Research and Development Solution Methodology of Informative Features Choosing for the RES Life Cycle Processes Analysis

Nevlyudov I.Sh., Andrusevich A.A., Khatnyuk I.S.

Abstract — **The methods of solving the problem of choosing the informative features to identify the RES functioning are offered by classifying the RES and life cycle processes in the feature space, and each of which has a value that allowed us to find a comprehensive criterion and formalize the selection process.**

I. Introduction

The functional problem of selecting informative features for monitoring the life cycle of radio electronic systems (RESs LC) can be solved within the framework of development methodology for dictionary attributes in the classification systems and the objects state identification [1-5]. In the working vocabulary one should use only those features that, on the one hand, are the most informative and that, on the other hand, may be available for measurement.

The description of the dictionary of features under conditions of constraints on the cost of observation hardware creation has some particulars. If the objects attributes are denoted by δ_j, j=1, 2,…, N, each object in the N-dimensional space of attributes can be represented as a vector $x=(x_1,x_2,...,x_N)$, its coordinates characterize the properties of objects.

To determine the measure of closeness or similarity between objects in N-dimensional attributes vector space a metric is introduced. The Euclidean metric can be used as

$$d^2(w_{pk},w_{ql})=\sum_{j=1}^{N}(x^j_{pk}-x_{ql}{}^j)^2, \tag{1}$$

$p,\ q=1,\ 2,\ ...,\ m;\quad k=1,\ 2,\ ...,\ k_p;\quad l=1,\ 2,\ ...,\ k_q,$

where $p,\ q=1,\ 2,\ ...,\ m;\quad k=1,\ 2,\ ...,\ k_p;\quad l=1,\ 2,\ ...,\ k_q,$

x^j_{pk} are the values of j-feature, k is an object, p is a class, i.e. the q is an object class, i.e. object w_{ql}.

As a measure of proximity between the objects of this class Ω_p, $p=1,\ 2,\ ...,\ m$, we will use the value

$$S(\Omega_p)=\sqrt{\frac{2}{k_p}\frac{1}{k_{p-1}}\sum_{k=1}^{k_p}\sum_{l=1}^{k_p}d^2(w_{pk},w_{pl})}\ , \tag{2}$$

that makes sense as class rms dispersion or the rms scatter of objects within class Ω_p, as a measure of proximity between a given pair of objects of classes Ω_p and Ω_q, $p,\ q=1,\ ...,\ m$, is the value

$$R(\Omega_p,\Omega_q)=\sqrt{\frac{1}{k_p\,k_q}\sum_{k=1}^{k_p}\sum_{l=1}^{k_q}d^2(w_{pk},w_{ql})}\ , \tag{3}$$

that makes sense of rms dispersion of a class of objects Ω_p and Ω_q.

A set of objects features that are used in the working vocabulary can be described as an N-dimensional vector $A=(\alpha_1,\alpha_2,...,\alpha_N)$, whose components takes the values 1 or 0 depending on whether there is or is not a possibility to determine the appropriate object feature.

Taking into account the *a* is the squared distance between two objects w_{pk} and w_{ql} the unequation given below will take the form

$$d^2(w_{pk},w_{ql})=\sum_{j=1}^{N}\alpha_j(x^{(j)}{}_{pk}-x^{(j)}{}_{ql})^2. \tag{4}$$

Consequently, the rms scatter of the class Ω_p and objects of classes Ω_p and Ω_q can be written as

$$S(\Omega_p)=\sqrt{\frac{2}{k_p}\frac{1}{k_p-1}\sum_{k=1}^{k_p}\sum_{l=1}^{k_p}\sum_{j=1}^{N}\alpha_j(x^{(j)}{}_{pk}-x^{(j)}{}_{pl})^2}\ , \tag{5}$$

$$R(\Omega_p,\Omega_q)=\sqrt{\frac{1}{k_p}\frac{1}{k_q}\sum_{k=1}^{k_p}\sum_{l=1}^{k_q}\sum_{j=1}^{N}\alpha_j(x^{(j)}{}_{pk}-x^{(j)}{}_{pl})^2}\ . \tag{6}$$

One can take into consideration that the cost of using the feature is proportional to the information contained in them, i.e. to such quantities of objects features that can be determined with their help. This assumption is fairly general.

