

Experience of convergence of methodologically various decisive rules for improvements of quality of identification of complex objects on analog signals

R A Krupchatnikov¹, M V Artemenko², N M Kalugina², A F Rybochkin²

¹Kursk State Agricultural I.I. Ivanov Academy 305021, 70, Karla Marksa ave, Kursk, 305021, Russia

²Southwest state university, 94, 50 years of October ave, Kursk, 305040, Russia

E-mail: artem1962@mail.ru

Abstract

The article deals with the author's experience of solving the problem of improving the quality of identification of complex objects by the recorded analog signals using recognition systems that implement the convergence of decision rules synthesized on the same samples using methodologically different approaches. The mathematical tools are considered: the statistical approach implemented by calculating the author's indicators of the system organization – analogues of statistical moments of the third and fourth orders; the synergetic approach-indicator latent variables are proposed to form on the basis of matrix transformations of the bi-spectral autocorrelation function of the recorded signal, the parameters of the time shift of which are adjusted in accordance with the Fourier spectrum of the reference classes of the objects. The results of testing the proposed methods on the example of solving a medical problem-preventive diagnosis of destructive lung conditions in asthmatic diseases are reduced. It is shown that the convergence of different rules allows to obtain identification solutions more correlated with independent external expertise than the results of their application separately.

1. Introduction

Quality control and correction of the states of complex technical biological and biotechnical objects determines the application of a system approach and specific methods of its identification as an element of a certain class, for which the content of the bases of shares of the acting agents is known [1]. According to the classical definition by Rastrigin L. A. the complex objects, first and foremost, characterized by: a lack of sufficient for the formalization of the mathematical description; a large number of feedbacks and non-target object processes that lead to additive superposition and modulation of the registered signals; the presence of auto control [2]; the stationary nature and temporal uniqueness (the phenomenon of the arrow of time by Prigogine I.).

These objects include: open-type technical systems, the elements which structures are in constant interaction with the external environment (vehicles, computer networks); physiological and functional systems that ensure the normal functioning and execution of the target functions of the body as a



whole (cardio, respiratory, musculoskeletal, etc. system); economic and social systems of different hierarchical levels; population; socio-technical landscapes, etc.

Since the functioning of complex objects (systems) systemically involved in the provision of the basic processes of human life and determines largely its safety, and to date in various information databases and sources accumulated experience in the management of various, more or less similar to each other objects, improving the quality of identification of the object, as a representative of a certain cluster, is a permanent current scientific and technical task [3, 4].

The transparency of the complex objects functioning causes a high degree of uncertainty which requires the development of adequate models and methods of decision what have got synergetic properties and based on the achievements of artificial intelligence and computer-information technologies. The behavior of a complex object from a cybernetic point of view, in general, is described by the change in the position in the space of the recorded features [5]. Features are represented as discrete and-or analog signals, direct and latent, scalar or matrix forms. The variety of forms leads to the need to apply different analytical paradigms and theories, methodologically and epistemologically, different from each other. Various expert councils, who designed to implement the optimal convergence, are involved to coordinate the final management conclusions [6]. Councils can be organized: by combining specialists (in this case, classical well-developed methods of expert analysis are used, by combining the means of "computer intelligence", by combining "representatives" of artificial and natural intelligence. The latter option is most preferable, because it allows to combine the advantages of objective and subjective [19] analyses in a reasoned and targeted way, especially when it is necessary to make insightful management decisions in the conditions of the situations of "gray" and "black" swans (Taleb N.).

According to the work [7], in general, the impact on the object is represented as: $FacInf = \{A, FI(t), TI, Rez\}$, where A - unambiguously determining the power of the impact factor, $FI(t)$ - effect function, TI - period of exposure, Rez - resonance coefficient, which assesses the possibility of resonance phenomena under the influence where $Rez \in [0,1]$, $Rez = 0$ - resonance phenomena are impossible, $Rez = 1$ - « surely arising ». As an indicator of Rez it is proposed to apply the function of belonging to the base variable of the square of the ratio of the object's natural frequency to the difference between the frequencies of its own and external influence (external or internal, integral or particular, of different physical nature).

