

Basic and Fine Structure of Pairwise Interactions in Gibbs Texture Models

Georgy Gimel'farb

Centre for Image Technology and Robotics
Department of Computer Science, Tamaki Campus
University of Auckland
Private Bag 92019, Auckland 1, New Zealand

Abstract. Gibbs models with multiple pairwise pixel interactions permit us to estimate characteristic interaction structures of spatially homogeneous image textures. Interactions with partial energies over a particular threshold form a basic structure that is sufficient to model a specific group of stochastic textures. Another group, referred here to as regular textures, permits us to reduce the basic structure in size, providing only a few primary interactions are responsible for this structure. If the primary interactions can be considered as statistically independent, a sequential learning scheme reduces the basic structure and complements it with a fine structure describing characteristic minor details of a texture. Whereas the regular textures are described more precisely by the basic and fine interaction structures, the sequential search may deteriorate the basic interaction structure of the stochastic textures.

Keywords: *image texture, Gibbs random field, interaction structure.*

1 Introduction

Spatially homogeneous image textures are represented as samples of a particular Gibbs random field by specifying a geometric structure and quantitative strengths, or Gibbs potentials, of multiple pairwise pixel interactions [2,3]. The interaction structure determines which pixels directly interact with a particular pixel in the sense that they effect conditional probabilities of grey levels in the pixel. The interacting pixels are usually called the neighbours, and the interactions are described by a neighbourhood graph [2].

The spatially homogeneous interaction structure is represented by several families of translation invariant pixel pairs, or second-order cliques of a neighbourhood graph, each clique family having its own potential. Generally, the potentials depend on grey level co-occurrences (GLC) in a pixel pair. The sum of the potentials over a clique family is the partial interaction energy that determines the contribution of the family to the probability of a particular image.

As shown in [3,4], the analytical first approximation of the maximum likelihood estimate (MLE) of the potential for a particular texture is proportional to the centered GLC histogram (GLCH) for the corresponding clique family in a given training sample of the texture. Therefore the characteristic interaction

structure can be chosen by comparing the analytical estimates of the partial energies for all clique families in a particular large search set.

Many spatially homogeneous textures belong to a specific class of stochastic textures that can be efficiently described by only a basic interaction structure consisting of interactions with the partial energies over a particular threshold. The threshold depends on the relative frequency distribution of the energies for the clique families in the search set [3,4].

Textures with a regular visual pattern can be only roughly modelled by such basic structure because they have also a fine structure of pairwise pixel interactions. The fine structure ranks below the basic structure in energy but describes visually important repetitive minor details. Generally, the probability distributions of the GLCs for the various clique families are statistically interdependent. But if the dependence between some families can be ignored, the clique families in the search set can be separated into the two groups: (i) the independent primary interactions with the top partial energies and (ii) the dependent secondary interactions with the lower energies obtained by a statistical interplay of the primary interactions. In this case the basic structure can be reduced in size and the fine structure can be recovered by an empirical sequential choice of the primary interactions that eliminates the secondary ones [5,6].

This paper compares the empirical sequential scheme of learning the interaction structure to the approximate analytical and the combined analytical-empirical sequential schemes. Textures that can be efficiently modelled by the sequentially chosen basic and fine structures form a specific group of regular textures differing from the stochastic textures. In the same time the sequential choice based on partial energies may result in worse interaction structures of stochastic textures with respect to the conventional thresholding of partial energies.

2 Search for the Interaction Structure

2.1 Basic Notation

Let $\mathbf{g} = [g_i : i \in \mathbf{R}; g_i \in \mathbf{Q}]$ be a digital greyscale image with a finite set of grey levels $\mathbf{Q} = \{0, 1, \dots, q_{max}\}$. Here, \mathbf{R} is a finite arithmetic lattice supporting the images. A spatially homogeneous structure $\mathbf{C} = [\mathbf{C}_a : a \in \mathbf{A}]$ of pairwise interactions between the pixels $i \in \mathbf{R}$ is specified by a particular subset of the clique families $\mathbf{C}_a = \{(i, j) : (i, j) \in \mathbf{R}^2; i - j = \text{const}_a\}$. Every family consists of the translation invariant cliques (i, j) with the fixed inter-pixel shift $i - j = \text{const}_a \equiv (\Delta x_a, \Delta y_a)$.

