in modelling natural textures show the utility of the proposed model.**

Keywords— **Texture, Markov/Gibbs random field, pairwise interaction, maximum likelihood estimate.**

I. Introduction

Generative models of Markov random fields (MRF) with Gibbs probability distributions (GPD) are promising for image modelling because of the facility to generate image samples by stochastic relaxation [10, 20, 34, 35, 39], to estimate parameters of the model from a training sample [21, 23, 25, 29, 32, 43], and to check the validity of the model directly by comparing the generated images with the learning samples visually and quantitatively. These models, in a more general form, were first introduced in statistical physics (see, for instance, [30]). The GPDs on finite lattices which are the most interesting for image modelling were investigated by Averintsev, Besag, Clifford, Dobrushin, and Hammersley (see [1, 3, 7, 16, 26] and the comprehensive surveys [14, 17]). The equivalence between the MRF and GPD under a positivity condition, that is, the existence of the GPD for any MRF with non-zero probabilities of all signal configurations in the lattice, and resulting factorization of the joint and conditional GPDs are due to the well-known theorems of Hammersley and Clifford [3, 26] and Averintsev [1] (see, also, [34]).

The factorization relates global features of the MRF to local interactions by specifying their geometric structure and probabilistic strengths. Such interaction in the images means that some local signal configurations are more probable than others in the lattice. The interaction structure is described by a neighborhood graph linking each pair of interacting pixels, called neighbors. The GPD is represented by a product of strictly positive factors, each defined on a corresponding complete subgraph, or clique of the pixels [3, 14, 10, 20, 28]. Each factor describes quantitatively the interaction strength in the corresponding clique. Usually, the factor is represented in an exponential form and the exponent is called a potential; the higher the potential, the more probable the signal configuration in the clique.

Extensive investigations of the Markov/Gibbs image models were initiated by Hassner and Sklansky [28], Cross and Jain [10], and Lebedev et al. [34]. These pioneering results were amplified, but directed to more practical problems of image restoration (noise removal) and segmentation by Derin [11-13], Geman and Geman [13, 20] and have been explored in many other publications (see, for example, [4, 5, 8, 15, 19, 22-25, 31, 32, 35-38, 40]). Comprehensive surveys of these works can be found in [18, 33, 41]. It should be noted that the traditional models borrowed from physics, such as the autobinomial or Gauss-Markov models, may not to be the best ones for describing particular texture types. For instance, the autobinomial model takes into account only nearest neighbors in the lattice and defines the interaction strength by a product of neighboring signal values. However, in the context of image modelling it is hard to understand why just the nearest neighbors hold advantages over the more distant ones or why the pair of neighboring gray levels (q, q) have the same interaction as $(1, q^2)$ and twice the interaction of $(\frac{q}{2}, q)$.

The Gauss-Markov model exploits many more interactions because of rather big (and usually square) windows used as cliques. The potential is proportional to a squared error of a weighted linear prediction of the gray level in the central pixel of the clique from the signals in other pixels. The model parameters (regression weights over the window $a \times a$) can be estimated, for instance, by minimizing a non-linear likelihood function of a^2 unknown weights [3]. Kashyap and Chellappa studied these models in-depth (see, for instance, [32, 33]) and introduced some parameter estimation schemes. Stochastic relaxation was used to generate the desired image textures. Cohen and Patel [9] modelled power spectral densities of the Gauss-Markov fields. These densities have an analytic representation in terms of the regression weights and are generated simpler than the images themselves due to a mutual independence of the spectral components. Then the images are formed by a discrete cosine transform of the generated densities. Both approaches show much better texture modelling than the autobinomial model, especially if the windows allow long-range interactions (say, $a = 13$ or more), but the larger the window, the more the computational difficulties of parameter estimation. Therefore, it is hard to involve arbitrarily long interactions. Moreover, to find the most characteristic interactions, one needs to exhaust and compare all the possible subwindows [32].

We can expect that the higher the number of signal interactions that can be determined from the image itself (that is, estimated as parameters of the GPD), the more effective will be the image modelling. Here, we present the Markov/Gibbs model of spatially uniform images that take account of multiple short- and long-range pairwise pixel interactions. It belongs to the exponential family of distributions [2] which are strictly log-concave (that is, strongly unimodal), under rather weak conditions [2, 29] that hold for the proposed model. This allows us, at least in principle, to estimate from a given learning sample both the interaction structure and strengths. The model exploits gray level difference histograms (GLDH) as sufficient statistics and, therefore, supports well-known applications of the GLDHs to describe textures [27].

Images with similar GLDHs are considered as belonging to the same type. We will call the images which can be generated successfully by this model *(spatially) uniform stochastic textures* to discriminate them from other more complex image types. This model is not parsimonious, relative to the Gauss-Markov class, in terms of the number of the parameters, but is much simpler as regards the potential for estimation and, in the main, as regards the choice of the most characteristic interactions.

