

IMAGE SYNTHESIS USING CONDITIONAL RANDOM FIELDS

E. Ahmadi[†], Z. Azimifar[‡], P. Fieguth[‡], Sh. Ayatollahi^{*}

[†]School of Electrical & Computer Engineering, Shiraz University, Shiraz, Iran

[‡]Systems Design Engineering, University of Waterloo, Waterloo, Canada

^{*}School of Petroleum & Chemical Engineering, Shiraz University, Shiraz, Iran
{e_ahmadi},{azimifar}@cse.shirazu.ac.ir

ABSTRACT

Methods of scientific imaging and image analysis have become pervasive in a great variety of fields, including the properties of porous media. To study the large-scale morphological properties of porous media, high resolution random (Monte Carlo) samples are required. The purpose of this paper is to propose a novel approach for the statistical synthesis of scientific images, based on the concept of Conditional Random Fields. We explore two different sets of potential functions are used to model the pore-structure characteristics, and Monte Carlo Markov chain methods are also used to sample the high resolution images from the trained model. The resulting images are of high quality, and show the performance of the proposed framework.

Index Terms— Graphical Models, Conditional Random Fields, Image Synthesis, Porous Media, Image Sampling

1. INTRODUCTION

Scientific images and image models have played a significant role in a wide variety of research fields, including chemistry, medicine and astrophysics. The effectiveness of these models has been validated and supported through sophisticated imaging tools including MRI, CT and ultrasound. A host of computer vision and image processing techniques (super-resolution, denoising, registration) have been developed to solve imaging problems.

One such field that has taken advantage of computer vision is porous media, the science of water-porous materials like wood, bone, cement, rocks and soil [1], of which two examples are shown in Fig. 1. These media have great significance in the fields of construction, medicine, petroleum engineering and environment. In order to study the permeability, porosity, transport properties of porous media, high resolution images are required [2]. Such images may be acquired using current imaging tools, however the viewing of interior parts of the medium requires cutting, polishing, and exposure to air, all of which act to alter the sample. On the other hand, 3D MRI images offer non-invasive 3D imaging, but have limited spatial resolution which resolve only the largest pores.



Fig. 1. Two examples of high resolution porous media images.

A possible alternative to acquiring high-resolution images, avoiding the problems of cutting and polishing, is based on image synthesis, the process of randomly creating new images from some form of image description or model.

Some past work has addressed synthesis of porous media images. Mohebi et. al. [3] used the posterior sampling framework proposed by Fieguth [4] to model high resolution images using a Gibbs Random Field (GRF) and Monte Carlo Markov Chain (MCMC) methods. To take advantage of the information present in a low-resolution image (e.g., MRI), the Gibbs model includes constraints based on the measurements.

Conditional Random Fields (CRF) are a state-of-the-art graphical model, extensively used in recent research [5]. The CRF as a discriminative model has reduced the complexity of joint probability modeling of generative models like Markov Random Fields (MRF). The CRF has outperformed previous graphical models in most machine learning and computer vision tasks [5, 6].

To meet our needs in the reconstruction of porous media images, a novel statistical approach based on CRF is proposed in this paper. The CRF is trained using a set consisting of high- and low-resolution images, such that for any presented low-resolution test sample the CRF will produce a high resolution image.

2. CONDITIONAL RANDOM FIELDS

Conditional random fields were originally proposed by Lafferty et. al. [6] for the segmenting and labeling of sequence data. The key concept is that the CRF *directly* models the *conditional* probability distribution, and not the *joint* distribution as in MRFs. Here follows a brief definition of the CRF.

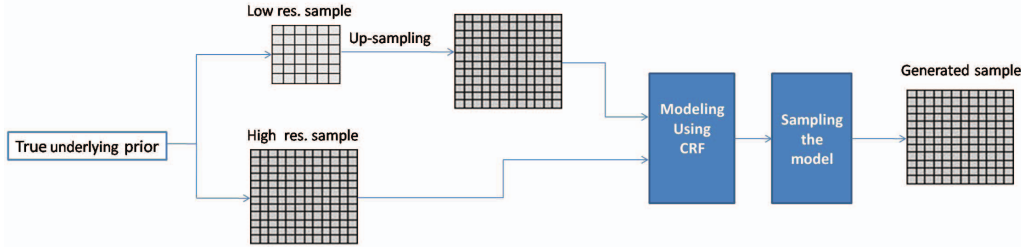


Fig. 2. Flow diagram of proposed framework: low and high resolution images are assumed to be statistically equivalent. The high resolution and cubic-interpolation upsampled images are presented to the CRF for training. The Gibbs sampler, right, generates new high samples from the constructed model.

