Texture classification in aerial photographs using multiscale and multilayer complex networks

Margarita N. Favorskaya[1], Angelina N. Zhukovskaya[1]

[1]Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia

Abstract

Texture classification using oriented complex networks considers the functional connections between topological elements and simulates the complex textures more accurately. In contrast to the classical spatial texture analysis, we offer a novel function of weights in complex networks and a classification method that takes into account the scaling and color of textures. For this, three complex networks represented R, G and B components are built, which provide invariance of color aerial photographs obtained at different times. Comparison of the classification results using the proposed multiscale complex networks and conventional texture analysis based on a statistical approach is given. Also we extended this approach on color aerial photographs using multilayer structure of complex network.

Keywords

Texture classification, complex networks, multilayer structure, multiscale invariance, aerial photograph.

1. Introduction

Texture analysis is a fundamental problem in computer vision tasks. Due to the fact that this is a constant research field since the 1960s, many different approaches have been developed that does not prevent from looking for new views on an old problem [1]. Any visible object of the real world has its own textural features at a certain scale, associated with local spatial variations such as color, orientation and intensity depending on lighting conditions. There are many definitions of texture in literature, which are employed in several areas. The main reason is that natural textures have a wide range of the properties, for example, from regularity to randomness, from homogeneity to heterogeneity. Such properties cannot be described in a unified manner. Depending on the task, texture is considered as primitives with specific spatial distributions [2], as a stochastic, possibly periodic, two-dimensional image field [3] or as simple color patterns. Nevertheless, spatial homogeneity, meaning statistical stationarity, is the most important property of texture, leading directly to self-similarity and texture interpretation as a fractal structure.

Extraction of texture features is usually performed at the initial stages of visual data processing that makes this procedure important from the point of view of future results. Conventional methods of texture analysis are grouped into four categories, such as statistical, spectral, structural and model-based methods. Recently, alternative approaches based on learning techniques have been developed. These approaches include bag-of-visual-words, cellular automata and

SDM-2021: All-Russian conference, August 24–27, 2021, Novosibirsk, Russia
favorskaya@sibsau.ru (M. N. Favorskaya); zhukowskaya.angelina@yandex.ru (A. N. Zhukovskaya)

complex networks. The use of complex networks (CNs) is one of the interesting and rapidly developing areas, when a texture is interpreted as a graph with certain topological properties and binary connections. This interpretation makes it possible to simulate the physical properties of natural textures.

The structure of this paper is as follows. In Section 2, the evolution of texture analysis methods is briefly reviewed. Section 3 describes the main propositions of CNs. The proposed method based on multiscale CNs is presented in Section 4. Further, the dataset and experimental results are reported in Section 5. Section 6 concludes the paper.

2. Evolution of texture analysis methods

Investigations in texture analysis began in the 1960s and continue to this day, because texture, as an inherent property of objects in the real world, is a determining factor in both the classification and modeling. Statistical texture analysis is one of the earliest methods based on estimating ordinal statistical moments, calculating gray-level co-occurrence matrices, and applying local binary patterns (LBPs). It should be noted that the use of LBPs proposed in the 1990s is a fast way to classify textures with fairly good results. This family of methods has led to numerous modifications, some of which have became difficult to implement [4]. Spectral methods are based on the calculation of Gabor filters, wavelet transforms and other spectral methods. In structural methods, texture is considered as a combination of small elements called textones that form a spatially structured pattern. These methods often use morphological decomposition. Model methods involve the construction of complex mathematical models and the estimation of their parameters, for example, fractal models and stochastic models based on Markov networks.

In addition to conventional approaches, alternative innovative methods are being developed based on the analysis of feature points, discriminant local features, cellular automata, deep neural networks and complex networks. Often these approaches use feature dictionaries in the form of bag-of-visual-words (BOVW). The main advantage of the methods based on the theory of complex networks is related to their ability to map relationships between structural elements of a texture. However, the reduction of the dimensionality of features and the search for new ways to build complex networks remain an area of interest.