The paper is organized as follows. In Section 2 we present initial assumptions about the gray-scale textures and introduce the spatially uniform Markov/Gibbs image model with multiple pairwise pixel interactions. Section 3 presents the learning scheme for estimating the model parameters. Experiments in

G. L. Gimel'farb is with the V.M.Glushkov Institute of Cybernetics (National Academy of Sciences of Ukraine), prospect Akademika Glushkova 40, Kiev-22, 252022, Ukraine). E-mail: gimel@image.kiev.ua

This research was done in Computer and Automation Research Institute (Hungarian Academy of Sciences) under Grant OTKA T014520.

generating natural textures and conclusions are given in Section 4.

II. Markov/Gibbs Model with Multiple Pairwise Interactions

A. Notation and assumptions

Let $\mathbf{X} = (X(i) : i \in \mathbf{R})$ be a 2-D lattice MRF with samples $\mathbf{x} = (x(i) : i \in \mathbf{R}; q = x(i) \in \mathbf{Q})$, where $\mathbf{R}$ is the finite 2-D lattice $\mathbf{R} = ((m,n) : m = 0, ..., M-1;\ n = 0, ..., N-1)$ of size $|\mathbf{R}| = M \cdot N$, and $\mathbf{Q} = \{0, 1, \ldots, q_{\max}\}$ is a finite set of signal values (gray levels, q) at lattice sites, or pixels, (m,n). The sample space for $\mathbf{x}$ is denoted by $\mathbf{SS}$.

We restrict our consideration only to spatially uniform isotropic and anisotropic textures whose global visual appearance depends mainly on the pairwise interactions between the signals. In this case the interaction structure is translation-invariant and represented by a first-order clique family containing the pixels themselves and a particular set of second-order clique families $\mathbf{K}_a = \{(i,j) \in \mathbf{R}^2 : i - j = (\mu_a, \nu_a)\}$ where $a \in \mathbf{A}$. Here, $\mathbf{A}$ is a set of indices and (μ_a, ν_a) denotes a relative displacement of the pixels in a clique. Any second-order family contains all pixel pairs that have the same relative pixel arrangement and differ by their absolute positions in the lattice.

Let $V(x(i))$ and $V_a(x(i), x(j))$ be potentials, or non-constant functions of the variables $x(i)$ such that their supports (i) or (i,j) are the cliques. The potential describes quantitatively the interaction strength in the clique. It is natural to assume a *shift-invariant strength* of the interactions of gray-scale textures because usually their visual appearance does not depend on the mean intensity, but is due to a pattern of intensity changes. In general, such invariance leads to a non-Markov Gibbs model. To stay within the Markov/Gibbs class, we specify potential values which are independent of these shifts. For the first-order family such a potential is zero-valued. For any second-order family the potential has the same value on all signal pairs that differ by a shift b: $\{V_a(q,u) \equiv V_a(q+b, u+b) : (q,u), (q+b, u+b) \in \mathbf{Q}^2\}$. Therefore, the second-order potentials depend only on the gray level differences: $\{V_a(q,u) \equiv V_a(d = q - u) : d \in \mathbf{D}\}$ where $\mathbf{D} = \{-q_{\max}, \ldots, 0, \ldots, q_{\max}\}$.

B. Representation of Spatially Uniform MRF Using Gray Level Difference Histograms

Under the above assumptions, the Markov/Gibbs image model takes the following form (see, also, [24, 25]):

$$\Pr(\mathbf{x}|\mathbf{V}) = Z_{\mathbf{V}}^{-1} \cdot \exp\left(\sum_{a\in\mathbf{A}} \sum_{(i,j)\in\mathbf{K}_a} V_a(x(i) - x(j))\right), \qquad (1)$$

and can be represented as the exponential family distribution:

$$\Pr(\mathbf{x}|\mathbf{V}) = Z_{\mathbf{V}}^{-1} \cdot \exp\left(\sum_{a\in\mathbf{A}} \sum_{d\in\mathbf{D}} V_a(d) \cdot H_a(d|\mathbf{x})\right). \qquad (2)$$

Here, $Z_{\mathbf{V}}$ is a normalizing factor (or the partition function [30]), $H_a(d|\mathbf{x}) = \sum\limits_{(i,j)\in\mathbf{K}_a} \delta(d - (x(i) - x(j)))$ is a component of the gray level difference histogram (GLDH) for the clique family $\mathbf{K}_a$, and $\delta(\ldots)$ denotes the Kronecker function.