Let $G = (S, E)$ be a graph such that Y is indexed by the vertices of G . Then (Y, X) is said to be a conditional random field if, when conditioned on X , the random variables Y_i obey the Markov property with respect to the graph: $P(Y_i|X, Y_{S-\{i\}}) = P(Y_i|X, Y_{N_i})$, where $S - \{i\}$ is the set of all nodes in G except i , N_i is the set of neighbors of the node i in G [5, 7]. Thus, CRF is a random field globally conditioned on observation X . The conditional distribution is defined as:

$$P(Y|X) = \frac{1}{Z(X)} \prod_{C_p \in C} \prod_{\Psi_c \in C_p} \exp\left(\sum_{k=1}^{K(p)} \lambda_{pk} f_{pk}(Y_c, X)\right)$$

$$Z(X) = \sum_Y \prod_{C_p \in C} \prod_{\Psi_c \in C_p} \exp\left(\sum_{k=1}^{K(p)} \lambda_{pk} f_{pk}(Y_c, X)\right) \quad (1)$$

where C_p is a clique template from the clique set C specified by the structure of the graph G , Ψ_c indicates a clique of type C_p , f_{pk} and λ_{pk} represent the k^{th} real-valued feature function defined on clique template C_p and the corresponding model parameter, respectively. The number of feature functions defined on clique template C_p is determined by $K(p)$. $Z(X)$ is a normalizing constant — the partition function — with respect to input X . The parameter λ_{pk} is the same for all cliques Ψ_c belonging to a common clique template C_p , an approach known as parameter tying.

There are different methods to estimate the parameters of CRF model [8]. Maximum Likelihood Estimation (MLE) approach is usually used for learning the parameters of the model. However, MLE requires the computation of the partition function $Z(X)$ which is a NP-hard problem. In this paper, pseudo-likelihood [8] is used for parameter estimation as an approximation of MLE. Assuming identically independent training data, the conditional log-likelihood is given by

$$\ell(\lambda) = \sum_{C_p \in C} \sum_{\Psi_c \in C_p} \sum_{k=1}^{K(p)} \lambda_{pk} f_{pk}(Y_c, X) - \log(Z(X)) \quad (2)$$

Because the log-likelihood function $\ell(\lambda)$ is concave, the parameters λ can be chosen so that the global maximum is obtained and the gradient or vector of partial derivatives with respect to each parameter λ_{pk} is zero. Differentiating $\ell(\lambda)$

Algorithm 1 : The sampling procedure

```

initialize a random image  $Y$ 
set the temperature  $T$  as  $T_0$ 
while stopping conditions are not met do
  decreasing the  $T$  slowly
  for each site  $s$  in  $Y$  do
     $y_s \leftarrow$  pick a sample from distribution  $P(y_s|y_{N_s}, x)$ 
    using Gibbs sampler
  end for
end while

```

with respect to parameter λ_{pk} gives

$$\frac{\partial \ell}{\partial \lambda_{pk}} = \sum_{\Psi_c \in C_p} f_{pk}(Y_c, X) - \sum_{\Psi_c \in C_p} \sum_{Y'_c} f_{pk}(Y'_c, X) P(Y'_c|X) \quad (3)$$

As the exact, analytical solution does not exist, the parameter is iteratively found by maximizing the log-likelihood function using gradient descent.

Given the test X the task of assigning labels to its pixels is inference, where the goal of inference is to find the maximum a posteriori (MAP) solution $Y^* = \operatorname{argmax}_Y P(Y|X)$. Instead of using the CRF method of inference, which is to some extent biased towards classification, here we propose to use a traditional sampling technique, such as the Gibbs sampler, for inference.

Once our conditional model is constructed and trained, we use the well-known Gibbs sampler [9], based on MCMC methods, to generate a sample from the modeled distribution [4]. In order to generate a probable sample, the Gibbs sampler is used along with simulated annealing, in which the sampler begins with a high temperature T_0 , then repeatedly scanning all of the lattice sites and sampling from the conditional $P(y_s|y_{N_s}, x)$, gradually decreasing the temperature. As is quite standard for simulated annealing, we use a geometric cooling schedule [10]:

$$T_{n+1} = \alpha^n T_0, 0 < \alpha < 1 \quad (4)$$

The pseudo-code of the sampling phase is depicted in Algorithm 1.