If the registration of the values of signs, defined as informative, is carried out continuously, we can talk about the analysis of some analog signals for the classification identification of both the object and its state, by calculating the values of the indicator variables and the application of certain decision rules. In most cases, these rules are well formalized in the form of products.

The specific signal from the object is characterized by spatiotemporal components reflected by spectral parameters (obtained, for example, by Fourier or Walsh transformations) and autocorrelation and autoregressive dependences [8], of different orders. Statistical analysis involves the use of values of particular and integral indicators (for example, indicators of the system organization in the analysis of the functioning of certain physiological systems [9]). Synergetic methodology involves the calculation of the indicators values which systematically reflecting the self-organization processes occurring in the object during the registration of the signal, in the form of changes in the structures and parameters of autocorrelation functions.

Thereby taking actuality of the problem by qualitative identification of the control object and the possibility of its solution by existing means of artificial intelligence, the purpose of the study is to increase the quality of the diagnostic classification of the object in the process of studying the structural components of the generated specific signal, by developing a convergence method of the application of statistical and synergetic results what approaches to the analysis of its space-time structure.

2. Results and Discussion

The classical (statistical) approach to the synthesis of decisive rules of the production type is to execution the following steps:

1. Formation of a set of features $\{\Omega\}$ what are characterizing the state and behavior of the object of study, the values of which can be observed and recorded, by expert analysis ways.

2. The selection of $\{\Omega\}$ the set of informative features $\{\Omega I\}$, which allows to use them as indicator variables of the crucial antecedents of production rules for the implementation of consequential that are certified by specialists.

3. Synthesis of antecedents, by determining on training samples specific ranges of values of the indicator variables included in the antecedents in the case of numerical values, it is recommended, after evaluating the distribution law, to calculate the confidence intervals with acceptable errors of the first and second types.

4. Evaluation on the examination sample of values (or functional definition) of confidence coefficients for decision rules.

The formation of $\{\Omega\}$ is specific for analog signals recorded by the sensors from the object, presupposes:

- first, the synthesis of additional, latent signals $S(t)$ and $iS(t)$:
 $dS(t) = \frac{\partial S(t)}{\partial t}$. $iS(t) = \int S(t)dt$, $S(t)$ – the recorded analog signal is represented in discrete form;

- secondly, the choice of informative frequency ranges-a lot of $\{DFR\}$ (for $S(t)$, $dS(t)$ and $iS(t)$);

- third, the definition of the composition of features for each of the ranges $\{DFR\}$ (for $S(t)$, $dS(t)$ and $iS(t)$).

Classical and proven methods and algorithms for distinguishing characteristics from $S(t)$ are presented in a number of foreign and domestic sources, for example, in the process of "blind" signal processing [8]. Exploration spectral analysis allows to select elements of the set $\{DFR\}$ (lower and upper bounds) using the following algorithm [7].

1. Registration of the signal from the object. Formation of training and control samples. Set of diagnostic frequency bands $\{DFR\} = \emptyset$.

2. Setting the frequency range randomly - df .

3. Determination of spectral characteristics in df (maximum and minimum values of the spectral function, the spectral area in the range) – $\{SV_{df}\}$.

4. Definition of diagnostic capabilities $\{SV_{df}\}$ by DSM method [30] with the fixation of the criterion of diagnostic possibilities (the criterion of diagnostic opportunities - CDO_{df}).

5. If $CDO_{df} \subset tr_{df}$ then df are included to $\{DFR\}$ and transition to item 7 is carried out. Otherwise, the check is performed - if the process of formation of $\{DFR\}$ is completed, then the transition to item 7, otherwise – the transition to item 3. (Here - tr_{df} , – user-defined designer acceptable threshold range for in df .)