Thus, the cost of using the features amount to

$$C=C(\alpha_1,...,\alpha_N)=\sum_{j=1}^{N}C_j\,\alpha_j\ , \tag{7}$$

where C_j is costs to determine the j-sign.

Manuscript received December 8, 2011.

Nevlyudov I.Sh. is with Kharkov national University of Radioelectronics (14 Lenin Avenue, Kharkov 61166, Ukraine, tel. (057) 702-14-86, E-mail: tapr@khture.kharkov.ua).

Andrusevich Anatoly Alexandrovich is with Aviation Krivoy Rog College (1 Tupolev St., Krivoy Rog 50045, Ukraine, tel. (0564) 27-56-79, E-mail: uchebotdel@kk.nau.edu.ua).

Khatnyuk Igor Stanislavovich is with Kharkov National University of Radioelectronics (14 Lenin Avenue, Kharkov 61166, Ukraine, tel. (057) 702-14-86, E-mail: tapr@khture.kharkov.ua)

As an indicator of quality or effectiveness of the designed recognition system we consider the functional, which depends in general on the function $S(\Omega_p)$, $R(\Omega_p,\Omega_p)$ and decision rule $L(w,\{w_g\})$

$$I = F[S(\Omega_p);\ R(\Omega_p,\Omega_q);\ L(w,\{w_g\})]. \quad (8)$$

Let the value $L(w,\{w_g\})$ be a measure of proximity between the recognizable object w and class Ω_g, $g = 1, 2, ..., m$, that is given by its objects $\{w_g\}$. As this measure of proximity let's consider the value

$$L(w,\{w_g\}) = \sqrt{\frac{1}{k_g}\sum_{g=1}^{k_g} d^2(w,w_g)}\ , \quad (9)$$

which is mean square distance between the object w and the classes objects Ω_p.

The decision rule consists in the following

$$w \in \Omega_g,\ \text{если}\ L(w,\{w_g\}) = \text{extr}\ L(w,\{w_i\}). \quad (10)$$

It is important to note that the decrease in $S(\Omega_p)$, "compression" of objects belonging to a given class, while increasing $R(\Omega_p,\Omega_q)$, i.e. "division" of objects belonging to different classes improve the quality of the recognition system. Therefore, improving the efficiency of the system will be linked to the achievement of functional extremum I.

II. Formulation of research problems

Formulation of research problems can be represented as follows.

Let the entire set of objects be divided into classes Ω_i, $i = 1, ..., m$, a priori all classes being described in the language of features x_j, $j = 1, ..., N$, and funds, whose value is equal to C_0, are allocated for creating observation hardware. It is required that a working vocabulary of features be built which provides the maximum possible efficiency of the system without exceeding the allocated funds.

Thus, the problem reduces to finding the condition extremum of the functional which looks like (8), i.e. to the definition of A which implements the

$$\underset{\alpha}{\text{extr}}\ I = \underset{\alpha}{\text{extr}}\,F\,[S(\Omega_p);\ R(\Omega_p,\Omega_q);\ L(w,\{w_g\})]$$

$$C = \sum_{j=1}^{N} C_j\alpha_j \le C_0. \quad (11)$$

Consider some special types of the functional (11). If the required efficiency of the recognition system can be achieved through a more compact arrangement of objects in each class, subject to certain conditions regarding the magnitude of $R(\Omega_p,\Omega_q)$, then the problem reduces to finding

$$\min_{\alpha}\ \max_{i=1,\dots,m}\ [S(\Omega_i)] \quad (12)$$

when

$$\sum_{j=1}^{N} C_j\,\alpha_j \le C_0\ \text{и}\ R(\Omega_p,\Omega_q) \ge R_0^{(pq)}. \quad (13)$$

If the required system performance can be achieved through the "bias" objects that belong to different classes under certain conditions with respect to the value $S(\Omega_i)$, $i = 1, \dots, m$, then the problem reduces to finding

$$\max_{\alpha}\ \min_{p,q=1,\dots,m}\ [R(\Omega_p,\Omega_q)] \quad (14)$$

when

$$\sum_{j=1}^{N} C_j\alpha_j \le C_0\ \text{и}\ S(\Omega_i) \le S_0^i. \quad (15)$$