3. THE PROPOSED FRAMEWORK

Because of measurement cost, time, and complexity, only a few training samples are available. For simplicity, we assume that the given images are normalized, such that we have one high resolution image H with pixel values in $\{-1, 1\}$ and one low resolution image L with intensities distributed over the interval $[-1, 1]$.

Porous media images are binary images representing pores (black pixels) and solids (white pixels). Our goal is to generate new samples with the same properties of H and L , simultaneously. A CRF is used in order to construct a model that characterizes the medium. From the Hammersley-Clifford theorem, the conditional distribution of H given L is defined as

$$P(H|L) = \frac{1}{Z(L)} \exp\left(\sum_i U_i(H_i, L) + \sum_i \sum_{j \in N_i} I_{ij}(H_i, H_j, L)\right) \quad (5)$$

where U_i is the unary potential associated with single site H_i , conditioned on L , whereas I_{ij} is the potential function, describing the conditional interaction between the neighboring sites H_i and H_j .¹ As the porous media images under consideration are homogeneous and isotropic, functions U_i and I_{ij} not a function of location ij therefore, with slight abuse of notation, we will write just U and I . Thus two clique templates are determined for the random field.

The clique potentials specify how local variables interact and how much the interaction contributes to the global distribution. Here, we propose two types of potential functions: one similar to the Ising model, used for segmentation in [5], a second based on a chordlength distribution.

3.1. Ising-directed potential functions

The Ising model is a popular binary model which is extensively used for modeling random fields [11]. The unary potential $U(H_i, L)$ is defined as

$$U(H_i, L) = \lambda_0 + \lambda_1 h_i l_i \quad (6)$$

where $\{\lambda_0, \lambda_1\}$ are the model parameters and h_i and l_i represent the pixel values corresponding to H_i and L_i , respectively. The potential function U specifies the properties of each pixel of image to be either pore or solid, ignoring the effects of neighboring pixels.

In the same way, the interaction potential function $I(H_i, H_j, L)$ can be written as:

$$I(H_i, H_j, L) = \lambda_2 + \lambda_3 h_i h_j |l_i - l_j| \quad (7)$$

where $\{\lambda_2, \lambda_3\}$ are the parameters of the CRF with respect to function I . It is expected that neighboring pixels have the same behavior and tend to have the same value, clearly with exceptions at an edge, where the neighboring sites are different.

¹A first-order neighborhood system is used so that sites $\{(k-1, l), (k+1, l), (k, l-1), (k, l+1)\}$ are the neighbors of (k, l) .

	Image 1	Image 2	Image 3
Ising PF	22%	17%	10%
Chordlength PF	11%	12%	6%

Table 1. The fraction of incorrectly reconstructed samples, inferred from the MSE between truth and reconstruction, for the three sample images of Fig. 3. The synthetic images are sampled using CRF models with two different proposed potential functions (PF): Ising and Chordlength.

3.2. Chord-directed potential functions

The chordlength distribution has been widely used for the characterization of pore structures in porous media [10, 3]. Motivated by the success of the chordlength approach in modeling, we propose unary and interaction potential functions as follows:

$$U(H_i, L) = \lambda_4 + \lambda_5 h_i l_i C_i$$

$$I(H_i, H_j, L) = \lambda_6 + \lambda_7 h_i h_j (l_i + l_j)(C_i + C_j) \quad (8)$$

where C_i and C_j are the relative lengths of chords corresponding to sites i and j , respectively. The relative chord associated with a general site s is computed as

$$C_s = \frac{C_{v,s}}{height} + \frac{C_{h,s}}{width} \quad (9)$$

where $C_{v,s}$ is the length of line segment starting at site s and is extracted vertically and downward to the first dissimilar site, based on pixel value (-1 or 1). In the same way, $C_{h,s}$ is the length of horizontal chord, and *height* and *width* normalize the chords to the size of the entire image.

The chordlength is far more discriminating than the Ising model, since the lengths and orientations of chords can say a great deal about the structure of pores and solids in a high resolution image, such that it is likely that images with long/short chords have large/chaotic pore structure.

The model is trained using images H and L , as described in Sec. 2. Finally, given a low resolution image, a high resolution image is sampled by running the Gibbs sampler. The generated images from the sampling phase obey the model and also well describe the porous media structures. The overall process of the proposed method is depicted in Fig. 2.

4. EXPERIMENTAL RESULTS

In order to evaluate the proposed method, we have applied it to three samples of porous media, with the underlying high resolution images having sizes of 128×128 . The selected porous media images were deliberately selected to have different pore structures and in order to demonstrate the robustness of our proposed model.