6. If $\{DFR\}$ does not significantly improve the values of the diagnostic criterion on the control sample, then the transition to item 3.

7. Ordering $\{DFR\}$ in ascending order of frequencies.

8. If there are restrictions on the number of ranges in $\{DFR\}$, then $\{DFR\}$ - either decreases - by eliminating the ranges according to the selected criterion, or increases - by going to item 2.

(Note that the algorithm allows for intersecting, but not matching, $\{DFR\}$ elements.)

As elements of the set $\{\Omega\}$ is proposed to use calculated in each band characteristics [10]: the minimum value of the spectrum $minDFR$, the maximum value of the spectrum $maxDFR$, the modal value of the spectrum $modDFR$, average power of signal $PDFR$, the mean and standard deviation - $msdDFR$, entropy - $entDFR$, negentropy - $nentDER$, the area of the spectrum, $SDFR$, pairwise relations – $maxDFR/minDFR$, $PDDFR/modDFR$, $msdDFR/modDFR$, $(maxDFR-minDFR)/modDFR$.

The formation of a set of informative features $\{\Omega I\}$ is recommended to be carried out depending on the data structure and volumes by the methods described in [11] using the Kendell concordance coefficient.

To determine statistically significant differences between objects belonging to different classes, it is proposed to use the following algorithm:

1. On the training sample, according to the recommendations [12], the discriminant function is identified as $Y(\omega_k) = Y(\omega_{k-1}) + (-1)^k \cdot H_{\omega_{k-1}}$, where $Y(\omega_k)$ - value of the identifier for a class ω_k , $H_{\omega_{k-1}}$ - step "inter-class differences" (it is recommended: $H_{\omega_i} = H_{\omega_j} = 10$), $k=0,1,\dots,K$, $Y(\omega_{-1})=0$, $H_{\omega_{-1}}=0$. On examination - thresholds are determined $Y(p)$ for decision rules and confidence coefficients equal to the values of diagnostic efficiencies in each alternative class ω_k .

2. Values of indicators of the system organization of object are defined (indicators of the system organization – $ISOas$, $ISOex$) by formulms [23]:

$$ISOas = \frac{\sum_{i=1}^{Qr} (p_i - \bar{p})^3}{Qr \cdot \sigma^3}, \quad ISOex = \frac{\sum_{i=1}^{Qr} (p_i - \bar{p})^4}{Qr \cdot \sigma^4}, \quad (1)$$

where: Qr – number of comparison implementations; σ^2 , \bar{p} – variance and mean, respectively, of the characteristic p_i from $\{\Omega I\}$.

3. Built membership function $K2_k$ on the media:

$$dISO_k^2 = (ISOas_k - ISOas_0)^2 + (ISOex_k - ISOex_0)^2 \quad (2)$$

(index k – corresponds to the k -th sign).

Functions $K2_k(dISO^2)$ determined by expert or analytical means. For the latter case, the authors propose empirically obtained formula:

$$K2(dISO_k) = \begin{cases} th(\pi \cdot \frac{dISO_k}{db_k}) \cdot \cos(\varphi), \text{ for a standard } Se_k \\ h(\pi \cdot \frac{dISO_k}{db_k}), \text{ in other case} \end{cases}, \quad (3)$$

where: $(\overline{ISOEas_0}, \overline{ISOEex_0})$ - reference state vector S_k , db_k - module $(\overline{ISOEas_k}, \overline{ISOEex_k})$, φ - the angle between the vectors $(\overline{ISOEas_k}, \overline{ISOEex_k})$ and $(\overline{ISOEas_0}, \overline{ISOEex_0})$.

4. Confidence is determined $K3_k$:

$$K3_k = (\sqrt{K1_k} + \sqrt{K2(dISO_k)} - \sqrt{K1_k \cdot K2(dISO_k)})^2 \quad (4)$$

5. The production decision rules of the form are synthesized:

"IF $K2(dISO_k) > 0.62$, then the object is in the S_k state with confidence $K3_k$."

