Information-Theoretical Technique for Optimizing Segmentation Quality

D. Murashov

Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, 119333, Vavilov street., 40, Moscow, Russia

Abstract

In this paper, a problem of image segmentation quality is considered. The problem of segmentation quality is viewed as selecting the best segmentation from a set of images generated by segmentation algorithm at different parameter values. A technique for selecting the best segmented image is proposed. Information redundancy measure is used as a criterion for optimizing segmentation quality. It is shown that proposed method for constructing the redundancy measure gives the extremal properties. Computing experiment confirmed that the segmented image corresponding to a minimum of redundancy measure produces the suitable dissimilarity when compared with the original image. The segmented image which was selected using the proposed criterion, gives the highest similarity with the ground-truth segmentations, available in the database.

Keywords: image segmentation; segmentation quality; redundancy measure; variation of information

1. Introduction

The paper deals with the problem of image segmentation quality. According to Haralik and Shapiro [1], segmentation is the process of partitioning image represented as a region Ω into n non-overlapping subregions Ω_1, Ω_2,..., Ω_n. The elements in subregions are grouped by some feature and differ from the elements of the adjacent areas. Formal definition of segmentation is given in [2]. Any of segmentation algorithms has one or more parameters. Parameters should be set in order to provide the best quality of the segmentation result. The problem of setting parameters is rather difficult. In this work, we formulate the problem of segmentation quality as follows. Suppose, for a given input image U we obtain a set of Q segmented images $V = \{V_1, V_2, ..., V_q, ..., V_Q\}$. It is necessary to choose image V_q providing minimum for a given performance criterion $M(U, V_q)$:

$$q_{\min} = \arg\min_q \left(M(U, V_q) \right),\ q = 1, 2, ..., Q.$$

When solving different tasks of image analysis, suitable quality criterion should be applied. This may be a visual evaluation of an expert or any quantitative measure. The results of segmentation are usually compared with an image partitioned manually and accepted as ground-truth [3]. The quality can be represented by parameters describing boundary detection error, region consistency, and segment covering. In papers [4, 3] the authors used precision-recall framework for comparing segment boundaries. In [5] Martin et al. proposed global and local consistency errors as the measures for comparing segments in images generated by segmentation algorithm and ground-truth segmentations. Some other measures for evaluating segmentation quality are discussed in works [6, 7].

If the segmentation operation is considered as clustering of pixels, then the set-theoretical, statistical, and information-theoretical measures proposed to compare data clustering results, are used [8]. The most commonly used are: chi-square measure; Rand Index [9] and its variants [10]; Fowlkes-Mallows measure [11]; mutual information and normalized mutual information [12]; variation of information [13]. These measures allows one to compare different versions of partitioning image into non-overlapping regions. In paper [3], the authors noted that the standard methodology for estimating efficiency of segmentation algorithms is not yet developed.

In paper [14] another approach is proposed. Parameters of the superpixel segmentation algorithm [15] were chosen depending on the result of estimating similarity of segmented and original images. As a measure of similarity the authors proposed to use weighted uncertainty index computed using the values of the normalized mutual information [12] between the color channels of the original and segmented images. Frosio and Ratner proposed to choose parameter value providing the best segmentation in terms of visual perception. The dependence of the uncertainty index on parameter value (and accordingly, the number of the subregions) is approximately monotonous (see [14]). Using expert estimations of segmentations at various parameter values, the areas of under-segmentation, over-segmentation, and optimal segmentation were defined in the space "parameter - uncertainty index". Then at the processing phase an iterative procedure was applied to obtain parameter value providing the best image partition. The drawbacks of this method are related with the necessity of the training procedure. Segmentation algorithm produces acceptable results only for those types of images that were involved in the training process.

If the segmentation method is developed according to the criteria of visual perception, then a model of the human visual system should be used. It is also preferable to have peaked or concave dependency of the segmentation system's performance index on the parameter value.

In work [16], a theoretical-information model of the human visual system is proposed. The model is based on Barlow hypothesis [17] about minimizing data redundancy at the early stages of signal processing in the human visual system.

