

# Modeling Image Textures by Gibbs Random Fields

Georgy Gimel'farb

*Centre for Image Technology and Robotics  
Department of Computer Science  
Tamaki Campus, The University of Auckland  
Private Bag 92019, Auckland 1, New Zealand  
g.gimelfarb@auckland.ac.nz*

---

## Abstract

Drawbacks of the traditional scenario of image modeling by Gibbs random fields with multiple pairwise pixel interactions are outlined, and a more reasonable alternative scenario based on Controllable Simulated Annealing is described. The latter scenario uses an analytic and stochastic approximation of Gibbs potentials to minimize a distance between the selected gray level co-occurrence or difference histograms for a given training sample and the simulated images.

*Key words:* Image modeling, Gibbs random field, stochastic texture, stochastic approximation, controllable simulated annealing.

---

## 1 Introduction

We address some practical aspects of probabilistic image modeling by Gibbs random fields [2,6,8,9]. The traditional scenario of modeling by Gibbs fields with multiple pairwise pixel interactions involves the following two stages [3,4]:

- (1) first analytic and subsequent stochastic approximation of the maximum likelihood estimate (MLE) of model parameters from a given training sample and
- (2) generation of images by pixel-wise stochastic relaxation using the Gibbs probability distribution (GPD) with the estimated parameters.

---

<sup>1</sup> This research was supported in part by the University of Auckland Research Committee grant XXXXX/9343/3414083.

The stochastic approximation produces a non-stationary Markov chain of images generated by pixel-wise stochastic relaxation under a specific schedule of changing the model parameters [9]. The schedule ensures that the parameters converge almost surely to the desired MLE in the limit, when the chain reaches an equilibrium, that is, the stepwise parameter changes tend to zero. Image generation with the fixed model parameters results in a stationary Markov chain of images having the estimated GPD in an equilibrium.

Unfortunately, there exist no statistical tests for an equilibrium of a Markov chain, and the theoretically justified schedules of stochastic approximation are impractically long because they are designed for the worst case [9]. Also, the stationary Markov chains obtained by stochastic relaxation have to be, at least, twice longer than the size of the parent population of all the images to ensure that the generated images have the desired GPD [5]. But only under such conditions one may expect that the obtained results possess the probabilistic features of a training sample.

Therefore in the traditional scenario we have to wait for an equilibrium, first, of a Markov chain of the images produced by stochastic approximation of the model parameters and, second, of a Markov chain generated by stochastic relaxation using the estimated parameters. In both cases there are no formal rules for testing whether the equilibrium has been actually reached, and the theoretically derived sizes of the chains are too big to be reached in practice.

In this paper we argue that the true model parameters do not constitute the goal of image modeling as the parameter estimation ranks next to the generation of images similar to a given “typical” training sample. Because the traditional scenario tends to do more than it is required for image modeling, it can be replaced by an alternative and more practicable scenario.

## 2 Basic notions and notation

Let  $\mathbf{R} = [(m, n) : m = 0, \dots, M - 1; n = 0, \dots, N - 1]$  denote a finite arithmetic 2D lattice supporting gray-scale images. In the subsequent text, we use a shorthand notation  $i$  for the pixels  $(m, n) \in \mathbf{R}$ . Let  $\mathbf{Q} = \{0, \dots, q_{\max}\}$  denote a finite set of gray levels.

Let  $\mathbf{g} = [g_i : i \in \mathbf{R}; g_i \in \mathbf{Q}]$  be a digital gray-scale image. We denote  $\tilde{\mathbf{g}}$  the reference image that represents all the images  $\mathbf{g}$  differing by only the gray range  $[g_{\min}, g_{\max}]$  where  $g_{\min} = \min_{i \in \mathbf{R}} g_i$  and  $g_{\max} = \max_{i \in \mathbf{R}} g_i$ . The reference image is obtained by mapping the initial gray range  $[g_{\min}, g_{\max}]$  of the image  $\mathbf{g}$  onto the whole range  $[0, q_{\max}]$  as follows:  $\tilde{g}_i = \frac{q_{\max}}{g_{\max} - g_{\min}} (g_i - g_{\min})$ .