In general, the form of recording such decisive rules is generalized to the form:

$$\text{"IF } Ans(\{ISO\})_{\omega_k}, \text{ then the object } \in \omega_k \text{ with confidence } U_{\omega_k} \text{"} \quad (\text{ISO rule})$$

Where: $Ans(\{ISO\})_{\omega_k}$ - the antecedent for the class ω_k whose arguments are the indicators of the ISO system organization.

To analyze the synergetic components of the analog signal, it is proposed to use the bispectral autocorrelation function, the values of which are calculated by the formula [8]:

$$R_{x,x+k,x+l} = \frac{1}{M} \sum_{i=0}^{M-1} x(i) \cdot x(i+k) \cdot x(i+l), \quad (5)$$

Where: $x(i)$ the i -th discrete of vector X ; M – the number of the discrete vector X ; k, l – «time shift»; $k=0,1, \dots, M/2$, $l=0,1, \dots, M/2$; X – a volume-constrained vector of values $S(t)$ or $dS(t)$ or $iS(t)$.

In order to optimize the process of obtaining the matrix R and its further processing, it is proposed to limit its dimension to ten rows and columns (time shifts), and to make the shift step self-organizing and dependent on the spectrum of the analyzed signal. In this way from formula (5), it is proposed to move to the formula:

$$R_{x,x+k_j,x+l_j} = \frac{1}{10} \sum_{i=0}^9 x(i) \cdot x(i+k_j) \cdot x(i+l_j) \quad , \quad (6)$$

where: k_j and l_j are elements of an array of integers ($j=0,1,\dots,9$) that are harmonic numbers of 10 local maxima of the spectral function.

The synthesis of decision rules is realized by the following method:

1. Spectral function is identified and are formed a vector of values of k_j and l_j .
2. Randomly selected from the training sample is one of the signals and a discrete vector X is formed.

3. The matrix of the bi-spectral autocorrelation function R_o is determined by the formula (6). The inverse matrix R^{-1} of the matrix R_o is calculated. The obtained R^{-1} matrix is defined as a support for the synthesis of the conditions of the decision rules in the future.

4. On the training sample of object-specific signals are formed by the formula (6) of the matrix B_l ($l=0,1\dots K, K$ - number of classes) which is a bi-spectral autocorrelation function. Indicator matrices C_i are calculated as a product of: $C_i=R^{-1}*B_i$. For each matrix C_i , the average values of the matrix squares are determined:

$$IP_i = \sqrt{\frac{1}{100} \sum_{i=1}^{10} \sum_{j=1}^{10} (C_{i,j})^2} \quad . \quad (7)$$

5. On the training sample for IP , confidence intervals are formed, which are typical for each alternative object class.

6. For each of the classes the decisive rules of the form are formed:

$$"IF \text{ Ans}(\{R\})_{\omega_k} , \text{ then the object } \in \omega_k \text{ with confidence } U_{\omega_k} " \quad , \quad (\text{rule R})$$

where: $\text{Ans}(\{R\})_{\omega_k}$ - the antecedent for the class ω_k formed by the bi-spectral autocorrelation function and shows the belonging (or non-belonging) of the value of the indicator variable IP to the corresponding confidence interval.

Since the obtained "ISO rules" and "R rules" reflect various methodological aspects of the analysis, it is proposed to synthesize convergent decision rules as follows. Let's denote by $Rul_{r,l}$ some decisive rule for class l . The general decision rule, in this case, is as follows:

$$"IF (\wedge_1^{Rr} (\text{Ans}_{\omega_k} (\text{II})_{Rul_r})) \text{ that object } \in \omega_k \text{ with confidence } (U \wedge)_{\omega_k} " \quad (\text{rule convergent})$$

where: $(\wedge_1^{Rr} (\text{Ans}_{\omega_k} (\text{II})_{Rul_r}))$ - the conjunction of the ancestor of the decision rules Rul_r obtained on the indicator variables II ; $(U \wedge)_{\omega_k}$ - confidence in the correlation (identification) of the object to the class ω_k , calculated by the formula:

