Research Journal of Applied Sciences 10 (8): 385-390, 2015
ISSN: 1815-932X

Concerning the Issue of Object Classification by Form Part 2

Vyacheslav Borisovich Fofanov and Alexey Nikolaevich Zhiznevsky
Kazan (Volga Region) Federal University, Kremlevskaya Street 18, Kazan, Russia

Abstract: In this second part of the research, the definition of form proposed in the first part is used for vehicle classification. The background information for this task solution is the images generated by optical-electronic systems of aircrafts. Obviously, the choice of a particular trait in particular the shape, starts with the study of its information content. In this study, under the informative form its ability to provide the correct classification of objects is meant. Therefore, the value of the information content is presented by the probability of correct classification. The presented results of computer experiments concerning the assessment of the information content for the mixture consisting of selected objects showed a high information content of a form sign. The set of objects is meant by the scene. Each object is a fragment of a homogeneous random field with its mean value and correlation function. It is assumed that an object correlation function tends to zero as the distance increases between random variables. It allows to evaluate the unknown arithmetic mean value of an object, calculated according to its image. From these assumptions, it follows that a scene from a formal point of view, is a locally homogeneous random field. The projection of an object, developed as the result of segmentation according to scene presentation usually differs from the previously prepared projection from the base of projections. This stems from the fact that some projection pixels may be erroneously classified to the background and some background pixels may also by erroneously classified to the object. The magnitude of discrepancy depends on the properties of the scene, determining its complexity. To assess the complexity, it is proposed to use the probability of correct projection form recognition, developed according to the scene image. The research provides, the results of computer experiments to assess the complexity of a scene, depending on the magnitude of a correlation radius and the signal/noise ratio. They show that despite the high potential of informative features of the form the classification result depends largely on the quality of the image that reflects the actual properties of the scene.

Key words: Classification by form, informative content of the form, the likelihood of an object correct classification, scene model, the search of interest areas, the segmentation of interest zone

INTRODUCTION

The shape of objects forming a scene is an important unmasking feature. The optoelectronic systems form an object projection image onto a plane perpendicular to the direction of observation. Therefore from a formal point of view, the classification of objects by form is reduced to the comparison of their projection form. One of possible approaches for the solution of this problem including the formalization of the form proposed in the first part of this study. In this second part, the use of this approach is considered to classify vehicles.

The source information for classifying is the scene images obtained from the aircraft at different angles to horizon. It is assumed that vehicles form a set (mixture) consisting of 3302 GAZ, KamAZ and UAZ 4310 469 vehicles. They are shown by Fig. 1-3. The information about the orientation of vehicles is not available.

Fig. 1: GAZ 3302

Fig. 2: KamAZ 4310

Corresponding Author: Vyacheslav Borisovich Fofanov, Kazan (Volga Region) Federal University, Kremlevskaya Street 18, Kazan, Russia

Fig. 3: UAZ 469

During the object projection classification the scenes constructed by pictures were compared to preliminary prepared projections from the projection base. The projection base constructed according to 3d models of real objects taken from the Internet was used for the experiments. When you create a database of projections, it is assumed that the object is in the center of the spherical coordinate system. The observation point is determined by the angle φ between the polar axis and the projection of an optical axis on the horizontal plane (azimuth), the angle θ between the optical axis and its projection on the horizontal plane (location angle) and the distance to the system center p.

The database stores the projection objects for 36 azimuth meanings (from 0-360° with an increment of 10°), 10 values of the place angle (from 0-90° with the increments of 10°) and a two-scale recording. The computer experiments were carried out using the image corresponding to two scales of recording. In one case, the side of a square pixel in the object plane makes 0.1 m in another one it makes 0.3 m. Thus for each type of objects the base stores 720 projections. The total number of projections in the database for three objects makes 2160.

INFORMATIVE CONTENT OF THE FORM

In this case, the mixture is composed of three types of objects visible at different location angles. We assume that all cars GAZ 3302 form G class that all vehicles KamAZ 4310 belong to the class named K and all the cars named UAZ 469 belong to the class U. The decision on the use of an object shape as a sign depends on the answer to the question about the extent to which a single projection of an object determines its class. The ability to determine the class of an object by its projections is called the informative content I of the form sign. It will be measured by the probability of correct classification for an object from a particular mixture. It's obvious that $0 \le I \le 1$.