Let  $\mathbf{C} = [\mathbf{C}_a : a \in \mathbf{A}]$  be a particular subset of families of translation invariant pixel pairs  $\mathbf{C}_a = \{(i, j) : i, j \in \mathbf{R}; i - j = \text{const}_a\}$  in the lattice. The subset  $\mathbf{C}$  describes the characteristic structure of pairwise pixel interactions.

Let  $\mathbf{V} = [\mathbf{V}_a : a \in \mathbf{A}]$  denote the Gibbs potential which specifies the quantitative strengths  $\mathbf{V}_a = \{V_a(q, q') : q, q' \in \mathbf{Q}\}$  of pairwise pixel interactions in every family  $\mathbf{C}_a$ . For every pixel pair  $(i, j) \in \mathbf{C}_a$ , the interaction strength depends on a particular signal co-occurrence  $[g_i = q, g_j = q']$  in the image  $\mathbf{g}$ .

The Gibbs image model with multiple pairwise pixel interactions to describe the translation invariant stochastic textures under admissible changes of the gray ranges is given by the following GPD [3,4]:

$$\Pr(\mathbf{g}|\mathbf{C}, \mathbf{V}) = \frac{1}{Z} \exp \sum_{a \in \mathbf{A}} \sum_{(i,j) \in \mathbf{C}_a} V_a(\tilde{g}_i, \tilde{g}_j) \quad (1)$$

where  $Z$  is the normalizing factor, or the partition function [2].

Let  $\mathbf{F}(\mathbf{g})$  denote the vector of the relative sample frequency distributions of signal co-occurrences in the chosen families  $\mathbf{C}$ . The total Gibbs energy per pixel in the exponent of Eq. (1), being rewritten as the dot product of the vector  $\mathbf{V}$  and the vector  $\mathbf{F}(\mathbf{g})$ , facilitates the parameter estimation from a given training sample [3]:

$$\frac{1}{|\mathbf{R}|} \sum_{a \in \mathbf{A}} \sum_{(i,j) \in \mathbf{C}_a} V_a(g_i, g_j) \equiv \mathbf{V} \bullet \mathbf{F}(\mathbf{g}) \quad (2)$$

where  $|\mathbf{Z}|$  denotes the cardinality of a set  $\mathbf{Z}$ .

### 3 Traditional modeling scenario

The GPD of Eq. (1) results in the following traditional modeling scenario:

- (1) Recover the characteristic interaction structure  $\hat{\mathbf{C}}$  and compute the first analytic approximation of the MLE of the Gibbs potential from a given training sample  $\mathbf{g}^\circ$ :

$$\mathbf{V}_0 = \lambda_0 (\mathbf{F}(\tilde{\mathbf{g}}^\circ) - \mathbf{M}_{\text{irf}}). \quad (3)$$

Here,  $\mathbf{M}_{\text{irf}}$  is the vector of the marginal probability distributions of signal co-occurrences for the independent random field (IRF). The factor  $\lambda_0$  is computed as a particular function of the vectors  $\mathbf{F}(\tilde{\mathbf{g}}^\circ)$  and  $\mathbf{M}_{\text{irf}}$  and statistical features of the IRF [3].

- (2) Refine the MLE of the potential by stochastic approximation that changes the potential at every step,  $t = 1, 2, \dots, T$ , as follows:

$$\mathbf{V}_t = \mathbf{V}_{t-1} + \lambda_t (\mathbf{F}(\tilde{\mathbf{g}}^\circ) - \mathbf{F}(\tilde{\mathbf{g}}_t)) \quad (4)$$

where every current image  $\mathbf{g}_t$  is generated by pixel-wise stochastic relaxation using the previous GPD  $\Pr(\mathbf{g}|\hat{\mathbf{C}}, \mathbf{V}_{t-1})$ . The refinement should be terminated when the potential  $\mathbf{V}_T$  is sufficiently close to the MLE  $\hat{\mathbf{V}}$ .

- (3) Generate by pixel-wise stochastic relaxation a stationary Markov chain of images having in an equilibrium the GPD  $\Pr(\mathbf{g}|\hat{\mathbf{C}}, \hat{\mathbf{V}})$ .

