Statistical comparison of images using Gibbs random fields

Bruno Rémillard * and Christian Beaudoin
Laboratoire de Recherche en Proababilités et Statistique
Université du Québec à Trois-Rivières
C.P. 500, Trois-Rivières (Québec), Canada, G9A 5H7
bruno_remillard@uqtr.uquebec.ca christian_beaudoin@uqtr.uquebec.ca

Abstract

The statistical tools used to compare images are often based on multivariate statistics and use images as observations. This approach does not take into account the spatial and dependence structures in the images. Here we follow another approach in the sense that the number of pixels in the images is the number of observations. We also take into account spatial and dependence structures by modelling images with Gibbs random fields with nearest-neighbors interactions. The idea of the proposed method of comparison of images is to model the images by Gibbs random fields with a finite number of parameters and to test equality of parameters between the pooled images of each group. We obtain a test statistic that is easy to calculate, and the limiting distribution of the statistic is a chi-square distribution. Applications to handwrittens signatures and pattern recognition are presented.

1 Introduction

In many applications one wants to compare images of the same object. For example, one is interested in comparing fingerprints, handwritten signatures, radar images, etc.

The statistical tools used to compare images are often based on multivariate statistics and use images as observations. This approach does not take into account the spatial and dependence structures in the images. Here we follow another approach in the sense that the number of pixels in the images is the number of observations. We also take into account spatial and dependence structures by modelling images with the so-called (finite) Gibbs random fields with nearest-neighbors interactions. These models are very popular in Statistical Physics. They have also been used by many authors for denoising images using a Bayesian approach; see for example [2] or [3].

*Work supported in part by the Fonds Institutionnel de Recherche and by the Natural Sciences and Engineering Research Council of Canada, Grants No. OGP0042137.

In what follows we assume that the images are comparable in the sense that they have the same size and the same orientations (e.g. [1]). The idea of the proposed method of comparison of images is to model the images by Gibbs random fields with a finite number of parameters and to test equality of parameters between the pooled images of each group.

In Section 2, images models with an arbitrary number of gray levels are introduced. Section 3 deals with methods for estimating the parameters of these images and an appropriate test statistic is defined. Section 4 is devoted to examples of applications for binary images, while the Appendix contains the proof of the asymptotic behavior of the proposed test statistic together with efficient Monte-Carlo Markov Chain methods for generating images.

2 Gibbs random fields

2.1 Notations

The set of pixels will be denoted by Λ. It is assumed that Λ is a rectangle of $\mathbb{Z}^2$. The "color" of the pixel at site x will be denoted by $\sigma(x) \in C \subset \{0,1,\ldots\}$, $x \in \Lambda$, where a white pixel has value 0, and a black pixel has value 1. Set $\sigma_A = C^A = \{\sigma(x); x \in A\}$. For example, for black and white images or binary images, $C = \{0,1\}$. For gray level images, $C = \{0,1,\ldots g-1\}$ and g is the number of gray levels. For sake of simplicity, extend every σ so that $\sigma(x) = 0$ whenever $x \notin \Lambda$.

If one wants to use statistics, one has to define an interesting parameterized family of distributions on the set of all possible images $\Omega = C^\Lambda$. The desirable features of random images are

- relatively few parameters;
- "Markov property": the law of the color of a pixel at site x given all other pixels should depend only on the colors of the pixels in a given neighborhood N_x of x;
- "stationarity": $N_x = x + N$, for any $x \in \Lambda$;

- each possible image has a positive probability.

One such family is given by Gibbs random fields introduced in the next subsection.

2.2 Gibbs random fields

Throughout the rest of the paper let $\pm e_1$, $\pm e_2$, $\ldots, \pm e_m$ be fixed vectors with integer components. Let $\alpha = (\alpha_1, \ldots, \alpha_{g-1})$ and $\beta = (\beta_1, \ldots, \beta_m)$. Set $\alpha_0 = 0$. These vectors will serve to define the neighbors of a pixel and α and β will be parameters of the family of laws for the images.