The deterministic tourist walk method, which belongs to the agent-based category, can be considered as a predecessor of CNs in texture analysis [5]. An agent visits pixels in accordance to predefined deterministic rule and memory size. Each walk included the transient part, where the agent walked freely, and the attractor in the form of a sequence of pixels, repeated during the walk. The description of texture was the joint distribution of transient times and attractor periods. Hereinafter, texture analysis using a combination of graph theory and partially self-avoiding deterministic walks was proposed by the same authors in [6]. First, an image was represented in the form of a regular graph. Then the regular graph was transformed into a graph with different properties and transitivity, which revealed different texture properties. The texture descriptor was presented as a histogram.

A combination of CNs and LBPs was introduced in [7]. The proposed local spatial pattern mapping (LSPM) method transformed CNs at different radial distances to LBPs, followed by concatenation of histograms, as is done in classification using conventional LBPs. The LSPM method described the uniformity of texture primitives and examined three Euclidean radial distances of LSPM. Experiments have confirmed that the classification results were highly dependent on a predefined set of thresholds during CNs construction.

In [8], the fusion of BOVW and CNs methods called BoVW-CN was proposed. This approach allowed to describe the keypoints of the image through a texture-based focus. Parameters such as mean centrality, number of communities, average degree, transitivity, average minimum path, motifs (small patterns of interconnections), and connectivity histogram were calculated for a limited number of salient points in the image, assuming that salient points fully reflect the properties of the texture.

Ribas et al. [9] proposed a combination of CNs and randomized neural networks (RNN) to obtain a texture description. First, a texture is simulated as a directed CN. Second, RNN is trained with information from the modeled CNs in order to consider the topological properties of the texture. The outputs of RNN provide texture signature. This method has demonstrated robustness to rotation of texture images.

CNs find their application not only in texture analysis, but also in image segmentation, when a cluster of nodes is related with image segmentation. The fundamental limitation of image segmentation based on CNs due to the excessive numbers of nodes in the network is overcome by the concept of super-pixels [10]. Mourchid et al. [11] proposed a framework that used a weighted region adjacency graph to represent an image as a network, where regions represented nodes in the network. This framework was based on community detection algorithms in graphs because networks are growing exponentially in size, variety and complexity.

From a brief literature review, we see that CNs are used in combination with other interesting techniques aimed at achieving invariance properties. The paper proposes a texture classification method for aerial photographs based on the use of complex networks calculated for various scales and refined the final classification.

3. Complex networks

Currently, the theory of complex networks is used in many fields, such as physics, sociology, biology, mathematics, computer science, medicine, ecology, linguistics, and others. Complex networks can be defined as graph structures with statistical mechanisms that give them an interdisciplinary nature. Complex networks demonstrate the flexibility and versatility of representing natural structures, including dynamic topology changes. In linguistics, CNs are suitable for automatic summarization due to strong correlation between the metrics of such networks and important text features. In [12], CNs were applied to select sentences for an extractive summary. The nodes of such CN corresponded to sentences, while the edges connected sentences that had common meaningful nouns. In computer vision, most of the problems are related to feature extraction and topological characteristics. In this sense, the CN-based descriptors can be used to classify objects and solve the recognition problems [13]. When describing a texture in terms of CNs, pixels are considered as nodes, and the connections between nodes determine the similarity. The main assumption is that different textures have different topologies of such networks, which are characterized by their own connectivity parameters. Such parameters are used as features in the classification of textures.