As shown in [5], to ensure the latter equilibrium, the Markov chain should be of length  $|\mathbf{Q}|^{2|\mathbf{R}|}$  which cannot be reached in practice even with the extremely small images such as  $6 \times 6 - 10 \times 10$  pixels. It follows that the traditional scenario seems to actually exploit only the similarity between most of the GPDs and the  $\delta$ -function, in particular, a very low variance of the total Gibbs energy in Eq. (2) around its expected value.

Under the GPD  $\Pr(\mathbf{g}|\hat{\mathbf{C}}, \hat{\mathbf{V}})$  over the parent population of all the images, the expected total energy per pixel is equal to the energy of a training sample:

$$\mathcal{E}\{\hat{\mathbf{V}} \bullet \mathbf{F}(\mathbf{g})\} = \hat{\mathbf{V}} \bullet \mathcal{E}\{\mathbf{F}(\mathbf{g})\} = \hat{\mathbf{V}} \bullet \mathbf{F}(\mathbf{g}^\circ). \quad (5)$$

Therefore, if the MLE of the model parameters is available, the generated images have mostly the energies, and consequently the relative sample frequency distributions  $\mathbf{F}(\mathbf{g})$  of signal co-occurrences, in the close proximity to the energy and similar distributions for the training sample. This permits us to replace the traditional modeling goal of getting a particular mathematical expectation  $\mathcal{E}\{\mathbf{F}(\mathbf{g})\} \approx \mathbf{F}(\mathbf{g}^\circ)$  of the relative frequency distributions by an alternative goal of a direct proximity between the generated and training distributions:  $\mathbf{F}(\mathbf{g}_T) \approx \mathbf{F}(\mathbf{g}^\circ)$ .

#### 4 Alternative modeling scenario

The alternative scenario combines the initial MLE-based parameter estimation of Eqs. (3) and (4) and the subsequent image generation into a single stochastic approximation process called *Controllable Simulated Annealing* (CSA) in [3]. The goal of the CSA is to generate samples having probabilities “around” the probability of a given training sample by minimizing, in the average, a distance between the relative sample frequency distributions of signal co-occurrences for the generated images and the training sample. These distributions are obtained by normalizing the gray level co-occurrence histograms (GLCHs) which are sufficient statistics for the Gibbs model of Eq. (1). Therefore the

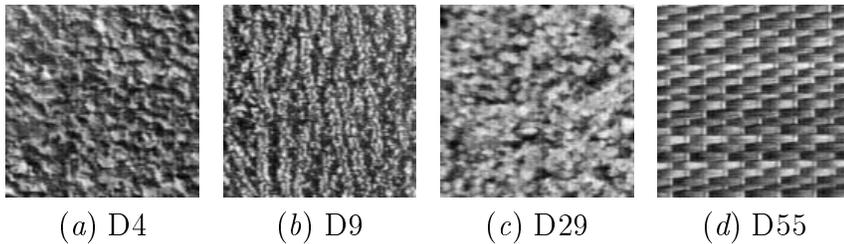


Fig. 1. Training samples of the natural image textures.

samples with the GLCHs distributed around the GLCHs of the training sample are of the main practical interest.

The alternative scenario results in both faster and better modeling (here, “better” means the smaller distance between the distributions, but the visual similarity is usually much higher, too). It is worth noting that the alternative scenario has an explicit and natural stopping rule which is absent in the traditional scenario. This stopping rule, based on a distance between the generated and training GLCHs, for instance, the chi-square distance, allows to guide and terminate, if necessary, the modeling process.

## 5 Experimental results

In these experiments we use the simplified Gibbs model of Eq. (1) with the potentials depending on gray level differences in the pixel pairs:  $V_a(g_i, g_j) = V_a(g_i - g_j)$ . This model has the gray level difference histograms (GLDHs) as sufficient statistics. In all the experiments both scenarios start the generation of the images from a sample of the IRF.

The four training reference samples of digitized image textures from [1] are shown in Figure 1. Figures 2 and 3 present results of generating samples of the texture D29 by the traditional and alternative scenarios. Plots under the samples show how the average chi-square distance, that is, the distance per family  $\mathbf{C}_a \in \hat{\mathbf{C}}$ , is changing at each step of generation. The average distance is computed between the GLDHs for the generated image and training sample.