For every (α, β), the probability of an image $\sigma = \sigma_\Lambda$ is defined by

$$P_{\Lambda,\alpha,\beta}(\sigma) = e^{H_{\Lambda,\alpha,\beta}(\sigma)} \Big/ Z_{\Lambda,\alpha,\beta},$$

where

$$H_{\Lambda,\alpha,\beta}(\sigma) = \sum_{x\in\Lambda} \alpha(\sigma(x)) + \sum_{l=1}^{m} \beta_l \sum_{x\in\Lambda} \sigma(x)\sigma(x+e_l),$$

and where $Z_{\Lambda,\alpha,\beta}$ is the normalizing constant defined by

$$Z_{\Lambda,\alpha,\beta} = \sum_{\omega\in\Omega} e^{H_{\Lambda,\alpha,\beta}(\omega)}.$$

Then one can check that for any finite subset A of Λ, one has

$$P_{\Lambda,\alpha,\beta}\left\{\sigma_A | \sigma_{\Lambda\setminus A}\right\} = e^{H_{\Lambda,\Lambda\setminus A}(\sigma_A)} \Big/ Z_{A,\sigma_{\Lambda\setminus A}},$$

where $Z_{A,\sigma_{\Lambda\setminus A}} = \sum_{\sigma_A\in\{0,1\}^A} e^{H_{A,\sigma_{\Lambda\setminus A}}(\sigma_A)}$, and

$$\begin{aligned} H_{A,\sigma_{\Lambda\setminus A}}(\sigma_A) &= \sum_{x\in A} \alpha(\sigma(x)) \\ &+ \sum_{l=1}^{m} \beta_l \sum_{\{x,x+e_l\}\subset A} \sigma(x)\sigma(x+e_l) \\ &+ \sum_{l=1}^{m} \beta_l \sum_{x\in A, x+e_l\in\Lambda\setminus A} \sigma(x)\sigma(x+e_l) \\ &+ \sum_{l=1}^{m} \beta_l \sum_{x\in A, x-e_l\in\Lambda\setminus A} \sigma(x)\sigma(x-e_l). \end{aligned}$$

In particular,
$P_{\sigma,\alpha,\beta,x}\{\omega(x) = j\}$

$$\begin{aligned} &= P_{\Lambda,\alpha,\beta}\{\omega(x) = j \mid \omega(y) = \sigma(y), \forall y \neq x\} \\ &= \frac{e^{\alpha_j + j\sum_{l=1}^{m} \beta_l v(\sigma,x,l)}}{\sum_{k=0}^{g-1} e^{\alpha_k + k\sum_{l=1}^{m} \beta_l v(\sigma,x,l)}}, \end{aligned} \qquad (2.1)$$

where $v(\sigma, x, l) = \sigma(x+e_l) + \sigma(x-e_l)$. The probabilities defined by (2.1) are called the local specifications of the Gibbs random field.

REMARK. The parameter β can be interpreted as a dependence parameter in the sense that if it equals zero, then the pixels at different sites are independent and

$$P\{\sigma(x) = j\} = e^{\alpha_j} \Big/ \left(\sum_{k=0}^{g-1} e^{\alpha_k}\right), \quad 0 \leq j \leq g-1.$$

3 Comparison of images

Using the Gibbs random fields models, we can compare images by first estimating the parameters of the respective groups and then perform an hypothesis test. Let us start with parameters estimation.

3.1 Parameters estimation

There are many methods that can be used to estimate the parameters of a statistical model, but the two most efficient ones are the maximum likelihood method and the maximum pseudo-likelihood method.

In our setting, the maximum likelihood method consists is finding the values (α, β) maximizing the probability $P_{\Lambda,\alpha,\beta}(\sigma)$, where σ is the observed image. Unfortunately since $Z_{\Lambda,\alpha,\beta}$ has not a closed form except in the independence case, it would take years to compute this constant for images having 10000 pixels. Therefore one cannot apply the maximum likelihood method directly. A similar method is proposed in [8]. To compute the constant term, the author relies on Monte-Carlo Markov Chain methods. It works as follows. Maximizing $P_{\Lambda,\alpha,\beta}(\sigma)$ is the same as maximizing $\log P_{\Lambda,\alpha,\beta}(\sigma)$. Hence using calculus it amounts to solve the following equations in (α, β):

$$\hat{p}_j = \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \mathbb{I}\{\sigma(x) = j\} = \frac{1}{|\Lambda|} \sum_{x\in\Lambda} P_{\Lambda,\alpha,\beta}\{\omega(x) = j\},$$

$1 \leq j \leq g-1$, and

$$\begin{aligned} \hat{\mu}_l &= \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \sigma(x)\sigma(x+e_l) \\ &= \frac{1}{|\Lambda|} \sum_{x\in\Lambda} E_{\Lambda,\alpha,\beta}\left(\omega(x)\omega(x+e_l)\right), \end{aligned}$$

$1 \leq l \leq m$.

