

# Image Inpainting by Cooling and Heating

David Gustavsson<sup>1</sup>, Kim S. Pedersen<sup>2</sup>, and Mads Nielsen<sup>2</sup>

<sup>1</sup> IT University of Copenhagen

Rued Langgaards Vej 7, DK-2300 Copenhagen S, Denmark

davidgsson@itu.dk

<sup>2</sup> DIKU, University of Copenhagen

Universitetsparken 1, DK-2100 Copenhagen Ø, Denmark

{kimstp,madsn}@diku.dk

**Abstract.** We discuss a method suitable for inpainting both large scale geometric structures and stochastic texture components. We use the well-known FRAME model for inpainting. We introduce a temperature term in the learnt FRAME Gibbs distribution. By using a fast cooling scheme a MAP-like solution is found that can reconstruct the geometric structure. In a second step a heating scheme is used that reconstruct the stochastic texture. Both steps in the reconstruction process are necessary, and contribute in two very different ways to the appearance of the reconstruction.

**Keywords:** Inpainting, FRAME, ICM, MAP, Simulated Annealing.

## 1 Introduction

Image inpainting concerns the problem of reconstruction of the image contents inside a region  $\Omega$  with unknown or damaged contents. We assume that  $\Omega$  is a subset of the image domain  $D \subseteq \mathbb{R}^2$ ,  $\Omega \subset D$  and we will for this paper assume that  $D$  form a discrete lattice. The reconstruction is based on the available surrounding image content. Some algorithms have reported excellent performance for pure geometric structures (see e.g. [1] for a review of such methods), while others have reported excellent performance for pure textures (e.g. [2,3,4]), but only few methods [5] achieve good results on both types of structures.

The variational approaches have been shown to be very successful for geometric structures but have a tendency to produce a too smooth solution without fine scale texture (See [1] for a review). Bertalmio et al [5] propose a combined method in which the image is decomposed into a structure part and a texture part, and different methods are used for filling the different parts. The structure part is reconstructed using a variational method and the texture part is reconstructed by image patch pasting.

Synthesis of a texture and inpainting of a texture seem to be, more or less, identical problems, however they are not. In [6] we propose a two step method for inpainting based on Zhu, Wu and Mumford's stochastic FRAME model (Filters, Random fields and Maximum Entropy) [7,8]. Using FRAME naively for

inpainting does not produce good results and more sophisticated strategies are needed and in [6] we propose such a strategy. By adding a temperature term  $T$  to the learnt Gibbs distribution and sampling from it using two different temperatures, both the geometric and the texture component can be reconstructed. In a first step, the geometric structure is reconstructed by sampling using a cooled - i.e. using a small fixed  $T$  - distribution. In a second step, the stochastic texture component is added by sampling from a heated - i.e. using a large fixed  $T$  - distribution.

Ideally we want to use the MAP solution of the FRAME model to reconstruct geometric structure of the damaged region  $\Omega$ . In [6] we use a fixed low temperature to find a MAP-Like solution in order to reconstruct the geometric structure. To find the exact MAP-solution one must use the time consuming simulated annealing approach such as described by Geman and Geman [9]. However to reconstruct the missing contents of the region  $\Omega$ , the true MAP solution may not be needed. Instead a solution which is close to the MAP solution may provide visually good enough results. In this paper we propose a fast cooling scheme that reconstruct the geometric structure and approaches the MAP solution. Another approach is to use the solution produced by the Iterated Conditional Modes (ICM) algorithm (see e.g. [10]) for reconstruction of the geometric structure. Finding the ICM solution is much faster than our fast cooling scheme, however it often fails to reconstruct the geometric structure. This is among other things caused by the ICM solutions strong dependence on the initialisation of the algorithm. We compare experimentally the fast cooling solution with the ICM solution.

To reconstruct the stochastic texture component the Gibbs distribution is heated. By heating the Gibbs distribution more stochastic texture structures will be reconstructed without destroying the geometric structure that was reconstructed in the cooling step. In [6] we use a fixed temperature to find a solution including the texture component. Here we introduce a gradual heating scheme.