The Gibbs potentials to generate the samples in Figures 2,*a* and 2,*b* are estimated by 300 steps of stochastic approximation, and the samples themselves are obtained by 300 steps of pixel-wise stochastic relaxation with the estimated model parameters. The sample in Figure 2,*c* is obtained in a similar way, but by using 2000 stochastic approximation and 300 stochastic relaxation steps.

In these three examples the convergence to the training GLDHs is too slow, and the final average chi-square distance is still about 7,000 – 10,000 after 300 steps of generation. Moreover, the first two samples even diverge from

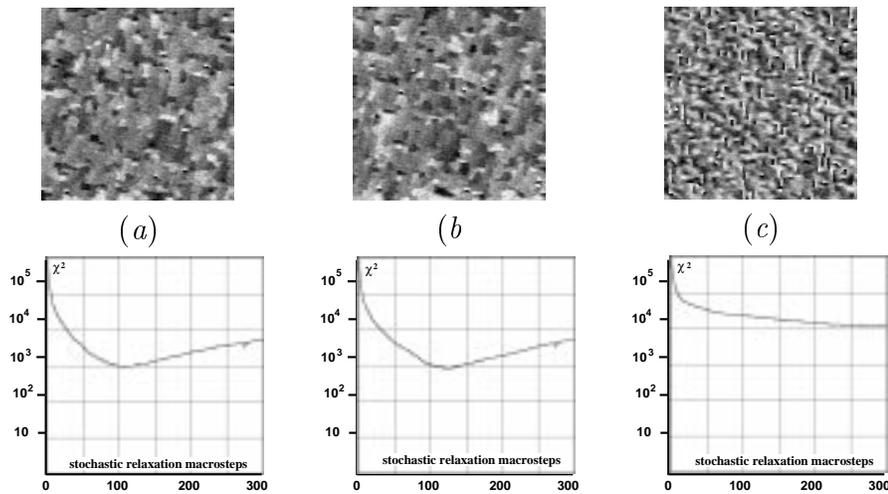


Fig. 2. Traditional modeling of the texture D29.

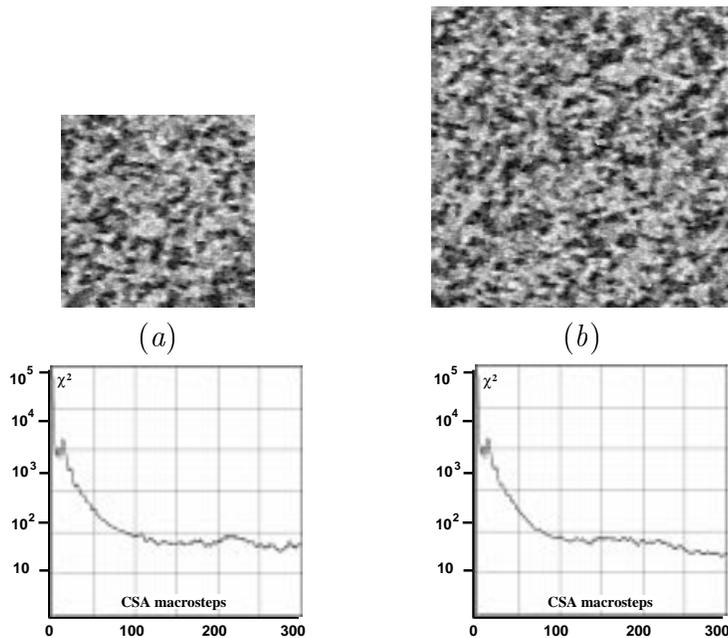


Fig. 3. CSA-based modeling of the texture D29.

the training sample after the first 100 or slightly more converging steps that gave the better average distance of about 1,000. Also, it is easily seen that the generated images have almost no visual resemblance to the training sample. The third example in Figure 2, *c* indicates that even after 2,000 steps of stochastic approximation for refining the potentials the created Markov chain of images is still very far from an equilibrium so that we cannot expect the better convergence at the subsequent generation stage.

