ANALYSIS OF THE COMPLEXITY OF ALGORITHMS FOR FINDING THE COEFFICIENTS OF THE MATHEMATICAL MODEL OF LOW-INTENSITY ELECTRORETINOSIGNAL

Pavlo Tymkiv

Ternopil National Ivan Puluj Technical University, Rus'ka str. 56, 46001, Ternopil, Ukraine; t_pavlo_o@ukr.net

Abstract: The complexity of methods of parametric identification is analyzed and their comparison is carried out at definition of coefficients of mathematical model of response of a retina of an eye at decrease in intensity of test light irritation. The algorithm of parametric identification of the mathematical model of retinal response based on direct complete search has a significant time complexity, which prevents rapid readjustment of the expert system in the case of remote, automated processing of low-intensity retinal response to diagnose the functional state of the body. Therefore, it is necessary to study the complexity of search algorithms and apply other approaches to solving problem of optimization or parameter identification. In this case, the criterion *K* of the optimality of the selection of coefficients will be the proximity of the simulated retinal response to the reference, pre-developed response.

Keywords: *electroretinosignal, parametric identification, complexity, optimization, low intensity.*

1. Introduction

With the development of scientific and technological progress, the negative man-made impact on the human body increases. There is a need for research to identify and early diagnosis of the risks of toxicity. A study of the Environmental Health Criteria 223 medical standard showed that the detection of risks (type of toxins, dose, duration of exposure, etc.) and early diagnosis of toxicity of the body is possible by assessing small characteristic changes in retinal response to light stimuli (electroretinosignals (ERS)). At the same time, the use of electroretinography helps to increase the level of objectivity of medical research, reduce the time of examination and the possibility of autonomous, remote use.

The obtained low-intensity retinal response is significantly distorted by noise due to the decrease in the ratio of signal energy to noise energy. At the same time, the complexity of processing is further increased due to the unknown or hidden nature of the toxin and its effect on the human body, which is reflected in changes in the amplitude-time parameters of waves or the emergence of new wave elements in ERS. For processing low-intensity ERS use optimal (morphological parameters of the signal) filtering, which in the case of autonomous and/or remote application, requires rapid reconfiguration, requires optimization and reduction of pre-processing time of the obtained low-intensity ERS.

2. Investigation of the complexity of methods for finding the coefficients of the mathematical model of low-intensity ERS

In computational complexity theory, the term "computational process complexity" is defined as the result of estimating the resources (usually time) required to perform an algorithm [1]. The limiting behavior of such complexity with increasing size of the problem will be called asymptotic time complexity.

Let *A* be the method of finding the coefficients of the mathematical model in the adaptive-recursive filter, and *N* be the dimension of the computational array of coefficients from the whole set of selection. Denote - a function that gives the upper limit of the maximum number of basic operations (addition, multiplication), which must perform method *A*, solving a problem of dimension *N*.

All methods are divided into polynomial and exponential, and some method *A* can be considered polynomial if it increases no faster than some polynomial from *N*. Otherwise, method A can be considered exponential. Typically, exponential problem-solving methods involve a complete direct search of all possible options, and due to the practical impracticability of solving such a problem, other approaches are developed for them. However, if there is a certain exponential algorithm for finding the optimal solution of such a problem, then another, more efficient (in the sense of less complexity) polynomial algorithm is used for practical implementation. In this case, there are not necessarily optimal solutions, but only acceptable for this type of problem or area of application of the solution (close to optimal). Such polynomial algorithms can differ significantly depending on the degree of the polynomial that approximates.

Consider one of the components of the concept of "complexity of the method" - time complexity [2,3]. By time complexity we mean the time spent by the algorithm of the method to obtain the final results. To record the time complexity, use the expression:

$$f_A(N) = O(g(N)) \quad (1)$$

where, $f_A(N)$ grows as $g(N)$ for N.

If there is a positive constant $C_{const} > 0$, such that

$$\frac{\lim_{N\to\infty} f_A(N)}{g(N)} = C_{const} \tag{2}$$

The estimate *O(g(N))* will be called the time asymptotic complexity of such an algorithm. In this case, the estimate *O(g(N))* for the function $f_A(N)$ can be used when the exact value $f_A(N)$ is unknown, and only the order of increasing time is known, which is spent when solving a problem with dimension *N* using this method *A*. Such exact values $f_A(N)$ will be depend on this implementation, while *O(g(N))* will be a characteristic of the method itself. In the case when the asymptotic time complexity of this method is equal $O(N^2)$ (then this algorithm will be called quadratic), with increasing *N*, the time to solve the problem increases by N^2. The exponential time complexity of a method in terms of complexity theory is written as $f_A(N) = O(k^N)$, where *k* – is an integer that is greater than one.