If a proper system performance can be achieved only by increasing the ratio of distances between the classes to the rms scatter within the classes of objects, then the problem reduces to finding

$$\max_{\alpha}\ \min_{p,q=1,\dots,m}\left[\frac{R^2(\Omega_p\ \Omega_q)}{S(\Omega_p)S(\Omega_q)}\right] \quad (16)$$

when

$$\sum_{j=1}^{N} C_j\alpha_j \le C_0. \quad (17)$$

III. The selection problem solution of informative features that characterize the state of the RES LC processes

The problem considered above is a generalization of a nonlinear programming problem. Optimality conditions for it can be formulated as follows: in order the vector C^0 to be an optimal strategy, it is necessary that there exist a scalar $\beta \ge 0$ and vector $\mu = \{\mu_1, ..., \mu_n\}$ which will be equal

$$\left.\begin{array}{l}
\left[\sum_{r=1}^{n}\mu_r\rho_r^j\right]\dfrac{dP_j(C_j^0)}{dC_j} = \beta,\quad j = 1, ..., N_p;\\
\sum_{j=1}^{N_p} C_j^0 = C_0\ ;\\
\sum_{r=1}^{n}\mu_r = 1,\ \mu_r = 0,\ \ \text{если}\ \sum_{j=1}^{N_p}\rho_r^j P_j(C_j^0) > W(C^0).
\end{array}\right\} \quad (18)$$

An introduction to the consideration of scalar β and vector μ increases the number of unknowns C_j^0, μ_r and β to a value $N_p + n + 1$. However, the number of equations equals the number of unknowns, since for every r or $\mu_r = 0$, or

$$\sum_{j=1}^{N_p}\rho_r^j\,P_j(C^0{}_j) = W(C^0). \quad (19)$$

Thus, the solution of the system (18) makes it possible to determine the composition of the working vocabulary signs and optimal allocation of costs to provide tools for observing the recognition system under the assumption of dependence of $P_j = P_j(C_j)$ and limitations to the total cost of these funds.

With the limitations associated with the ability to use a dictionary of all attributes, the problem arises of choosing a limited list (up to 2-3 characters). Here it is possible navigate in the location of the individual components of feature vector with respect to the boundaries of performance facilities monitoring.

Since for the boundary value of the parameter y^j_{rp}, the end of the vector X must be located on the boundary of operability, it is necessary to satisfy the equality

$$x^i{}_{rp} = a^i_j y^j{}_{rp}. \tag{20}$$

A correlation coefficient r_{ij} between the parameters may be an additional criterion for selection under statistical estimation. Since the maximum correlation coefficient provides a maximum amount of information

$$J(y^j) = H(y^i) - H(y^j / y^i), \tag{21}$$

contained in the parameter y^i. Here $H(y^i)$ is the initial entropy; $H(y^j / y^i)$ is the conditional entropy of the object after the measurement of parameter y^j.

The use of a binary correlation algorithm allows for the participation of decision makers (DM) to formalize and automate the processes of input, processing, and recognition of an image.

IV. RES LIFE CYCLE PROCESSES STATE IDENTIFICATION

The solution of the RES life cycle identification, involves creation the rules that define the RES state.

The signs, that allow to distinguish the state of an object under monitoring, are efficiency indicators, which will have a given value or an extreme value for the selected state. To identify the RES state from the observed parameters in the monitoring process it is necessary to select a set of parameters, in which the value of performance indicators will have given or extreme values.

The objects of observation - the parameters and characteristics of RES - can be considered as points of vector and functional spaces. For all pairs of points in the set Q, there is a binary relation of the comparative effectiveness: point x is more efficient than point y then and only then when (x,y) є Φ or in another writing x є y. The problem of allocating the core is solved by ensuring the RES life cycle – a set of maximal elements of the variables X to a binary relation $\Phi : X^* = \mathrm{Max}\ (Q, \Phi)$. It is assumed that the solution of the problem exists i.e. set X* is not empty. In many problems, we may assume that the solution is a set X* which consists of one element, and the relationship between the elements is established by means of functionals $\Lambda(x)$. For instance, the point x is more effective than point y, when $\Lambda(x) < \Lambda(y)$ or $\Lambda(x) > \Lambda(y)$. It can be shown that in problems of determining the effective pixels $x_0 \in X^*$ with constraints $x \in Q_1$, the functional $f = \lambda\Lambda'(x_0)$, where Frechet derivative $\Lambda'(x_0)$ at the point x_0 is support functional to Q_1, at the point x_0 (i.e. $(f, x_0) < (f, x)$ for all $x \in Q_1$.