The paper has the following structure. In section 2 FRAME is reviewed, in section 2.1 filter selection is discussed and in section 2.2 we explain how FRAME is used for reconstruction. Inpainting using FRAME is treated in section 3. In section 3.1 a temperature term is added to the Gibbs distribution, the ICM solution and fast cooling solution is discussed in sections 3.2 and 3.3. Adding the texture component by heating the distribution is discussed in section 3.4. In section 4 experimental results are presented and in section 5 conclusion are drawn and future work is discussed.

## 2 Review of FRAME

FRAME is a well known method for analysing and reproducing textures [8,7]. FRAME can also be thought of as a general image model under the assumptions that the image distribution is stationary. FRAME constructs a probability distribution  $p(I)$  for a texture from observed sample images.

Given a set of filters  $F^\alpha(I)$  one computes the histogram of the filter responses  $H^\alpha$  with respect to the filter  $\alpha$ . The filter histograms are estimates of marginal distributions of the full probability distribution  $p(I)$ . Given the marginal distributions for the sample images one wants to find all distributions that have the same expected marginal distributions, and among those find the distribution with maximum entropy, i.e. by applying the maximum entropy principle. This distribution is the least committed distribution fulfilling the constraints given by the marginal distributions. This is a constrained optimisation problem that can be solved using Lagrange multipliers. The solution is

$$p(I) = \frac{1}{Z(\Lambda)} \exp\left\{-\sum_i \sum_\alpha \lambda_i^\alpha H_i^\alpha\right\} \quad (1)$$

Here  $i$  is the number of histogram bins in  $H^\alpha$  for the filter  $\alpha$  and  $\Lambda = \{\lambda_i^\alpha\}$  are the Lagrange multipliers which gives information on how the different values for the filter  $\alpha$  should be distributed. The relation between  $\lambda_i^\alpha$ :s for different filters  $F^\alpha$  gives information on how the filters are weighted relative to each other.

An Algorithm for finding the distribution and  $\Lambda$  can be found in [7]. FRAME is a generative model and given the distribution  $p(I)$  for a texture it can be used for inference (analysis) and synthesis.

## 2.1 The Choice of Filter Bank

We have used three types of filters in our experiments: The delta filter, the power of Gabor filters and Scale Space derivative filters. The delta, Scale Space derivative and Gabor filters are linear filters, hence  $F^\alpha(I) = I * F^\alpha$ , where  $*$  denotes convolution. The power of the Gabor filter is the squared magnitude applied to the linear Gabor filter.

The Filters  $F^\alpha$  are:

- Delta filter - given by the Dirac delta  $\delta(x)$  which simply returns the intensity at the filter position.
- the power of Gabor filters - defined by  $|I * G_\sigma e^{-i\omega x}|^2$ , where  $i^2 = -1$ . Here we use 8 orientations,  $\omega = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4}$  and 2 scales  $\sigma = 1, 4$ , in total 16 Gabor filters have been used.
- Scale space derivatives - using 3 scales  $\sigma = 0.1, 1, 3$  and 6 derivatives  $G_\sigma, \frac{\partial G_\sigma}{\partial x}, \frac{\partial G_\sigma}{\partial y}, \frac{\partial^2 G_\sigma}{\partial x^2}, \frac{\partial^2 G_\sigma}{\partial y^2}, \frac{\partial^2 G_\sigma}{\partial x \partial y}$ .

For both the Gabor and scale space derivative filters the Gaussian aperture function  $G_\sigma$  with standard deviation  $\sigma$  defining the spatial scale is used,

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).$$

Which and how many filters should be used have a large influence on the type of image that can be modelled. The filters must catch the important visual appearance of the image at different scales. The support of the filters determines

a Markov neighbourhood. Small filters add fine scale properties of the image, while large filters add coarse scale properties of the image. Hence to model properties at different scales, different filter sizes must be used. The drawback of using large filters is that the computation time increases with the filter size. On the other hand large filters must be used to catch coarse scale dependencies in the image.