According to the classification given by D. Knut, there are both upper and lower estimates of complexity. In this case, the upper estimate of complexity will be determined from the expression:

$$O(f(N)) \ - \ \{g(N) \mid \exists \ \ C > 0 \ \ i \ \ N_0 > 0 : \mid g(N) \mid g(N) \ \ \le \ \ Cf(N), \forall N \ge N_0\} \tag{3}$$

Thus, the lower estimate of complexity will be determined by the expression:

$$\Omega(f(N)) \ - \ \{q(N) \mid \exists C > 0 \ \ i \ \ N_0 > 0 : \mid g(N) \ \ \ge \ \ Cf(N), \forall N \ge N_0\} \tag{4}$$

The following effective complexity estimates are defined as follows:

$$\theta(f(N)) \ - \ \{g(N) \mid \exists C_1, C_2, \ \ N_0 \ \ > \ \ 0 : C_2 f(N) \ \ \le \ \ g(N) \ \ \le \ \ C_2 f(N), \forall \ \ \ N \ge N_0\} \tag{5}$$

In this case, *O(f(N))* will be used to indicate the upper estimate of the growth rate of such functions or to indicate the totality of all functions that grow no faster than the function *f(N)*. If the method is executed for the specified time *O(f(N))*, it means that the time for its implementation can be limited at the top by the value of the function *O(f(N))* for all inputs of dimension n. Execution time in the worst situation is also limited by the function O (f (N)), in this situation it is called the execution time of the algorithm.

Estimation $\Omega(f(N))$ is used to indicate lower estimates of the growth rate of functions or to indicate the set of all functions that grow slightly slower than the function *f(N)*.

The estimate *q(f(N))* is used to indicate functions of the same order as the function *f(N)*. This method is necessary to describe the "optimal" methods. If a certain method of solving the problem processes inputs with data of size *n* for such time cn^2, where *c* – is a certain constant, and the time complexity of such an algorithm is $O(n^2)$, for the whole number *n*, except for a finite (i.e. empty) set, non-negative values. Such a record $O(n^2) \cap \Omega(n)$ will denote a class of functions that have a velocity of not less than *n*, but not more than n^2.

Spatial and temporal complexity, i.e. functions in relation to the size of the problem, are two fundamental estimates of efficiency in the analysis of methods. We will consider the complexity of computational methods as an estimate of the resources (often time) required for the operation of the algorithm of this method.

The complexity of the algorithms of the methods will be measured by the required resources, i.e. the duration of calculations or the required amount of memory. The time complexity measured in this way is independent of implementation. Estimation of numerical computational complexity qualitatively demonstrates the influence of the amount of input data on time and the amount of memory [4,5].

The mathematical model of the reference low-intensity ERG in each cycle n = 0,1,2,… can be represented by summing:

$$a_2 s_{n-2} + a_1 s_{n-1} + s_n = \xi_n \tag{6}$$

where, s_{n-1}, s_{n-2} – initial conditions [6,7].

In the prototype method, sequential direct selection of coefficients (parameters) a_1, a_2, and initial (baseline) conditions s_{n-1}, s_{n-2} evaluated the proximity of the simulated $\hat{s}_n$ ERS to the previously known S_n retina ERS response (for example, in a database displaying known response types) by the value of the criterion:

$$\kappa = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(s_n - \hat{s}_n)^2} \tag{7}$$

where *N* – is the number of values S_n and $\hat{s}_n$. And the difference between the value of the reference and simulated low-intensity electroretinogram (ERG) is defined as the root mean square error (RMS) of the simulation, which will serve as a measure of the choice between the prototype method and the improved method when using binary classification. The block diagram of the prototype method is shown in Fig.1.

