Development of informative neighborhood selection technology for modeling texture images

E. Biryukova[1], R. Paringer[1,2], A. Kupriyanov[1,2]

[1]*Samara National Research University, 34 Moskovskoe Shosse, 443086, Samara, Russia*
[2]*Image Processing Systems Institute – Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, 151 Molodogvardeyskaya st., 443001, Samara, Russia*

Abstract

The paper proposes a method for constructing an informative neighborhood for modeling texture images. To describe the characteristic features of textures used assumptions underlying model representation texture images described by using a Markov random field. The results of the conducted experimental researches confirm that application of the developed approach allows to reduce the dimensionality of the features space while preserving the reliability of the classification.

Keywords: Markov random field; Gaussian Markov field; texture image classification; co-occurrence matrix; causal neighborhood

1. Introduction

Texture analysis widespread in the processing of various types of images. However, despite the fact that even in 1979 Haralick noted that the methods of distinguishing textures are developed individually for each specific case [1], there is no clear definition of texture or a particular concept in solving problems analysis of texture images.

Haindl wrote that the texture is a surface property, which is the spatial information contained in the object's surface [2].

The literature describes three approaches to texture analysis [1, 3, and 4]:

– A statistical approach, wherein the set of features used to provide texture image characteristics.

– Structural modeling allows us to consider texture as two-dimensional images composed of many primitives or subpatterns that are arranged accordance with a certain rule.

– Stochastic modeling suggests that the texture is the realization of a stochastic process that is characterized by certain parameters. This approach allows you to get good results for the generation of realistic natural texture images using Markov random fields [5].

To classification texture images, we will apply the model image as a realization of a random Markov field, that is, a stochastic approach to texture analysis. Great contribution to the development of this model has made by Haralick, who introduced the statistical and structural approaches to the description of texture [6] and suggested using of features based on the matrix of mutual probability distribution. The proposed is the gray level co-occurrence matrix [1]. It describes the spatial relationships of brightness pairs of texture elements.

2. Representation of the image according to the model of the Markov random field

Introduction of stochastic models and random fields models have led to the development of image reconstruction algorithms, segmentation, modeling and texture classification. In particular, Markov random fields is very useful for modeling spatial relationships, as well as for the study of stochastic interaction between the observed values, including the analysis of medical images and interpretation of remote sensing images [7].

The theory of Markov random field (MRF) provides a convenient and consistent method for modeling communication between dependent entities, such as image pixels and correlated features. Convenience is achieved due to the characteristic mutual influence among such objects, when using conditional distribution of MRF. The practical use of the model Markov random field obtained thanks to the theorem of equivalence between MRF and the Gibbs distribution, which was introduced by Hammersley and Clifford in 1971 [8]. This is because the joint distribution required for most applications, but the conclusion of the joint distribution of the conditional is very difficult for MRF. Equivalence theorem of Markov random fields and Gibbs points out that the joint distribution of MRF is the simplest form of the Gibbs distribution.

We will consider the model of a Gaussian Markov random field (GMRF), which is a particular case of MRF, where the value of the pixel in the position (i, j) is statistically independent of neighboring pixels. This means that the model takes into account the spatial interaction between the various components within each color component, and interaction of [9]. Image is represented on a rectangular lattice S = M * N with p number of bands.

Let $X(i,j) = [x_1(i,j)x_2(i,j) \dots x_p(i,j)]$ is a vector in a texture region R. It is assumed that the vector at a position (i, j) represents the linear combination of the color components of neighboring pixels and additive Gaussian noise. Let $\mu_1, \mu_2 \dots \mu_p$ denote the mean color intensity, and $e_1, e_2 \dots e_p$ the spatial interaction of pixels and v_{xy} be the expected value of $e_x e_y$. x, y takes on the values from 1 to p. Let ϕ_{xy} the associated parameters of the model and $\sum$ the co-occurrence matrix.

Spatial interaction of color pixels is defined as:

CEUR-WS.org/Vol-1901/paper7.pdf

$$e_1(i,j) = (x_1(i,j) - \mu_1) - \sum_{(m,n)\in N_{11}} \phi_{11}(m,n)(x_1(i+m,j+n) - \mu_1) - \sum_{(m,n)\in N_{12}} \phi_{12}(m,n)(x_2(i+m,j+n) - \mu_2) - \cdots - \sum_{(m,n)\in N_{1p}} \phi_{1p}(m,n)\big(x_p(i+m,j+n) - \mu_p\big).$$

Similarly it is defined for $e_2(i,j), e_3(i,j) \dots e_p(i,j)$. The generalized form is given by:

$$e_k(i,j) = (x_k(i,j) - \mu_k) - \sum_{(m,n)\in N_{k1}} \phi_{k1}(m,n)(x_1(i+m,j+n) - \mu_1) - \sum_{(m,n)\in N_{p2}} \phi_{k2}(m,n)(x_2(i+m,j+n) - \mu_2) - \cdots - \sum_{(m,n)\in N_{kp}} \phi_{kp}(m,n)\big(x_p(i+m,j+n) - \mu_p\big), k = \overline{1,p},$$

where N_{xy} denote neighboring pixels. If x=y, then the neighboring pixels will correspond to the same color component. Otherwise, the neighboring pixels are of the other components.