Figure 3 shows that the desired textures can be simulated much easier and faster by the alternative CSA-based scenario. It should be noted that all the steps in both scenarios have the same computational complexity. For exam-

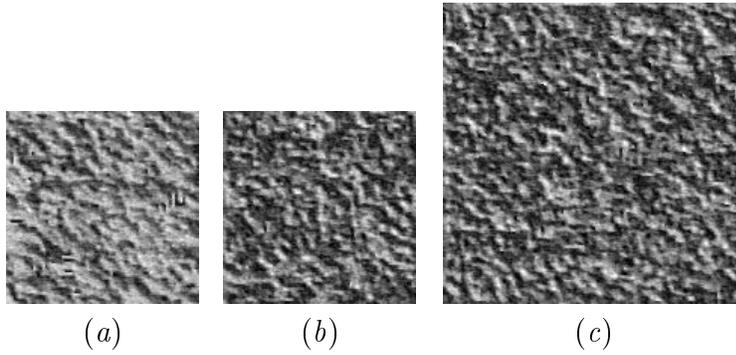


Fig. 4. Traditional (a) and CSA-based (b,c) modeling of the texture D4; the training sample is presented in Figure 1.

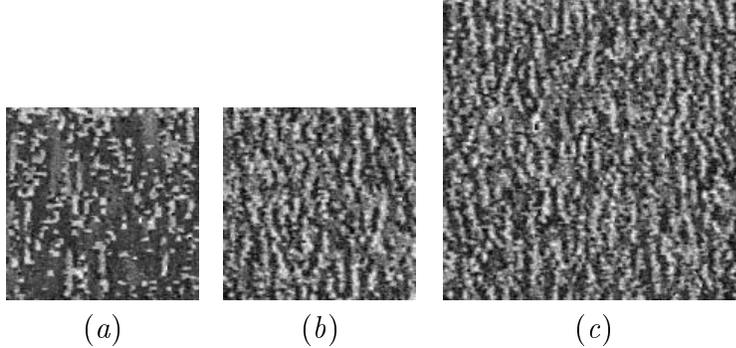


Fig. 5. Traditional (a) and CSA-based (b,c) modeling of the texture D24; the training sample is presented in Figure 1.

ple, only 300 CSA steps result in the considerably better convergence of the generated samples to the training ones in Figures 3,a and 3,b. The final average chi-square distances between the generated and training GLDHs for these two samples are about two orders smaller (30 – 70) than for the traditional scenario that performs in total 600 or 2300 steps. It is easily seen that the final images are visually similar to the training sample in Figure 1,(c).

Figures 4 – 7 permit us to compare results of simulating four more textures from [1] by both the traditional and alternative scenarios. The training samples for Figures 4 – 6 are given in Figure 1. In all the cases, except for the samples D55,c and D93,b–d, the traditional scenario performs 500 steps of stochastic approximation to refine the potentials and then 300 stochastic relaxation steps to generate the image (1000 and 500 steps for D55,a and 10,000 and 500 steps for D93,b–d, respectively). The alternative scenario uses only 300 CSA steps, except that the samples D93,h–j are obtained by 600, 1000, and 10,000 CSA steps, respectively. These latter samples which are very similar to the sample D93,g obtained by only 300 CSA steps, suggest that the CSA is quite stable after reaching a particular proximity to a training sample.

Figure 8 presents the training samples and the images simulated by the CSA for the two natural textures from [7]. Figure 9 indicates that the 300 steps

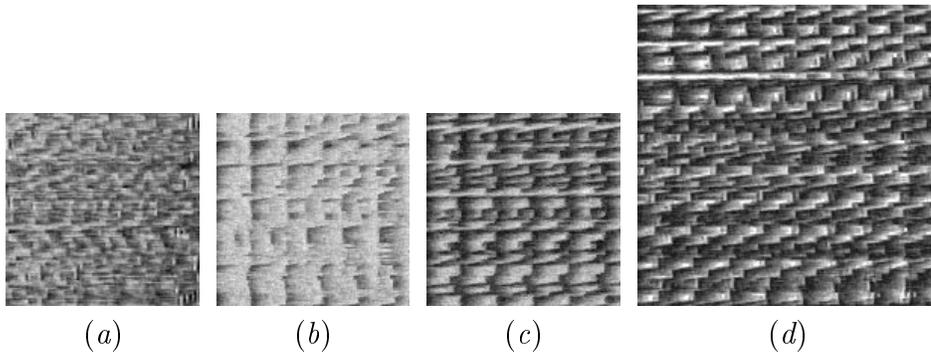


Fig. 6. Traditional  $(a,b)$  and CSA-based  $(c,d)$  modeling of the texture D55; the training sample is presented in Figure 1.

