DOI 10.1007/s10559-018-0010-7
Cybernetics and Systems Analysis, Vol. 54, No. 1, January, 2018

HEURISTIC CRITERION FOR CLASS RECOGNITION BY SPECTRAL BRIGHTNESS

A. I. Arkhipov,[1†] N. M. Glazunov,[2] and A. V. Khyzhniak[1‡]

UDC 519.21:681.3

Abstract. *The authors consider the problem of recognition of a class of objects by the results of multispectral measurements (spectral brightness of signals) and available spectral and statistical characteristics of the given classes. On the basis of probabilistic and statistical considerations, as well as quantization of continuous distributions, the heuristic recognition criterion is proposed. Based on the criterion, the heuristic method of recognition is presented. Modifications of the method are proposed to improve its reliability and efficiency.*

Keywords: *class of recognition objects, probability density function, normal distribution, multivariate density, variance, standard deviation, discrete distribution, alternative hypothesis, heuristic recognition criteria.*

INTRODUCTION

In the paper, we will represent a heuristic criterion for recognition of a class of objects by the results of measurement of their spectral brightness and known mean values of spectral brightness of reference classes of objects. Methods of probability theory and mathematical statistics [1–3] are actively applied to solve optimization problems, in statistical theories of identification, recognition, and training [4–6], as well as in number theory [7]. These theories use methods and algorithms of probability theory and mathematical statistics to acquire information and derive hypotheses that are tenable to some extent. One of the problems is handling data from multispectral GPS surveys [8] in order to class given objects among already known ones. Objects in the classes are described by appropriate attributes.

Practical needs and theoretical studies initiate new statements of recognition problems, form new problems, which in turn stimulate the development of other methods of their research. In the present paper, we will apply probability and statistical methods to represent heuristic recognition criterion for classes of objects of the specified recognition problem. The heuristic reasons being used are partially caused by analogies between number theory and probability theory (see for example [7, 9] and references therein). In brief, the content of the study is as follows: vector $\overline{X} = (X_1, X_2, \ldots, X_n)$ is current vector of averaged sampled values (of spectral brightness) of a random sample with unknown value of class k. Vectors $\overline{L_k}$ in the amount m are fixed vectors of averaged sampled values (of spectral brightness) of random samples from m known (reference) classes. A class with unknown value of k belongs to one of the reference classes. Based on $\overline{X}$ and $\overline{L_k}$, this value k can be calculated by the proposed heuristic method. In Sec. 1, the necessary concepts are introduced and problem statement is given; Sec. 2 represents the heuristics of the criterion with application of statistical hypotheses and quantization; Sec. 3 describes heuristic recognition criterion on the basis of measurements (of spectral brightness of signals) and quantization, heuristic recognition method based on this criterion, and modifications of the method in order to increase the recognition reliability and make the computations more efficient.

[1]State Institution "Scientific Centre for Aerospace Research of the Earth, Institute of Geological Science, National Academy of Sciences of Ukraine," Kyiv, Ukraine, [†]*aiarh19443@gmail.com*; [‡]*avsokolovska@i.ua*. [2]National Aviation University, Kyiv, Ukraine, *glanm@yahoo.com*. Translated from Kibernetika i Sistemnyi Analiz, No. 1, January–February, 2018, pp. 105–110. Original article submitted February 13, 2017.

 1060-0396/18/5401-0094

1. PROBLEM STATEMENT

Let there be *m* classes of objects subject to recognition, which are defined by *m* training samples containing N_k objects, respectively, where *k* is current class number. Let *n* be the number of measured values of attributes, for example, the quantity of spectral channels (zones) being used and *i* be current attribute number, for example, the number of specific spectral channel (zone). Let $(X_1, X_2, \dots, X_n)$ be averaged result of the measurement of values of attributes of the class of objects being recognized, for example, averaged result of measurement of spectral brightness (brightness) of the class of objects being recognized, with respect to the set of spectral channels being used. For class k, vector $L_k = (L_{k,1}, L_{k,2}, \dots, L_{k,n})$ is the result of current measurements of brightness of *n* attributes (for example, with respect to *n* spectral channels). We assume that vector components are independent and brightness randomly varies from class to class, taking values near their average values. The quantity $L_{k,i,s}$ means the result of measurement of object *s* belonging to class *k* with respect to the *i*th channel. Let us assume that the mean value of random variable $L_{k,i}$ is equal to

$$\bar{L}_{k,i} = \frac{1}{N_k} \sum_{s=1}^{N_k} L_{k,i,s},$$

for values $i = 1, \dots, n$ and combine them in vector $\overline{L_k} = (\bar{L}_{k,1}, \bar{L}_{k,2}, \dots, \bar{L}_{k,n})$.