Obvious relations for the GLDHs, namely that $\sum\limits_{d\in\mathbf{D}} H_a(d|\mathbf{x}) = |\mathbf{K}_a| \quad \forall \mathbf{x} \in \mathbf{SS};\ a \in \mathbf{A}$, yield a unique representation of the GPD in (1) and (2) by centering the potentials; it adds the following constraints to (1):

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

This can be derived also from the concept of a relative Hamiltonian in [17]. The GPD in (2) is also invariant to the corresponding centering of the GLDHs.

Let $\mathbf{V} = \{V_a(d) : d \in \mathbf{D};\ a \in \mathbf{A}\}$ and $\mathbf{H}_{\mathrm{cn}}(\mathbf{x}) = \{H_{\mathrm{cn},a}(d|\mathbf{x}) : d \in \mathbf{D};\ a \in \mathbf{A}\}$ be the $(G + |\mathbf{A}|)$-component vectors of the centered potentials and centered GLDHs, respectively, and let $\bullet$ denote the dot product. Here, $G = 2 \cdot q_{\max} \cdot |\mathbf{A}|$. As a result of the centering, the vectors $\mathbf{V}$ and $\mathbf{H}_{\mathrm{cn}}(\mathbf{x})$ are in the same G-dimensional vector subspace $\mathcal{R}^G$. The GPD $\Pr(\mathbf{x}|\mathbf{V}) = Z_{\mathbf{V}}^{-1} \cdot \exp(\mathbf{V} \bullet \mathbf{H}_{\mathrm{cn}}(\mathbf{x}))$ in (2) is the regular exponential family distribution with the minimal canonical parameter $\mathbf{V}$ and minimal sufficient statistic $\mathbf{H}_{\mathrm{cn}}(\mathbf{x})$ if and only if (a) the vector $\mathbf{V}$ is in $\mathcal{R}^G$ and (b) the components of the vector $\mathbf{H}_{\mathrm{cn}}(\mathbf{x})$ are affinely independent (cf. Corollary 8.1 and p.116 in [2]; see, also, [21, 29]); that is, if the condition $\mathbf{V} \bullet \mathbf{H}_{\mathrm{cn}}(\mathbf{x}) = \mathrm{const}, \quad \forall \mathbf{x} \in \mathbf{SS}$, implies that all potential values are equal to zero ($\mathbf{V} = \mathbf{0}$ and $\mathrm{const} = 0$). This holds for the model in (2) and (3) as can be shown explicitly, by identifying particular image pairs $\mathbf{x}$, $\mathbf{x}'$ such that the differences $\mathbf{H}_{\mathrm{cn}}(\mathbf{x}) - \mathbf{H}_{\mathrm{cn}}(\mathbf{x}')$ form an orthogonal G-dimensional basis in $\mathcal{R}^G$.

Similar linear representations of the GPD exponent, but with rather different terms, can be found in many publications, in particular, in [20, 21]. A similar form of the GPD has been introduced in the convolved form of an aura matrix which does not discriminate between cliques from the different families but with the same signal configuration values [19, 40]. The GPD in (2) describes the model more precisely by revealing explicitly the structure and strength of the pairwise pixel interactions.

In this model we assume that all pairwise clique families except the given set $\mathbf{A}$ have zero-valued potentials. The assumption of zero-valued potentials in all clique families ($\mathbf{V} = \mathbf{0}$) leads to a model of the independent random field (IRF) with equiprobable signal values in the pixels. In other words, the model in (2) differs from the IRF only in the given set $\mathbf{A}$ of clique families.

The log-likelihood function $LF(\mathbf{V}|\mathbf{x}^\circ) \equiv \frac{\ln \Pr(\mathbf{x}^\circ|\mathbf{V})}{|\mathbf{R}|}$ of the potential vector $\mathbf{V}$ for a given learning sample $\mathbf{x}^\circ$ is strictly concave and has a unique maximum if and only if the vector $\mathbf{H}_{\mathrm{cn}}(\mathbf{x}^\circ)$ is not on a boundary of a domain of the vectors $\mathbf{H}_{\mathrm{cn}}(\mathbf{x})$ in $\mathcal{R}^G$ (see [2], [29] for details). The interior of the $(G + |\mathbf{A}|)$-dimensional hyperrectangular domain containing the GLDH vectors is given by the following conditions:

$$0 < F_a(d|\mathbf{x}^\circ) < 1 \quad \forall d \in \mathbf{D};\ a \in \mathbf{A}, \qquad (4)$$

where $F_a(d|\mathbf{x}^\circ) = \frac{H_a(d|\mathbf{x}^\circ)}{|\mathbf{K}_a|}$ denotes the marginal sample relative frequency of the gray level difference for a given learning sample $\mathbf{x}^\circ$ obtained by normalizing the GLDH. In $\mathcal{R}^G$ this hyperrectangular domain is restricted to a hypertetrahedral domain of the vectors $\mathbf{H}_{\mathrm{cn}}(\mathbf{x})$. However, such a mapping preserves the interior or boundary positions of these vectors, as can be easily shown by using barycentric co-ordinates of these positions in the latter domain. In other words, the MLE of the potentials for the model in (1) and (2) exists if the conditions of (4) hold. It is easy to show that the conditional maximum of the likelihood function $LS(\mathbf{V}|\mathbf{x}^\circ)$, given the constraints of (3), is obtained in the point where the first-order partial derivatives of this function with respect to the components of the parameter $\mathbf{V}$ are equal to zero.