In this work, basing on principle of minimizing data redundancy [16], we propose to use a measure of information redundancy as a segmentation quality criterion. We show that a particular method of forming information-theoretical criterion provides it with extremum. In order to demonstrate that the segmented images corresponding to the minima of the redundancy measure yield acceptable dissimilarity with the original images and ground-truth segmentations, we conduct an experiment on images taken from Berkeley Segmentation Dataset BSDS500 [3].

2. Optimization of Segmentation Quality

Let the initial and segmented images be the input and the output of a stochastic information system. Levels of lightness in images are the continuous random variables U and V with probability mass functions of $p(u)$ and $p(v)$, where u and v are the values of U and V, respectively. Operation of segmentation can be represented by an information channel model:

$$V = F(U + \eta), \tag{1}$$

where U is an input signal, V is a channel output, F is a transformation function, and η is a channel noise. We assume that noise η is Gaussian random variable with zero mean value and variance σ_η^2; variables V and η are independent.

We propose to use a redundancy measure as a criterion of segmentation quality. The redundancy measure is defined as follows [16]:

$$R = 1 - \frac{I(U,V)}{C(V)}, \tag{2}$$

where $I(U;V)$ is a mutual information between the system input and output, $C(V)$ is a channel capacity. We take $C(V) = H(V)$, where $H(V)$ is an entropy of the output. Then, taking into account the fact $I(U;V) = H(V) - H(V \mid U)$, expression (2) takes the form:

$$R = \frac{H(V \mid U)}{H(V)}, \tag{3}$$

where $H(V \mid U)$ is a conditional entropy of the output V under condition that the input is equal to U.

We will show that the redundancy measure of the segmentation system described by the model (1-3) depends on number of segments and can have a minimum.

Probability mass function of the output may be represented by a sum

$$p(v) = \sum_{k=1}^{K} P(v_k)\delta(v - v_k), \tag{4}$$

where $P(v_k)$ is a probability of lightness value v_k assigned to pixels of a segment having number k, $\delta(v - v_k)$ is a delta-function, K is a number of segments in the output image.

To find analytic dependence $R(K)$, we will use a continuous version of model (1). Taking into account expression (4), differential entropy of the output can be written as follows:

$$H(V) = -\int_{-\infty}^{+\infty} p(v)\log p(v)dv = -\sum_{i=1}^{K}\left[P(v_i)\log P(v_i)\right]. \tag{5}$$

Let all values v_i be equiprobable: $P(v_i) = 1/K$. Then it follows from (5) that

$$H(V) = \log K. \tag{6}$$

Next, we shall find an expression for differential conditional entropy $H(V \mid U)$. Conditional entropy $H(V \mid U)$ is a measure of information about signal noise η measured at the system output. In this case, we may take [18]:

$$H(V \mid U) = H(\eta). \tag{7}$$

Differential entropy of the Gaussian noise is equal to [18]:

$$H(\eta) = \frac{1}{2}\left[\log e + \log(2\pi\sigma_\eta^2)\right], \tag{8}$$

where σ_η^2 is a variance of the system noise.

We assume that the probability mass function of the input image lightness is represented as a Gaussian mixture model of K components, which may overlap partially. The components of the mixture correspond to the segments of the output image V. Areas of component overlapping generate noise η. The overlapping areas are formed by pixels of U having the same lightness values, but related to different segments in image V. Substituting (6)-(8) into (3), we get the following expression for redundancy measure:

$$R(K) = \frac{\log e + \log(2\pi\sigma_\eta^2)}{2\log K}. \tag{9}$$

It follows from (9) that the redundancy measure depends linearly on logarithm of system noise variance and inversely on logarithm of number of produced segments K. Function (9) have minimum at a point $K_{\min}$ if the noise variance σ_η^2 is close to zero at small K and rapidly grows when K increases. Computing experiments confirmed that such behavior of the noise variance is taking place.

Taking into account dependency of the redundancy measure R on number of segments K, the best segmented image should be selected in the following way. Suppose that chosen segmentation algorithm has parameter t affecting the partitioning. The input image U is segmented at different parameter values. As a result, a set of Q segmented images $\mathcal{V} = \{V_1, V_2, ..., V_Q\}$ is obtained. Next, for input image U and each of the segmented images $V_q, q = 1, 2, ..., Q$, the redundancy measure R is computed. We choose image V_q providing minimum to R : $R(V_q) = R_{\min}$. Image V_q partitioned into $K_{\min}$ segments fits parameter value $t = t_{\min}$.