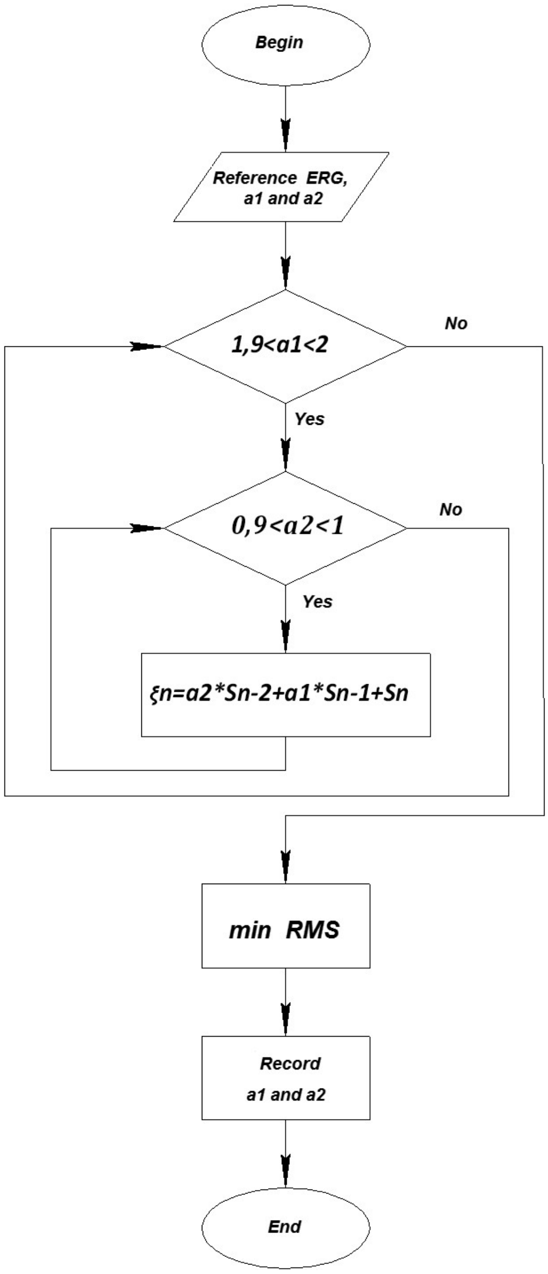


Fig. 1. Algorithm for finding coefficients by direct directed search

In this case, the complexity of this method of estimating the parameters of the model will be determined by:

$$C = n_{a_1} \cdot n_{a_2} \cdot n_{S_{n-1}} \cdot n_{S_{n-2}} \cdot n_{\tau} \tag{8}$$

where, *n* • – the number of values, coefficients a_1, a_2, the number of initial conditions s_{n-1}, s_{n-2}, the number of ERS samples. The value of all numbers *n* • depends on the value of the predetermined criterion (*K*) and can be determined, for example, by the number of smaller digits *m* • in the binary code of the corresponding coefficients, or the initial conditions and the number of retinal response samples. Hence, the complexity of the method is exponential, type *O(2M)* [8], where:

$$M = m_{a1} + m_{a2} + m_{Sn-1} + m_{Sn-2} + m_{\tau} \tag{9}$$

which necessitates the improvement of the method. In [9] developed an improved method for finding the coefficients of the computational mathematical model of low-intensity ERS, the block diagram of which is shown in Fig. 2.

