Synthesis and analysis of doubly stochastic models of images

Konstantin K. Vasiliev[1] and Nikita A. Andriyanov[1]

[1]Ulyanovsk State Technical University, Severny Venets, 32, 432027, Russia
nikita-and-nov@mail.ru,
WWW home page: http://tk.ulstu.ru/

Abstract. The problem of describing inhomogeneous images by the models with varying parameters is considered in the paper. Synthesis of doubly stochastic images based on autoregressive random fields was performed. In addition, the possibility of simulating various images is shown under taking into account the choice of parameters and methods of their transformation. We investigate the characteristics of inhomogeneous images and get the "image-correlation" dependences. Particular attention is paid to the problems of estimating the doubly stochastic models' parameters. We describe a technique that allows one to form a doubly stochastic model with its implementation for real images. We also consider algorithms for detecting point anomalies against a background of doubly stochastic signals and images. An optimal detection algorithm has been developed. The algorithm provides a gain in comparison with a detector based on the autoregressive models. The results of practical application of the elaborated algorithms to real images are presented.

Keywords: Doubly stochastic models, inhomogeneous images, statistical analysis of images, model parameters estimation, anomalies detection, random processes and fields, image processing

1 Introduction

Classical regression with constant correlation coefficients is widely used in scientific studies [1–7]. However, the constant coefficients do not allow describing real satellite data, which are characterized by spatial heterogeneity. Indeed, to simulate the image to be similar to the real one, its formation on a multidimensional grid can not be performed at each point by the same model. Therefore, the problem of regression models "dynamization" is so urgent. One of the approaches to ensuring such "dynamization" is a recurrent change in the parameters of the model.

2 Synthesis of doubly stochastic model

Consider the synthesis of a doubly stochastic model both on the basis of statistical modeling and on the basis of real data.

2.1 The idea of doubly stochastic model

Let the modeling of the image take place in accordance with the three-stage simulation algorithm [8]. The first stage involves generation of a component of a homogeneous random field (RF) ϱ (basic RF). In the second stage, it is needed to perform transformation of the obtained RF ϱ to make the RF with correlation parameters $\{\rho_{\boldsymbol{j}}, \boldsymbol{j} \in (j_1, j_2, ..., j_M)\}$, where M is a parameter describing the dimensions of the simulated image. These parameters characterize the correlation between the generated pixel of the image and its neighbors, similarly to the usual autoregression (AR) model. Finally, using RF with the parameters $\rho_{\boldsymbol{j}}$, we can get the main image.

To form the basic RF, we may use different models of RF [14]. Let us take, for example, the presentation of a two-dimension RF $X=\{x_{\boldsymbol{i}}, \boldsymbol{i} \in \Omega\}$, that we can get by presentation of first order AR model. If we use this model, we need to get the basic RF, and then transform its values into a row correlation $\{\rho_{x_{\boldsymbol{ij}}}, i = 1, 2, ..., M_1, j = 1, 2, ..., M_2\}$ and column correlation $\{\rho_{y_{\boldsymbol{ij}}}, i = 1, 2, ..., M_1, j = 1, 2, ..., M_2\}$

$$\begin{aligned}\tilde{\rho}_{xi,j} &= r_{1x}\tilde{\rho}_{xi-1,j} + r_{2x}\tilde{\rho}_{xi,j-1} - r_{1x}r_{2x}\tilde{\rho}_{xi-1,j-1} + \varsigma_{xi,j},\\ \tilde{\rho}_{yi,j} &= r_{1y}\tilde{\rho}_{yi-1,j} + r_{2y}\tilde{\rho}_{yi,j-1} - r_{1y}r_{2y}\tilde{\rho}_{yi-1,j-1} + \varsigma_{yi,j},\end{aligned} \tag{1}$$

where $\varsigma_{xi,j}$ and $\varsigma_{yi,j}$ are the two-dimensional RF of independent Gaussian random values (RV) with zero means and variances

$M\{\varsigma^2_{xi,j}\} = \sigma^2_{\varsigma x}$,

$\sigma^2_{\varsigma x} = \sigma^2_{\rho_x}\left(1 - r^2_{1x}\right)\left(1 - r^2_{2x}\right)$,

$M\{\varsigma^2_{yi,j}\} = \sigma^2_{\varsigma y} = \sigma^2_{\rho_y}\left(1 - r^2_{1y}\right)\left(1 - r^2_{2y}\right)$;

$\sigma^2_{\rho_x}$ and $\sigma^2_{\rho_y}$ define the variances of the basic random fields of correlation parameters for the row and column, respectively.