3. Computing Experiment

A method for choosing the best variant of segmentation is applied to the superpixel algorithm SLIC (Simple Linear Iterative Clustering) [19] supplemented with the post-processing procedure. In order to merge superpixels into homogeneous regions corresponding to objects in the original image, a two-step post-processing procedure is proposed.

At the first step neighboring superpixel areas are merged. For making a decision on merging, a threshold decision rule is used. This rule allows merging if the following inequality is taking place:

$$d_c(C_i, C_j) \le t, \tag{10}$$

where $d_c(C_i, C_j)$ is the distance between centers of adjacent superpixels with numbers i and j in the selected color space; t is a threshold value.

In the experiments, we used 25 images from the Berkeley Segmentation Dataset BSDS500 [3] transformed to CIE *Lab* color space. The experiment includes three stages. One of the test images is shown in Figure 1.

Fig. 1. A test image from BSDS500 dataset.

At the first stage, each of the test images is segmented using algorithm SLIC and post-processing procedure at different values of parameter t. Each of the images generates a set of Q segmented images $\mathcal{V} = \{V_1, V_2, ..., V_Q\}$. For input image U and each of the segmented images $V_q, q = 1, 2, ..., Q$ the redundancy measure R is computed. To involve all color channels, we use the weighted version of the redundancy R:

$$R_w(U, V_q) = \frac{R_L H_L(U) + R_a H_a(U) + R_b H_b(U)}{H_L(U) + H_a(U) + H_b(U)}, \tag{11}$$

where R_i is the redundancy measure determined in color channel $i \in \{L, a, b\}$ of images U and V_q; H_i is the entropy of the color channel i of the input image.

We apply SLIC algorithm and the first step of post-processing procedure to all test images. For each of the test images a set of segmented images is generated at initial superpixel size $a = 16$ pixels, parameter $m = 2$ (for details see [19]), and threshold values t changing in the range $0 \le t \le 3.6$ with increment equal to 0.2. Relationship between threshold t and number of segments K in images V_q generated by one of the test images is shown in Figure 2.

For each test image and related set of segmented images we computed the weighted redundancy measure R_w. Dependency of measure R_w on number of segments K for the test image shown in Figure 1 is depicted in Figure 3. Minimum of R_w is reached at $K = 55$ corresponding to threshold value $t = 2.6$.

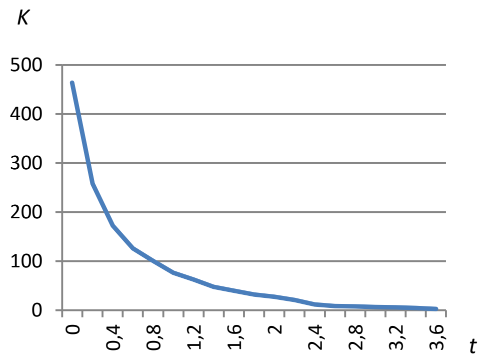


Fig. 2. Relationship between threshold value t and number of segments K.

At the second stage, segmentation quality is estimated. We estimate the amount of information about the input image, which was lost in segmentation process. For this purpose we compare the set of Q segmented images with the input image U using normalized version of variation of information proposed in [13] for comparing clusterings. This metric was also used in [3] for comparing segmented images. Here we use the weighted index based on this metric:

$$VI_w(U,V_q) = \frac{VI_L H_L(U) + VI_a H_a(U) + VI_b H_b(U)}{H_L(U) + H_a(U) + H_b(U)}, \qquad (12)$$

$$VI_i(U,V_q) = \frac{H_i(U) + H_i(V_q) - 2I_i(U,V_q)}{H_i(U,V_q)}, \qquad (13)$$

where $VI_w(U,V_q)$ is the weighted variation of information; VI_i is the distance between color channels i of images U and V_q; I_i is their mutual information; $H(U,V_q)$ is the joint entropy.

In order to estimate the distance between the input and the segmented images, we compute weighted normalized variation of information (12-13). The curve representing $VI_w(U,V_q)$ as the function of number of segments is depicted in Figure 3 by dashed line. One can see that distance between the input and segmented image decreases when K grows and become nearly stable at $K_{\min}$ corresponding to minimal redundancy value.