The co-occurrence matrix is defined as follows:

$$\sum = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1p} \\ v_{21} & v_{22} & & v_{2p} \\ \vdots & & \ddots & \vdots \\ v_{p1} & v_{p2} & \cdots & v_{pp} \end{pmatrix}.$$

The expected value v_{kl} is represented as:

$$v_{kl} = E[e_k e_l] = \frac{1}{M_R} \sum_{(i,j)\in R} e_k(i,j) e_l(i,j).$$

Having described all the terms, the probability density function of $X(i,j)$ is found to be:

$$P(X(i,j)|R) = \frac{1}{((2\pi)^P\ |\Sigma|)^{\frac{1}{2}}}\ \exp\{\frac{-1}{2}(e_1(i,j)e_2(i,j) \dots e_p(i,j)) \sum (e_1(i,j)e_2(i,j) \dots e_p(i,j))^t\}.$$

3. Choice of informative neighborhood

Winkler in "Image Analysis, Random Fields and Dynamic Monte Carlo Methods" [10] writes the restoration images and modeling textures with random fields, in detail the finite random fields, including MRF applies Monte Carlo methods for Markov chains. Chohen for example textile fabrics control automation task [11] solves the problem of detection and localization of various kinds of defects, which uses Gaussian Markov random field and the non-causal neighborhood. Kovtun in [12] proposes a model image, a feature of which is that the segmentation and each texture are set independent random fields. His work is an attempt to highlight the problem of texture segmentation of the general class of problems of generation and modeling of Markov random fields.

Thus, in [3, 5, 7-12] is said about using the Markov random field model to describe and generate texture images. One of the parameters of the described model is the probability distribution of the brightness of neighboring pixels. In this case, the choice of the neighboring pixels, in works devoted to this subject, using non-causal neighborhood (Fig. 1).

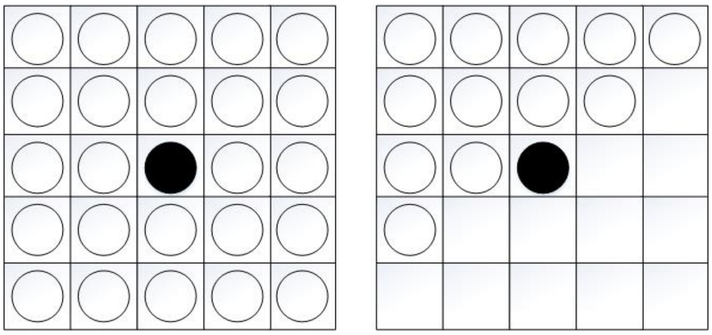
The paper proposes a new method of selecting the informative neighborhood to describe the characteristics of the texture.

The following algorithm can represent description of the main stages of the technology of selecting the informative neighborhood:

1. Choosing raw data: the shape of the neighborhood, a set of features calculated from the surroundings and the separate images on the textural classes.
2. Calculation features of the selected neighborhood for each image. Form the initial sample.
3. Calculate individual separability criteria for each feature [13]. We assess informative features, based on the value criterion [14].
4. Excluded from the original sample with features of lower value separability criterion.
5. We exclude from the neighborhood of the pixels corresponding to the non-informative characters.

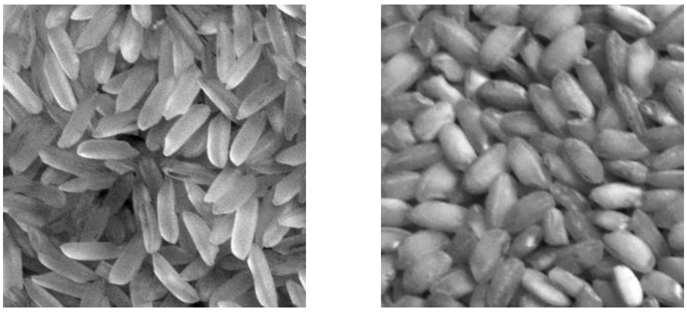
Thus, the remaining pixels constitute informative neighborhood.