Parameters $\tilde{\rho}_{xi,j}$ and $\tilde{\rho}_{yi,j}$ allow us to adjust the geometric characteristics of objects imitated on the image while we are simulating the basic RF. The closer their values to unity, the larger the area in the image is occupied by such an object.

Meanwhile, choosing the method of the basic RF transformation into a set of correlation coefficients, it is possible to provide acceptable covariance functions (CF) of the images. Usually, the selected conversion method may contain some desired correlation structure of the original image. Thus, the proposed algorithm can be used to simulate images that are close (in probability) to real images from satellites.

The averages of RF $\varsigma_{xi,j}$ and $\varsigma_{yi,j}$ are equal to zero; so, the average for $\tilde{\rho}_{xi,j}$ and $\tilde{\rho}_{yi,j}$ (1) is also equal to zero. In view of the foregoing, we restrict ourselves to such a choice of the transformation of the values of the basic RF, at which the value of the RF $\tilde{\rho}_{xi,j}$ and $\tilde{\rho}_{yi,j}$ will be increased by a constant at every point, *i.e.* mathematical expectations m_{ρ_x} and m_{ρ_y} will be the following:

$$\begin{aligned}\rho_{xi,j} &= \tilde{\rho}_{xi,jbase} + m_{\rho_x},\\ \rho_{yi,j} &= \tilde{\rho}_{yi,jbase} + m_{\rho_y},\end{aligned} \tag{2}$$

where index "base" characterizes the value of the basic RF.

Figure 1 shows the way to imitate an image whose correlation parameters change according to the transformation of another simulated image based on the first-order AR model, so called the Habibi model.

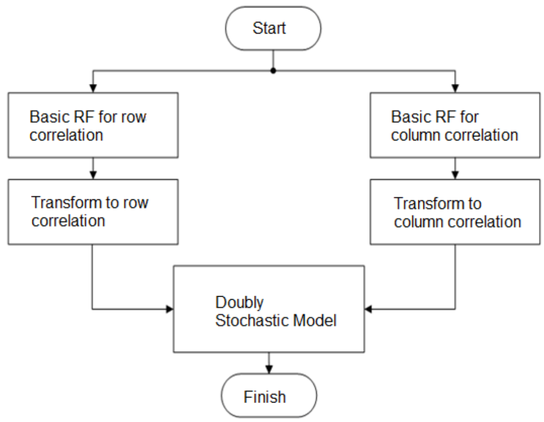

Fig. 1. Algorithm for simulating a doubly stochastic image

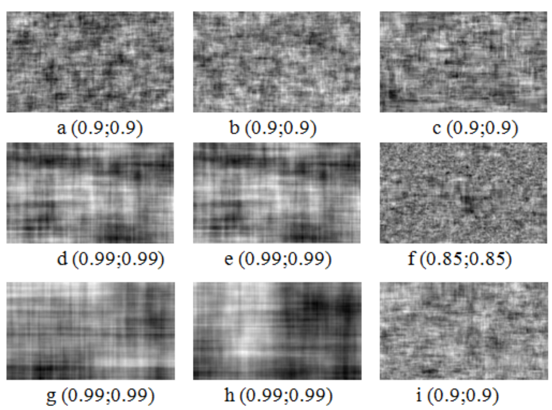

Fig. 2. Images generated by a doubly stochastic model based on the first-order AR

Figure 2 shows an example of using the algorithm according to the scheme of Fig. 1 for generating doubly stochastic RF, whose parameters vary in accordance with the equations of the AR. Here, (a, d, h) are presentations of the basic RF

$\{\rho_{xi,j}\}$; (b, e, i) are presentations of the basic RF $\{\rho_{yi,j}\}$; (c, f, i) are presentations of the doubly stochastic model based on AR model.