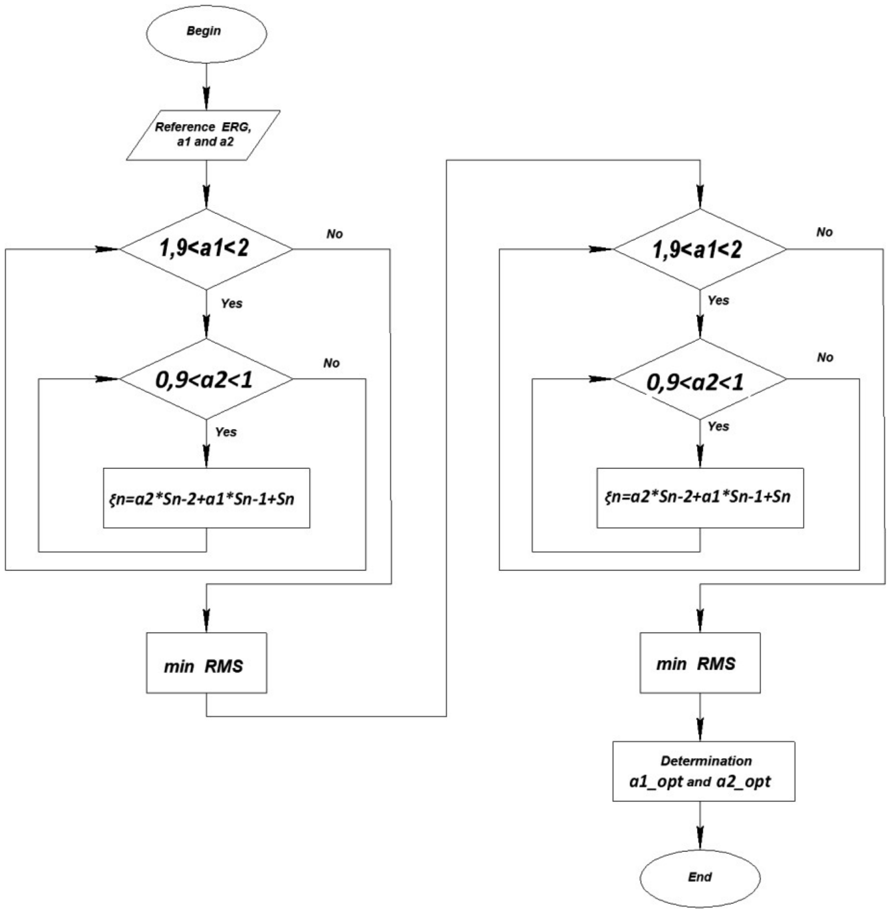


Fig. 2. Optimized method for determining the coefficients of the Hooke-Jeeves search strategy.

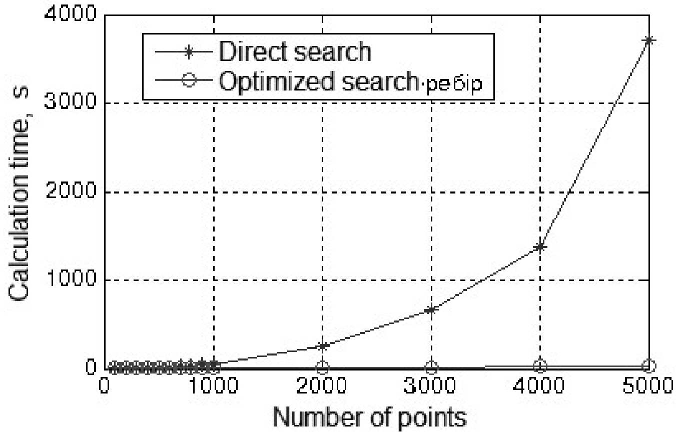
To compare and evaluate the prototype method and the improved method, we determine the time of selection of coefficients by the prototype method and the improved method at different numbers of search points (table 1).

Table 1

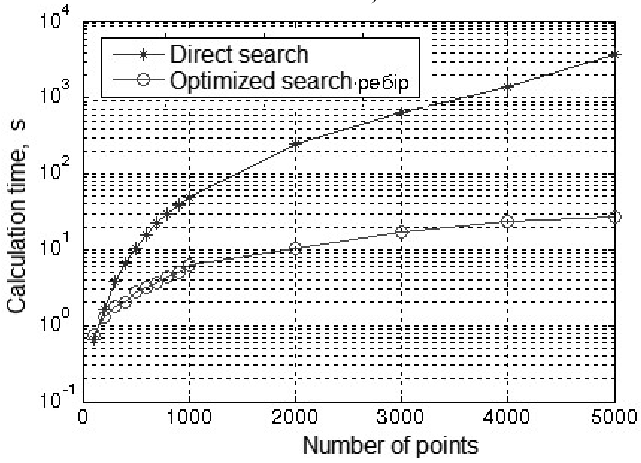
Comparison of calculation time of algorithm of direct full directed search and search on the basis of Hooke-Jeeves search strategy

Number of points, *N*	**Processing time with direct directional search algorithm, *s***	**Processing time with the algorithm according to Hook-Jeeves strategy, s**
100	0,6698	0,7666
200	1,6306	1,3215
300	3,7987	1,7511
400	6,7140	2,0514
500	10,5308	2,7493
600	15,7425	3,1958
700	22,2913	3,6477
800	29,5568	4,4972
900	37,9793	5,0008
1000	48,1306	5,8352
2000	245,9232	10,4593
3000	658,2834	16,9708
4000	1365,1457	23,1177
5000	3702,8461	30,1974

Graphically, the results of the comparison are shown in Fig. 3



a)



b)

Fig. 3. The dependence of the search time of the coefficients by the prototype method and the improved method on the number of points (a - linear scale, b - semi-logarithmic scale)

However, the time advantage of the improved method of finding coefficients requires research and comparison of their accuracy. In this case, the measure of accuracy will be the RMS modelling of low-intensity ERS according to pre-determined coefficients by two methods.

As can be seen from Fig. 3, direct search has polynomial complexity. And the optimized search on the basis of Hooke-Jeeves search strategy, allowed to reduce time of definition of algorithm. However, its implementation is somewhat limited due to shortcomings: The algorithm is based on cyclic motion in coordinates, which can lead to the degeneration of the algorithm into an infinite sequence of searches through research without searching for a sample. When using the method, it is assumed that the objective function under study is unimodal, i.e. has only one optimum in the study interval. This means that the solution may be a local minimum, not a global one, if there are several minima in the study area.