Experimental technology study carried out based on texture images “Kylberg Texture Dataset v. 1.0”[15]. Consider the application of technology to the two classes of images and rice1 rice2 selected database. Figure 2 shows examples of the images under consideration.



a) b)

Fig. 1. Example of the surrounding area: (a) the non-causal, (b) causal.



a) b)

Fig. 2. Example images: (a) - rice1, (b) - rice2.

To distinguish the classes of texture images, we used statistical features calculated by the formula:

$$\lambda(\Delta x, \Delta y, \mathrm{n}) = \frac{\sum \left| f(x,y) - f(x+\Delta x,\ y+\Delta y) \right|^n}{N},$$

where f is the image intensity function, N is the number of image pixels. In the following research we used the features λ, calculated at Δx, Δy = 0, ±1, ±2, n = 1, 2, 3. Because the features are symmetrical used causal neighborhood.

Individual criteria of separability for each feature were calculated (Figure 3).

For a sample consisting of n elements, divided into classes g and comprising a p features separability criterion is calculated using the following formulas:

$$J = tr((\boldsymbol{T})^{-1}\boldsymbol{B}),$$

where $\boldsymbol{T} = \boldsymbol{B} + \boldsymbol{W}$.

$\boldsymbol{B}$ – is the intergroup dispersion matrix. The elements of this matrix are calculated according to the formula:

$$b_{ij} = \sum_{k=1}^{g} n_k(\overline{x}_{ik} - \overline{x}_i)(\overline{x}_{jk} - \overline{x}_j), i,j = \overline{1,p}\ ,$$

$\boldsymbol{W} -$ is the intragroup dispersion matrix. The elements of the matrix are calculated according to the formula:

$$w_{ij} = \sum_{k=1}^{g}\sum_{m=1}^{n_k}(x_{ikm} - \overline{x}_{ik})(x_{jkm} - \overline{x}_{jk}),\ \ i,j = \overline{1,p},$$

x_{ikm} – is the value of the $i\text{-}th$ feature for the $m\text{-}th$ element of k class, $\overline{x}_{ik} = {}^{1}\!/_{n_k} \sum_{m=1}^{n_k} \overline{x}_{ikm}$ – is the mean value of the $i\text{-}th$ feature of k class, $\overline{x}_i = {}^{1}\!/_{n} \sum_{k=1}^{g} n_k \overline{x}_{ik}$ – is the mean value of the $i\text{-}th$ feature in all the classes, and n_k is the number of elements in k class.

0.74	0.77	0.82	0.77	0.55
0.58	0.74	0.85	0.61	
0.58	0.61	●		
0.53				

Fig. 3. Mean value of the separability criterion.

The higher the value of the criterion is, the more the separability of the classes grows.

After calculating the individual criteria for separability, features with a low value criterion were excluded. Analysis of the feature space, led to the conclusion that some of the neighboring pixels carry information about the features of the texture (pixel information are highlighted in Figure 3). It was excluded from the neighborhood of the pixels corresponding to non-informative features (calculated at (Δx = 2, Δy = 2), (Δx = -2, Δy = 1), (Δx = 1, Δy = 1), (Δx = -2, Δy = 0), (Δx = -1, Δy = 0), (Δx = -2, Δy = -1)). Thus, we resins are informative neighborhood new form. Modified neighborhood for the test classes is shown in Figure 4.

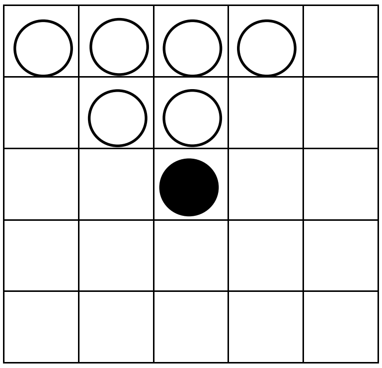


Fig. 4. Modified neighborhood.

To study the effectiveness of the technology was evaluated the quality of the selected neighborhood. The evaluation was conducted by calculating the clustering error on the based of k-means algorithm, where the centers of the starting classes used as initial conditions [16]. Under the error of clustering is understood the proportion of images that were not attributed to their class. Clustering error in the case of using features calculated by causal neighborhood was 0.21, the modified 0.19, which confirms the information content of the modified neighborhood.