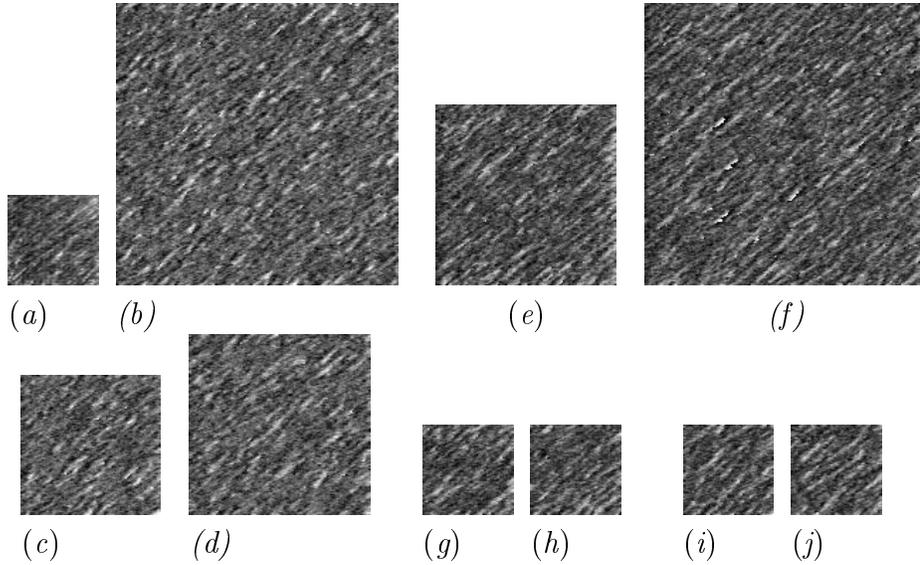


Fig. 7. Training sample  $64 \times 64$   $(a)$  and traditional  $(b - d)$  and CSA-based  $(e - j)$  modeling of the texture D93.

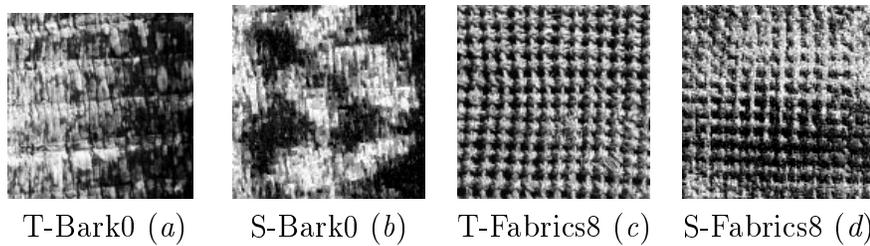


Fig. 8. Training and CSA-simulated samples of the natural textures Bark 0 and Fabrics 8.

of the CSA-based scenario, used in Figure 8, are considerably superior to the 2000 steps of stochastic approximation to refine the potentials and the subsequent 600 steps of stochastic relaxation with the fixed refined potentials to generate the images. In this latter case there is almost no convergence of the generated GLDHs to the training ones, and hence the visual resemblance be-