Once these probabilitities and expectations are estimated using Monte-Carlo methods (see appendix), one can use recursive methods to solve the equations. However the calculation time is enormous. To overcome this difficulty, we propose to use the maximum pseudo-likelihood method. Instead of maximizing $P_{\Lambda,\alpha,\beta}(\sigma)$, one maximizes the product L of the local specifications, which is called the pseudo-likelihood, that is

$$L = L_\sigma(\alpha,\beta) = \prod_{x\in\Lambda} P_{\sigma,\alpha,\beta,x}\{\omega(x)=\sigma(x)\}. \quad (3.2)$$

Note that maximizing L is the same as maximizing

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_\sigma(\alpha,\beta) = \frac{1}{|\Lambda|}\log L \\ &= \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\log P_{\sigma,\alpha,\beta,x}\{\omega(x)=\sigma(x)\}. \quad (3.3)\end{aligned}$$

REMARK. Under the independence hypothesis, that is $\beta = 0$, the maximum likelihood and maximum pseudo-likelihood estimates of α are the same and are given by $\hat{\alpha}_j = \log(\hat{p}_j/\hat{p}_0)$, where $\hat{p}_j$ is the proportion of pixels having color j, $0 \le j \le g-1$.

Using calculus, it is easy to see that the values of $(\hat{\alpha},\hat{\beta})$ that maximize the pseudo-likelihood must satisfy the following equations:

$$\frac{\partial\mathcal{L}}{\partial\alpha_j} = \hat{p}_j - \frac{1}{|\Lambda|}\sum_{x\in\Lambda}P_{\sigma,\alpha,\beta,x}\{\omega(x)=j\} = 0, \quad (3.4)$$

$1 \le j \le g-1$, and

$$\frac{\partial\mathcal{L}}{\partial\beta_l} = m_l - \frac{1}{|\Lambda|}\sum_{x\in\Lambda}E_{\sigma,\alpha,\beta,x}\{\omega(x)\}v(\sigma,x,l) = 0, \quad (3.5)$$

where $m_l = \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sigma(x)v(\sigma,x,l)$, $1\le l\le m$.

Note that the Hessian matrix $H = H_\sigma(\alpha,\beta)$ of $\mathcal{L}$ satisfies $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^\top & H_{22}\end{pmatrix}$, where

$$\begin{aligned}(H_{11})_{ij} &= \frac{\partial^2\mathcal{L}}{\partial\alpha_i\partial\alpha_j} \\ &= \begin{cases} -\frac{1}{|\Lambda|}\sum_{x\in\Lambda}P_{\sigma,\alpha,\beta,x}\{\omega(x)=i\} \\ \quad\times[1-P_{\sigma,\alpha,\beta,x}\{\omega(x)=i\}], & i=j, \\ \frac{1}{|\Lambda|}\sum_{x\in\Lambda}P_{\sigma,\alpha,\beta,x}\{\omega(x)=i\} \\ \quad\times P_{\sigma,\alpha,\beta,x}\{\omega(x)=j\}, & i\ne j,\end{cases}\end{aligned}$$

$1 \le i,j \le g-1$,

$$\begin{aligned}(H_{12})_{ij} &= \frac{\partial^2\mathcal{L}}{\partial\alpha_i\partial\beta_j} \\ &= -\frac{1}{|\Lambda|}\sum_{x\in\Lambda}P_{\sigma,\alpha,\beta,x}\{\omega(x)=i\}v(\sigma,x,j) \\ &\quad\times[i - E_{\sigma,\alpha,\beta,x}\{\omega(x)\}],\end{aligned}$$

$1 \le i \le g-1$, $1 \le j \le m$, and

$$\begin{aligned}(H_{22})_{ij} &= \frac{\partial^2\mathcal{L}}{\partial\beta_i\partial\beta_j} \\ &= -\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\left[E_{\sigma,\alpha,\beta,x}\{\omega^2(x)\}\right. \\ &\quad \left.-E^2_{\sigma,\alpha,\beta,x}\{\omega(x)\}\right]v(\sigma,x,i)v(\sigma,x,j),\end{aligned}$$

$1 \le i,j \le m$.