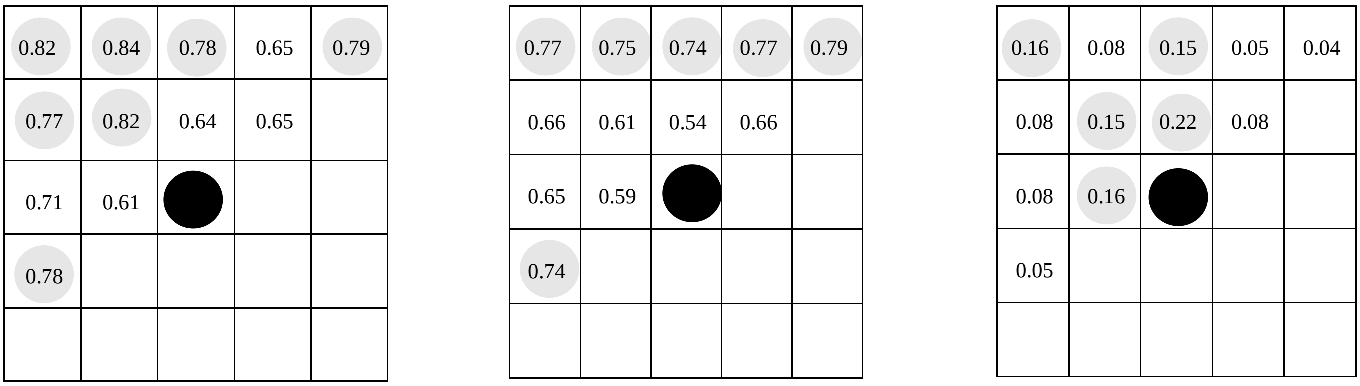
Table 1 shows the values of the clustering error in the case of features, calculated using the causal neighborhood and modified to distinguish other classes of images from the selected base textures.

Table 1. The values of the clustering errors when using different neighborhoods.

compared classes	causal neighborhood	modified neighborhood
blanket1, and canvas1	0.03	0.03
scarf1, and scarf2	0.18	0.16
Linseeds, and sesameseeds	0.46	0.40

As Table 1 shows that clustering error value using the modified neighborhood does not exceed the error value using a causal neighborhood, which indicates the information content received surroundings, and hence the effectiveness of the proposed technology.

Figure 5 shows the mean values of separability criteria for the cases considered in Table 1, the modified neighborhoods are highlighted in color.



blanket1 and canvas1

scarf1 and scarf2

linseeds and sesameseeds

Fig. 5. The mean values of the separability criterion for different classes.

4. Conclusion

The paper presents the technology of choice informative neighborhood, which has shown to be effective for the considered classes of texture images. The features space and the clustering error were reduced by reducing the number of neighboring pixels. The proposed technique can be used to modeling texture images, wherein for the calculation of the model parameters using a Markov random field neighborhood.

Acknowledgements

This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 15-29-03823, # 15-29-07077, # 16-41-630761, # 17-01-00972); by the ONIT RAS program # 6 "Bioinformatics, modern information technologies and mathematical methods in medicine" 2017.

References

[1] Haralick RM. Statistical and structural approaches to texture. Proceedings of the IEEE 1979; 67(5): 786–804.
[2] Haindl M. Texture synthesis. CWI Quarterly 1991; 4: 305–331.
[3] Dubes RC, Jain AK. Random field models in image analysis. Journal of Applied Statistics 1989; 16(2): 131–164.
[4] Ahuja N, Rosenfeld A. Mosaic models for textures. IEEE Transactions on Pattern Analysis and Machine Intelligence 1981; PA MI 3: 1–10.
[5] Plastinin AI. The method of forming features texture images based on Markov models. Dis. kan. those. Sciences, Samara, 2012.
[6] Haralick RM. Statistical and structural approaches to the description of textures. TIIRE 1979; 5: 98–118.
[7] Li SZ. Markov Random Field in Image Analysis. Springer-Verlag, 2009; 362 p.
[8] Hammersley JM, Clifford P. Markov field on finite graphs and lattices, unpublished, 1971
[9] Mridula J. Feature Based Segmentation of Colour Textured Images using Markov Random Field Model. Master of Technology, Odisha, India, 2011.
[10] Winkler G. Image Analysis, Random Fields and Dynamic Carlo Methods Monte. Novosibirsk: Branch "Geo" Publishing RAS, 2002; 343 p.
[11] Cohen FS, Fan Z, Attali S. Automated Inspection of Textile Fabrics Using Textural Models. IEEE Transactions on Pattern Analysis and Machine Intelligence 1991; 8(13): 803–808.
[12] Kovtun IV. Texture image segmentation based on Markov random fields. Control systems and machines 2003; 4: 46–55.
[13] Fukunaga K. Introduction to statistical pattern recognition theory. Moscow: Nauka, 1979; 270 p.
[14] Biryukova E, Paringer R, Kupriyanov A. Development of the effective set of features construction technology for texture image classes discrimination. CEUR Workshop Proceedings 2016; 1638: 263–269. DOI: 10.18287/1613-0073-2016-1638-357-363.
[15] Kylberg G. Kylberg Texture Dataset v. 1.0, 2014. URL: http://www.cb.uu.se/~gustaf/texture.
[16] Mandel ID. Cluster analysis. Moscow: Finance and Statistics, 1988.