Let us define a texture as a two-dimensional pixel structure. In grayscale images, each pixel has an intensity value, which is an integer in the range $g = 0, \ldots, L$, $L = 255$. Let $I(x, y) = g$, $x = 1, \ldots, M$ and $y = 1, \ldots, N$ be the intensity of the pixel with coordinates (x, y) in image I. Let us build a graph $G = (V, E)$, in which each pixel $I(x, y)$ is a node $v_{x,y} \in V$ and connections between two pixels $I(x, y)$ and $I(x', y')$ are defined by an undirected edge $e \in E$, $e = (v_{x,y}, v_{x',y'})$, where Euclidian distance is no more that radius r:

$$E = \left\{ e = \left(v_{x,y}, v_{x',y'} \right) \in E | dist \left(v_{x,y}, v_{x',y'} \right) \leq r \wedge I(x, y) < I(x', y') \right\}, \tag{1}$$

where $dist(v_{x,y}, v_{x',y'})$ is an Euclidian distance between two pixels, defined by following expression:

$$dist(v_{x,y}, v_{x',y'}) = \sqrt{(x - x')^2 + (y - y')^2}.$$

Each edge has a weight $w(e)$ computed by (2):

$$w(e) = dist^2(v_{x,y}, v_{x',y'}) \cdot f^2 \frac{|I(x, y) - I(x', y')|}{L} \quad \forall e = (v_{x,y}, v_{x',y'}) \in E. \tag{2}$$

Weights are often normalized to the range $[0, 1]$ for two cases $r = 1$ and $r > 1$. In literature, we can find different functions $w(e)$, for example [9]. We suggest this function in the form of (3).

$$w(e) = \begin{cases} \dfrac{|I(x, y) - I(x', y')|}{L} & \text{if } r = 1, \\ \dfrac{dist^2(v_{x,y}, v_{x',y'}) - r}{2r^2} \cdot \dfrac{|I(x, y) - I(x', y')|}{L} & \text{if } r > 1. \end{cases} \tag{3}$$

As one can see, the multiplier F

$$F = \frac{dist^2(v_{x,y}, v_{x',y'}) - r}{2r^2 \cdot L} \tag{4}$$

is a constant for different values r. Thus, we can pre-calculate these values and store them in a special matrix. According to (4), value F increases with a larger value r. We can interpret (4) as an inverse filter with a minimum value at $r = 1$.

Note that the network constructed according to (1)–(3) has the same number of connections, in other words such network is a regular graph. A regular graph is not a complex network and requires further transformations. The simplest transformation is to use the threshold value t for the original set of edges E in order to form a subset $E_t \subseteq E$, where each edge $e \in E_t$ has a weight $w(e)$ equal to or less than the value t. As a result of this transformation, a new network $G_t = (V, E_t)$ is formed, which is an intermediate stage in the evolution of the network:

$$E^* = \{e \in E | w(e) \leq t\} \tag{5}$$

Figure 1 depicts an example of creating a complex network using a sliding window with sizes 7×7 pixels.

Due to the wide variety of textures, it is difficult to say immediately which values of r and t will be the best in terms of CN stability. Thus, the learning process of CN is an experimental selection of these parameters. By applying a set of threshold values t, $t \in T$, to the original a

55	107	29	7	34	80	43
62	24	37	12	23	48	129
92	3	65	89	73	32	212
126	59	48	100	105	34	240
97	184	52	64	125	18	208
82	109	50	71	39	242	168
56	18	31	24	209	154	45

b

			0.12			
	0.22	0.09	0.09	0.11	0.15	
	0.14	0.14	0.04	0.11	0.10	
0.03	0.04	0.20		0.02	0.07	0.18
	0.12	0.19	0.14	0.10	0.12	
	0.03	0.07	0.03	0.09	0.21	
			0.10			

c

			0.12			
	0.22	0.09	0.09	0.11	0.15	
	0.14	0.14	0.04	0.11	0.10	
0.03	0.04	0.20		0.02	0.07	0.18
	0.12	0.19	0.14	0.10	0.12	
	0.03	0.07	0.03	0.09	0.21	
			0.10			

Figure 1: Creating complex network: a — original image as the nodes in the graph; b — edges of the nodes, $r < 3$; c — complex network, $t < 0.11$.

network G, we can study the behavior of texture properties based on calculations of statistical features. The degree (connectivity) of a node v determines the number of edges connected to this node. Knowing the degree, we can build a histogram and calculate the mean, contrast, energy, and entropy. These features have a significant role in the description of the texture because they describe local properties (at the level of neighboring nodes) and global properties (at the level of histogram). Moreover, it is possible to simulate the dynamic behavior of a complex network by transforming the original network and then calculating its specific properties [14].