The standard deviation is supposed to be equal to

$$\sigma_{k,i} = \left(\frac{1}{N_k - 1} \sum_{s=1}^{N_k} (L_{k,i,s} - \bar{L}_{k,i})^2 \right)^{1/2}.$$

Condition 1. Suppose that vector $(X_1, X_2, \dots, X_n)$ is random and its coordinates are independent, normally distributed random variables.

Remark 1. If we do not suppose a priori that the distribution of coordinates of vector $(X_1, X_2, \dots, X_n)$ is normal, then to check whether empirical values of coordinates correspond to the normal law, we can use sampled values and Pierson's χ^2-criterion or the Smirnov–Kolmogorov criterion [2, 3]. For the Sato–Tate distribution, such research is presented in [9] in detail.

Assume that Condition 1 of the distribution normality is satisfied. In this case, the probability density of random variable $X_{k,i}$ is

$$p_{k,i}(x_i) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{k,i}} e^{\frac{(x_i - \bar{L}_{k,i})^2}{2\sigma_{k,i}^2}}.$$

Under the same condition, multivariate density of vector $(X_{k,1}, X_{k,2}, \dots, X_{k,n})$ is defined by the formula

$$p_k(x_1, \dots, x_n) = p_{k,1}(x_1) \cdots p_{k,n}(x_n) = \left(\frac{1}{\sqrt{2\pi}} \right)^{n/2} \prod_{i=1}^{n} \left(e^{\frac{(x_i - \bar{L}_{k,i})^2}{2\sigma_{k,i}^2}} \frac{1}{\sigma_{k,i}} \right). \tag{1}$$

Problem. Let vector $\bar{X} = (X_1, X_2, \dots, X_n)$ be averaged sampled value of random sample with unknown value of class k. It is required to specify a criterion that allows determining the number k with the maximum confidence. In other words, for parameter k, the required criterion should find its most probable value $P_k = P_k(\bar{X})$. For example, if the class being recognized is a field occupied with vegetative culture (culture) with number k, then P_k is the probability that this field is actually occupied with culture with number k.

It should be unambiguously known that the considered class pertains to one of the classes with numbers $1, \dots, m$. The main problem is to determine the algorithm of calculating $P_1, \dots, P_m$ in terms of respective initial parameters $L_{k,i}$, $\sigma_{k,i}$, and L_k. This problem involves certain difficulties caused by continuous distribution of $L_{k,i}$.

The heuristics of the problem solution is described in the next section.

2. CRITERION HEURISTICS

Let us consider the case where all the classes k are different ($k = 1, \dots, m$). Let $\overline{L}_1, \dots, \overline{L}_m$ be sampled values that characterize classes $1, \dots, m$, respectively. Thus, we have m competing hypotheses H_k, out of which it is necessary to choose the one that corresponds the most to the sampled value $\overline{X} = (X_1, X_2, \dots, X_n)$.

Remark 2. Hypotheses $H_1, \dots, H_m$ are alternative, i.e., from $m_1 \neq m_2$ it follows that the statement $(\overline{L}_{m_1} = \overline{X}) \& (\overline{L}_{m_2} = \overline{X})$ is false.

Remark 3. If $\overline{L}_k = \overline{X}$, then one and only one of the competing hypotheses $H_1, \dots, H_m$ is true.

Let us define event $\mathbf{A}_k$ by the equality $\overline{L}_k = \overline{X}$, i.e., the event is that random variable takes value $\overline{X}$. Then events $\mathbf{A}_k$ are pairwise incompatible ($k = 1, \dots, m$) and it is certainly known that one of them took place. We consider that event $\mathbf{A} = \mathbf{A}_1 + \dots + \mathbf{A}_m$ has nonzero probability. Justifiability of event $\mathbf{A}_k$ is defined by the conditional probability $P(\mathbf{A}_k | \mathbf{A})$, i.e., the probability that event $\mathbf{A}_k$ took place provided that event $\mathbf{A}$ took place. Denote by $P_k = P_k(\overline{X})$ the probability that sampled value $\overline{X}$ belongs to class k.