If P(G), P(K) and P(U) are the a priori probabilities of classes (the shares of objects for the corresponding classes in the mixture) and P(G|G), P(K|K) and P(U|U) are the probabilities of a correct classification for G, K and U objects, respectively, the information content of I (the probability of correct mixture object classification) takes the following form:

$$I = P(G)P(G|G)+P(K)P(K|K)+P(U)P(U|U)$$

Table 1: Relative frequency of correct classification for G, K and U object classes depending on the elevation angle and resolution

	Resolution (m)					
	P(U|U)		P(K|K)		P(G|G)	
Elevation angle	0.1	0.3	0.1	0.3	0.1	0.3
0	0.939	0.975	1.000	0.997	0.964	1.000
10	0.878	0.967	1.000	1.000	0.953	0.994
20	0.964	0.992	1.000	1.000	0.897	0.919
30	0.986	0.986	1.000	1.000	0.911	0.939
40	0.994	0.975	1.000	0.994	0.969	0.944
50	0.992	0.997	1.000	0.997	0.847	0.983
60	0.992	0.994	1.000	0.997	0.994	1.000
70	1.000	1.000	1.000	1.000	0.997	1.000
80	1.000	0.997	1.000	0.997	1.000	1.000
90	1.000	1.000	1.000	1.000	1.000	1.000
Average	0.974	0.988	1.000	0.998	0.953	0.978

To estimate the probabilities of P(G|G), P(K|K) and P(U|U), we use the projection base. It is assumed that the azimuth φ, he elevation angle θ and the distance ρ to the object at the time of observation are known. An object orientation is defined by the angle ψ. At the time of observation, there is no information about it. Therefore, the angle ψ during the calculation is considered as a random value with a uniform distribution, taking the values from 0-350° inclusive with the increment of 10°. The shape of each projection from the base is compared to the method specified in (to test the hypothesis of homogeneity) with 108 base projections with the same resolution and elevation angle. It is obvious that the shape of the compared projection may coincide with the projections of different class objects. Therefore, the number of matches is counted for each class. The object belonged to the class for which the number of matches was the largest one. If there are several such classes any one of them is selected. The object classification is recognized as a proper one if the greatest number of matches falls on the projection of its class. At that the number of matches with the projections of the remaining classes is smaller. As an estimate of an object correct classification probability the ratio of an object projection number which ensure its correct classification to the total number of object projections equal to 36 (at a given location angle).

Table 1 shows the relative frequency of correct classification of objects for each class, depending on the elevation angle θ and resolution, calculated according to the results of fifty experiments. The level of significance (probability of true hypothesis rejection) in all experiments was = 0.05. The last line of the table shows the relative frequencies averaged for each column. It corresponds to

Table 2: Evaluation of informative content of the form feature at object classification from a mixture (resolution of 0.1 m)

	P(K)										
P(U)	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	0.953	0.958	0.962	0.967	0.972	0.976	0.981	0.986	0.991	0.995	1.000
0.1	0.955	0.96	0.964	0.969	0.974	0.979	0.983	0.988	0.993	0.997	
0.2	0.957	0.962	0.967	0.971	0.976	0.981	0.985	0.99	0.995		
0.3	0.959	0.964	0.969	0.973	0.978	0.983	0.987	0.992			
0.4	0.961	0.966	0.971	0.975	0.98	0.985	0.99				
0.5	0.964	0.968	0.973	0.978	0.982	0.987					
0.6	0.966	0.97	0.975	0.98	0.984						
0.7	0.968	0.972	0.977	0.982							
0.8	0.97	0.974	0.979								
0.9	0.972	0.977									
1.0	0.974										

Table 3: Evaluation of informative content of the form feature at object classification from a mixture (resolution of 0.3 m)

	P(K)										
P(G)	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	0.978	0.980	0.982	0.984	0.986	0.988	0.990	0.992	0.994	0.996	0.998
0.1	0.979	0.981	0.983	0.085	0.987	0.989	0.991	0.993	0.995	0.997	
0.2	0.980	0.982	0.984	0.986	0.988	0.990	0.992	0.994	0.996		
0.3	0.981	0.983	0.985	0.987	0.989	0.991	0.993	0.995			
0.4	0.982	0.984	0.986	0.988	0.990	0.992	0.994				
0.5	0.983	0.985	0.987	0.989	0.991	0.993					
0.6	0.984	0.986	0.988	0.990	0.992						
0.7	0.985	0.987	0.989	0.991							
0.8	0.986	0.988	0.990								
0.9	0.987	0.989									
1.0	0.988										

the case when the location angle is a random variable with a uniform distribution, taking the values from 0-90° inclusive with the increment of 10°. According to the data specified in Table 1, the most favorable elevation angle for observation is the angle of 90°.