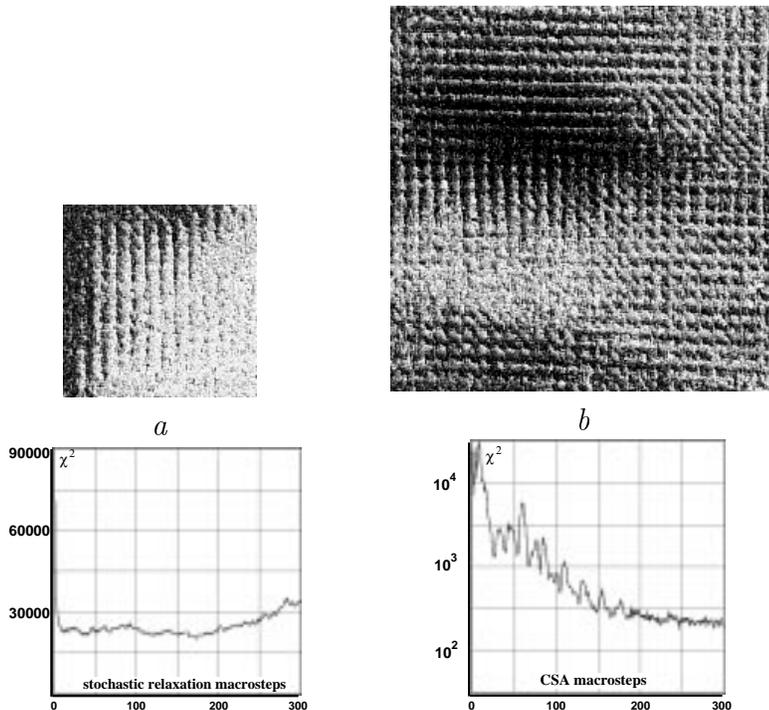


Fig. 9. Traditional and CSA-based modeling of the texture Fabrics 8 in Figure 8, (c): (a) the image  $128 \times 128$  generated by the traditional scenario (2,000 steps of potential refinement and 300 steps of stochastic relaxation) and (b) the image  $256 \times 256$  obtained by the CSA-based scenario (300 steps).

tween the training and simulated images is very poor. The CSA gives faster and much better results although the simulated images demonstrate continuous transitions between the original and inverted patterns of the same texture learned from the training sample. These transitions are unavoidable for the simplified Gibbs model that takes account of only the gray level differences if the GLDHs for a training sample are almost symmetric with respect to inversion  $q \rightarrow q_{\max} - q$  of the gray range, and the texture Fabrics 8 in Figure 8 has just such a feature.

Figure 10 shows the convergence of the CSA in terms of the average chi-square distance for the texture Bark 0. This texture is not translation-invariant, so that our modeling can only approximate it with the translation-invariant textures having the closely similar GLDHs.

In all these experiments the average chi-square distance between the generated and training GLDHs for the alternative scenario obtained after a few hundred CSA steps is about two orders smaller than for the traditional scenario that performs much more simulation steps of the same computational complexity. Moreover, if the potential refinement by stochastic approximation is terminated too early, then the traditional scenario may demonstrate even a divergence of this distance after a relatively short converging time.

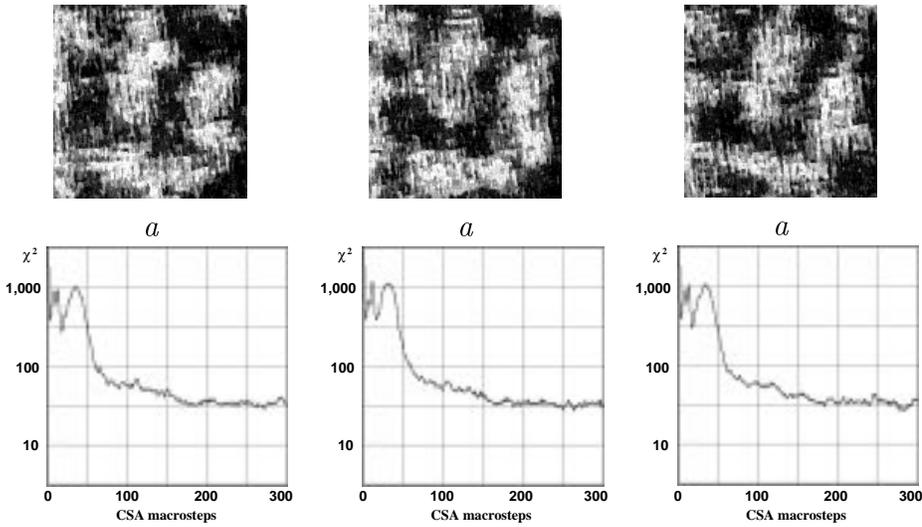


Fig. 10. CSA modeling of the texture Bark 0 in Figure 8,(a): (*a1 - c1*) the samples  $128 \times 128$  generated by 300 CSA steps, (*a2 - c2*) the convergence plots for these samples.