REMARK. Let $\zeta_{\omega,\sigma,x}$ be the $(g-1+m)$-dimensional vector defined by

$$(\zeta_{\omega,\sigma,x})_i = \begin{cases} I\!I\{\omega(x)=i\}, & 1\le i\le g-1, \\ \omega(x)v(\sigma,x,i+1-g), & g\le i\le g-1+m,\end{cases} \quad (3.6)$$

where $I\!I_A$ is the indicator function of the set A, that is $I\!I_A(y) = \begin{cases} 1 & \text{if } y\in A \\ 0 & \text{otherwise}\end{cases}$. One can see that the gradient of $\mathcal{L}$ is given by

$$\begin{aligned}\nabla\mathcal{L} &= \left(\frac{\partial\mathcal{L}}{\partial\alpha_1},\ldots,\frac{\partial\mathcal{L}}{\partial\alpha_{g-1}},\frac{\partial\mathcal{L}}{\partial\beta_1},\ldots,\frac{\partial\mathcal{L}}{\partial\beta_m}\right) \\ &= \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\{\zeta_{\sigma,\sigma,x} - E_{\sigma,\alpha,\beta,x}(\zeta_{\omega,\sigma,x})\} \\ &= (\hat{p},m) - \frac{1}{|\Lambda|}\sum_{x\in\Lambda}E_{\sigma,\alpha,\beta,x}(\zeta_{\omega,\sigma,x}). \quad (3.7)\end{aligned}$$

Moreover the Hessian matrix H of $\mathcal{L}$ can also be written in the form

$$\begin{aligned}H &= -\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\left[E_{\sigma,\alpha,\beta,x}(\zeta_x\zeta_x^\top)\right. \\ &\quad \left.-E_{\sigma,\alpha,\beta,x}(\zeta_x)E_{\sigma,\alpha,\beta,x}(\zeta_x^\top)\right]. \quad (3.8)\end{aligned}$$

It follows from representation (3.8) that H is negative definite provided the linear span generated by the vectors $\{(v(\sigma,x,1),\ldots,v(\sigma,x,m)); x\in\Lambda\}$ has dimension m.

3.2 Estimation of parameters for a group of images

If we have $k \ge 1$ independent images $\sigma_1,\ldots,\sigma_k$ of the same object, one can estimate the common parameters (α,β) by finding the roots of the pooled function $\mathcal{L}^{(k)}(\alpha,\beta)$ defined by

$$\mathcal{L}^{(k)}(\alpha,\beta) = \sum_{i=1}^{k}\mathcal{L}_{\sigma_i}(\alpha,\beta),$$

where $\mathcal{L}_{\sigma_i}$ is the function defined by (3.3) calculated with the image σ_i.

3.3 Test statistic

Suppose one has k images of an object and one observes j new images numbered $k+1$ to $k+j$. To compare images, one just have to test the null hypothesis $H_0 : (\alpha_{old}, \beta_{old}) = (\alpha_{new}, \beta_{new})$, that is the parameters of the new j images are the same as the one of the k images.

The usual pseudo-likelihood ratio test has not a nice asymptotic distribution, so we propose a related statistic R defined by

$$R = \frac{|\Lambda|}{\frac{1}{j}+\frac{1}{k}}(\hat{\theta}-\theta^\star)^\top \hat{I}(\hat{V})^{-1}\hat{I}(\hat{\theta}-\theta^\star),$$

where $\hat{\theta}$ and $\theta^\star$ are respectively the estimations of $\theta = (\alpha, \beta)$ for the k old images and the new j images, $S_{\sigma,x,\theta} = \zeta_{\sigma,\sigma,x} - E_{\sigma,\alpha,\beta,x}(\zeta_{\omega,\sigma,x})$, $x \in \Lambda$,

$$\hat{I} = \frac{1}{k|\Lambda|}\sum_{i=1}^{k}\sum_{x\in\Lambda} S_{\sigma_i,x,\hat{\theta}}\, S^\top_{\sigma_i,x,\hat{\theta}},$$

$$\hat{J}_l = \frac{1}{k|\Lambda|}\sum_{i=1}^{k}\sum_{x\in\Lambda} S_{\sigma_i,x,\hat{\theta}}\, S^\top_{\sigma_i,x+e_l,\hat{\theta}},$$