From the given sample $\mathbf{x}^\circ$, the set $\{\mathbf{K}_a : a \in \mathbf{A}\} \leftrightarrow \{(\mu_a, \nu_a) : a \in \mathbf{A}\}$ of the clique families and the vector $\mathbf{V}$ of the potential values for each clique family need to be estimated.

III. Maximum Likelihood Parameter Estimation

A. Log-Likelihood Function

The first-order partial derivatives of the log-likelihood function $LF(\mathbf{V}|\mathbf{x}^\circ)$, at the point $\mathbf{V}$, or components of the gradient, have the form (see, also, [2]):

$$\frac{\partial LF(\mathbf{V}|\mathbf{x}^\circ)}{\partial V_a(d)} = \rho_a \cdot (F_a(d|\mathbf{x}^\circ) - \mathcal{E}\{F_a(d|\mathbf{x})|\mathbf{V}\}), \tag{5}$$

where $\rho_a = \frac{|\mathbf{K}_a|}{|\mathbf{R}|}$ and $\mathcal{E}\{\ldots|\mathbf{V}\}$ is the expectation of the marginal sample relative frequency, or the marginal probability of the gray level difference $MP_a(d|\mathbf{V})$ if the images $\mathbf{x} \in \mathbf{SS}$ have the GPD of (1) or (2) with parameter $\mathbf{V}$. Therefore, the gradient lies also in $\mathcal{R}^G$.

The matrix $\mathbf{MS}$ of second-order partial derivatives of the log-likelihood function with respect to the components of the parameter $\mathbf{V}$ is proportional to the covariance matrix for the GLDHs taken with the negative sign [2].

For the model in (2) and (3), the following system of equations holds at the unique maximum point:

$$\forall a \in \mathbf{A};\ d \in \mathbf{D} \quad F_a(d|\mathbf{x}^\circ) = MP_a(d|\mathbf{V}). \tag{6}$$

The conditions of Eq. (4) allow us to use the normalized GLDHs of the learning sample to check the validity of the model and suitability of the learning sample itself in representing the given texture. The model contains $|\mathbf{A}|$ clique families and $G = 2 \cdot q_{\max} \cdot |\mathbf{A}|$ potential values to be estimated from the given learning sample with $|\mathbf{R}| = M \cdot N$ initial signals. To exploit the asymptotic consistency of the MLEs, it is essential that the number of these parameters should be much less than the number of signals: $G << |\mathbf{R}|$.

If the conditions of (4) do not hold, we need to tailor the textures to the model by reducing the number of potential values and/or the number of clique families. In particular, the potential values can be equalized for different, but visually similar, signal configurations, or the number of gray levels can be reduced by quantization of the gray range. Sometimes, the images can be demagnified to reduce the number of clique families.

B. Initial Estimates of the Potential Values

Let $\mathbf{V} = \mathbf{0}$ be the zero point, that is, $\mathbf{0} = \{V_a(d) = 0 : a \in \mathbf{A};\ d \in \mathbf{D}\}$, in the space of the parameters $\mathbf{V}$. This point corresponds to the singular case of the MRF, or equivalently to the IRF with independent and equiprobable signals over the lattice.

The very weak interdependencies of the pairwise signal differences in the IRF allow us to approximate the matrix $\mathbf{MS}$ by a diagonal matrix with components which are proportional to the variances of the GLDHs taken with the negative sign: $\mathrm{Var}(H_a(d|\mathbf{x})|\mathbf{0}) = |\mathbf{K}_a| \cdot \phi(d)$ where $\phi(d) = MP(d) \cdot (1 - MP(d))$ denotes the variance of the estimator of the marginal probability $MP(d) = \frac{1 + q_{\max} - \mathrm{abs}(d)}{(1 + q_{\max})^2}$ of the gray level difference d for the IRF and $\mathrm{abs}(d)$ denotes the absolute value of d.