A partial interaction energy $E_a(\mathbf{g})$ of a clique family \mathbf{C}_a in an image \mathbf{g} is

$$E_a(\mathbf{g}) = \sum_{(i,j) \in \mathbf{C}_a} V_a(g_i, g_j) = \mathbf{V}_a \bullet \mathbf{H}_a(\mathbf{g}) \quad (1)$$

where $\mathbf{V}_a = [V(q, s) : (q, s) \in \mathbf{Q}^2]$ is a Gibbs potential for the clique family \mathbf{C}_a with values depending on the GLCs (q, s) , $\mathbf{H}_a(\mathbf{g})$ is the GLC histogram (GLCH) collected in the image \mathbf{g} over the family \mathbf{C}_a , and \bullet denotes the dot product.

Let $\mathbf{V} = [\mathbf{V}_a : a \in \mathbf{A}]$ and $\mathbf{H}(\mathbf{g}) = [\mathbf{H}_a(\mathbf{g}) : a \in \mathbf{A}]$ denote the potential vector and the GLCH vector, respectively. The GLC-based Gibbs image model with multiple pairwise pixel interactions [4]:

$$\Pr(\mathbf{g}|\mathbf{C}, \mathbf{V}) = \frac{1}{Z_{\mathbf{C},\mathbf{V}}} \exp(E(\mathbf{g})) \tag{2}$$

relates the probability of every sample g to its total interaction energy $E(\mathbf{g})$:

$$E(\mathbf{g}) = \sum_{a \in \mathbf{A}} E_a(\mathbf{g}) = \mathbf{V} \bullet \mathbf{H}(\mathbf{g}). \tag{3}$$

As shown in [4], the analytical first approximation of the MLE of the potential \mathbf{V}_a is proportional to the difference between the GLCH $\mathbf{H}_a(\mathbf{g}^\circ)$ and the expected uniform GLCH $\mathbf{H}_{a,\text{irf}}$ for the samples of the independent random field (IRF), or what is the same, to the centered GLCH. The IRF is described by the Gibbs model of Eq. (2) with zero-valued potentials $\mathbf{V}_a = \mathbf{0}$.

Therefore, characteristic interaction structure can be recovered by comparing the analytical estimates of the relative partial interaction energies

$$e_a(\mathbf{g}^\circ) = \mathbf{H}_a(\mathbf{g}^\circ) \bullet (\mathbf{H}_a(\mathbf{g}^\circ) - \mathbf{H}_{a,\text{irf}}) \tag{4}$$

for a large search set \mathbf{W} of possible clique families \mathbf{C}_a . In all the experiments below, the search set \mathbf{W} contains 3280 clique families with the inter-pixel shifts in the range $-40 \leq \Delta x_a, \Delta y_a \leq 40$.

2.2 Basic Structure via Thresholding the Partial Energies

In the simplest case, the basic interaction structure can be learnt by comparing relative partial energies $e_a(\mathbf{g}^\circ)$ to a threshold that depends on the frequency distribution of all the energies for the search set \mathbf{W} . Such a structure is sufficient to simulate many natural image textures called stochastic textures in [3,4].

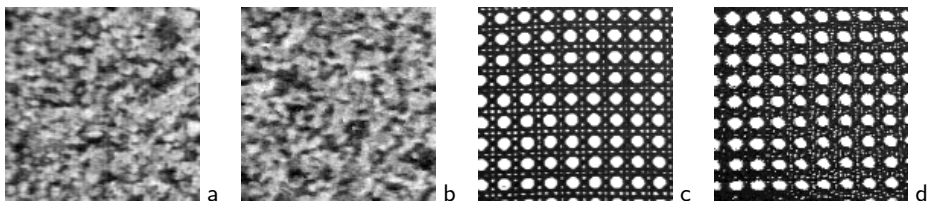


Fig. 1. Training and simulated samples D29 (a, b) and D101 (c, d) with the basic structures learnt by thresholding the analytically estimated partial energies.

Figure 1 shows, for example, the training and simulated samples 128×128 of the textures D29 “Beach sand” and D101 “Cane” [1]. The basic interaction structures (11 and 39 clique families, respectively) were learnt by using the threshold

$\theta = E_m + 4\sigma$ where E_m and σ are the mean energy and the standard deviation, respectively. Such a basic structure results in the visually good simulation of the stochastic texture D29. But the minor repetitive details of the more regular texture D101 are not described at all, and the simulation gives only a very rough approximation of the original visual pattern.