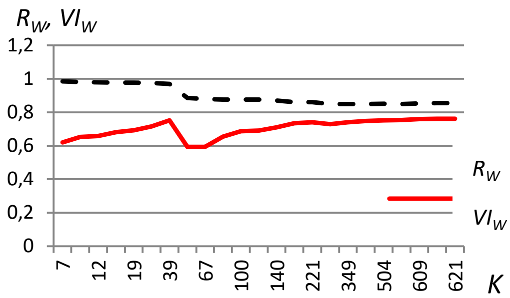


Fig. 3. Dependency of redundancy R_w and normalized variation of information VI_w on number of segments K for image shown in Figure 1.

At the third stage, using the weighted index (12) based on metric (13), we compare a set of Q segmented images with the ground-truth segmentations V_s^{GT}, $s = 1, 2, ..., S$, (S is a number of ground-truth segmentations for a test image U) available in BSDS500 dataset. The result of comparing segmentations of image shown in Figure 1 is represented in Figure 4 as the curves reflecting relationship between normalized variation of information $VI_w(V_s^{GT}, V_q)$, $q = 1, 2, ..., Q$, $s = 1, 2, ..., S$, and number of segments K in images V_q. It can be seen from Figure 4 that for the majority of the ground-truth segmentations, the distance

$VI_w(V_s^{GT}, V_q)$ is minimal when image V_q is partitioned into 55 segments. This V_q gives minimum to redundancy measure R_w. Taking into account the fact that ground-truth segmentations were produced manually, we can conclude that the proposed technique allows one to obtain the best segmentation in terms of visual perception.

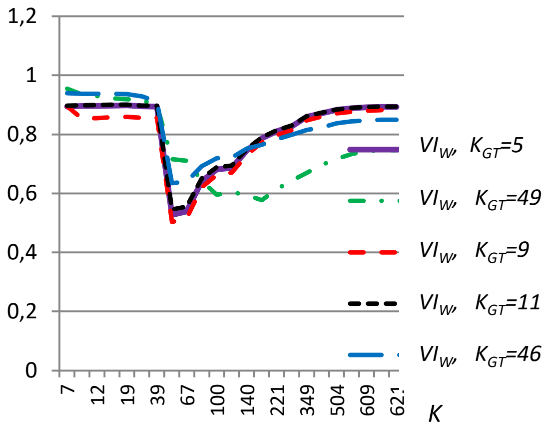


Fig. 4. Normalized variation of information $VI_w(V_s^{GT}, V_q)$ computed for segmented images V_q and ground-truth segmentations with different number of segments K_{GT}.

Ground-truth segmentation of image shown in Figure 1 and segmented image fitting of the redundancy measure minimum, are depicted in Figure 5.

(a) (b)

Fig. 5. Segmented and ground-truth images: (a) segmented image from Figure 1, $K = 55$; (b) ground-truth segmentation, $K_{GT} = 9$.

To show the efficiency of the proposed technique, the following relative difference is introduced:

$$\Delta K_{rel} = \frac{K_{\min} - K_{\min}^{GT}}{K_{\max}}, \tag{14}$$

where $K_{\min}$ is a number of segments corresponding to $R_{\min}$; $K_{\min}^{GT}$ is a number of segments in image V_q, which corresponds to the minimum of distance $VI_w(V_t^{GT}, V_q)$; $K_{\max}$ is the highest possible number of segments in images V_q obtained from input image U. For example, for image shown in Figure 1 $K_{\min} = 55$, $K_{\max} = 621$, and $K_{\min}^{GT} = 181$ for the ground-truth segmentation with number of segments $K_{GT} = 49$; $K_{\min}^{GT} = 55$ for other ground-truth segmentations (see Figure 4). Histogram of ΔK_{rel} values computed for 25 test images and 125 ground-truth segmentations (5 ground-truth segmentations per each of the test images) is depicted in Figure 6. Figure 6 shows a sufficiently large group of test images such that magnitude of ΔK_{rel} is near zero. The ground-truth segmentations of these images are close enough in the sense of measure (12-13) to segmentations, which minimize redundancy R_w.