Thus, the problem of analyzing the results of observations in the monitoring process is reduced to the determination of support functionals at points of observation, which makes it possible to estimate the deviation of the observed points from effective points.

In terms of functional analysis [4,5]: let Q be some set in linear topological space E, E' be the dual space, $x_0 \in Q$ is an extreme point of Q, K_b is the cone of possible directions in Q at the point x_0, K_k is the cone of tangent directions for Q in x_0. If the set of linear functionals, which are reference to the Q at the point x_0, denotes the Q^*, then Q*={ $f \in E', f(x) \ge f(x_0)$ for all $x \in Q$, i.e. support functional and the extreme point $x_0 \in Q$ provide an opportunity to allocate a set Q. It can be shown that if Q is closed convex set, then $Q^* = K_k{}^*$, i.e. forms a cone formed by the set of linear functionals reference to Q in x_0. The cone of tangent directions can be determined from the Frechet derivatives of the operators (convex function), which connect the sets of parameters and performance indicators.

Let us consider the methods of finding K^* for ways to specify K using different functionals.

Example 1. In the case of determination Q using affine sets: $E = E_1 \times E_2$; E_1, E_2 are the linear topological space, a performance characteristics set is defined E_2, D is a linear operator from E_1 to E_2, $K = \{x \in E$, $x = (x_1, x_2) : Dx_1 = x_2\}$, $K^* = \{f \in E', f = (f_1, f_2) :$ $f_1 = -D^* f_2\}$, and as a reference separating function one can use the expression

$$f(x) = (-D^* f_2, x_1) + (f_2, x_2) = -(f_2, D^* x_1 - x_2).$$

Application of this function to separate the sets in the parameter space and the formulation of rules, that establish a correspondence between the sets of parameters and values of performance indicators, can provide the identification of states in the process of monitoring RES LC.

Example 2. A functional $\Lambda(x)$ in the linear space E has a derivative $\Lambda'(x_0, g)$ at the point x_0 in the direction of g, there is

$$\lim_{\varepsilon \to +0} \frac{\Lambda(x_0 + \varepsilon h) - \Lambda(x_0)}{\varepsilon} = f(x_0, g). \tag{22}$$

The functionals are correctly decreasing at any point and allow us to find the cone of decreasing directions. The functional $\Lambda(x)$, which is a set in a Banach space E, is differentiable (or Frechet differentiable) at x_0, if there exists a linear functional $f \in E'$ such that for all $g \in E$

$$\Lambda(x_0 + g) = \Lambda(x_0) + (f, g) + o(\|g\|) . \quad (23)$$

If $\Lambda(x)$ is differentiable at the point x_0, then $F(x)$ is correctly decreases at the point x_0 and $K = \{g : (\Lambda'(x_0), g) < 0\}$. K is decreasing direction cone of the functional $\Lambda(x)$ at the point x_0, $\Lambda(x)$ (satisfies a Lipschitz condition in a neighborhood of $x_0 \in E$, $\Lambda(x)$ is differentiable at x_0 in any direction, and $f(x_0,g)$ as a function of g is convex if $g \in K$ and $\Lambda'(x_0,g)$ $g \in E$, E is a Banach space, $\Lambda(x)$ satisfies a Lipschitz condition in the neighborhood of x_0 (for some $\varepsilon_0 > 0$ $|\Lambda(x_1) - \Lambda(x_2)| \le \beta\|x_1 - x_2\|$ for all $\|x_1 - x_0\| \le \varepsilon_0$, $\|x_2 - x_0\| \le \varepsilon_0$) and) $\Lambda'(x_0,g) < 0$ will be satisfied, then $\Lambda(x)$ is correctly decreasing at x_0, and $K = \{g : \Lambda'(x_0,g) < 0\}$.

Example 3. In the case of the set, which is not defined by a functional. If Q is a convex set, then the decreasing direction set K_b at x_0 takes the form

$$K_b = \{\lambda(Q^0 - x_0), \lambda > 0\}$$

(i.e. $K_b = \{g : g = \lambda(x - x_0),\ x \in Q^0, \lambda > 0\}$).