Typically, texture signatures are evaluated using Linear Discriminant Analysis (LDA). The LDA method is a supervised learning method and, therefore, requires a preliminary determination of the number of classes. However, other classifiers are also used, for example, based on distance metrics or histogram analysis by transforming a complex network into local spatial patterns [7].

Various modifications of complex networks are known, for example, CNs for the categorization of textures, dynamic texture recognition, face recognition, texture classification using BOVW, etc.

4. Proposed method for texture classification

The developed method of texture classification takes into account different image scales. Scale invariance is achieved by building a pyramid of images in such a way that the texture is analyzed by sliding windows of sizes 31×31, 15×15 and 7×7 pixels. Also we chose the values of parameters r and t during the CN learning. Aerial photographs are high resolution images, thus the number of extracted texture samples will be sufficient for supervised learning. During testing, parameters such as mean M, contrast C, energy E and entropy H are calculated for each analyzed patch using (6)–(9), where $z_{i,j}$ is the intensity value, $p(z_{i,j})$ is the number of pixels with values $z_{i,j}$, M is the size of the sliding window in one direction, L is the number of intensity levels, $L = 255$.

$$M = \frac{1}{M \times M} \sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}, \tag{6}$$

$$C = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(z_{i,j}) z_{i,j}^2, \tag{7}$$

$$E = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p^2(z_{i,j}), \tag{8}$$

$$H = -\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(z_{i,j}) \log_2 p(z_{i,j}). \tag{9}$$

One of the interesting but difficult to calculate texture parameters is the fractal dimensionality. The main idea is to verify that CNs consist of self-repeating patterns at all scales, using fractal methodologies. This is a special issue for further investigation. Note that fractal dimensionality is a distinctive feature of natural and man-made objects.

Analysis of color aerial photographs leads to triple increase in the number of features in RGB color space. In this case, we build multilayer CNs close to the approach presented in [15].

Aerial photographs, as a rule, contain a limited number of textures. Therefore, the classification stage can be simplified using the Minkowski distance [16] or the maximum likelihood method as an estimation function. The study is aimed at comparing the results of texture classification in the spatial domain with and without the factor of scale invariance of aerial photographs.

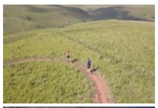
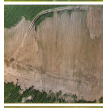
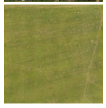
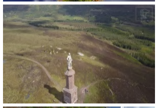
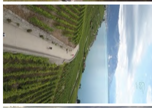
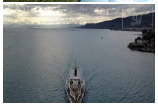
5. Experimental results

Since we could not find a publicly available aerial texture dataset, we used 12 video sequences obtained from a dataset "Drone Videos DJI Mavic Pro Footage in Switzerland" [17] and aerial photographs from dataset "Aerial Textures" [18]. The dataset "Drone Videos DJI Mavic Pro Footage in Switzerland" can be applied to a variety of computer vision tasks and has no ground-truth labels. This dataset includes 18 short video sequences in various natural environments with and without people present. The total size is 1.71 GB. We used frames from 6 video sequences. Dataset "Aerial Textures" is a library, which consists of 145 high quality textures in extremely high resolution of up to 400 million pixels, suitable for general usage, and includes images of roads, cross road walks, fields and pavements at different times of season and frozen water features. This library provides photorealistic texturing of large areas with close views. We used about 40 high resolution photographs from "Aerial Textures Field Summer" and "Aerial Textures Road Summer". The total size of "Aerial Textures JPG, Field Summer" and "Aerial Textures JPG, Road Summer" is 10.9 GB. The main parameters of video sequences and aerial photographs are depicted in Table 1.