The truncated Taylor's series expansion of the log-likelihood function about $\mathbf{0}$ is, using the gradient formula in (5),

$$\begin{aligned} &LF(\mathbf{V}|\mathbf{x}^\circ) \cong LF(\mathbf{0}|\mathbf{x}^\circ) + \\ &\lambda \cdot \sum_{a \in \mathbf{A}} \rho_a^2 \cdot \sum_{d \in \mathbf{D}} (F_a(d|\mathbf{x}^\circ) - MP(d))^2 - \\ &\tfrac{\lambda^2}{2} \cdot \sum_{a \in \mathbf{A}} \rho_a^3 \cdot \sum_{d \in \mathbf{D}} \phi(d) \cdot (F_a(d|\mathbf{x}^\circ) - MP(d))^2 \end{aligned} \tag{7}$$

where the scaling factor λ defines the step from the zero point $\mathbf{0}$ along the gradient direction and $\rho_a = \frac{|\mathbf{K}_a|}{|\mathbf{R}|}$; $a \in \mathbf{A}$.

Thus, initial estimates of the potential values can be computed analytically by maximizing the function in (7) with respect to the factor λ: $\forall d \in \mathbf{D};\ a \in \mathbf{A}$

$$V_{a,[0]}(d) = \lambda_{[0]} \cdot \rho_a \cdot (F_a(d|\mathbf{x}^\circ) - MP(d)), \tag{8}$$

where

$$\lambda_{[0]} = \frac{\sum\limits_{a \in \mathbf{A}} \rho_a^2 \cdot \sum\limits_{d \in \mathbf{D}} (F_a(d|\mathbf{x}^\circ) - MP(d))^2}{\sum\limits_{a \in \mathbf{A}} \rho_a^3 \cdot \sum\limits_{d \in \mathbf{D}} \phi(d) \cdot (F_a(d|\mathbf{x}^\circ) - MP(d))^2} \tag{9}$$

The Markov/Gibbs model under consideration is non-uniform at the borders of the lattice and the cardinalities $|\mathbf{K}_a$ of the second-order clique families are slightly less than the lattice cardinality $|\mathbf{R}|$ (that is, the factors $\rho_a < 1$). However, as the lattice expands ($|\mathbf{R}| \to \infty$), these cardinalities approach the lattice cardinality ($\rho_a \to 1\ \forall a \in \mathbf{A}$). Thus, the initial estimates in (8) are almost independent of the lattice size.

C. Heuristic Search for Clique Families Describing the Interaction Structure

The initial potential estimates allow us to search for clique families which best describe the local pairwise pixel interactions in a given learning sample. The estimates in (8) show that the strength of the pairwise interaction depends on the departure from the IRF in any given clique family. The families with a weak interaction have potential estimates which are close to the zero point and thus can be excluded from the model because of their small influence on the MRF.

Thus, the following heuristic solution of this problem can be proposed. We exhaust all possible pairwise cliques in a given search window of possible relative shifts between the pixels in the clique $|\mu_a \le \mu_{\max}$; $|\nu_a| \le \nu_{\max}$ and reconstruct the particular structure of the pairwise pixel interaction in the given image sample by comparing the distances between the initial estimates of the potentials in (8) and the zero point in the parameter space. In other words, we compute the distance between the normalized GLDHs, or the sample marginal relative frequencies, $\{F_a(d|\mathbf{x}^\circ) :\ d \in \mathbf{D}\}$ for all clique families a in the window, and the triangle distribution $\{MP(d) :\ d \in \mathbf{D}\}$ for the IRF. The families representing the most characteristic spatially uniform pairwise interactions can be found by a proper thresholding of these distances. We can therefore consider that all such models have the same interaction structure (corresponding to the given search window) and differ only by the potentials for these cliques: non-zero values for the characteristic cliques and zero values for all the other families. This feature simplifies a comparison of different uniform stochastic textures.

Experiments with several simulated and natural textures allow us to propose the following heuristic search strategy, giving rather good results in describing their interaction structures:

(i) Compute GLDHs for all clique families in the search window;

(ii) Compute the distances between the normalized GLDHs and triangle distribution;

(iii) Compute the mean distance MD and standard deviation σ of the latter distances;

(iv) Compute the threshold $Thr = MD + k \cdot \sigma$ (in our experiments, $k = 3$ or 4);

(v) Choose the clique families whose distances exceed this threshold to represent the characteristic structure of the signal interactions.

Of course, this strategy has been determined empirically and should be improved by more rigorous theoretical investigations.