Example 4. $P(x)$ is an operator from E_1 in E_2, that is differentiable in a neighborhood of x_0, $P(x_0) = 0$. $P'(x)$ is continuous in the neighborhood of x_0, and $P'(x_0)$ displays E_1 for all E_2 (i.e. linear equation $P'(x_0)g = b$ has a solution g for every $b \in E_2$), then the set of tangent directions K to set $Q = \{x : P(x) = 0\}$ at x_0 is the subspace $K = \{g : P'(x_0)g = 0\}$.

When the $P'(x_0)E_1 \ne E_2$, one can only state that $K \subset \{g : P'(x_0)g = 0\}$.

Example 5. Let $x \in R^m$, $Q = \{x : G_i(x) = 0,\ i = 1, ..., n\}$, where $G_i(x)$ are functions continuously differentiable in a neighborhood of x_0, $G_i(x_0) = 0$, $i = 1,...,n$, and vectors $G_i'(x_0)$, are linearly independent. Then $K = \{g \in R^n : (G_i'(x_0), g) = 0, i = 1,...,n\}$.

Here $E_1 = R^m$, $E_2 = R^n$, $P(x) = (C_1(x), \ldots, G_n(x))$, $P'(x_0)$ is matrix $m \times n$, i column is equal to $G_i'(x_0)$.

Example 6. In the process of monitoring it is necessary to determine whether the effective value of the function- RES characteristics $w(z)$ is provided and if in the simplest case the extreme value of an differentiable objective function is provided for one variable, for which it is necessary to check whether the derivative is zero at the observed value of the parameter. For multi-dimensional objective functions and their arguments, this problem may be conceded as part of set theory and functional analysis.

Formalizing in the observation problem of optimal setting, as one of the RES life cycle processes, lies in the fact that it is necessary to estimate optimal function of the process setting $v(z) \in M$, where z is the parameter determining the numerical value of the required characteristic $w(z)$ of a setting object to provide such a phase trajectory, which ensures the equality $w(0) = c$, $w(Z) = d$ and extreme values of the integral functional $\int_0^Z \Phi(w(z), v(z), z)dz$, in the case of a connection, given by the differential equation $\frac{dw(z)}{dz} = \varphi(w(z), v(z), z)$.

In problems that requires maximum compliance of an optimised characteristic to some desired characteristic, minimum mean square deviation criterion is used

$$W_2(X) = \overline{(Y(X) - Y^*)^2}\ , \quad (24)$$

where Y^* is the value characteristic desired or required by the technical project.

For the characteristic, which is given by a discrete set of points, the objective function

$$W_2(X) = \frac{1}{N}\sum_{i=1}^{N}\gamma_i(Y(X,p_i) - Y_i^*)^2, \quad (25)$$

where N is the number of sampling points of the independent variable p; $Y(X,p_i)$ is the value of the optimized performance in the i-th point of the sampling interval; γ_i – optimized characteristic weight coefficient values that reflects the importance of the i-th point compared with other points (usually, $0 < \gamma_i > 1$).

In some optimization problems it is necessary to ensure the excess or not the excess of some given level optimized characteristics. These criteria of optimality are implemented by the following functions:

– to provide excess of a given level

$$W_3(X) = \begin{cases} 0 & \text{at} \quad Y(X) \ge Y_H^*, \\ (Y - Y(X))^2 & \text{at} \quad Y(X) < Y_H^*; \end{cases} \quad (26)$$

– to provide unexcess of a specified level

$$W_4(X) = \begin{cases} 0 & \text{at } Y(X) \le Y_B^*, \\ (Y(X) - Y_B^*)^2 & \text{at } Y(X) > Y_B^*, \end{cases} \quad (27)$$

where Y_H^*, Y_B^* are the feasible region lower and upper bounds for the characterization of Y(X).

If it is necessary that an optimized characteristic be within some permissible zone (boundary), a combination of two previous optimization criteria are used

$$W(X) = \begin{cases} 0 & \text{at} \quad Y_H^* \le Y(X) \le Y_B^*, \\ (Y(X) - Y_B^*)^2 & \text{at} \quad Y(X) > Y_B^*, \\ (Y_H^* - Y(X))^2 & \text{at} \quad Y(X) < Y_H^*. \end{cases} \quad (28)$$

For those cases when it is necessary to implement only the shape of the curve while ignoring a constant vertical bias, there the shift criterion is used.

$$W_6(X) = \sum_{i=1}^{N} \gamma_i (Y_i^* - Y(X, p_i) - Y_{cp})^2, \qquad (29)$$

where $Y_{cp} = \frac{1}{N} \sum_{i=1}^{N} (Y_i^* - Y(X, p_i))$.