Table 2 and 3 show (as illustrated) the estimates of probabilities for a proper classification (informative content) of an object for the blend with the chosen values of a priori probabilities P(G), P(K), P(U) and two resolutions. The probabilities of the last row of Table 1 were used as the probabilities P(G|G), P(K|K) and P(U|U). The entries to the Table 2 and 3 are represented by a priori probabilities P(U) and P(K). The third probability P(G) is determined from the following equation:

$$P(G)+P(K)+P(U) = 1$$

The comparison of data specified in Table 2 and 3 shows that at the triple deterioration of resolution (which is equivalent to a range increase) the informative content of a form sign decreases no more than by 0.025. Thus, the shape is sufficiently reliable telltale sign for the classification of objects from a specified mixture.

SCENE MODEL

The initial information about the vehicles (further referred to as objects) is a scene presentation. Each scene object is determined by a finite subset A of the integer grid Z^2 called its projection and a set of random variables $\xi_A = (\xi_\alpha)_{\alpha\in A}$ with the finite set values $Y = \{0, 1, ..., |Y|-1\}$. The family type $x_A = (x_\alpha)_{\alpha\in A}$, $x_\alpha \in Y$ is called an object ξ_A presentation and the family:

$$Y^A = \{x_A = (x_\alpha)_{\alpha\in A}: x_\alpha \in Y,\ \alpha\in A\}$$

is called the set of its all presentations. It is assumed that the object ξ_A is the fragment of a homogeneous random field with an average value m_A and covariance dispersion function K_A for which the conditions of Slutsky's ergodic theorem are fulfilled. Let (Kramer and Leadbetter, 1969):

$$B(\alpha, r) = \{z = (z_1, z_2)\in Z^2: |z_1-\alpha|,\ |z_2-\alpha_2| \le r\}$$

square neighborhood with center at the point $\alpha\in Z^2$ and the radius r, owned by A, $n = |B(\alpha, r)|$. Slutsky's theorem shows that the arithmetic mean $\bar{x}_\alpha$ of the form:

$$\bar{x}_\alpha = \frac{1}{n}\sum_{z\in B(\alpha,r)} x_z$$

is a consistent estimator of the mean value m_A. The set of all objects with non intersecting projections, the sum of which is equal to Z^2 is called a locally uniform scene.

In computer (numerical) experiments homogeneous random fields $(\xi_z)_{z\in Z^2}$ are established with the moving summation along a square neighborhood with the radius $\hat{r}$ at the family of independent random variables $(\xi_z)_{z\in Z^2}$. That is:

$$\xi_z = \frac{1}{\left|B(z,\hat{r})\right|} \sum_{t \in B(z,\hat{r})} \zeta_t, z \in Z^2$$

All random variables ζ_z have the same distribution with zero mean-value. For any t = (t_1, t_2)$\in Z^2$, the correlation function meaning:

$$R(t) = \left(E\xi_z\xi_{z+t}\right) / \sqrt{\left(E\xi_z^2\right)\left(E\xi_{z+t}^2\right)}$$

Of this field is calculated easily. Indeed:

$$R(t) = \begin{cases} \dfrac{\left|B(0,\hat{r}) \cap B(t,\hat{r})\right|}{(2\hat{r}+1)^2}, t \in B(0,2\hat{r}) \\ 0, t \in Z^2 / B(0,2\hat{r}) \end{cases}$$

Let d is the Euclidean distance on Z^2. The points z and t from Z^2 will be called the neighboring ones or the neighbors if d(z, t) = 1. Let call the point z from $A \subset Z^2$, a limit one if it has the neighbor t from Z^2/A. The number of boundary points of the set A will be called its boundary and denoted by the symbol Fr(A). Let Q is the square on Z^2. The family $\xi_Q = (\xi_z)_{z \in Q}$ is called the zone of interest for the object ξ_A if $A \subset Q/Fr(Q)$ if is a random field fragment with the mean $m_{Q/A}$ and the dispersion $\sigma_{Q/A}^2$ and if $m_A \neq m_{Q/A}$. At $m_A > m_{Q/A}$ an object is called a bright spot and at $m_A < m_{Q/A}$, it is called a dark spot. Apparently, the first mention of such objects appeared in (Cook and Rosenfeld, 1970; Pickett *et al.*, 1970).