2.3 Empirical Sequential Learning

The above thresholding may produce basic structures of larger size than one needs for describing the characteristic visual features of a texture if some GLC distributions over clique families with the top partial energies can be considered as statistically independent. Below, such families will be referred to as the primary ones. Let \mathbf{C}_α and \mathbf{C}_β be the primary clique families with the sizable energies $E_\alpha(\mathbf{g}^\circ) > \theta$ and $E_\beta(\mathbf{g}^\circ) > \theta$, respectively. Then they can give rise to a lower but still sizable energy $E_\gamma(\mathbf{g}^\circ)$ for the secondary family \mathbf{C}_γ , such that $\text{const}_\gamma = \text{const}_\alpha + \text{const}_\beta$, although the latter family may not take part in the basic structure. It is evident that the straightforward thresholding cannot detect a fine interaction structure describing minor but visually characteristic regular details of a texture if their interaction energies are lower than the energies of secondary interactions produced by the primary basic ones.

Empirical sequential learning, proposed first by Zalesny [5,6], reduces the basic structure to only the primary interactions and recovers the fine structure by repeating iteratively the texture simulation and structure selection steps. We shall restrict our consideration to the specific type of sequential learning that is based on the relative partial energies of interactions. At each iteration t , a new image sample $\mathbf{g}^{[t]}$ is simulated under a current interaction structure $\mathbf{C}^{[t]}$. Then the GLCHs $\mathbf{H}_a(\mathbf{g}^\circ)$ for a given training sample \mathbf{g}° are compared to the GLCHs $\mathbf{H}_a(\mathbf{g}^{[t]})$ for the simulated sample, and the clique family with the maximum relative partial energy

$$e_a(\mathbf{g}^\circ) = \mathbf{H}_a(\mathbf{g}^\circ) \bullet (\mathbf{H}_a(\mathbf{g}^\circ) - \mathbf{H}_a(\mathbf{g}^{[t]})) \quad (5)$$

is selected to be added to the current structure.

In principle, all the statistical interplay between the primary and secondary energies is taken into account by simulation so that both the basic and fine structures of the minimum size are expected to be found. But it should be noted that image simulation with a fixed interaction structure and potentials results in a set of different samples such that their GLCHs approach the GLCHs for a given training sample only in average. Therefore the obtained basic and especially fine structures will reflect also a particular sequence of simulated images, and the same training sample may produce notably different interaction structures.

Figures 2–4 show the textures simulated after learning the interaction structure by the empirical sequential search using the training samples D29 and D101 in Figure 1. The primary structure found for the texture D29 contains only the two clique families with the inter-pixel shifts $[0, 1]$ and $[1, 0]$. The additional fine structure includes the families with very low energies. As a result, they are chosen rather arbitrary, and the resulting structure is unsuitable for simulating the

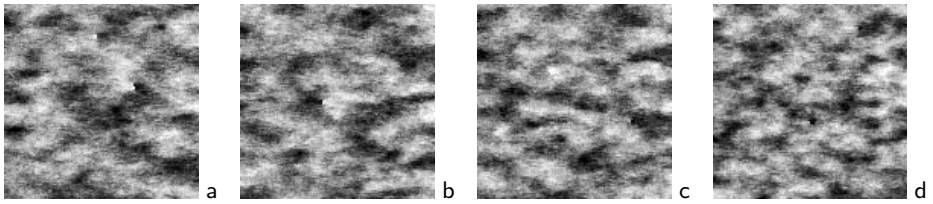


Fig. 2. Texture D29 simulated with 4 (a), 5 (b), 7 (c), and 11 (d) clique families found by the sequential empirical choice of the family with the top relative partial energy. Notice that these images do not mimic the initial visual pattern of Figure 1,a.