Analysis of the results (Fig. 2) shows that application of a doubly stochastic model of a RF allows imitating images with given correlation properties. Calculating the variances of the obtained RF, we can conclude that they are close to the given ones, and the implementation of the doubly stochastic RF is stationary. The additional advantage of this approach is that the formation of small images does not require large computational costs.

Figure 3 shows examples of images generated by the doubly stochastic models with different average values of correlation coefficients.

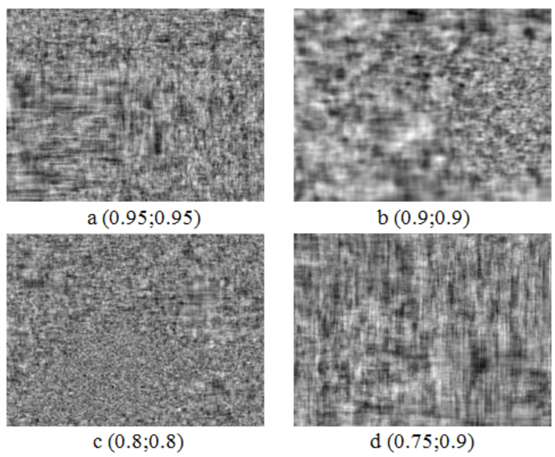

Fig. 3. Doubly stochastic images

Investigation of the correlation properties of the basic image $\varrho = \{\rho_{\boldsymbol{i}}, \boldsymbol{i} \in \Omega\}$ has shown that on its basis it is possible to obtain objects of different geometric shapes against the background of the main image. At the same time, by increasing the dimension of N-dimensional rectangular grid Ω, we can make a model for RF of higher orders. This provides a description of the images, which are inherent in a fairly complex terrain.

Obviously, changing the parameters allows us to form very different images. In this case, different algorithms for parameter transformation are possible. Thus, the simplest models of images were obtained.

2.2 Synthesis using parameters estimation

The problem of estimating parameters is reduced to estimating the fields $\rho_{xi,j}$ and $\rho_{yi,j}$ in model (1) [9,10]. In the considered case, in order to identify the parameters of such a model, it is necessary to evaluate the statistical parameters

m_{ρ_x}, m_{ρ_y}, $\sigma^2_{\rho_x}$ and $\sigma^2_{\rho_y}$. To do this, we shall use the sliding window method. Using some simple mathematical transformations, we can obtain the following estimates of the mathematical expectation, variance, and correlation by row and column at each point, proceeding from the necessity of presentation of an AR of the second order

$$m_{z(i+\frac{N-1}{2})(j+\frac{N-1}{2})} = \frac{1}{N^2} * \sum_{l=i}^{i+N-1} \sum_{k=j}^{j+N-1} X_{lk},$$

$$\sigma^2_{z(i+\frac{N-1}{2})(j+\frac{N-1}{2})} = \frac{1}{N^2-1} * \sum_{l=i}^{i+N-1} \sum_{k=j}^{j+N-1} (X_{lk} - m_{zlk})^2,$$

$$\rho_{xz(i+\frac{N-1}{2})(j+\frac{N-1}{2})} = \frac{1-\sqrt{1-\left(\frac{\sum_{l=i}^{i+N-2}\sum_{k=j}^{j+N-1}(X_{lk}-m_{zlk})*(X_{(l+1)k}-m_{z(l+1)k})}{(N-1)*(N)*\sigma^2_{z(i+\frac{N-1}{2})(j+\frac{N-1}{2})}}\right)^2}}{\frac{\sum_{l=i}^{i+N-2}\sum_{k=j}^{j+N-1}(X_{lk}-m_{zlk})*(X_{(l+1)k}-m_{z(l+1)k})}{(N-1)*(N)*\sigma^2_{z(i+\frac{N-1}{2})(j+\frac{N-1}{2})}}},$$

$$\rho_{yz(i+\frac{N-1}{2})(j+\frac{N-1}{2})} = \frac{1-\sqrt{1-\left(\frac{\sum_{l=i}^{i+N-1}\sum_{k=j}^{j+N-2}(X_{lk}-m_{zlk})*(X_{l(k+1)}-m_{zl(k+1)})}{(N-1)*(N)*\sigma^2_{z(i+\frac{N-1}{2})(j+\frac{N-1}{2})}}\right)^2}}{\frac{\sum_{l=i}^{i+N-1}\sum_{k=j}^{j+N-2}(X_{lk}-m_{zlk})*(X_{l(k+1)}-m_{zl(k+1)})}{(N-1)*(N)*\sigma^2_{z(i+\frac{N-1}{2})(j+\frac{N-1}{2})}}}.$$

where z-index is introduced to descript the observed data.