$1 \le l \le m$, and $\hat{V} = \hat{I} + 2\sum_{l=1}^{m} J_l$.

It follows that if the conditions in Subsection 6.1 are satisfied, then for large values of $|\Lambda|$, R behaves asymptotically like a Chi-square random variable with $g - 1 + m$ degrees of freedom.

The p-value of a test statistic having value R is thus approximated by $P(\chi^2_{g-1+m} > R)$. The null hypothesis is rejected when that p-value is too small.

4 Examples of applications

For the binary images treated in this section, we have chosen a model with eight neighbors at every interior sites, as seen in figure 1.

4	3	2
5	0	1
6	7	8

4	3	2
1	0	1
2	3	4

Figure 1: Neighbors of pixel at site 0 (left), pairs of neighbors (right).

In the case of binary images , $g = 2$, $(\alpha, \beta) = (\alpha, \beta_1, \beta_2, \beta_3, \beta_4)$, and the local specifications are given by:

$$P_{\sigma,\alpha,\beta,x}\{\omega(x) = \sigma(x) \mid \omega(y) = \sigma(y), \forall y \neq x\}$$

$$= 1\Big/\left(1 + e^{(1-2\sigma(x))\left[\alpha+\sum_{i=1}^{4}\beta_i\{\sigma(x+e_i)+\sigma(x-e_i)\}\right]}\right),$$

where $e_1 = (1,0)$, $e_2 = (1,1)$, $e_3 = (0,1)$, $e_4 = (-1,1)$, $e_{i+4} = -e_i$, $1 \le i \le 4$ are the eight neighbors of $(0,0)$; therefore the eight neighbors of x are $x + e_i$, $i = 1,\ldots,8$.

4.1 Application of the method with simulated images

For this simulation, images have been generated using the Gibbs sampler method explained in Subsection 6.2 of the Appendix. We generated 1000 pairs of 100×100 images with different parameters and we compared the t wo images of each pair by calculating our proposed test statistic R, and we rejected the null hypothesis that the parameters are the same if $R > 11.071$. This is the decision rule for a 5% level test.

The results are given in the table below. One can see that the chi-square approximation of the statistic is quite good since the rejection rate is 5.8% for images with the same parameters, which is very close to the true value of 5%. For the first image, the value of θ is always $(-0.3, 0.1, -0.4, 0.2, 0.05)$.

θ for second image	Proportion of rejections
$(-0.30, 0.1, -0.4, 0.2, 0.05)$	0.058
$(-0.35, 0.1, -0.4, 0.2, 0.05)$	0.256
$(-0.40, 0.1, -0.4, 0.2, 0.05)$	0.675
$(-0.30, 0.2, -0.4, 0.2, 0.05)$	0.810

These results indicates that the test has a very power since small differences in parameters are easily detected.

4.2 Comparison of handwritten signatures

For the second application, we considered handwritten signatures from the same person and we compared the first one with the other ones. Using a level of 5%, only the last one was rejected, having a value $R = 17.92$. The associated p-value is 0.0031. One can see that the image is somewhat different from the other ones so it is not surprising to reject the null hypothesis in that case.

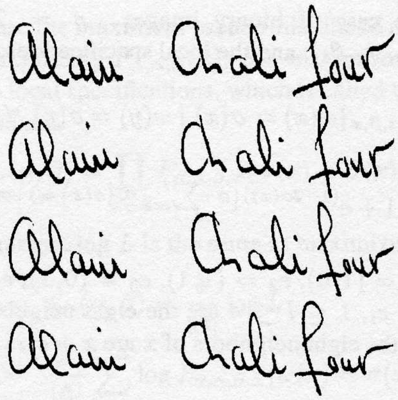

Figure 2: Four handwritten signatures of Alain Chalifour.