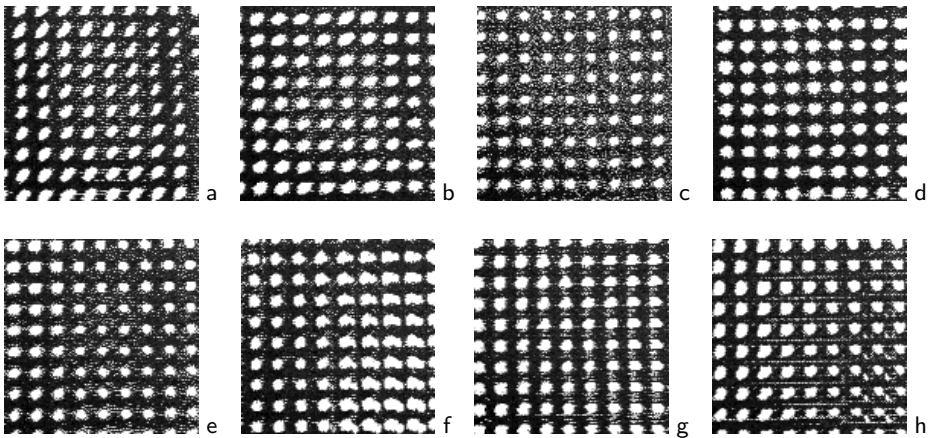


Fig. 3. Texture D101 simulated with 7 (a) – 14 (h) clique families found by the sequential empirical choice of the family with the top relative partial energy.

texture samples that possess the visual similarity to the training sample D29 in Figure 1,a. The images in Figure 2, in contrast to the simulated sample in Figure 1,b, differ much from the training sample even when the overall interaction structure is of the same or greater size than the basic structure recovered by thresholding the energies.

But as shown in Figures 3 and 4, the overall structure found sequentially for modelling the texture D101 with a repetitive visual pattern, both contains less clique families and represents better the fine details than the basic structure found by thresholding the energies. The sequential choice of a single top clique family with the highest relative energy of Eq. (5) proposed in [5,6] forms the basic structure of about 16 clique families and the fine structure of 4–6 clique families (compare Figure 1,d to Figures 3 and 4).

Similar results in Figure 5 are obtained twice faster by choosing the two top families at each step. But the structures of similar size obtained by choosing more than two clique families per iteration give somewhat worse simulation results (see

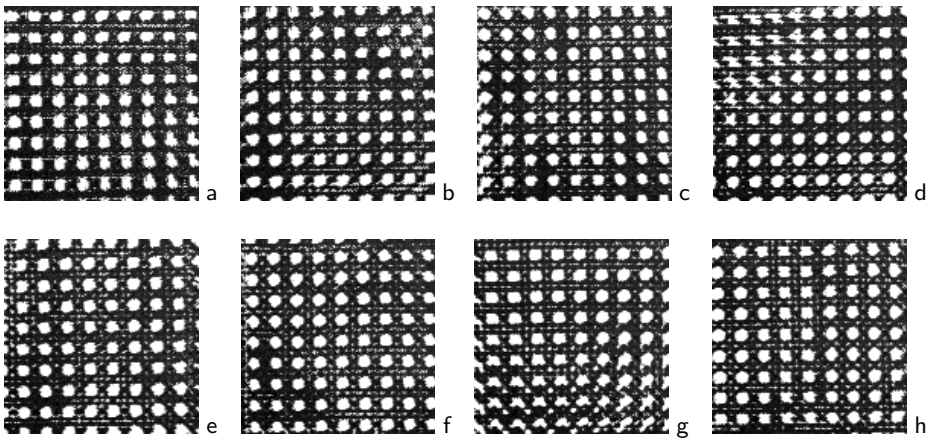


Fig. 4. Texture D101 simulated with 15 (a) – 22 (h) clique families found by the sequential empirical choice of the family with the top relative partial energy.

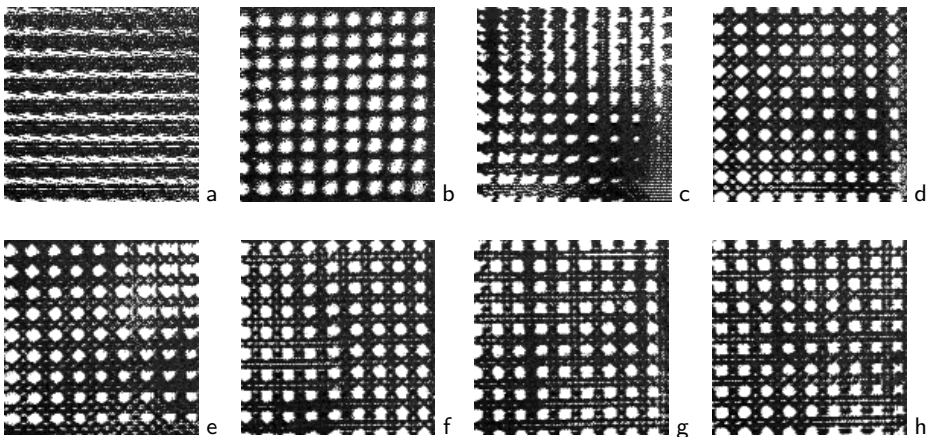


Fig. 5. Texture D101 simulated with 4 (a) – 12 (e), 16 (f), 18 (g), and 22 (h) clique families found by the sequential empirical choice of the top two families.