Gabor filters are orientation sensitive and have been used for analysing textures in a number of papers and are in general suitable for textures (e.g. [11,12]). By carefully selecting the orientation  $\omega$  and the scale  $\sigma$ , structures with different orientations and scales will be captured.

It is well known from scale space theory that scale space derivative filters capture structures at different scales. By increasing  $\sigma$  in the Gaussian kernel, finer details are suppressed, while coarse structures are enhanced. By using the full scale-space both fine and coarse scale structures will be captured [13].

## 2.2 Sampling

Once the distribution  $p(I)$  is learnt, it is possible to use a Gibbs sampler to synthesise images from  $p(I)$ .  $I$  is initialised randomly (or in some other way based on prior knowledge). Then a site  $(x, y)_i \in D$  is randomly picked and the intensity  $I_i = I((x, y)_i)$  at  $(x, y)_i$  is updated according to the conditional distribution [14,10]

$$p(I_i|I_{-i}) \tag{2}$$

where the notation  $I_{-i}$  denotes the set of intensities at the set of sites  $\{(x, y)_{-i}\} = D \setminus (x, y)_i$ . Hence  $p(I_i|I_{-i})$  is the probability for the different intensities in site  $(x, y)_i$  given the intensities in the rest of the image. Because of the equivalence between Gibbs distributions and Markov Random Fields given a neighbourhood system  $N$  (the Hammersley-Clifford theorem, see e.g. [10]), we can make the simplification

$$p(I_i|I_{-i}) = p(I_i|N_i) \tag{3}$$

where  $N_i \subset D \setminus (x, y)_i$  is the neighbourhood of  $(x, y)_i$ . In the FRAME model, the neighbourhood system  $N$  is defined by the extend of the filters  $F^\alpha$ .

By sampling from the conditional distribution in (3),  $I$  will be a sample from the distribution  $p(I)$ .

## 3 Using FRAME for Inpainting

We can use FRAME for inpainting by first constructing a model  $p(I)$  of the image, e.g. by learning from the non-damaged part of the image,  $D \setminus \Omega$ . We then use the learnt model  $p(I)$  to sample new content inside the damaged region  $\Omega$ . This is done by only updating sites in  $\Omega$ . A site  $(x, y)_i \in \Omega$  is randomly picked and updated by sampling from the conditional distribution given in (3). If the

site  $(x, y)_i$  is close (in terms of filter size) to the boundary  $\partial\Omega$  of the damaged region, then the filters get support from both sites inside and outside  $\Omega$ . The sites outside  $\Omega$  are known and fixed, and are boundary conditions for the inpainting. We therefore include a small band region around  $\Omega$  in the computation of the histograms  $H^\alpha$ . Another option would have been to use the whole image  $I$  to compute the histogram  $H^\alpha$ , however this has the downside that the effect of updates inside  $\Omega$  on the histograms are dependent on the relative size ratio between  $\omega$  and  $D$ , causing a slow convergence rate for small  $\Omega$ .

### 3.1 Adding a Temperature Term $\beta = \frac{1}{T}$

Sampling from the distribution  $p(I)$  using a Gibbs sampler does not easily enforce the large scale geometric structure in the image. By using the Gibbs sampler one will get a sample from the distribution, this includes both the stochastic and the geometric structure of the image, however the stochastic structure will dominate the result.

Adding an inverse temperature term  $\beta = \frac{1}{T}$  to the distribution gives

$$p(I) = \frac{1}{Z(\Lambda)} \exp\left\{-\beta \sum_{\alpha} \sum_i \lambda_i^\alpha H_i^\alpha\right\}. \quad (4)$$

In [6] we proposed a two step method to reconstruct both the geometric and stochastic part of the missing region  $\Omega$ :

1. Cooling: By sampling from (4) using a fixed small temperature  $T$  value, structures with high probability will be reconstructed, while structures with low probability will be suppressed. In this step large geometric structures will be reconstructed based on the model  $p(I)$ .
2. Heating: By sampling from (4) using a fixed temperature  $T \approx 1$ , the texture component of the image will be reconstructed based on the model  $p(I)$ .