4.3 Pattern recognition

To illustrate again the method, let us compare each of the last three letters in Figure 3 with the first one.

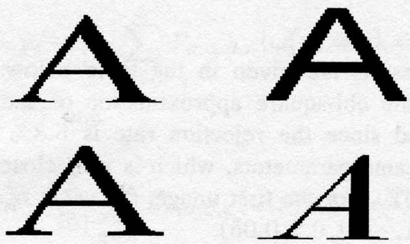

Figure 3: Four different fonts for letter A.

The respective values of the test statistic are: 104.46, 102.66, and 44.63. Therefore each time the null hypothesis is rejected, using a level of 5%.

5 Conclusion

We proposed a new statistical method for comparing two groups of images. The idea is to model images using parameterized families of probability distributions and to test equality of parameters of the respective groups. This method has the advantage of being easy to implement and also have a very good discriminating power, as seen by simulations and comparisons of real binary images. We believe that our method could be very usefull in comparing radar images with true images in geomatics, or to find fingerprints that match well.

6 Appendix

In this appendix, we first prove that the limiting distribution of the test statistic is chi-square and then we state some efficient ways of generating Gibbs random fields.

6.1 Asymptotic behavior of the test statistic

Let $\theta = (\alpha, \beta)$ be fixed and let μ be an ergodic stationary Gibbs probability measure with local specifications given by (2.1). Set $S_{\sigma,x,\theta} = \zeta_{\sigma,\sigma,x} - E_{\sigma,\alpha,\beta,x}(\zeta_{\omega,\sigma,x})$, $x \in \Lambda$.

Since we used the maximum pseudo-likelihood method for estimating parameters, it is tempting to use the difference of the log-pseudo-likelihoods as a test statistic as it is done in [6, Theorem, p. 158] for the maximum likelihood method. However, because of the dependence in the models, it does not work. Instead we can use a quadratic form based on the differences of parameters in the two groups.

To make it clear, set

$$\hat{I} = \frac{1}{k|\Lambda|} \sum_{i=1}^{k} \sum_{x \in \Lambda} S_{\sigma_i,x,\hat{\theta}}\, S^{\top}_{\sigma_i,x,\hat{\theta}},$$

and

$$\hat{J}_l = \frac{1}{k|\Lambda|} \sum_{i=1}^{k} \sum_{x \in \Lambda} S_{\sigma_i,x,\hat{\theta}}\, S^{\top}_{\sigma_i,x+e_l,\hat{\theta}},$$

$1 \leq l \leq m$. Further set $\hat{V} = \hat{I} + 2\sum_{l=1}^{m} J_l$.

If $I = E\left(S_{\sigma,x,\theta} S^{\top}_{\sigma,x,\theta}\right)$ and

$$V = I + \sum_{l=1}^{m} E\left(S_{\sigma,x,\theta} S^{\top}_{\sigma,x+e_l,\theta}\right) + \sum_{l=1}^{m} E\left(S_{\sigma,x+e_l,\theta} S^{\top}_{\sigma,x,\theta}\right).$$

It follows from [5] that $|\Lambda|^{1/2}\left(\hat{\theta} - \theta\right) \approx N_l(0, I^{-1}VI^{-1}/k)$. Let $\theta^\star$ estimated using the j new images. Then $|\Lambda|^{1/2}(\theta^\star - \theta) \approx N_l(0, I^{-1}VI^{-1}/j)$.

Since $\hat{I}$ and $\hat{V}$ converges almost surely to I and V respectively, it follows that

$$\frac{|\Lambda|}{\frac{1}{j} + \frac{1}{k}} (\hat{\theta} - \theta^\star)^{\top} \hat{I}(\hat{V})^{-1} \hat{I}(\hat{\theta} - \theta^\star)$$

converges in law to a random variable having chi-square distribution with $r = g - 1 + m$ degrees of freedom.