Figure 6). Also, it should be noted that the visual quality of simulation does not steadily increase with the structure size. As follows from Figures 3–6, the quality may even degrade after adding a clique family and then be restored after adding one–two more families.

2.4 Analytical Sequential Learning

Assuming the probability distributions of the GLCs for the primary clique families are statistically independent, the secondary interactions with relatively high

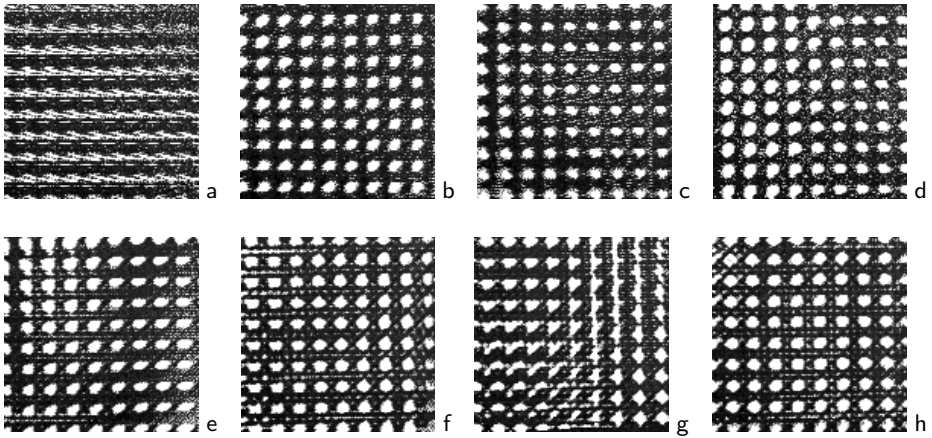


Fig. 6. Texture D101 simulated with 4 (a) – 32 (h) clique families found by the sequential empirical choice of the top four families.

energies can be approximately taken into account by recomputing each secondary GLCH, $\mathbf{H}_\gamma(\mathbf{g})$, using each currently chosen primary GLCH, $\mathbf{H}_\alpha(\mathbf{g})$, and the corresponding previous GLCH $\mathbf{H}_\beta(\mathbf{g})$ such that $\text{const}_\gamma = \text{const}_\alpha + \text{const}_\beta$. In this case all the GLCHs in the search set \mathbf{W} can be analytically updated after adding to a current interaction structure the next clique family with the maximum relative energy of Eq. (5) with respect to the training sample. Such an analytical estimation does not take account of all the statistical interplay of the families but only approximates the actual distribution of the relative energies.

Figures 7,a–b, and 8,a–b demonstrate the grey-coded actual and analytically computed distributions of the partial energies of Eq. (4) for the textures D29 and D101 over the search set \mathbf{W} . Here, each square box of size 4×4 pixels represents a particular inter-pixel shift $\text{const}_a = (\Delta x, \Delta y)$; $-40 \leq \Delta x, \Delta y \leq 40$. The energy distributions for the textures D29 and D101 are computed, respectively, with the 10 and 15 sequentially chosen primary clique families. The corresponding interaction structures as well as the more detailed structures with the 22 clique families are shown in Figures 7,c–d and 8,c–d. The low-energetic fine structure of the texture D29 does not represent specific visual features and is obviously arbitrary, as distinct from the regular fine structure of the texture D101.

The sequential analytical scheme results in a sufficiently accurate approximation of the actual interaction energies. Therefore it can be used for reducing the size of the basic structure with respect to the like structures obtained by thresholding the energies of Eq. (4). The samples D101 in Figure 9,a–c simulated with the analytically found interaction structures containing 15–25 clique families are very similar to the sample D101 in Figure 1,d simulated with the 39 families. The texture in Figure 9,d simulated with the 32 clique families reflects also some fine visual details but to the lesser extent than the samples in Figures 4,e–h, 5,d–h, and 6,h, obtained by the empirical sequential learning.