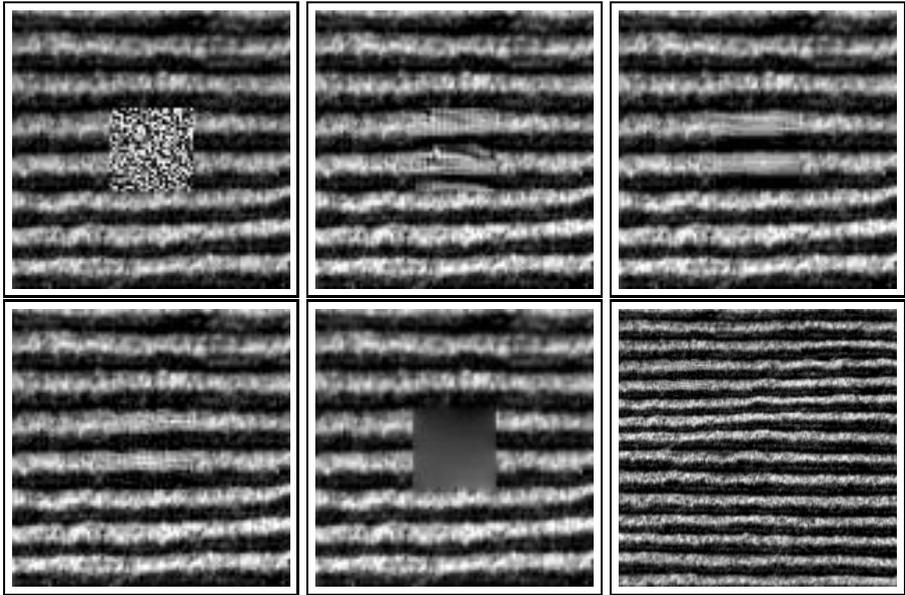
In the first step the geometric structure is reconstructed by finding a smooth MAP-like solution and in the second step the texture component is reconstructed by adding it to the large scale geometry.

In this paper we propose a novel variation of the above discussed method. We consider two cooling schemes and a gradual heating scheme which can be considered as the inverse of simulated annealing.

### 3.2 Cooling - The ICM Solution

Finding the MAP solution by simulated annealing is very time consuming. One alternative method is the Iterated Conditional Modes (ICM) algorithm. By letting  $T \rightarrow 0$  (or equivalently letting  $\beta \rightarrow \infty$ ) the conditional distribution (3) will become a point distribution. In each step of the Gibbs sampling one will set the new intensity for a site  $(x, y)_i$  to

$$I_i^{\text{new}} = \arg \max_{I_i} p(I_i | I_{N_i}). \quad (5)$$



**Fig. 1.** From top left to bottom right: a) the image containing a damaged region b) the ICM solution c) the fast cooling solution d) adding texture on top of the fast cooling solution by heating the distribution e) total variation (TV) solution and f) the reconstructed region in context (can you find it?)

This is a site-wise MAP solution (i.e. in each site and in each step the most likely intensity will be selected). This site-wise greedy strategy is not guaranteed to find the global MAP solution for the full image. The ICM solution is similar but not identical to the high  $\beta$  sampling step described in [6]. The ICM solution depends on initialisation of the unknown region  $\Omega$ . Here we initialise by sampling pixel values identically and independent from a uniform distribution on the intensity range.

### 3.3 Cooling - Fast Cooling Solution

The MAP solution for the inpainting is the most likely reconstruction given the known part of the image  $D \setminus \Omega$ ,

$$I^{\text{MAP}} = \arg \max_{I_{i \forall (x,y), i \in \Omega}} p(I \mid I(D \setminus \Omega), \Lambda) . \tag{6}$$

Simulated annealing can be used for finding the MAP solution. Replacing  $\beta$  in (4) with an increasing (decreasing) sequence  $\beta_n$  called a cooling (heating) scheme. Using simulated annealing one starts to sample using a high temperature  $T$  and slowly cooling down the distribution (4) by letting  $T \rightarrow 0$ . If  $\beta_n$

is increasing slowly enough and letting  $n \rightarrow \infty$  then simulated annealing will find the MAP solution ( see e.g. [9,10,14]). Unfortunately simulated annealing is very time consuming.