D. *Refining Estimates of Potentials by Stochastic Approximation*

After computing the initial estimates in (8) of the potential functions, we refine them by solving the system of equations (6) using a multi-step stochastic approximation (StA) closely related to a technique proposed in [43]. The StA can be implemented because of the possibility of generating samples under the given GPD by means of pixelwise stochastic relaxation. The Metropolis relaxation algorithm [39] is used in our experiments. Let us refer to as the (macro)step the successive stochastic relaxation pass round all the $|\mathbf{R}|$ pixels in the lattice, without repetition, under equiprobable random choice of each next pixel [23, 39]. Each step $t = 1, 2, \ldots$ of the StA includes the following operations:

(*i*) Generate an image $\mathbf{x}_{[t]}$ from a previous one $\mathbf{x}_{[t-1]}$ by stochastic relaxation under the GPD $\Pr(\mathbf{x}|\mathbf{V}_{[t-1]})$ with the current parameter estimate (the image $\mathbf{x}_{[0]}$ is an IRF sample);

(*ii*) Update the parameter estimates using the normalized GLDHs for the generated image and a contracted StA-step along the current approximation of the gradient in (5):

$$\forall d \in \mathbf{D};\ \ a \in \mathbf{A} \qquad (10)$$
$$V_{a,[t]}(d) = V_{a,[t-1]} + \lambda_{[t]} \cdot \left(F_a(d|\mathbf{x}^\circ) - F_a(d|\mathbf{x}_{[t]})\right);$$

(*iii*) Check the quality of these estimates by computing a distance between the goal GLDHs for the learning sample $\mathbf{x}^\circ$ and current ones for the generated image $\mathbf{x}_{[t]}$;

(*iv*) Terminate the process if the current quality reaches the given threshold (the case of good model matching) or if the number of steps exceeds the given limit due to poor quality of the estimates (the case of poor model matching).

There is rather wide scope in the choice of possible schedules for contracting successive steps during the StA process. The schedule ensuring almost sure convergence of the process in (10) to the desired MLE is deduced in [43]. However, in practice, as mentioned in [43], this theoretical contraction is too slow to achieve convergence in a reasonable time. In the experiments below we used the following two heuristic schedules:

$$\lambda_{[t]} = \frac{c_0 + 1}{c_1 + c_2 \cdot t} \cdot \lambda_{[0]} \ \text{ or } \ \lambda_{[t]} = \frac{c_0 + 1}{c_1 + c_2 \cdot t} \cdot \lambda_{[t]}^{\max}, \qquad (11)$$

where c_0, c_1, c_2 are user-specified control parameters and $\lambda_{[t]}^{\max}$ is an estimate of the factor λ obtained by maximizing the truncated Taylor's series for the likelihood function LF in (7) in the neighborhood of the point $\mathbf{V}_{[t-1]}$. In the second case we use the rough approximation of the current covariance matrix by the diagonal one that resembles the initial diagonal matrix at step $t = 1$ except for replacing the unknown marginals with the current sample relative frequencies: $\mathrm{Var}(H_a(d|\mathbf{x})|\mathbf{V}) = |\mathbf{K}_a| \cdot F_a(d|\mathbf{x}_{[t]}) \cdot \left(1 - F_a(d|\mathbf{x}_{[t]})\right)$. The first schedule in (11) is similar to the one found empirically and used in [43], but in our experiments the final results were better, in general, for the second variant with $c_0 = c_1 = c$ and $c_2 = 1$. The convergence speed depended noticeably on the control parameter: small values $c \in [0 \ldots 5]$ gave slower but basically monotone convergence; larger values $c \in [10 \ldots 30]$ usually resulted in faster but rather oscillating convergence (of a hyper-relaxation type). In spite of the overall good results of the above StA-learning these heuristic schedules need more theoretical and experimental studies.

This StA-learning scheme can be regarded also as a self-adjustable algorithm of image generation (in a broad sense, a type of simulated annealing [20, 43]), because each StA-step changes the potential estimates so as to obtain a better approximation of the GLDHs of the given learning sample with the histograms for the current image $\mathbf{x}_{[t]}$. The image formed finally by StA-learning resembles the desired texture more closely than the images generated with the learnt potentials in the GPD and ordinary stochastic relaxation.

IV. Experimental Results and Conclusions

Our experiments used the following natural textures, including some from Brodatz [6]: "Fur of Baboon" (50×50 pixels), "Bark of Tree" (a digitized fragment 120×60 of the photoimage D12 from [6]), "Blood Cells" (60×30), "Pressed Cork" (a fragment 100×100 of the photoimage D4 from [6]), "Wood Grain" (65×65). Figure 1 shows the learning samples, the learnt clique families, and the images generated by the Metropolis relaxation [39] under the GPD of (1) with the learnt clique families and potentials. Images with 16 gray levels (that is, $q_{\max} = 15$) were generated and used for estimation. In these examples 2, 8, 2, 6, and 7 clique families were chosen, respectively, and 75, 255, 75, 195, and 255 potential values were estimated for them using the proposed learning scheme. The interaction maps in Figure 1(b) represent each clique family by two square boxes with coordinates μ_a, ν_a and $-\mu_a, -\nu_a$ with respect to the origin $(0,0)$ indicated by a black mark. The darker the box, the more characteristic the family. In particular, the chosen clique families have intra-clique pixel displacements $(1,0)$ and $(2,0)$ for the texture "Fur of Baboon" and $(3,1)$ and $(-3,2)$ for the texture "Blood Cells".