In order to get an object projection by image, we use the approach proposed in (Fofanov, 2007; Fofanov *et al.*, 2007). In accordance with it a zone of interest is developed initially. Then, the segmentation for a zone of interest is performed, the result of which is an object projection. During the third step the comparison of projections by form is performed and the decision on vehicle belonging to one of the vehicle classes is taken.

ZONES OF INTEREST SEARCH EFFICIENCY

During the first stage of decoding for each object it is necessary to build a zone of interest. To do this, the method described in (Aleev *et al.*, 2011) is used. It is obvious that the success of this operation depends not only on the method applied but also on the properties (the complexity) of a scene. To describe the complexity, it is proposed to use the signal/noise ratio and the correlation radius. The value k of the signal/noise ratio for the object ξ_A within the zone of interest ξ_Q is calculated according to the following formula:

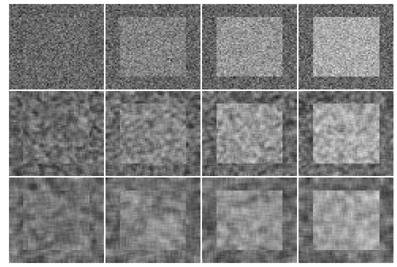

Fig. 4: Zone of interest example images

$$k = \frac{\left|m_A - m_{Q/A}\right|}{\sigma_A + \sigma_{Q/A}}$$

For the scenes obtained by sliding summation along a square neighborhood of the radius $\hat{r}$, the correlation between the pixels ξ_z and ξ_t at $|z-t| > 2\hat{r}\sqrt{2}$ is absent. Thus, $2\hat{r}\sqrt{2}$ is used as a correlation radius.

To illustrate these properties Fig. 4 shows one image of 12 areas of interest. They are united in a table of three rows and four columns. In the top row, there are the images of the scenes with $\hat{r}=0$. Further such a scene is called Bernoullian one. In accordance with the definition they form a family of mutually independent random variables. In the second and third lines, there are the images of the scenes with $\hat{r}=1$ and $\hat{r}=2$, respectively. For all the images of the first column k = 0.25 of the second column -0.75 of the third column 1.25 and of the fourth column 1.75. As we expected, the visibility of objects increases with the growth of k and decreases with the increase of $\hat{r}$.

At the deciphering instead of unknown average m_A and $m_{Q/A}$ their statistical estimates are used computed with the radius r within a neighborhood. In order to identify the dependence of correct classification probability concerning the square plots of scenes (prospective zones of interest) on k and r the computer experiments are performed using locally homogeneous random fields. The projection of the object was the square with a side of 40 pixels, the party l of a zone of interest is = 60 pixels. The experimental results for the scenes with $\hat{r}=0,1,2$ are presented in Table 4.

From the above data it follows that the smoothing radius r increase in all experiments leads to the increased probability of correct classification for a zone of interest, starting from a certain k value. For the scenes with $\hat{r}=0$, the value k = 0.13 for the scenes with 1 $\hat{r}=1$, the value k = 0.25 and for the scenes with $\hat{r}=2$, the value k = 0.50. At that the decrease of correct classification probability associated with the correlation radius increase is offset by the increase of signal/noise ratio.

Table 4: Estimates of correct classification probability for a zone of interest

	Signal/noise ratio k								
Radius r	0.13	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Bernoulli scene ($\hat{r}$ = 0)									
1	0	0.4	1	1	1	1	1	1	1
2	0.1	0.5	1	1	1	1	1	1	1
3	0.1	1	1	1	1	1	1	1	1
4	0.3	1	1	1	1	1	1	1	1
5	0.8	1	1	1	1	1	1	1	1
Locally uniform scene ($\hat{r}$ = 1)									
1	0.1	0.2	0.3	0.7	0.7	1	1	1	1
2	0.3	0.3	0.2	0.9	1	1	1	1	1
3	0.2	0.3	0.4	0.9	1	1	1	1	1
4	0.1	0.4	0.8	1	1	1	1	1	1
5	0	0.6	0.9	1	1	1	1	1	1
Locally uniform scene ($\hat{r}$ = 2)									
1	0	0	0.1	0.7	0.7	0.8	0.9	1	1
2	0	0	0.2	0.7	0.7	1	1	1	1
3	0	0.1	0.3	0.8	1	1	1	1	1
4	0	0	0.4	0.9	1	1	1	1	1
5	0.1	0.1	0.7	1	1	1	1	1	1