Proposition 1. To solve the formulated problem, i.e., to determine the class of sampled value $\overline{X}$, the following is necessary:

- calculate the values of probabilities P_k;
- among the calculated values $P_1, \dots, P_m$, choose the maximum $P_{k_{\max}}$; this value $k = k_{\max}$ is the number of the required class.

Quantization. Let us quantize continuous distributions with discrete ones, i.e., select discrete references. Thereby we will locally approximate these distributions by constants. For a point $\overline{X} = (X_1, X_2, \dots, X_n)$ of the Euclidean n-dimensional space, we will construct full-sphere B_δ of radius δ centered at this point:

$$B_\delta(\overline{X}) = \left\{ t_i | \sqrt{\sum_{i=1}^{n} (t_i - X_i)^2} \le \delta \right\}.$$

Thus, if $\overline{L}_k = \overline{X}$, then $\overline{L}_k \in B_\delta(\overline{X})$ since random vector $\overline{L}_k$ will certainly fall within the full-sphere $B_\delta(\overline{X})$ if it gets at its center. Denote event $\overline{L}_k \in B_\delta(\overline{X})$ as $\mathbf{B}_k$.

Remark 4. Put $t = (t_1, \dots, t_n)$, $dt = dt_1 \cdots dt_n$. Probability $\mathrm{P}(\mathbf{B}_k)$ of event $\mathbf{B}_k$ is $\mathrm{P}(\mathbf{B}_k) = \int_{B_\delta(\overline{X})} \cdots \int p_k(t)\, dt$.

Here, $p_k(t) = \prod_{i=1}^{n} p_{k,i}(t_i)$ since random variables $X_{k,i}$ are mutually independent.

Remark 5. If event $\mathbf{A} = \mathbf{A}_1 + \dots + \mathbf{A}_m$ takes place and events $\mathbf{A}_k$, $k = 1, \dots, m$, are pairwise incompatible, then

- event $\mathbf{B} = \mathbf{B}_1 + \dots + \mathbf{B}_m$ has nonzero probability;
- for sufficiently small δ, events $\mathbf{B}_1, \dots, \mathbf{B}_m$ are also pairwise incompatible.

Proposition 2. Under these conditions,

$$\mathrm{P}(\mathbf{B}_k | \mathbf{B}) = \frac{p_k(\overline{X})}{\sum_{s=1}^{m} p_s(\overline{X})}.$$

Proof. Since for sufficiently small δ events $\mathbf{B}_1, \dots, \mathbf{B}_m$ are also pairwise incompatible according to Remark 5, $\mathbf{B}_k \mathbf{B}_l = 0$ for $k \neq l$ is an impossible event and $\mathbf{B}_k \mathbf{B}_k = \mathbf{B}_k$. From here, $\mathrm{P}(\mathbf{B}_k \mathbf{B}) = \mathrm{P}(\mathbf{B}_k \mathbf{B}_1 + \dots + \mathbf{B}_k \mathbf{B}_m) = \mathrm{P}(\mathbf{B}_k)$. Thus,

$$\mathrm{P}(\mathbf{B}_k | \mathbf{B}) = \frac{\mathrm{P}(\mathbf{B}_k)}{\sum_{s=1}^{m} \mathrm{P}(\mathbf{B}_s)}. \tag{2}$$

Since functions $p_k(t)$ are continuous, the approximate equality

$$\mathrm{P}(\mathbf{B}_k) = \int_{B_\delta(\overline{X})} \cdots \int p_k(t)dt \approx p_k(\overline{X}) \int_{B_\delta(\overline{X})} \cdots \int dt = p_k(\overline{X}) V_\delta$$

takes place for a sufficiently small δ, where V_δ is the volume of a full-sphere of radius δ in n-dimensional Euclidean space. Substituting this result into formula (2) yields

$$\mathrm{P}(\mathbf{B}_k | \mathbf{B}) \approx \frac{\mathrm{P}(\mathbf{B}_k)}{\sum_{s=1}^{m} \mathrm{P}(\mathbf{B}_s)} = \frac{V_\delta p_k(\overline{X})}{V_\delta \sum_{s=1}^{m} p_s(\overline{X})} = \frac{p_k(\overline{X})}{\sum_{s=1}^{m} p_s(\overline{X})}.$$

3. HEURISTIC RECOGNITION CRITERION

Standards are classes of objects with known description. The objects being recognized, whose dimensions approximately coincide with dimensions of standards are located without overlapping on a plane (or nearly plane) surface. The result of measurements of the object being recognized is compared with descriptions of standards and a decision is made whether the object belongs to the given class.