6.2 Monte-Carlo Markov Chain methods

It is sometimes impossibly difficult to generate directly observations with a given law. The idea behind Monte-Carlo Markov Chain methods is the following: one generates a Markov chain where the stationary measure is exactly the desired law. Then one runs the Markov chain for some time and one takes the last observation of the Markov chain. A nice introduction of these methods is [7]; see also [4].

We will now describe the algorithms we used for the generation of gray levels images.

6.2.1 Metropolis-Hastings algorithm

Let $\omega_n = \sigma \in \Omega = \{0,1\}^\Lambda$ be the configuration at time n. Then the next configuration ω_{n+1} at time $n+1$ is determined in the following way:

- choose a pixel x at random in Λ;
- choose color $j \in \{0,1,\ldots,g-1\}\backslash j_0$ at random, where $j_0 = \omega_n(x)$;
- choose a number U in the interval $(0,1)$;
- if $U \leq e^{\alpha_j - \alpha_{j_0} + (j-j_0)\sum_{l=1}^{m} \beta_l v(\omega_n, x, l)}$, then $\omega_{n+1}(x) = j$ and $\omega_{n+1}(y) = \omega_n(y)$, pour tout $y \neq x$, where $v(\sigma, x, l) = \sigma(x+e_l) + \sigma(x-e_l)$; otherwise $\omega_{n+1} = \omega_n$.

6.2.2 Gibbs sampler algorithm

For this algorithm, one must enumerate the points in Λ in as periodic way, that is $x_1, x_2, \ldots$, in such a way that $\Lambda = \{x_1,\ldots,x_p\}$ and $x_{n+p} = x_n$, where $p = |\Lambda|$. The new configuration ω_{n+1} is determined in the following way:

- choose U at random in the interval $(0,1)$;
- if $P_{\omega_n,\alpha,\beta,x_{n+1}}\{(\omega(x_{n+1}) \leq j-1\} < U \leq P_{\omega_n,\alpha,\beta,x_{n+1}}\{(\omega(x_{n+1}) \leq j\}$, for some $0 \leq j \leq g-1$, set $\omega_{n+1}(x_{n+1}) = j$ and $\omega_{n+1}(y) = \omega_n(y)$ for all $y \in \Lambda \setminus \{x_{n+1}\}$. Here $P_{\omega_n,\alpha,\beta,x_{n+1}}\{(\omega(x_{n+1}) = j\}$ is given by formula (2.1).

REMARK. Note that the initial configuration in not important. However there are two interesting choices: $\omega_0 = 0$, or the color of each pixel of ω_0 is determined independently with $P(\omega_0(x) = j) = e^{\alpha_j} \Big/ \sum_{k=0}^{g-1} e^{\alpha_k}$.

In the case $\beta = 0$, the Gibbs sampler algorithm is superior to the Metropolis-Hastings algorithm since the image ω_p has the right law, which is not necessarily the case for the Metropolis-Hastings algorithm because of the sampling with replacement of the pixels in the latter case.

Acknowledgments

References

[1] Andrews, H.C. (1972), Some unitary transformations in pattern recognition and image processing, *Inform. Processing*, Vol. 71, pp. 155-160.

[2] Besag, J. (1986). On the statistical analysis of dirty pictures (with discussion). *J. Roy. Statist. Soc. Ser. B* **48** 259–302.

[3] Geman, D. (1990), Random fields and inverse problems in imaging. In *Ecole d'été de probabilités de St-Flour XVIII, Lect Notes Math.* 1427, pp. 117-193.

[4] Hastings, W.K. (1970), Monte Carlo sampling methods using Markov Chains and their applications, *Biometrika*, Vol. 57, pp. 97-109.

[5] Janzura, M. (1997), Asymptotic results in parameter estimation for Gibbs random fields, *Kybernetika*, Vol. 33, pp. 133–159.

[6] Serfling, R.J. (1980). *Approximation theorems of mathematical statistics*. Wiley series in probability and mathematical statistics. *John Wiley and Sons.*

[7] Siddhartha, G., and Greenberg, E. (1995), Understanding the Metropolis-Hastings algorithm, *The American Statistician*, Vol. 49, pp. 327-335.

[8] Younès, L. (1991), Maximum likelihood estimation for Gibbsian fields, Spatial statistics and imaging, *IMS Lect. Notes, Monogr. Ser.* 20, pp. 403-426.