4. Rough Evaluation of Application Domain

It follows from Figure 6 that the proposed technique combined with SLIC segmentation algorithm cannot provide ΔK_{rel} defined by (16) close to zero for all of the test images. Segmentation quality depends on image content and properties of the applied segmentation algorithm. It is necessary to find a criterion for assessing possibility to obtain good segmentations in terms of visual perception. Criterion value should be computed using characteristics extracted from input image. We propose to use average entropy of image channels in CIE *Lab* space as a criterion:

$$H_{avg}(U) = \frac{H_L(U) + H_a(U) + H_b(U)}{3},$$

where H_i is the entropy of the color channel i of the input image, $i \in \{L, a, b\}$. In Figure 7 a diagram plotted in axes $(H_{avg}, |\Delta K_{rel}|)$ is shown. The values of H_{avg} and ΔK_{rel} are computed from test images used in the previous section. One can see from Figure 7 that SLIC algorithm equipped with postprocessing procedure and redundancy measure as a segmentation quality criteria, gives $|\Delta K_{rel}| < 0.1$ for input images having $H_{avg} < 3.5$ or $H_{avg} > 3.85$. Computing the value of H_{avg} does not demand input image segmentation.

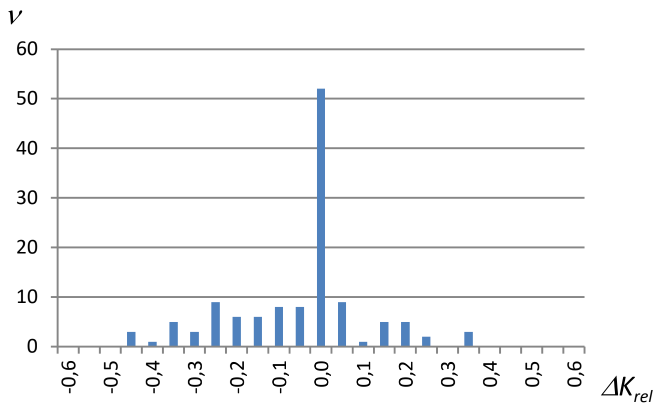
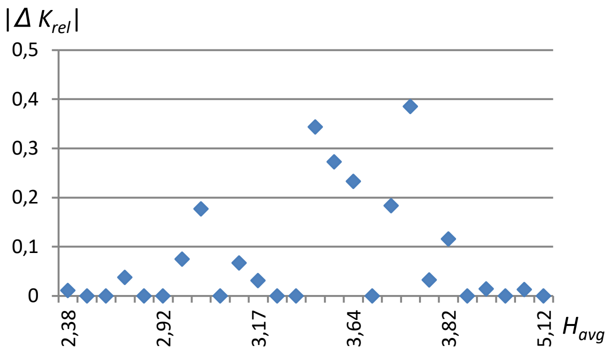


Fig. 6. Histogram of ΔK_{rel} values computed for 25 test images and 125 ground-truth segmentations; ν is a frequency of occurrence of particular ΔK_{rel}



value.

Fig. 7. Relation between H_{avg} and ΔK_{rel}.

5. Conclusion

The problem of image segmentation quality was considered. The problem of segmentation quality was studied as a task of selecting the best segmentation from a set of images generated by segmentation algorithm at different parameter values.

A technique based on theoretical-information criterion was proposed for selecting the best segmented image. We proposed to use information redundancy measure as a performance criterion. It was shown that the proposed way of constructing the redundancy measure provides the performance criterion with extremum. Computing experiment was conducted using images from the Berkeley Segmentation Dataset. The experiment confirmed that the segmented image corresponding to a minimum of redundancy measure, produced the acceptable information dissimilarity when compared with the original image. The segmented image which was selected using the proposed criteria, gives the minimal distance from the majority of ground-truth segmentations available in BSDS500 database.

For generating sets of partitioned images with different number of segments we used SLIC segmentation algorithm with the post-processing procedure. The approximate criterion for evaluating possibility to obtain good segmentation is given. The proposed technique of optimizing segmentation quality can be combined with other segmentation algorithms having parameters for tuning.

The future research will be aimed at the improving segmentation noise model and more precise evaluating the boundaries of application domain.

Acknowledgements

The research was supported in part by the Russian Foundation for Basic Research (grants No 15-07-09324 and No 15-01-04671).

References

[1] Haralick, R. M. Image segmentation techniques / R. M. Haralick, L. G. Shapiro // Computer vision, graphics, and image processing.—1985. —Vol. 29, no. 1. —P. 100–132.