It should be noted that the relations found for the mathematical expectation and variance remain valid for other versions of the doubly stochastic model.

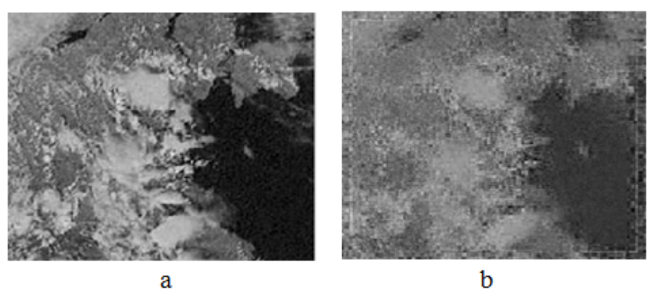

Fig. 4. Real and simulated images

Figure 4 shows how basing on estimates obtained by using the sliding window, we can represent the original image (a) as an implementation of the model of images with varying parameters (b).

3 Anomalies detection

We solve the problem of detecting anomalies for simulated and real images.

3.1 Point anomalies detection on the simulated data

We shall use the Neyman-Pearson approach, according to which we choose the detection rule, that ensures the maximum probability of correct detection. This is provided by the fact that the probability of a false alarm does not exceed a predetermined value P_{F0}. Thus, the optimal detection rule (in the sense of the Neyman-Pearson test) maximizes the probability

$P_D = 1 - P_M = \int_{G_1} ... \int \omega(\boldsymbol{z}/H_1) d\boldsymbol{z}$

with an additional restriction

$\int_{G_1} ... \int \omega(\boldsymbol{z}/H_0) d\boldsymbol{z} = P_{F0}$.

We shall calculate the characteristics of the detection of a point signal with level S, which can appear at a discrete point in time i. We write down the decisive rule [11–13] as follows:

If $\lambda = S^T(P_E + V_n)^{-1}(\boldsymbol{z_S} - \hat{\boldsymbol{x}}_E) > \lambda_0$, then there is the signal S,

Else if $\lambda = S^T(P_E + V_n)^{-1}(\boldsymbol{z_S} - \hat{\boldsymbol{x}}_E) \leq \lambda_0$, then there is not the signal S.

Then it can be reduced to the form

If $\lambda = \frac{S(z_i - \hat{x}_{Ei})}{\sigma_n^2} > \lambda_0$, then there is the signal S,

Else if $\lambda = \frac{S(z_i - \hat{x}_{Ei})}{\sigma_n^2} \leq \lambda_0$, then there is not the signal S, where σ_n^2 is the noise variance.

On the next step, we find the mathematical expectations and variances of the left part of the detection rule under conditions of presence and absence of a signal

$M\{\lambda/H_1\} = \frac{S^2}{\sigma_n^2}$,

$M\{\lambda/H_0\} = 0$,

$\sigma_\lambda^2 = D\{\lambda/H_1\} = D\{\lambda/H_0\} = \frac{S^4}{\sigma_n^4}(\sigma_n^2 + \sigma_e^2)$, where σ_e^2 is parameter that defines variance of main RF.

Then the probabilities of false alarm and missed target can be written as follows:

$P_F = 0.5 - \Phi_0(\frac{\lambda_0}{\sigma_\lambda})$ is the false alarm probability,

$P_M = 0.5 - \Phi_0(\frac{h - \lambda_0}{\sigma_\lambda})$ is the missed target probability, where Φ_0 is Laplace function, $h = S^2/\sigma_n^2$.

So, to calculate the probability of correct detection it is necessary to subtract from the unit the probability of the missed target. We find the threshold value λ_0 considering that the probability of the false alarms is $P_{F0} = 0.001$.