To reconstruct  $\Omega$ , the true MAP solution may not be needed, instead a solution which is close to the MAP solution may be enough. We therefore adopt a fast cooling scheme, that does not guarantee the MAP solution. The goal is to reconstruct the geometric structure of the image and suppress the stochastic texture.

The fast cooling scheme used in this paper is defined as (in terms of  $\beta$ )

$$\beta_{n+1} = C^+ \cdot \beta_n \quad (7)$$

where  $C^+ > 1.0$  and  $\beta_0 = 0.5$ .

### 3.4 Heating - Adding Texture

The geometric structures of the image will be reconstructed by sampling using the cooling scheme. Unfortunately the visual appearance will be too smooth, and the stochastic part of the image needs to be added.

The stochastic part should be added in such a way that it does not destroy the large scale geometric part reconstructed in the previous step. This is done by sampling from the distribution (4) using a heating scheme similar to the cooling scheme presented in previous section and using the solution from the cooling scheme as initialisation.

The heating scheme in this paper is

$$\beta_{n+1} = C^- \cdot \beta_n \quad (8)$$

where  $C^- < 1.0$  and  $\beta_0 = 25$ .

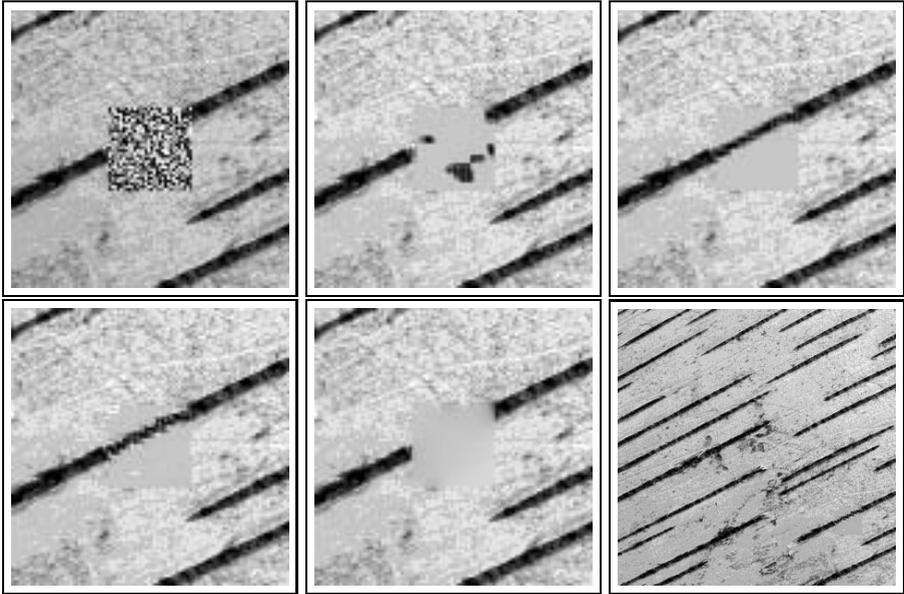
By using a decreasing  $\beta_n$ , value finer details in the texture will be reproduced, while coarser details in the texture will be suppressed.

## 4 Results

Learning the FRAME model  $p(I)$  is computational expensive, therefore only small image patches have been used. Even for small image patches the optimisation times are at least a few days. After the FRAME model has been learnt, inpainting can be done relatively fast if  $\Omega$  is not too large.

The dynamic range of the images have been decreased to 11 intensity levels for computational reasons. The images that have been selected includes both large scale geometric structures as well as texture.

The delta filter, 16 Gabor filters and 18 scale space derivative filters have been used in all experiments and 11 histogram bins have been used for all filters (see section 2.1 for a discussion).



**Fig. 2.** From top left to bottom right: a) the image containing a damaged region b) the ICM solution c) the fast cooling solution d) adding texture on top of the fast cooling solution by heating the distribution e) total variation (TV) solution and f) the reconstructed region in context (can you find it?)