The high-resolution images are actual microscopic images, however due to the cost and complexity of acquiring measurements from multiple instruments, the low resolution

images are generated by down-sampling the corresponding high resolution samples, here down-sampled by simple averaging by a factor of $d = 8$. Clearly the choice of d affects the amount of information remaining in L to guide the reconstruction of H . The high resolution and down-sampled images are shown in first two rows of Fig. 3.

For each image pair we have performed to reconstruction experiments, based on the two proposed potential functions. Given a low resolution sample, the high resolution one is sampled based on the sampling method developed in Sec. 2. The low resolution sample used for training is the same one given to the model for sampling.

The reconstructed images are shown in Fig. 3. A visual inspection of the resulting images proves the success of the proposed method. Furthermore, the effect of the CRF potential function can be better understood, in comparing the poorer reconstructions of the Ising model, a weak local model with very limited descriptive power, relative to the fairly detailed reconstructions returned by the chordlength model. In many cases details not visible in L are palusibly reconstructed in H .

To supplement the visual evaluation of the results, the Mean Squared Error (MSEs) between the ground-truth H and reconstructed image R can be measured:

$$MSE(H, R) = \frac{1}{S} \sum_s (h_s - r_s)^2 \quad (10)$$

where S measures the number of pixels in H or R , and h and r are the pixel intensities in H and R , respectively. Because the images are binary, a given reconstructed pixel is either correct, $(h - r)^2 = 0$, or incorrect, $(h - r)^2 = 4$. Therefore $MSE/4$ directly measures the fraction of pixels which are reconstructed incorrectly, as reported and quantified in Table 1.

5. CONCLUSIONS

In this paper, we proposed a graphical random field-based model for the characterization of porous media images. The model is trained using high and low resolution images. The well-known Gibbs sampler is used to generate the artificial samples from the model. The introduced energy functions are simple, easy to interpret and descriptive. The experimental results illustrate the success of the approach, in that pore and solid structures are reconstructed from low-resolution inputs.

6. REFERENCES

- [1] A. Kaestner, E. Lehmann, and M. Stampanoni, "Imaging and image processing in porous media research," *Advances in Water Resources*, 2008.
- [2] A. Mohebi and P. Fieguth, "Posterior sampling of scientific images," *ICIAR*, 2006.
- [3] A. Mohebi, P. Fieguth, and M. A. Ioannidis, "Statistical fusion of two-scale images of porous media," *Advances in Water Resources*, 2009.
- [4] P. Fieguth, "Hierarchical posterior sampling for images and random fields," *ICIP*, 2003.
- [5] S. Kumar and M. Hebert, "Discriminative fields for modeling spatial dependencies in natural images," *NIPS*, 2003.
- [6] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," *Proc. ICML*, 2001.
- [7] R. Klinger and K. Tomanek, "Classical probabilistic models and conditional random fields," Dortmund University of Technology, 2007.
- [8] S. Vishwanathan, N. Schraudolph, M. Schmidt, and K. Murphy, "Accelerated training of conditional random fields with stochastic meta-descent," *ICML*, 2006.
- [9] S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution and the bayesian restoration of images," *IEEE Trans. PAMI*, 1984.
- [10] S. K. Alexander, P. W. Fieguth, and E. R. Vrscay, "Hierarchical annealing for random image synthesis," *EMM-CVPR*, 2003.
- [11] G. Winkler, "Image analysis, random fields, and markov chain monte calro methods," Springer-Verlag, 2003.

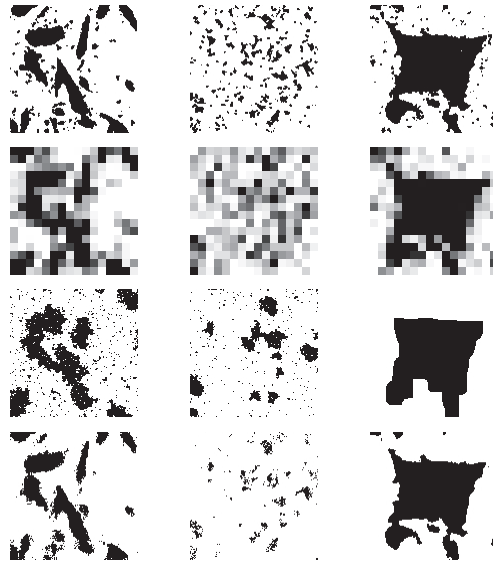


Fig. 3. The ground-truth images are shown in the first row. The second row contains generated artificial samples by down-sampling. The third and fourth rows correspond to reconstructed samples using Ising-directed and chord-directed potential functions, respectively.