Figures 5a and 5b show graphs of the correct detection probability depending on the signal-to-noise ratio q in the forecast using one and two observations, respectively. It is seen that the detection on the basis of a doubly stochastic RF model is more effective than detection based on the classic AR model. In Fig. 5 the solid line is the theoretical probability for the algorithm based on the doubly stochastic RF model, the dashed line is the experimental probability for the algorithm based on the doubly stochastic RF model, the dot-and-dashed line is the probability for the algorithm based on the AR model.

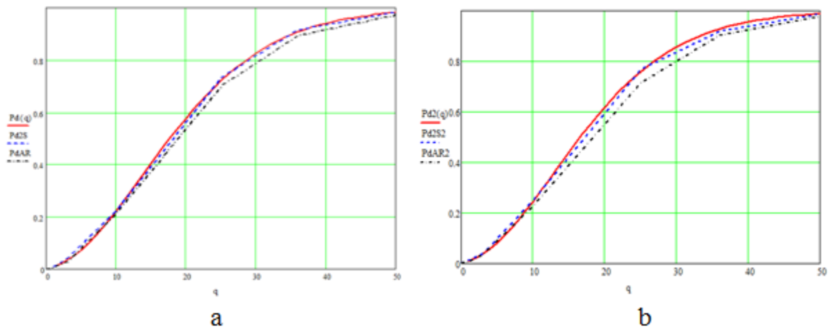

Fig. 5. Probability of correct detection with a prediction based on one (a) and on two (b) observartions

Analysis of the graphs in Fig. 5 shows that the detection efficiency increases in the case of using the doubly stochastic RF model and the larger volume of information for forecasting. The gain on 0.5 correct detection probability level is about 5-6% if we use two observations for prediction.

3.2 Extensive anomalies detection on real data

Now let us compare the work of two detectors of anomalies constructed on the basis of the doubly stochastic model (Algorithm 1) and on the basis of the autoregressive model (Algorithm 2). In this case, the detection will be performed on the real images obtained from the LandSat-8 satellite. Studies are conducted for three images. At the same time, in each image, 4 regions are selected where an anomaly can be located. It is worth to note that the areas are selected based on the structure of the images being examined with taking into account the greater and less heterogeneity and the fact that the detection procedures are performed not for the entire image, but only for these areas.

Figure 6 shows an example of one of the processed images with signals located in different parts of the images. The picture also reflects the probabilities of correct detection obtained using two algorithms. The dimensions of all images are 250x250. The images are distorted by the white Gaussian noise with a single dispersion. The size of the square is 4x4, the radius of the circle is 2. The signal-to-noise ratio is 1. The statistics are taken out 500 times. We also provide the probability of correct detection. The probabilities for Algorithm 1 are at the left , the probabilities for Algorithm 2 are at the top.

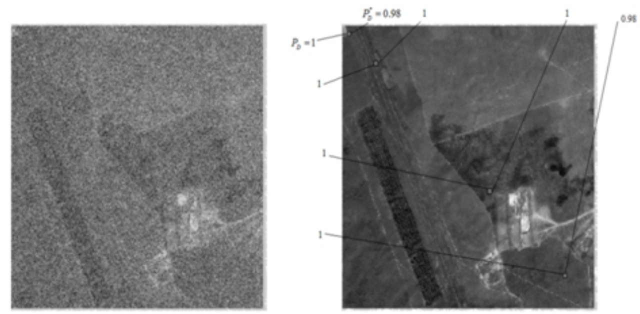

Fig. 6. Noisy (left) and source images (right) with probabilities of correct detection of a square signal

Figure 7 shows the gain of Algorithm 1 with respect to Algorithm 2 for the magnitude of the threshold signal at the probability of correct detection of 0.5 and the probability of false alarm 0.001, which corresponds to the threshold $L = 3.1\sigma_z^2$.

Shape/Image	Position 1	Position 2	Position 3	Position 4
Square on Image 1	1	1	0	4
Circle on Image 1	5	2	0	2
Square on Image 2	18	3	13	4
Circle on Image 2	23	4	3	5
Square on Image 3	11	7	4	10
Circle on Image 3	16	5	7	7

Fig. 7. Percent gains of the detection algorithm based on the doubly stochastic model in comparison with the detection algorithm based on the AR model

Analysis of the results shows that the algorithm based on the doubly stochastic model works better than the algorithm based on a simple autoregressive model and provides reliable detection of the signal in 90-95% of cases.