Proposition 3. Recognition criterion C_k has the form

$$C_k = p_k(\overline{X}),$$

where p_k are defined by formulas (1).

Heuristic Substantiation. Applying formula (2) of Proposition 2 and letting δ tend to zero, we obtain $\mathrm{P}(\mathbf{A}_k | \mathbf{A}) = \lim_{\delta \to 0} \mathrm{P}(\mathbf{B}_k | \mathbf{B}) = \dfrac{p_k(\overline{X})}{\sum_{s=1}^{m} p_s(\overline{X})} = R_k$. For the given $\overline{X}$, the inequality $0 < \sum_{s=1}^{m} p_s(\overline{X}) \le 1$ holds; the sum $\sum_{s=1}^{m} p_s(\overline{X})$ is a constant, hence, instead of R_k we can use $C_k = p_k(\overline{X})$, which gives the required criterion. Calculate values $C_1, \ldots, C_k$ and choose their maximum $C_{k_{\max}}$. This value $k = k_{\max}$ is the number of the required class.

Remark 6. To make the calculations of $\sigma_{k,i}$ more efficient, it is possible to replace them with root-mean-square averaging with respect to one of the parameters, for example, averaging with respect to spectral channels (zones), i.e., instead of value $\sigma_{k,i}$ use $\sigma_k = \sqrt{\dfrac{\sigma_{k,1}^2 + \cdots + \sigma_{k,n}^2}{n}}$.

The root-mean-square averaging can be carried out with respect to both parameters k and i, which denote the number of the class and spectral channel, respectively. As a result, instead of $\sigma_{k,i}$ it is possible to use

$$\sigma = \sqrt{\frac{\sum_{k=1}^{m} \sum_{i=1}^{n} \sigma_{k,i}^2}{mn}} = (mn)^{-1/2} \sqrt{\sum_{k=1}^{m} \sum_{i=1}^{n} \sigma_{k,i}^2}.$$

Let us remind that first-kind error for this class with number k is that the hypothesis H_k about identity of the class being recognized to the sample with number k was rejected, while this event took place.

Remark 7. An analysis of test calculations of the results of n-channel measurements by the proposed criterion has shown that the probability of a first-kind error for the classes under study varies within 5–20% (significance level α). To decrease the probability of making unreliable decisions for each class k, it is possible to choose significance level α_k, which should take values within the required limits ($\alpha_k < \alpha$).

CONCLUSIONS

We have presented a heuristic criterion of recognition of a class of objects in the set of classes of objects on the basis of the results of measurements of the recognized class of objects and available spectral and statistical characteristics of standards of this recognition problem. With the use of this criterion, we have proposed a heuristic recognition method based on the results of multispectral measurements (spectral brightness of signals) and quantization, as well as its modification in order to increase recognition reliability and make calculations efficient.

REFERENCES

1. I. N. Kovalenko, Probability Calculation and Optimization [in Russian], Naukova Dumka, Kyiv (1989).
2. A. N. Kolmogorov, Selected Scientific Works, Vol. 2, Probability Theory and Mathematical Statistics [in Russian], Mat. Institut im. V. A. Steklova RAN, Nauka, Moscow (2005).
3. B. V. Gnedenko, A Course in Probability Theory [in Russian], Nauka, Moscow (1988).
4. P. S. Knopov and E. J. Kasitskaya, Empirical Estimates in Stochastic Optimization and Identification, Kluwer Acad. Publ., New York (2002).
5. A. R. Webb and K. D. Copsey, Statistical Pattern Recognition, John Wiley, New York (2011).
6. V. N. Vapnik, Statistical Learning Theory, Wiley, New York (1998).
7. A. I. Arkhipov, E. E. Bayadilov, and V. N. Chubarikov, "About Carlson's abscissa in the moment problem of the Riemann zeta function," Doklady RAN, Vol. 392, No. 1, 10–11 (2003).
8. V. I. Lyal'ko and M. A. Popov (eds.), Satellite Methods of Mineral Exploration [in Russian], Karbon-Ltd, Kyiv (2012).
9. N. M. Glazunov, "Arithmetic modeling of random processes and *r*-algorithms," Cybern. Syst. Analysis, Vol. 48, No. 1, 17–25 (2012).