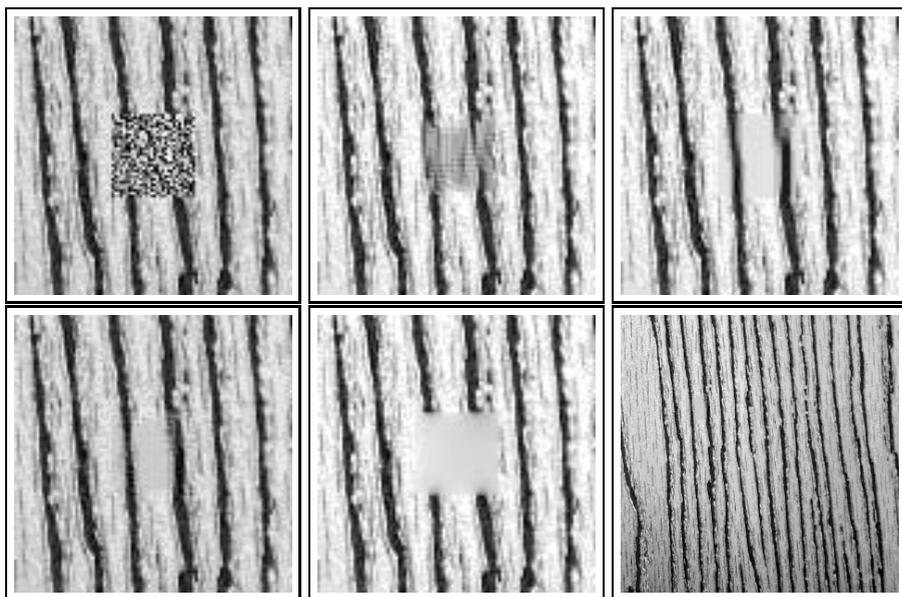
In the cooling scheme (7), we use  $\beta_0 = 0.5$ ,  $C^+ = 1.2$  and the stopping criterion  $\beta_n > 25$  in all experiments. In the heating scheme (8), we use  $\beta_0 = 25$ ,  $C^- = 0.8$  and the stopping criterion  $\beta_n < 1.0$ .

Each figure contains an unknown region  $\Omega$  of size  $30 \times 30$  that should be reconstructed. Figure 1 contains corduroy images, figure 2 contains birch bark images and figure 3 wood images. Each figure contains the original image with the damaged region  $\Omega$  with initial noise, the ICM and fast cooling solutions and the solution of a total variation (TV) based approach [1] for comparison.

The ICM solution reconstruct the geometric structure in the corduroy, but fails to reconstruct the geometric structure in both the birch and the wood images. This is due to the local update strategy of ICM, which makes it very sensitive to initial conditions. If ICM starts to produce wrong large scale geometric structures it will never recover.

The fast cooling solution on the other hand seem to reconstruct the geometric structure in all examples and does an even better job than the ICM solution for the corduroy image. The fast cooling solutions are smooth and have suppressed the stochastic textures. Because of the failure of ICM we only include results on heating based on the fast cooling solution.

The results - image d) - after the heating are less smooth  $\Omega$ 's, but it is still smoother than  $I \setminus \Omega$ . The total variation (TV) approach produce a too smooth solution even if strong geometric structures are present in all example.



**Fig. 3.** From top left to bottom right: a) the image containing a damaged region b) the ICM solution c) the fast cooling solution d) adding texture on top of the fast cooling solution by heating the distribution e) total variation (TV) solution and f) the reconstructed region in context (can you find it?)

## 5 Conclusion

Using FRAME to learn a probability distribution for a type of images gives a Gibbs distribution. The boundary condition makes it hard to use the learnt Gibbs distribution as it is for inpainting; it does not enforce large scale geometric structures strongly enough. By using a fast cooling scheme a MAP-like solution is found that reconstruct the geometric structure. Unfortunately this solution is too smooth and does not contain the stochastic texture. The stochastic texture component can be reproduced by sampling using a heating scheme. The heating scheme adds the stochastic texture component to the reconstruction and decrease the smoothness of the reconstruction based on the fast cooling solution.