Analysis of data presented in Fig. 7 shows that a similar situation is in terms of ensuring at least the same efficiency. But in some cases, a significant gain is retained in the case of processing other heterogeneous satellite images. The probability of correct detection depends not only on the shape and size of the signal itself, but, also on the brightness values in its closest neighborhood. In this sense, a more universal algorithm is an algorithm based on doubly stochastic models.

4 Conclusion

Thus, application of the doubly stochastic models provides an adequate description of heterogeneous images, and the detection algorithms developed for such models lead to a gain in the detection of point signals of the order of 5-10%. We also can improve the correct detection probability by introducing more points for prediction or by increasing the signal level. We have investigated the modeling data, but the performed algorithms can become useful tool to detect various anomalies in the real multispectral images. In addition, the processing of real images also yielded gains in the detection of extended anomalies of different shapes on the average of the order of 10-15% with respect to the signal-to-noise ratio. It is worth to note that the novelty is offered by the proposed model with varying parameters, since it, unlike autoregressive models, makes it possible to generate images with different correlation properties. A comparison of the proposed algorithms based on a doubly stochastic model was performed with known autoregressive algorithms.

Acknowledgements. The study was supported by RFBR, project 16-41-732027.

References

1. Kashyap, R.L., Rao, A.R.: Construction of dynamic stochastic models from experimental data. Trans. from the English, ed. by V.Pugachev. Moscow. Nauka, 384 (1983)
2. Shalygin, A.S., Palagin, Yu.A.: Applied methods of statistical modeling. Mechanical Engineering, 320 (1986)
3. Spektor, A.A.: Digital image processing. Novosibirsk, NSTU, 350 (2002)
4. Vasiliev, K.K. et al: Applied theory of random processes and fields. Ulyanovsk, UlSTU, 256 (1995)
5. Vasiliev, K.K., Popov, O.V.: Autoregressive models of random fields with multiple roots. Proceedings of the 4th Conf. “Image recognition and image analysis: new information technologies.”, Novosibirsk, 4.1, 258–260 (1998)
6. Vasiliev, K.K., Sluzhivy, M.N.: Mathematical modeling of communication systems. Ulyanovsk, UlSTU, 170 (2010)
7. Boxing, J., Jenkins, G.: Analysis of time series. Forecast and management. Trans. from the English, ed. by V.P. Pisarenko. Moscow, Mir. 1, 406 (1974)
8. Vasil’ev, K. K., Dement’ev, V. E., Andriyanov, N. A.: Doubly stochastic models of images. Pattern Recognition and Image Analysis. 25(1), 105–110 (2015)
9. Vasiliev, K.K., Dementiev, V.E., Andriyanov, N.A.: Estimation of the parameters of doubly stochastic random fields. Radiotekhnika. 7, 103–106 (2014)
10. Vasiliev, K.K., Dementiev, V.E., Andriyanov, N.A.: Analysis of the effectiveness of estimating the changing parameters of a doubly stochastic model. Radiotekhnika. 6, 12–15 (2015)

11. Andriyanov, N.A.: Investigation of the efficiency of correct detection of signals against the background of one-dimensional doubly stochastic random processes. Technical Sciences: Problems and Perspectives Materials of the IV International Scientific Conference, 37-40 (2016)
12. Vasiliev, K.K., Dementiev, V.E., Andriyanov, N.A.: Detection of extended signals on a background of doubly stochastic images. Radiotekhnika. 9, 23–27 (2016)
13. Andriyanov, N.A., Vasiliev, K.K., Dementiev, V.E.: Anomalies detection on spatially inhomogeneous polyzonal images. CEUR Workshop Proceedings. Volume 1901, 2017 International Conference Information Technology and Nanotechnology. Session Image Processing, Geoinformation Technology and Information Security, IPGTIS-ITNT 2017; Samara; Russian Federation, 10-15 (2017)
14. Vasiliev, K.K., Dementiev, V.E., Andriyanov, N.A.: Filtration and restoration of satellite images using doubly stochastic random fields. CEUR Workshop Proceedings, "REIT 2017 - Proceedings of the 1st International Workshop on Radio Electronics and Information Technologies", 1814, 10-20 (2017)