We created our own dataset as a set of texture patches belonging to the classes "Forest", "Meadow", "Field", "Road", "Mountains", "River", "Sand", and "Sky". Each class includes 40–60 patches, thus our dataset is balanced. The total number of patches is 420 patches of 192 $\times$ 192 pixels, which were divided into training set and test set as 70% and 30%, respectively. Examples of patches are depicted in Figure 2.

Texture analysis was performed in "The R Project for Statistical Computing". The used sliding windows had sizes 31×31, 15×15 and 7×7 pixels. During the experiments, we adapted the parameters of r and t (see Figure 1) for each scale and each class. Such manual setting is a disadvantage of using CNs. However, this approach is suitable for solving the problem with

Table 1
Main parameters of test frames and photographs.

Drone Videos DJI Mavic Pro Footage in Switzerland				Aerial Textures		
Caption	First frame	Resolution	Number of frames	Caption	Photograph	Resolution
Berghouse Leopard.mp4		1280×720	1073	AET FIELD Summer 03		8000×8000
Bluemlisal Flyover.mp4		1280×720	957	AET FIELD Summer 10		8000×6000
Isles of Glencoe.mp4		1280×720	899	AET FIELD Summer 20		12000×10000
DJI_0501.mov		3840×2160	232	AET ROAD Summer 33		10000×1700
DJI_0574.mov		3840×2160	928	AET ROAD Summer 45		16000×9000
DJI_0596.mov		3840×2160	1015	AET ROAD Summer 55		12000×6000

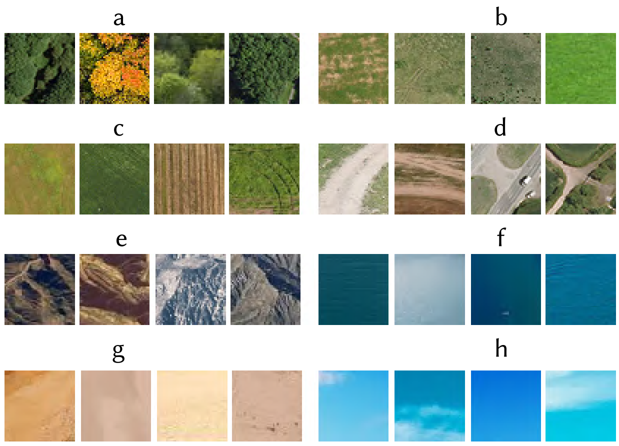


Figure 2: Examples of patches from eight classes: "Forest", "Meadow", "Field", "Road", "Mountains", "River", "Sand", "Sky" (a–h respectively).

limited classes. Since textures belonging to the same class are not homogeneous, the structures of CNs are different for different scaling. We calculated mean M, contrast C, energy E and entropy H for three scales and three color components (R, G and B) for each class of texture using CNs based on the training set. For classification, we used the Minkowski distance.

We compared our results with the conventional texture analysis based on a statistical approach similar to texture analysis done in [19] using normalized energy E_n, relative smoothness R_m and normalized entropy H_n. The formulae for estimating normalized homogeinity E_n, relative smoothness R_m and normalized entropy H_n are provided by (10)–(12).

$$E_n = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p^2(z_{i,j}) \Big/ \log_2 L, \tag{10}$$

$$R = 1 - \frac{1}{1 + \sum\limits_{i=0}^{L-1} (z_i - m)^2 p(z_i) \Big/ (L-1)^2}, \quad R_m = \begin{cases} -logR & \text{if } R > 0 \\ 10 & \text{if } R = 0 \end{cases}, \tag{11}$$

$$H_n = H \Big/ \log_2 L. \tag{12}$$

Table 2
Comparative results of texture classification.