$$(U \wedge)_{\omega_k} = 1 - \max(U_{\omega_k, Rul_{r,l}}) \cdot \prod_{r=1}^R (1 - a_r) - \prod_{r=1}^R (a_r \cdot (1 - U_{\omega_k, Rul_{r,l}})) / a_r = \begin{cases} 1, & \text{if the rule was applied } Rul_{r,l} \\ 0, & \text{if the rule was not applied } Rul_{r,l} \end{cases} \quad (8)$$

Application (8) improves the quality of object identification by increasing the value of the confidence function in case of convergence of decision rules. To eliminate the influence of "weakly-confident" decision rules, in the formula (8) of the binary variable a_r assign a zero value if the confidence of the rule is below a certain level defined by the user.

As an experimental basis for checking the effectiveness of the proposed method of convergence of decision rules, the medical problem of identification of various destructive asthmatic forms in a patient

was chosen [11] (I — stage of compensation, II — stage of increasing ventilation disorders or stage of decompensation, III stage — hypercapnic coma). The experimental material representing sound signals was formed on the basis of international medical information databases. Various decision rules are obtained, for example:

- If $(I_p \subset 1 \pm 0.2)$, then the state of the respiratory system are normal and asthma is absent with a confidence of 0.96;

- If $(I_p \subset 2,1 \pm 0.6)$, then the state of the respiratory system are normal and asthma is absent with confidence of 0.92 (stage I);

- If the I_p values do not belong to the specified intervals, additional examination with a confidence of 0.95 is required to determine the state of the respiratory system.

To calculate these indicators as informative, $\{\Omega I\} = \{modDFR, PDFR, msdDFR, entDFR, SDFR, maxDFR/minDFR, PDDFR/modDFR, msdDFR/modDFR\}_k$. ($k=2$ – two frequency bands are selected) were selected. The calculated values of ISO_{as} and ISO_{ex} system organization indicators for the classes "healthy" and "asthma" were: $ISO_{as_0} = 0.38 \pm 0.15$, $ISO_{ex_0} = 1.9 \pm 0.32$, $ISO_{as_1} = -0.1 \pm 0.2$, $ISO_{ex_1} = 0.98 \pm 0.41$. According to the given data, the corresponding production rules are constructed and the values of the confidence functions are determined.

Options for convergent application of the rules are presented in table 1 - the values of diagnostic efficiency. The values of the indicators set by an expert doctor when listening to sound signals are highlighted in brackets.

Table 1. Confidence in the correct diagnostic solution.

Apply rules		Patient condition			
(rule ISO)	(rule R)	«healthy»	I stage	II stage	III stage
no	no	0.04	0.06	0.16	0.05
no	yes	0.96	0.92	0.86	0.95
yes	no	0.88	0.94	0.84	0.94
yes	yes	0.995	0.995	0.978	0.997
Medical expert		(0.97)	(0.96)	(0.95)	(0.98)

Analysis of the data in the table allows us to draw the following conclusions:

1. The results of the application of the obtained decisive rules are mutually consistent, which emphasizes their acceptability for operation in real conditions;
2. Convergence of decision rules allows to improve the quality of preventive diagnostic identification of the object of study;
3. The convergence results of several decision rules have a stronger correlation with external, independent expertise than the application of each of them separately.

3. Conclusion

The results of the research allowed to develop a method of convergence of the decision rules of identification of the object of study on the specific analog signals issued by it, synthesized methodologically different approaches. The basis of the method is the combination of antecedent one-way consequents and the modification of the values of the confidence coefficient, which is more correlated with the objective data of expert identification.