These results allow us to conclude that the Markov/Gibbs image model with multiple pairwise pixel interactions is promising for simulating spatially uniform textures as well as in discriminating between them. Of course, uniform textures which can be modelled efficiently by the proposed model form lower-level type (micro)textured images. Nevertheless, a reasonable number of natural textures belongs to this type, as well as many artificial ones, and this justifies the use of the model in image modelling and processing.

Acknowledgments

The author would like to acknowledge the invaluable help of Anil K. Jain in improving the content and the clarity of the paper and to thank Mikhail Schlesinger and Alexey Zalesny for fruitful discussions and comments. The author is very grateful to anonymous reviewers and the Associate Editor for the rigorous analyses and multiple corrections taken into account in the paper.

References

[1] M. B. Averintsev, "Description of Markov Random Fields Using Gibbs Conditional Probabilities", *Probability Theory and Its Applications*, vol. **XVII**, no. 1, pp.21-35, 1972. In Russian.

[2] O. Barndorff-Nielsen, *Information and Exponential Families in Statistical Theory*, Wiley, 1978.

[3] J. E. Besag, "Spatial Interaction and the Statistical Analysis of Lattice Systems", *J. Royal Stat. Soc. London*, vol. **B36**, pp.192-236, 1974.

[4] J. Besag, "On the Statistical Analysis of Dirty Pictures", *J. Royal Stat. Soc. London*, vol. **B48**, pp.259-302, 1986.

[5] J. Besag, "Toward Bayesian Image Analysis", *Advances in Applied Statistics: Statistics and Images: 1*. A Supplement to *J. of Applied Statistics*, vol. **20**, no. 5/6, pp.107-119, 1993.

[6] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, New York: Dover Publications, 1966.

[7] D. Brook, "On the Distinction between the Conditional Probability and the Joint Probability Approaches in the Specification of Nearest-Neighbor Systems", *Biometrika*, vol. **51**, no. 3-4, pp.481-483, 1964.

[8] F. S. Cohen and D. B. Cooper, "Simple Parallel Hierarchical and Relaxation Algorithms for Segmenting Noncausal Markovian Fields", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. **2**, pp.195-219, 1987.

[9] F. S. Cohen and M. A. S. Patel, "Modeling and Synthesis of Images of 3D Textured Surfaces", *CVGIP: Graphical Models and Image Processing*, vol. **53**, no. 6, pp.501-510, 1991.

[10] G. R. Cross and A. K. Jain, "Markov Random Field Texture Models", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **5**, no. 1, pp.25-39, 1983.

[11] H. Derin and W. S. Cole, "Segmentation of Textured Images Using Gibbs Random Fields", *Computer Vision, Graphics, and Image Processing*, vol. **35**, no. 1, pp.72-98, 1986.

[12] H. Derin and H. Elliot, "Modelling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **9**, no. 1, pp.39-55, 1987.

[13] H. Derin, H. Elliot, R. Cristi, and D. Geman, "Bayes Smoothing Algorithm for Segmentation of Images Modelled by Markov Random Fields", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **6**, no. 6, pp.707-720, 1984.

[14] H. Derin and P. A. Kelly, "Discrete-Index Markov-Type Random Processes", *Proc. IEEE*, vol. **77**, no. 10, pp.1485-1510, 1989.

[15] P. A. Devijver, "Hidden Markov Mesh Random Field Models in Image Analysis", *Advances in Applied Statistics: Statistics and Images: 1.* A Supplement to *J. of Applied Statistics*, vol. **20**, no. 5/6, pp.187-227, 1993.

[16] R. L. Dobrushin, "Gibbs Random Fields for the Lattice Systems with Pairwise Interaction", *Functional Analysis and Its Applications*, vol. **2**, no. 4, pp.31-43, 1968. In Russian.

[17] R. L. Dobrushin and S. A. Pigorov, "Theory of Random Fields", *Proc. of the 1975 IEEE-USSR Joint Workshop on Information Theory*, December 15-19, 1995, Moscow, USSR. New York: IEEE, pp.39-49, 1976.

[18] R. C. Dubes and A. K. Jain, "Random Field Models in Image Analysis", *J. of Applied Statistics*, vol. **16**, no. 2, pp.131-164, 1989.

[19] I. M. Elfadel and R. W. Picard, "Gibbs Random Fields, Cooccurrences, and Texture Modeling", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **16**, no. 1, pp.24-37, 1994.

[20] 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, pp.721-741, 1984.

[21] B. Gidas, "Parameter Estimation for Gibbs Distributions from Fully Observed Data", *Markov Random Fields: Theory and Applications* (R. Chellappa and A. Jain, Eds), Boston: Academic Press, pp.471-483, 1993.