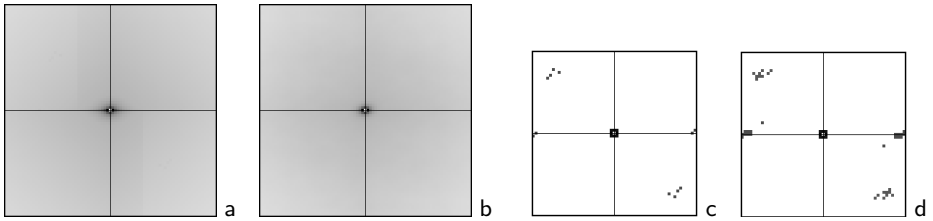


Fig. 7. Estimated with 10 clique families (a) and actual (b) energies for the D29 model and the interaction structures with 10 (c) and 22 (d) analytically chosen families.

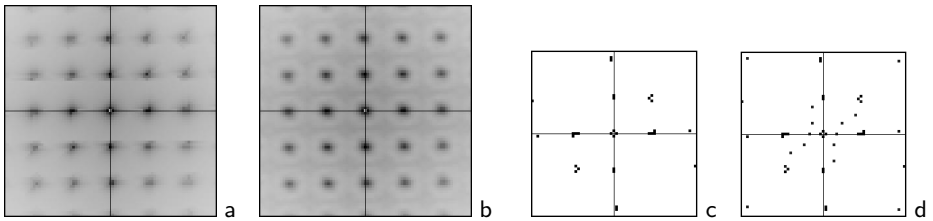


Fig. 8. Estimated with 15 clique families (a) and actual (b) energies for the D101 model and the interaction structures with 15 (c) and 22 (d) analytically chosen families.

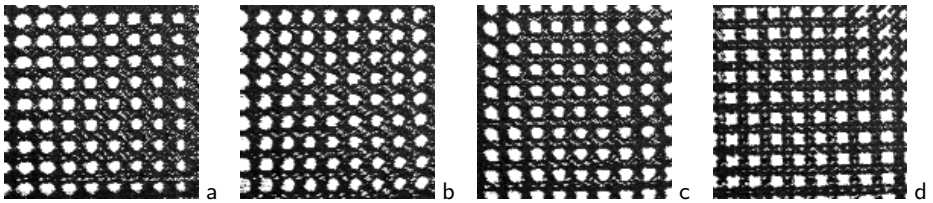


Fig. 9. Texture D101 simulated with 15 (a), 19 (b), 25 (c), and 32 (d) analytically chosen clique families.

As concerning the stochastic texture D29, the analytical sequential search for the top relative partial energies has the same drawbacks as the empirical one. Figures 1, a–b, 7, and 10 show that the basic structure recovered by thresholding produces much better simulation results even when the analytically chosen structure is of larger size.

The possible reason is that the assumed statistical independence of the GLC distributions for the primary interactions does not hold for this texture so that the exclusion of the secondary interactions by using the relative energies of Eq. (5) is not justified. In such a case the search for the reduced basic structure and additional fine can only deteriorate the actual basic structure.

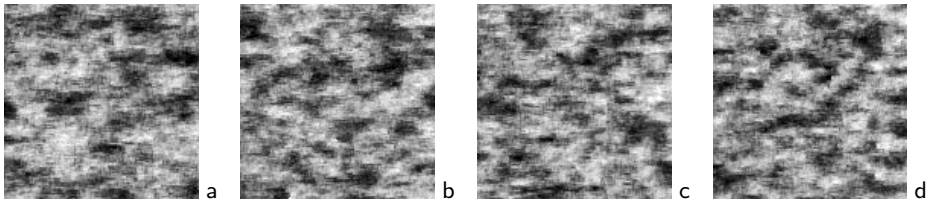


Fig. 10. Texture D29 simulated with 4 (a), 10 (b), 15 (c), and 22 (d) analytically chosen clique families.

2.5 Combined Sequential Learning

As follows from the above experiments, the sequential learning can produce the efficient interaction structures only if our assumption about the independent primary GLC distributions has a reasonable fit to the textures under consideration. If this assumption holds, let these latter be called the regular textures.