For the static approach of the synthesis of adequate decision rules it is recommended to apply the indicators of the system organization, similar to the particular moments of 3 and 4 orders. In the analysis of the synergetic properties of the object as an indicator variable is recommended to use the average sum of the squares of the elements of the diagnostic matrix, which is the product of two matrices: the bi-spectral autocorrelation function of the object and the inverse matrix of this function for some object belonging to the reference class. Self-organization of matrix elements calculation, expressed in the determination of time shifts when calculating the values of the autocorrelation

function, by analyzing the spectrum of the signal of the reference class, allows to reduce the resources of the computational process and increase its efficiency.

The developed methods are well formalized for use in the design of decision support systems.

Decision-making systems - areas of theoretical and practical application of research results, automatic control and pattern recognition, diagnostics of complex objects.

4. Acknowledgments

The work was carried out within the framework of the federal target program "Research and development in priority areas for the development of the scientific and technological complex of Russia for 2014-2020" and in accordance with the scientific direction of the South-West State University "Development of medical and environmental information technologies" and financial support for the grant No.11643GU/2017 as part of the contest "UMNIK"(cod 0033037).

References

- [1] Rassel S, Norvig P 2018 *Iskusstvennyj intellekt: sovremennyyj podhod* (Moskva: Vil'jams)1408.
- [2] Zhdanov A A 2009 *Avtonomnyj iskusstvennyj intellekt* (Moskva: BINOM)
- [3] Guchuk V V 2018 Applied aspects of the organization of an interactive mode for increasing the security of functioning of control systems of complicated objects *European Science*. **No 1 (33)**. 18-21
- [4] Semenov S S , Voronov E M , Poltavskij A V, Krjanov A V 2016 *Metody prinjatija reshenij v zadachah ocenki kachestva i tehničeskogo urovnja slozhnyh tehničeskikh sistem* (URSS)
- [5] Viner N 2001 *Chelovek upravljajushhij*. (1. Kibernetika i obshhestvo. 2. Tvorec i robot. 3. Stat'ja "Kibernetika i obshhestvo" URSS)
- [6] Ushakov V N, Lavrov N G, Ushakov A V 2014 *Konstruirovanie reshenij v zadache o sbližhenii stacionarnoj upravljajemoj sistemy* *Tr. IMM UrO RAN*, **20, No 4** 277–286
- [7] Artjomenko M V, Borodin M V, Kalugina N M, Kurochkin A G 2017 *Identifikacija faktorov vozdeystvija na obekt samoorganizacionnym analizom spektra harakterizujushhego signala* *Informacionno-izmeritel'nye i upravljajushhie sistemy* **T. 15. No 5** 8-12.
- [8] Kravchenko V F 2007 *Cifrovaja obrabotka signalov i izobrazhenij v radiofizicheskikh prilozhenijah* (Moskva: FIZMATLIT)
- [9] Artemenko M V, Dronova T A, Kalugina N M, Bojcov A A 2017 *Primenenie pokazatelej sistemnoj organizacii v avtomatizacii diagnostiko-terapevtičeskogo processa* *Izvestija Jugo-Zapadnogo gosudarstvennogo universiteta. Serija: Upravlenie, vychislitel'naja tehnika, informatika. Medicinskoe priborostroenie*. **T 7 No 2 (23)** 122-134.
- [10] Artemenko M V, Pisarev M V 2016 *Synthesis and study of mathematical models of biomedical signals methods of selforganization modeling by the example of the photoplethysmogram* *International Journal of Applied and Fundamental Research*. **5** 31-35.
- [11] Artemenko M V, Chernetskaia I E, Kalugina N M, Shekina E N 2018 *Bootstrap and counter-Bootstrap approach for formation of the cortege of informative indicators by results of measurements* *Jurnal of Physics: Conference Series* **998** 120-126
- [12] Voroncov I M, Shapovalov V V, Sherstjuk Ju M 2006 *Zdorov'e. Opyt razrabotki i obosnovanie primeneniya avtomatizirovannyh sistem dlja monitoringa i skrinirujushhej diagnostiki narushenij zdorov'ja* (SPb.: «IPK «Kosta»)