Therefore, the application of genetic algorithms for solving the optimization problem is promising in terms of finding the optimal values of the coefficients of the mathematical model of low-intensity ERS.

Genetic algorithm is a method that reflects the natural evolution of problem-solving methods, and especially optimization problems. Genetic algorithms are search procedures based on the mechanisms of natural selection and inheritance. They use the evolutionary principle of survival of the most adapted individuals. They differ from traditional optimization methods by several basic elements. In particular, genetic algorithms:

a) process not values of parameters of the task, and their coded form;

b) carry out search of the decision proceeding not from a single point, and from their some population;

c) use only the objective function, not its derivatives or other additional information;

d) apply probabilistic rather than deterministic selection rules.

These four properties, which can also be formulated as parameter coding, population operations, the use of minimum task information, and randomization of operations, result in the stability of genetic algorithms and their superiority over other widely used technologies.

3. Conclusions

It was found that the processing of the retinal response to test light stimulation with reduced intensity, methods regulated by the ISCEV standard (bandpass filter) was ineffective due to the reduction of the ratio of retinal response energy to noise energy in the selected ERS. And the use of coordinated filtration (in the value of the Norse filter) or optimal filtration in the Kolmogorov-Winner value (with the criterion of optimum minimum deviation) to process such a response is further complicated by an unknown change in retinal response or the appearance of new segments when exposed to toxins.

The study of the method of parametric identification of the model of low-intensity retinal response on the basis of direct directed search, requires a significant amount of time, i.e. great time complexity. Therefore, there are works in which the possibilities of using the method of parametric identification based on the Hooke-Jeeves algorithm are investigated.

The application of the optimization method based on the Hooke-Jeeves algorithm has disadvantages, which in some cases limits its application in expert systems for processing low-intensity ERS. Therefore, the use of genetic algorithms is promising. However, will require additional research and comparisons of complexity, statistical research and determination of the reliability of the decision on the choice of method for determining the coefficients of the mathematical model of low-intensity retinal response ERS

References

1. Papadimitriou, C. H. (1994). *Computational complexity*. Addison-Wesley.

2. Cobham, A. (1965). The intrinsic computational difficulty of functions. *Logic, methodology and philosophy of science*, Proceedings of the 1964 International Congress. Amsterdam, The Netherlands.

3. Tymkiv, P.O., Yavorskyi, B.I. (2011). Parametrychna identyfikatsiia matematychnoi modeli elektroretynosyhnalu. *XV naukova konferentsiia*.. Ternopil, Ukraine.

4. Tymkiv, P.O., Zabytivskyi, V.P., Yavorskyi, B.I. (2016). Syntez filtru Kalmana dlia opratsiuvannia nyzko intensyvnoho elektroretynosyhnalu. *Vymiriuvalna ta obchysliuvalna tekhnika v tekhnolohichnykh protsesakh*, *1*. 168-176.

5. Tymkiv, P.O. (2018). Veryfikatsiia udoskonalenoho metodu vyznachennia koefitsiientiv filtru Kalmana u nyzko intensyvnii elektroretynohrafii. *FUNDAMENTALNI TA PRYKLADNI PROBLEMY SUChASNYKh TEKhNOLOHII, Proceeings of the International Conference*. Ternopli, Ukraine.

6. Tkachuk, R., Yavorskyi, B. (2009). Metod pobudovy biotekhnichnoi systemy dlia otsiniuvannia elektroretynohram z pidvyshchenymy virohidnistiu ta efektyvnistiu. *Visnyk TDTU*, *14*(3), 102-110.

7. Tkachuk, R. A. (2009). Optymalna obrobka elektroretynosyhnalu dlia vyznachennia formy elektroretynohramy. *Vymiriuvalna tekhnika ta metrolohiia*, *70*, 9-13.

8. Bakulev, P.A. (2004). *Radyolokatsyonnye systemy: uchebnyk dlia vuzov*. Moskva: Radyotekhnyka.

9. Tymkiv, P.O., Demchuk, L.B. (2020). Application of Hooke-Jeeves algorithm for electroretinosignal processing, *Innovatsii partnerskoi vzaiemodii osvity, ekonomiky ta sotsialnoho zakhystu v umovakh inkliuzii ta prahmatychnoi reabilitatsii sotsiumu, Proceedings of the International Conference*. Kamianets-Podilskyi, Ukraine.