The form of objective function effects the computing process important characteristics and, convergence of the optimization process. Derivatives signs of the objective function for the controlled parameters are not constant throughout the feasible area, this circumstance leads to the ravine of the character (for example, problems of circuit design), which leads to high computing cost and requires special attention to the choice of optimization method.

Another feature of the objective functions is that they usually multiextremal and along with the global minimum have local minima.

A general class of problems identifying the set of efficient solutions consists of multi-criteria optimization problems. They are characterized by the fact, that binary relation on the set of alternatives is associated with a set of indicators, forming an efficiency vector criterion. This binary relation is generated by a variety of ways. Thus, if the

$$W(x) = (W^1(x), \ldots, W^m(x)) \qquad (30)$$

is vector criterion on the set of X, than binary relation can be Pareto ratio, or ratio of Slater. In other cases, a binary relation on X is defined by a system of preferences of the decision maker (DM). It is assumed that the main source of information is a person who has information sufficient to make a (unique) solution. Identification of the system of DM preferences represents one of the main problems in solving multiobjective problems. Usually a procedure for identifying the decision maker preferences are based on the language of vector evaluations of alternatives.

Algorithms based on scalarization – reducing to a parametric family of scalar optimization problems – are most illustrative.

IV. Conclusion

The scientific result of the paper is as follows: methods of informative features selection were developed that solve the problems for monitoring the RES life cycle, by classifying the states of RES and the life cycle processes in the feature space, each of them having a particular significance. This has allowed to find comprehensive criteria and formalize the selection process. Heuristic methods for selecting the criteria for the use of prototypes and basic information priorities are proposed, when there is an insufficient number of a priori data for correct classification.

The methods of solving the problem of choosing informative features of RES LC processes were researched that will allow a meaningful description of the processes for implementing RES LC decision-making procedures in the man-machine systems.

References

[1] *Podinovskii V.V., Nogin V.D.* Pareto-optimal solutions of the multicriteria problems. M.: Fizmatlit, 2007. 255p., (in Russian).

[2] *Vishnekov A.V., Ivanova E.M., Safonov I.E.* Methods of decision making in distributed CAD / CAM / CAE systems. Methods for complex estimation of options in terms of certainty of the initial information // Quality and CALS-technologies. 2006. № 3(11). P. 35-40 (in Russian).

[3] *Kozhevnikov A.M.* Methods of CALS-technologies to optimize the choice of electrical and thermal elektroradio products regimes // ITPP. 2000. № 3. P.23-26, (in Russian).

[4] *Biryukov S. I.* Optimization. Some theoretical aspects. M.: MZ-Press, 2003. 246 p. (in Russian).

[5] *Vasilyev V.I.* Recognition systems. Reference book - Kiev.: Naukova Dumka, 1983. 422 p., (in Russian).

Nevlyudov Igor Shakirovich, Doctor of Technical Sciences, professor, head of TPRA Department of Kharkiv National University of Radio Electronics. Research interests: instrumentation technology, flexible manufacturing systems, robotics. Address: 14 Lenin Avenue, Kharkov 61166, Ukraine, tel. (057) 702-14-86. E-mail: tapr@khture.kharkov.ua

Andrusevich Anatoly Alexandrovich, Director of Aviation Krivoy Rog College. Research interests: instrumentation technology, process monitoring, production of electronic systems, improving the reliability of electronic systems. Address: 1 Tupolev St., Krivoy Rog 50045, Ukraine, tel. (0564) 27-56-79. E-mail: uchebotdel@kk.nau.edu.ua

Khatnyuk Igor Stanislavovich, Postgraduate student of TPRA Department, Kharkov National University of Radio Electronics. Research interests: instrumentation technology, designing of flexible PCBs, improving the reliability of electronic systems functioning. Address: 14 Lenin Avenue, Kharkov 61166, Ukraine, tel. (057) 702-14-86. E-mail: tapr@khture.kharkov.ua