Class	Size of patch	The conventional texture analysis		The proposed CNs analysis	
		Mean TR, %	Mean HTER, %	Mean TR, %	Mean HTER, %
Forest	7×7	94.02	6.87	**96.12**	4.65
	15×15	95.32	5.79	**97.02**	3.84
	31×31	95.92	5.68	**97.81**	3.07
Meadow	7×7	92.96	7.17	**94.26**	5.85
	15×15	93.52	7.59	**96.76**	4.72
	31×31	93.36	6.61	**96.90**	4.15
Field	7×7	94.46	6.11	**95.28**	5.84
	15×15	94.98	5.89	**96.41**	4.92
	31×31	95.02	5.64	**97.08**	4.98
Road	7×7	**95.98**	5.08	95.65	5.85
	15×15	96.32	4.23	**97.16**	3.70
	31×31	96.46	4.01	**97.89**	3.24
Mountains	7×7	97.24	3.97	**98.41**	3.05
	15×15	98.02	3.72	**98.92**	2.39
	31×31	98.32	3.02	**99.08**	2.27
River	7×7	**94.02**	6.52	93.20	8.08
	15×15	**94.19**	5.19	93.48	6.83
	31×31	95.22	4.05	**95.91**	4.25
Sand	7×7	84.91	14.82	**86.90**	13.96
	15×15	85.44	14.39	**88.23**	11.02
	31×31	87.28	10.04	**88.82**	11.49
Sky	7×7	**90.54**	9.71	89.29	10.86
	15×15	**91.49**	8.19	90.02	9.62
	31×31	92.18	7.87	**92.73**	7.39

For texture classification was used artificial neural network of direct propagation with three hidden layers. The comparative results are shown in Table 2, where TR is True Recognition and HTER is Half Total Error Rate, which combines the False Rejection Rate (FRR) and the False Acceptance Rate (FAR).

The results presented in Table 2 show that the CNs approach classifies the homogeneous textures close to statistical approach. This can be explained by the fact that CNs build almost regular graphs on homogeneous textures. Mean TR values are in the ranges 84.91–98.32 and 87.29–99.08 for conventional texture analysis and proposed CNs analysis, respectively (in Table 2, the best true recognition results are highlighted in bold). Images with explicit fractal structures provide the best recognition accuracy. In our case, there are “Mountains” and “Forest”. Also good results were achieved due to the color differences of the class samples in this limited dataset, despite the available color variations within the same class. The main advantages of our approach are the ability to reflect the physical nature of textures and low computational costs at the training step due to the use of simple metrics for classification. Multiscaling increases the recognition accuracy slightly and decreases the HTER values in a statistical sense, primarily due to the fractal properties of natural textures.

6. Conclusions

In this research, we study the application of CNs theory for texture classification in aerial photographs. We propose a new function for calculating the weights of the edges and introduce a CN training based on multiscale and multilayer properties. For experiments, own dataset was created from the video sequences “Drone Videos DJI Mavic Pro Footage in Switzerland” and photographs of the “Aerial Textures”. It includes a set of texture patches belonging to the classes “Forest”, “Meadow”, “Field”, “Road”, “Mountains”, “River”, “Sand”, and “Sky”. The conducted experiments show that despite the CNs approach classifies the natural textures close to statistical approach, the proposed CNs reflect the physical nature of textures and have low computational costs at the training step due to the use of simple metric for classification (Minkowski distance). In general, the results of true recognition are 1–3% higher compared to conventional texture analysis.

References

[1] Lofstedt T., Brynolfsson P., Asklund T., Nyholm T., Garpebring A. Gray-level invariant Haralick texture features // PLoS ONE. 2019. Vol. 14. No. 2. P. e0212110.1–e0212110.18.

[2] Haralick R.M., Shanmugam K. Dinstein I. Textural features for image classification // IEEE Trans. Systems, Man and Cybernetics. 1973. Vol. 3. P. 610–621.