The empirical sequential learning outperforms the faster analytical scheme as concerning the fine interaction structure of a regular texture. But the reduced basic structures recovered empirically or analytically are very similar so that the sequential learning can be accelerated by combining the both approaches.

Figure 11 shows the results of simulating the texture D101 when the reduced basic interaction structure with 15 clique families is first found analytically (see Figures 8 and 9) and then is appended with the fine structure of 1–8 clique families by the empirical learning. It is evident that the purely empirical and the combined analytical–empirical sequential learning produce very similar final results, but the latter approach is much faster than the former one.

3 Conclusions

These and other our experiments (as well as experiments [6] in empirical sequential learning based on the chi-square distances between the GLCHs) suggest that modelling of spatially homogeneous textures with the Gibbs model of Eq. (2) must take account of possible statistical dependences between the clique families that form the characteristic interaction structure. Stochastic textures introduced in [3,4] have basic structures of only weakly interdependent primary interactions so that no interaction with a sizeable partial energy can be considered as the secondary one and excluded from the basic structure.

Regular textures differ from the stochastic ones in that they can be modelled by the reduced basic and the additional fine structures. The initial basic structure contains both the strongly and weakly interdependent interactions with a sizable energy. Assuming that only the top-energetic interactions are the independent primary ones, the basic structure is reduced in size by the empirical or analytical sequential exclusion of the dependent secondary interactions. Then the fine structure is recovered in the like way by the empirical sequential learning.

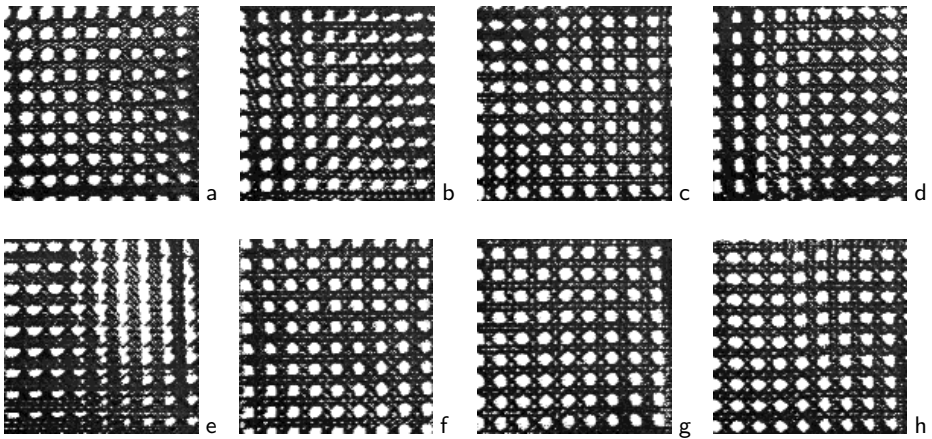


Fig. 11. Texture D101 simulated with the interaction structure containing 15 analytically found clique families and the additional 1 (a) – 8 (h) families found by the sequential empirical choice of the top family.

The sequential learning extends the range of image textures that can be modelled by multiple pairwise pixel interactions but it does not replace the simple energy thresholding for the stochastic textures. Also, the sequential learning schemes, as well as the parallel thresholding of partial energies, have still no theoretically justified rules for choosing an adequate size of the interaction structure. Thus the number of clique families to be used in the Gibbs model of a particular texture is selected, mainly, on the experimental base. Our experiments and experiments in [4,6] show that many natural spatially homogeneous images are of the stochastic or regular type. But a vast majority of images are outside these types and should be modelled by other means.

References

1. Brodatz, P.: Textures: A Photographic Album for Artists and Designers. Dover Publications, New York (1966)
2. Chellappa, R., Jain A. (eds): Markov Random Fields: Theory and Application. Academic Press, Boston (1993)
3. Gimel'farb, G.L.: Texture modeling by multiple pairwise pixel interactions. IEEE Trans. Pattern Analysis Machine Intell. **18** (1996) 1110–1114
4. Gimel'farb, G.L.: Image Textures and Gibbs Random Fields. Kluwer Academic Publishers, Dordrecht (1999)
5. Zalesny, A.V. Personal communication. (1998)
6. Zalesny, A.V. Analysis and synthesis of textures with pairwise signal interactions. Project 50542 (EC ACTS 074 "Vanguard"). Techn. Rept. Katholieke Univ., Leuven (1999)