[22] G. L. Gimel'farb, "Gibbs Random Fields and Compound Bayesian Decisions at the Lower Level of Digital Image Processing", *Pattern Recogn. Image Anal.: Adv. Math. Theory and Applications in the USSR*, vol. **1**, no. 1, pp.39-49, 1991.

[23] G. L. Gimel'farb and A. V. Zalesny, "Low-Level Bayesian Segmentation of Piecewise-Homogeneous Noisy and Textured Images", *Int. J. Imaging Systems Technology*, vol. **3**, no. 3, pp.227-243, 1991.

[24] G. L. Gimel'farb and A. V. Zalesny, "Probabilistic Models of Digital Region Maps Based on Markov Random Fields with Short- and Long-Range Interaction", *Pattern Recogn. Lett.*, vol. **14**, no. 5, pp.789- 797, 1993.

[25] G. L. Gimel'farb and A. V. Zalesny, "Markov Random Fields with Short- and Long-Range Interaction for Modelling Gray-Scale Textured Images", *Proc. 5th Int. Conf. on Computer Anal. of Images and Patterns (CAIP'93)*, Budapest, Hungary, Sept. 13-15, 1993, *Lecture Notes in Computer Science 719*, Berlin: Springer-Verlag, pp.275-282, 1993.

[26] J. M. Hammersley and D. C. Handscomb, *Monte-Carlo Methods*, London: Methuen, 1964.

[27] R. M. Haralick, "Statistical and Structural Approaches to Textures", *Proc. IEEE*, vol. **67**, no. 8, pp.786-804, 1979.

[28] M. Hassner and J. Sklansky, "The Use of Markov Random Fields as Models of Textures", *Computer Graphics and Image Processing*, vol. **12**, no. 4, pp.357-370, 1980.

[29] M. Jacobsen, "Existence and Unicity of MLE in Discrete Exponential Family Distributions", *Scand. J. Statist.*, vol. **16**, pp.335-349, 1989.

[30] A. Isihara, *Statistical Physics*, New York: Academic Press, 1971.

[31] N. Karssmeijer, "A Relaxation Method for Image Segmentation Using a Spatially Dependent Stochastic Model", *Pattern Recognition Letters*, vol. **11**, no. 1, pp.13-23, 1990.

[32] R. L. Kashyap and R. Chellappa, "Estimation and Choice of Neighbors in Spatial Interaction Models", *IEEE Trans. Information Theory*, vol. **29**, no. 1, pp.60-72, 1983.

[33] R. L. Kashyap, "Image Models", *Handbook on Pattern Recognition and Image Processing* (T.Y.Young and K.-S.Fu, Eds), Orlando: Academic Press, pp.247-279, 1986.

[34] D. S. Lebedev, A. A. Bezruk, and V. M. Novikov, *Markov Probabilistic Model of Image and Picture*, Preprint (Inst. of Information Transmiss. Probl., Acad. Sci. USSR), Moscow: VINITI, 1983. In Russian.

[35] D. S. Lebedev, *Statistical Theory of Video Data Processing*, Moscow: Moscow Inst. of Physics and Technology (MFTI), 1988. In Russian.

[36] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic Solution of Ill-Posed Problems in Computational Vision", *J. Amer. Stat. Assoc.*, vol. **82**, no. 397, pp.76-89, 1987.

[37] J. L. Marroquin, "Deterministic Bayesian Estimation of Markov Random Fields with Applications to Computational Vision", *Proc. First Int. Conf. on Computer Vision*, London, England, June 8-11, 1987. Washington, D. C.: IEEE Computer Soc. Press, pp.597-601, 1987.

[38] J. Marroquin, "Deterministic Interaction Particle Models for Image Processing and Computer Graphics", *CVGIP: Graph. Models and Image Process.*, vol. **55**, no. 5, pp.408-417, 1993.

[39] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equations of State Calculations by Fast Computing Machines", *J. Chemical Physics*, vol. **21**, pp.1087-1091, 1953.

[40] R. W. Picard and I. M. Elfadel, "Structure of Aura and Cooccurrence Matrices for the Gibbs Texture Model", *J. Math. Imaging and Vision*, vol. **2**, no. 1, pp.5-25, 1992.

[41] M. Tuceryan and A. K. Jain, "Texture Analysis", *Handbook of Pattern Recognition and Computer Vision* (C.H.Chen, L.F.Pau, and P.S.P.Weng, Eds). Singapore: World Publishing, pp.235-276, 1993.

[42] M. Wasan, *Stochastic Approximation*, Cambridge University Press, 1969.

[43] L. Younes, "Estimation and Annealing for Gibbsian Fields", *Annales de l'Institut Henri Poincare*, vol. **24**, no.2, pp.269-294, 1988.