A possible continuation of this approach is to replace the MAP-like step with a partial differential equation based method and a natural choice is the Gibbs Reaction And Diffusion Equations (GRADE) [15,16], which are build on the FRAME model.

We decompose an image into a geometric component and a stochastic component and use the decomposition for inpainting. This is related to Meyer's [17,18] image decomposition into a smooth component and a oscillating component (belonging to different function spaces). We find it interesting to explore this theoretic connection with variational approaches.

## Acknowledgements

This work was supported by the Marie Curie Research Training Network: Visiontrain (MRTN-CT-2004-005439).

## References

1. Chan, T.F., Shen, J.: Variational image inpainting. *Communications on Pure and Applied Mathematics* vol. 58 (2005)
2. Efros, A.A., Freeman, W.T.: Image quilting for texture synthesis and transfer. In: *Proceedings of SIGGRAPH '01, Los Angeles, California, USA* (2001)
3. Efros, A.A., Leung, T.K.: Texture synthesis by non-parametric sampling. In: *IEEE International Conference on Computer Vision, Corfu, Greece*, pp. 1033–1038 (1999)
4. Bonet, J.S.D.: Multiresolution sampling procedure for analysis and synthesis of texture images. In: *Computer Graphics, ACM SIGGRAPH*, pp. 361–368 (1997)
5. Bertalmio, M., Vese, L., Sapiro, G., Osher, S.: Simultaneous structure and texture image inpainting. *IEEE Transactions On Image Processing* 12(8), 882–889 (2003)
6. Gustavsson, D., Pedersen, K.S., Nielsen, M.: Geometric and texture inpainting by gibbs-sampling. In: *SSBA07* (2007)
7. Zhu, S.C., Wu, Y.N., Mumford, D.: Filters, random fields and maximum entropy (frame): To a unified theory for texture modeling. *International Journal of Computer Vision* 27(2), 107–126 (1998)
8. Zhu, S.C., Wu, Y.N., Mumford, D.: Minimax entropy principle and its application to texture modelling. *Neural Computation* 9(8), 1627–1660 (1997)
9. Geman, S., Geman, D.: Stochastic relaxation, gibbs distribution, and the bayesian restoration of images. *IEEE Transaction PAMI* 6, 721–741 (1984)
10. Winkler, G.: *Image Analysis, Random Fields, and Markov Chain Monte Carlo Methods*. In: *Number 27 in Stochastic Modelling and Applied Probability*, Springer, Heidelberg (2006)
11. Bigun, J.: *Vision with Direction - A Systematic Introduction to Image Processing and Computer Vision*. Springer, Heidelberg (2006)
12. Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using gabor filters. *Pattern Recogn.* 24(12), 1167–1186 (1991)
13. ter Haar Romeny, B.M.: *Front-End Vision and Multi-Scale Image Analysis: Multi-Scale Computer Vision Theory and Applications*, written in Mathematica. *Computational Imaging and Vision*, vol. 27. Kluwer Academic Publishers, Dordrecht (2003)
14. Liu, J.S.: *Monte Carlo Strategies in Scientific Computing*. Springer Series in Statistics. Springer, Heidelberg (2004)
15. Zhu, S.C., Mumford, D.: Prior learning and gibbs reaction-diffusion. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 19(11), 1236–1250 (1997)
16. Zhu, S.C., Mumford, D.: Grade: Gibbs reaction and diffusion equation - a framework for pattern synthesis, denoising, image enhancement, and clutter removal. *IEEE Trans. PAMI* 19(11), 1627–1660 (1997)
17. Meyer, Y.: *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. American Mathematical Society (AMS), Boston, MA, USA (2001)
18. Aujol, J.F., Gilboa, G., Chan, T., Osher, S.: Structure-texture image decomposition — modeling, algorithms, and parameter selection. *International Journal of Computer Vision* 67(1), 111–136 (2006)