## 6 Concluding remarks

These and similar experiments show that the CSA-based modeling scenario is more practicable in texture simulation than the traditional scenario based on the prior parameter estimation and subsequent image simulation by stochastic relaxation. The traditional scenario demonstrates too slow convergence to the desired images and has no theoretically justified stopping rules at each stage of simulation. The CSA-based scenario has more practicable goal of approximating the GLCHs or GLDHs for a training sample by the similar histograms for the generated images, and this goal can be easily reached by the conventional stochastic approximation.

The textures described by the Gibbs model of Eq. (1) and simulated with a high degree of accuracy by the CSA-based scenario are assigned in [3] to a particular class of stochastic textures. A large body of experiments show that many natural homogeneous textures belong to this class: actually, the CSA-based simulation of 38 textures from [1] and 165 textures  $128 \times 128$  from [7] has shown that more than 30% of them belong to the stochastic textures. To be certain, let us identify by eye which textures in Figure 11 are natural.

Here, every upper right texture is natural and the two others are simulated by the CSA, but all three images are visually quite similar. Of course, vastly more textures which are outside the class of stochastic textures should be modeled by other means and scenarios.

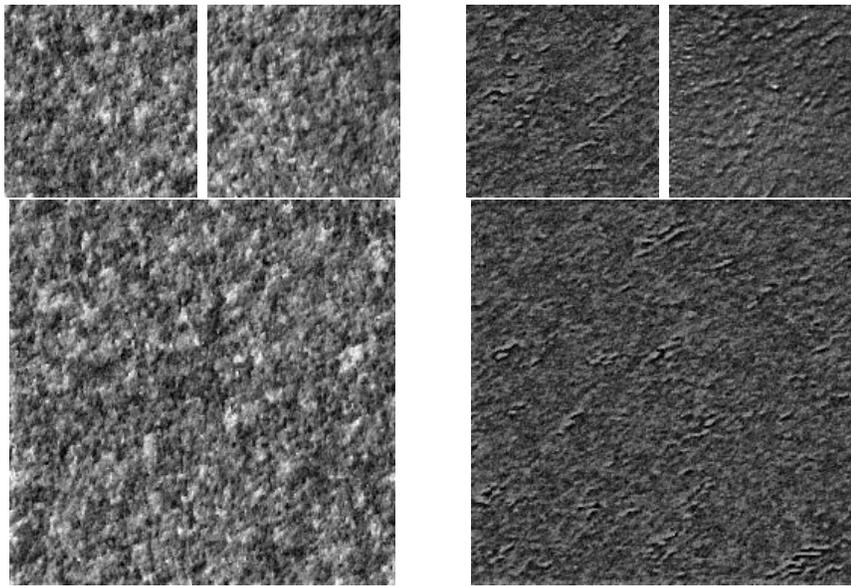


Fig. 11. Textures Fabrics 15 and Sand 2 from the MIT “VisTex” database [7].

## References

- [1] Brodatz, P., 1966. Textures: A Photographic Album for Artists and Designers. Dover Publications, New York.
- [2] Chellappa, R., Jain, A. (Eds), 1993. Markov Random Fields: Theory and Application. Academic Press, Boston.
- [3] Gimel'farb, G., 1996. Texture modeling with multiple pairwise pixel interactions. IEEE Trans. on Pattern Analysis and Machine Intelligence 18, 721–741.
- [4] Gimel'farb, G., 1998. On comparing two scenarios for probabilistic image modeling. In: Proc. Intern. Conf. on Image and Vision Computing New Zealand'98, 16-20 November 1998, Auckland, New Zealand. Uniprint, Auckland, 334–339.
- [5] van Laarhoven, P. J. M., 1988. Theoretical and Computational Aspects of Simulated Annealing. Stichting Mathematisch Centrum, Amsterdam.
- [6] Li, S. Z., 1995. Markov Random Fields: Theory and Application. Academic Press, Boston.
- [7] Pickard, R., Graszyk, C., Mann, S., Wachman, J., Pickard, L., Campbell, L., 1995. VisTex Database. MIT Media Laboratory, Cambridge, Mass.
- [8] Winkler, G., 1995. Image Analysis, Random Fields and Dynamic Monte Carlo Methods. Springer, Berlin.
- [9] Younes, L., 1988. Estimation and annealing for Gibbsian fields. Annales de l'Institut Henri Poincaré 24, 269–294.