Table 5: Estimates of correct classification probability for a projection form

	Signal/noise ratio k												
Radiusr r	0.13	0.25	0.5	0.75	1	1.25	1.5	1.75	2	3	4	5	6
Bernoulli scene ($\hat{r}$ = 0)													
1	0	0.00	0.08	0.18	0.40	0.68	0.70	0.82	0.96	1.00	0.96	1.00	1.00
2	0	0.02	0.08	0.16	0.28	0.70	0.72	0.64	0.84	0.98	1.00	0.94	0.94
3	0	0.00	0.00	0.04	0.26	0.56	0.60	0.68	0.74	0.94	0.96	0.90	1.00
Locally uniform scene ($\hat{r}$ = 1)													
1	0	0.00	0.00	0.00	0.00	0.06	0.14	0.18	0.4	0.8	0.98	1.00	0.98
2	0	0.00	0.00	0.00	0.00	0.00	0.10	0.10	0.08	0.42	0.64	0.84	0.88
3	0	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.12	0.36	0.58	0.74
Locally uniform scene ($\hat{r}$ = 2)													
1	0	0.00	0.00	0.00	0.00	0.06	0.14	0.18	0.40	0.80	0.98	1.00	0.98
2	0	0.00	0.00	0.00	0.00	0.00	0.10	0.10	0.08	0.42	0.64	0.84	0.88
3	0	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.12	0.36	0.58	0.74

ZONE OF INTEREST SEGMENTATION EFFICIENCY

At the second stage of decoding, the segmentation of interest zone is performed in order to obtain an object projection. To do this, the method of spot was used presented in (Fofanov and Zhiznevskii, 2012). It should be borne in mind that the combining (pixel by pixel) of the projection, built according to an object image with the projection of the same object located in the database using the shifts and turns is not always possible. It may be turned out that in the course of segmentation some pixels of the projection will be referred (erroneously) to the background. On the other hand, some background pixels will be attached (also by error) to an object. It is obvious that at large differences between the projections an object will be classified in a wrong way.

To measure the differences between the projections, it is proposed to use the probability of a projection form correct classification developed according to an object image. If the shape of a projection, built according to the image of an object, matches with its projection form of the projection base, the classification is considered to be correct one. Otherwise, it is considered as an incorrect one. To identify the probability of correct classification dependence on the size of k, r and $\hat{r}$, the computer experiments were performed. They calculate the relative frequency of projection form correct classification, built according to the image of an object. The results are shown in Table 5.

SUMMARY

The preliminary data obtained during the studies showed that during the automatic classification of vehicles the use of a form feature provides the correct classification probability close to one for the scenes with a relatively low signal/noise ratio.

CONCLUSION

The performed studies for a shape informative content showed that the probability of a correct classification for vehicles by projection shape is close to one. However at the solution of real-world problems a decisive influence on this probability was performed by the quality of segmentation which depends on the complexity of a scene. The study presents the evaluations of correct classification probability based on the signal/noise ratio and the correlation radius of a locally uniform scene.

ACKNOWLEDGEMENT

The research is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

REFERENCES

Aleev, R.M., S.A. Martynov and V.B. Fofanov, 2011. Remarks on Searching Zones of Interest in Locally Uniform Scene. Pattern Recognition and Image Analysis, 21 (2): 212-215.

Cook, C.M. and A. Rosenfeld, 1970. Size Detectors Proc. IEEE, Letters, 58 (12): 1956-1957.

Fofanov, V.B., 2007. Formalization of a scene in the issue of multizone images decoding. Optical J., 74 (3): 51-54.

Fofanov, V.B., A.V. Demchenko and R.F. Kuleev, 2007. Interpretation of multispectral images: methods and results. Optical J., 74 (3): 55-59.

Fofanov, V.B. and A.N. Zhiznevskii, 2012. Segmentation of Regions of Interest on Locally Homogeneous Scenes. Pattern Recognition and Image Analysis, 22 (2): 257-264.

Kramer, G. and M. Leadbetter, 1969. Stationary random processes. The properties of selective functions and their applications: Trans. from English. M.: Mir.

Pickett, R.M., B.C. Lipkin and A. Rosenfeld, 1970. Visual Analysis of Texture in the Detection and Recognition of Objects Picture Processing and Psychopictorics New York: (Eds.) Academic Press, pp: 289-308.