[2] Gonzalez, R. C. Digital image processing: Pearson prentice hall / R. C. Gonzalez, R. E. Woods // Upper Saddle River, NJ. —2008

[3] Arbelaez, P. Contour detection and hierarchical image segmentation / P. Arbelaez, M. Maire, C. Fowlkes, J. Malik // IEEE transactions on pattern analysis and machine intelligence. —2011. —Vol. 33, no. 5. —P. 898–916.

[4] Martin, D. R. Learning to detect natural image boundaries using local brightness, color, and texture cues / D. R. Martin, C. C. Fowlkes, J. Malik // IEEE transactions on pattern analysis and machine intelligence. —2004. —Vol. 26, no. 5. —P. 530–549.

[5] Martin, D. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics / D. Martin, C. Fowlkes, D. Tal, J. Malik // Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on / IEEE. —Vol. 2. —2001. —P. 416–423.

[6] Zhang, H. Image segmentation evaluation: A survey of unsupervised methods / H. Zhang, J. E. Fritts, S. A. Goldman // Computer vision and image understanding. —2008. —Vol. 110, no. 2. —P. 260–280.

[7] Csurka, G. What is a good evaluation measure for semantic segmentation? / G. Csurka, D. Larlus, F. Perronnin, F. Meylan // BMVC. —Vol. 27. — 2013. —P. 32.1–32.11.

[8] Comparing Clusterings – An Overview: Rep. : 2006-04 / Universit¨at Karlsruhe (TH); Executor: S. Wagner, D. Wagner : 2007. — URL: http://digbib.ubka.uni-karlsruhe.de/volltexte/1000011477.

[9] Rand, W. M. Objective criteria for the evaluation of clustering methods / W. M. Rand // Journal of the American Statistical association. — 1971. —Vol. 66, no. 336. —P. 846–850.

[10] Unnikrishnan, R. A measure for objective evaluation of image segmentation algorithms / R. Unnikrishnan, C. Pantofaru, M. Hebert // Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05) - Workshops - Volume 03. — CVPR ’05. — Washington, DC, USA : IEEE Computer Society, 2005. — P. 34–41. — URL: http://dx.doi.org/10.1109/CVPR.2005.390.

[11] Fowlkes, E. B. A method for comparing two hierarchical clusterings / E. B. Fowlkes, C. L. Mallows // Journal of the American statistical association. —1983. —Vol. 78, no. 383. —P. 553–569.

[12] Ana, L. N. F. Robust data clustering / L.N.F. Ana, A.K. Jain // Computer Vision and Pattern Recognition, 2003. Proceedings. 2003. IEEE Computer Society Conference on / IEEE. —Vol. 2. — 2003. —P. 128–133.

[13] Meilă, M. Comparing clusterings: an axiomatic view / M. Meilă // Proceedings of the 22nd international conference on Machine learning / ACM. — 2005. —P. 577–584.

[14] Frosio, I. Adaptive segmentation based on a learned quality metric. / I. Frosio, E. R. Ratner // VISAPP’2015.—Vol. 1.— 2015. —P. 283–292.

[15] Felzenszwalb, P. F. Efficient graph-based image segmentation / P. F. Felzenszwalb, D. P. Huttenlocher // Int. J. Comput. Vision. — 2004. —Vol. 59, no. 2. —P. 167–181. —URL: http://dx.doi.org/10.1023/B:VISI.0000022288.19776.77.

[16] Atick, J. J. Towards a theory of early visual processing / J.J. Atick, A.N. Redlich // Neural Computation.—1990.—Vol. 2, no. 3. — P. 308–320.

[17] Barlow, H. B. Possible principles underlying the transformations of sensory messages / H.B. Barlow. — 1961. — P. 217–234.

[18] Haykin, S. Neural Networks: A Comprehensive Foundation. / S. Haykin. —2nd edition. —Upper Saddle River, NJ, USA : Prentice Hall PTR, 1998.

[19] Achanta, R. SLIC superpixels compared to state-of-the-art superpixel methods / R. Achanta, A. Shaji, K. Smith, et al. // IEEE transactions on pattern analysis and machine intelligence. —2012. —Vol. 34, no. 11. —P. 2274–2282.