[3] Cross G.R., Jain A.K. Markov random field texture models // IEEE Trans. Pattern Anal. Mach. Intell. 1983. Vol. 5. No. 1. P. 25–39.

[4] Liu L., Fieguth P., Guo Y., Wang X., Pietikäinen M. Local binary features for texture classification: Taxonomy and experimental study // Pattern Recognition. 2017. Vol. 62. P. 135–160.

[5] Backes A.R., Gonçalves W.N., Martinez A.S., Bruno O.M. Texture analysis and classification using deterministic tourist walk // Pattern Recognition. 2010. Vol. 43. No. 3. P. 685–694.

[6] Goncalves W.N., Backes A.R., Martinez A.S., Bruno O.B. Texture descriptor based on partially self-avoiding deterministic walker on networks // Expert Systems with Applications. 2012. Vol. 39. P. 11818–11829.

[7] Thewsuwan S., Horio K. Texture classification by local spatial pattern mapping based on complex network model // International Journal of Innovative Computing, Information and Control. 2018. Vol. 14. No. 3. P. 1113–1121.

[8] de Lima G.V.L., Saito P.T.M., Lopes F.M., Bugatti P.H. Classification of texture based on Bag-of-Visual-Words through complex networks // Expert Systems with Applications. 2019. Vol. 133. P. 215–224.

[9] Ribas L.C., de Mesquita Sá Junior J.J., Scabini L.F.S., Bruno O.M. Fusion of complex networks and randomized neural networks for texture analysis // Pattern Recognition. 2020. Vol. 103. P. 107189.1–107189.10.

[10] Cuadros O., Botelho G., Rodrigues F., Neto J.B. Segmentation of large images with complex networks // 25th Conference on Graphics, Patterns and Images (SIBGRAPI). IEEE, Ouro Preto, Brazil, 2012. P. 1–8.

[11] Mourchid Y., Hassouni M.El., Cherifi H. A general framework for complex network-based image segmentation // Multimed. Tools. Appl. 2019. Vol. 78. P. 20191–20216.

[12] Antiqueira L., Oliveira O.N.Jr., da Fontoura Costa L., das Graças Volpe Nunes M. A complex network approach to text summarization // Information Sciences. 2009. Vol. 179. No. 15. P. 584–599.

[13] Chalumeau T., Costa L.F., Laligant O., Meriaudeau F. Texture discrimination using hierarchical complex networks // In: Damiani E., Yétongnon K., Schelkens P., Dipanda A., Legrand L., Chbeir R. (Eds). Signal Processing for Image Enhancement and Multimedia Processing / Multimedia Systems and Applications Series. Springer, Boston, MA. 2008. Vol. 31. P. 543–550.

[14] Costa L. da F., Rodrigues F.A., Travieso G., Villas Boas P.R. Characterization of complex networks: a survey of measurements // Advances in Physics. 2007. Vol. 56. No. 1. P. 167–242.

[15] Scabini L.F.S., Condori R.H.M., Goncalves W.N., Bruno O.M. Multilayer complex network descriptors for color-texture characterization // CoRR arXiv preprint. 2018. arXiv:1804.00501v1. P. 1–20.

[16] Popescu D., Ichim L. Image recognition in UAV application based on texture analysis // In: Battiato S., Blanc-Talon J., Gallo G., Philips W., Popescu D., Scheunders P. (Eds) Advanced Concepts for Intelligent Vision Systems (ACIVS 2015) / Lecture Notes in Computer Science. Springer, Cham, 2015. Vol. 9386. P. 693–704.

[17] Drone Videos DJI Mavic Pro Footage in Switzerland. Available at: https://www.kaggle.com/kmader/drone-videos (accessed June 20, 2021).

[18] Aerial Textures. Available at: https://rendering.ru/aerial-textures.html (accessed June 20, 2021).

[19] Favorskaya M.N., Buryachenko V.V. Patch-based adaptive image and video watermarking schemes on handheld mobile devices // Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIV-2/W1-2021. 2021. P. 47–53.