Exercise 3: Stochastic sampling (Due 16/5/2017)*

Statistical Methods in Image Processing 048926

Analytic Exercises (15 points)

The Metropolis Sampling algorithm can be used to draw samples from a Gibbs distribution $p(x) = \frac{1}{z}\exp\{-u(x)\}$.

Metropolis Algorithm:

Initialize X^0

For $k = 0$ to $K-1$ {

/*Generate proposal*/

$W \sim q(w|X^k)$ /*Use proposal distribution to generate W*/

/*Accept or reject proposal*/

$\alpha \leftarrow \min\left\{1, e^{-[u(W)-u(X^k)]}\right\}$ /*Compute acceptance probability*/

$X^{k+1} \leftarrow \begin{cases} W & \text{with probability } \alpha \\ X^k & \text{with probability } 1-\alpha \end{cases}$

}

1. Let $W = X_k + Z$ where Z is independent of X_k and has a symmetric density $p_z(z)$. Show that the proposal density $q(w|x)$ obeys the symmetry condition $q(w|x) = q(x|w)$, which is required in order for the metropolis sampler to converge to the distribution of X.

2. Assume that X is discrete and denote the probability of accepting the proposal at the kth step by $\alpha(w|x^k) = \min\{1, \exp\{-[u(w)-u(x_k)]\}\}$. Show that the transition probabilities for the Metropolis algorithm are given by

$$p(x^{k+1}|x^k) = \begin{cases} q(x^{k+1}|x^k)\alpha(x^{k+1}|x^k) & \text{if } x^{k+1} \neq x^k, \\ 1 - \sum_{w \neq x^k} q(w|x^k)\alpha(w|x^k) & \text{if } x^{k+1} = x^k. \end{cases}$$

*Please send your solutions to Tamar

Matlab Exercises (85 points)

1) Gibbs Sampler

(a) The Ising model is a pairwise MRF model for **binary images** with 4-connected neighborhoods, defined as

$$p(X) \propto \exp\left\{-\beta \sum_{\{r,s\}\in\mathcal{C}} \psi_c(x_r, x_s)\right\},$$

where

$$\psi_c(x_r, x_s) = \delta(x_s \neq x_r) = \begin{cases} 1 & \text{if } \quad x_r \neq x_s, \\ 0 & \text{if } \quad x_r = x_s. \end{cases}$$

Write an explicit expression for the conditional distribution of the ith pixel given the rest of the image, $p(X_i = x_i | X_{V\setminus i})$.

(b) Implement a function for drawing the i-th pixel given the rest of the image according to the Ising model.

(c) Implement a function for drawing an image X of size 50×50 using the Gibbs sampling method:

i. Initialize X to be a random binary image.

ii. Select ordering of pixels.

iii. Update the image X^k pixel-wise. Draw each pixel according to the Ising distribution using the function from **1**)b.

iv. Repeat for $K = 100$ times. Plot the current image every 10 iterations.

For the Gibbs distribution, the boundaries of the image have a significant impact. Examine the effect of padding the image with zeros, ones, and circular padding. Draw an image using each of the above, with $\beta = 0, 0.25, 1, 2$. Discuss the effect of β and the padding values on the characteristics of the resulting images.

(d) Draw again images using this model, this time initialize X to be all zeros. How does the initialization influence the drawing process?

2) Metropolis Sampler

The attachment of this exercise includes the parameters of two trained MRF image models: one with product of student-t cliques, as suggested in [1], and one with Gaussian Scale Mixture (GSM) cliques, as suggested in [2]. Each student-t expert and each Gaussian in the mixture is associated with a 3×3 filter and several additional parameters, which are all supplied in the mat file. Please see the appendix file for detailed explanations.

(a) Implement a function for drawing a variable w_i from a normal distribution with standard deviation σ_w and mean x_i:

$$q(w_i|x_i) = \frac{1}{2\pi\sigma_w^2} \exp\left\{\frac{1}{2\sigma_w^2}(w_i - x_i)^2\right\}.$$

(b) We would like to determine whether w_i is a more probable value for the i-th pixel of an image X, than the original pixel value x_i. How can we do this for a Gibbs distribution $p(x) \propto \exp\{-\sum_c \psi(x_c; \theta)\}$?

(c) Implement a function for calculating the probability density of the i-th pixel to equal x_i. Use the student-t model with the attached filters and parameters:

$$p(x) \propto \prod_c \phi_c(x_c; \theta),$$

where

$$\phi_c(x_c; \theta) = \prod_{m=1}^{M} (1 + \frac{1}{2}(J_m^T x_c)^2)^{-\alpha_m}.$$

(You will not need to calculate the product over all the cliques c. Why?)

(d) Implement a function for calculating the probability density of the i-th pixel to equal x_i. Use the GSM model with the attached filters and parameters:

$$p(x) \propto \prod_c \phi(x_c; \theta),$$

where

$$\phi_c(x_c; \theta) = \prod_{m=1}^{M} \sum_{n=1}^{N} \alpha_{mn} \frac{1}{2\pi s_n} \exp\left\{\frac{1}{2s_n}(J_m^T x_c)^2\right\},$$

(Again, calculating the product for all the cliques c is not needed)

(e) Use the Metropolis sampler to draw an image X of size 50×50 from the MRF model with the product of student-t cliques:

i. Initialize X to be random.

ii. Select ordering of pixels.

iii. Update the image X^k pixel-wise.

A. for each pixel x_i draw $w_i \sim \mathcal{N}(x_i, \sigma_w^2)$

B. if w_i is more probable then replace x_i with w_i

C. else, replace x_i with w_i with probability $\alpha = \frac{p(w_i)}{p(x_i)}$

iv. Repeat $K = 100$ times. Plot the image every 10 iterations.

Use zero padding at the boundaries.

(f) Use the Metropolis sampler to draw an image X of size 50×50 from the MRF model with the GSM cliques. Use $K = 100$ iterations. Plot the image every 10 iterations.

(g) Draw again images using these two models, with two different σ_w values. How does σ_w influence the process?

(h) Draw again images using these two models, but this time initialize X to be zeros. How does the initialization influence the process?

(i) Compare and discuss the results of the two models.

References

[1] Roth, S., Black, M.J.: Fields of experts: A framework for learning image priors. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume 2. (2005) 860–867

[2] Schmidt, U., Gao, Q., Roth, S.: A Generative perspective on MRFs in low-level vision. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE (2010) 1751–1758