ИЗВЕСТИЯ КАФЕДРЫ UNESCO ГУАП «ДИСТАНЦИОННОЕ ИНЖЕНЕРНОЕ ОБРАЗОВАНИЕ»

Сборник статей

Выпуск 9

BULLETIN OF THE UNESCO CHAIR "DISTANCE EDUCATION IN ENGINEERING" OF THE SUAI

Collection of the papers

Issue 9

РОССИЯ, САНКТ-ПЕТЕРБУРГ   2024   SAINT PETERSBURG, RUSSIA

ISA EMEA&Pakistan (The International Society of Automation) and SUAI (Saint Petersburg State University of Aerospace Instrumentation) have organized the Twentieth ISA Europe, Middle East, Africa & Pakistan student paper competition (ISA EMEA&Pakistan SPC-2024). Papers of professors and the best students were included into this issue of the Bulletin of the UNESCO chair "Distance education in engineering" of the SUAI. Papers can be interesting for students, postgraduate students, professors and specialists.

It is my pleasure to congratulate the ISA St. Petersburg (Russia) Section and the St. Petersburg State University of Aerospace Instrumentation {SUAI) on successfully completing the XX ISA EMEA&Pakistan Student Paper Competition.

1 applaud the students who contributed their time, knowledge, and expertise to prepare a paper for this competition. 1 sincerely thank the review committee, including ISA Europe, Middle East, and Africa District student volunteers, who selected the papers for publication and awards.

Our community of students are engineers and leaders of the future. ISA welcomes their contributions in helping us achieve our vision of "creating a better world through automation", and our mission of "advancing technical competence by connecting the automation community to achieve operational excellence." ISA will continue to play an important role in students' continuing education and professional development.

1 extended *my* best wishes to all students, lectures, and committee members in the XX ISA EMEA&Pakistan Student Paper Competition.

наступающим новым годом!

Sincerely,

Prabhu Soundarrajan 2024
ISA Society President

I would like to extend congratulations to the ISA Russia Section, ISA EMEA&Pakistan District, and The Saint Petersburg State University of Aerospace Instrumentation (SUAI) for successfully organizing the 20th ISA International Student Paper Competition (The XX ISA EMEA&Pakistan SPC 2024).

As an educator and a member of ISA for over 50 years, I never tire of the opportunity to share with students the amazing challenges and personal rewards that a career in automation can bring. ISA is proud to have the opportunity to nurture the next generation of automation professionals.

We look forward to continuing the close relationship we have established between ISA, the Russia Section, District 12, and the SUAI. Through distance learning classes on project management and ongoing international online forums, we are developing new understandings in the technical, cultural, and personal arenas.

Congratulations to those who developed papers for this volume and to the advisory committee who had the difficult task of making paper selections.

Sincerely,

Gerald W. Cockrell

Gerald W. Cockrell
ISA Former President

Standards

Certification

Education & Training

Publishing

Conferences & Exhibits

ISA
67 Alexander Drive
P.O. Box 12277
Research Triangle Park, NC 27709
PHONE (919) 549-8411
FAX (919) 549-8288
E-MAIL info@isa.org
www.isa.org

AUTOMATION OF VERIFICATION OF THE PROPERTIES OF MATERIALS OF ADDITIVE MANUFACTURING WITH CARBON FIBER REINFORCED PRODUCTION OF FDM

Aleksandr Chabanenko

Saint Petersburg State University of Aerospace Instrumentation
E-mail: *Chabalexandr@gmail.com*

Abstract. *The article presents the results of mechanical tests of models made using two 3D printing technologies, layer-by-layer fused deposition (FDM) and continuous carbon fiber (CFF). Both technologies use materials based on Polylactide (PLA) or nylon (PA) reinforced with carbon fibers. The work includes both uniaxial tensile testing of the materials under study and metrological measurements of surfaces obtained using two 3D printing technologies. The test results showed a significant impact of the type of technology on the strength of the constructed models and on the quality of the technological surface layer. After analyzing the parameters of the primary profile, roughness and waviness, it can be clearly stated that the quality of the technological surface layer is much better for CFF models compared to FDM technology. In addition, the tensile strength of parts made of carbon fiber-rich material is much higher for specimens made with CFF technology compared to FDM*

3D printing technologies have been known since the 1980s. At this time, the first program based on CAD modeling and the first 3D printers were created. Among the most commonly used technologies are fused deposition modeling (FDM) and equivalent fused filament generation (FFF) methods. These technologies are popular because of their simplicity, the affordability of professional machines, and low material costs.

In addition, 3D printing contributes to the development of composite models with new innovative characteristics. Due to improved accuracy, FDM technology can be used to print enclosures for electronic equipment [1].

Today, models are made of various materials such as PLA (polylactic acid), ABS (acrylonitrile butadiene styrene), nylon (PA) and many of their modifications. In this way, FDM technology has found wide application in various industries, not only because of the advantages mentioned, but also because of the ease of post-processing of components, such as heat treatment, chemical treatment, machining, polishing, painting or coating.

Changing the composition of materials for FDM is popular, especially the addition of various substances to improve the mechanical properties. These additives may include components that reduce the hardness of the material to 30 Shore, flame retardants for applications in the aviation industry, and additives that enhance mechanical properties such as carbon fibers and fiberglass. Some additives provide electrical conductivity and have magnetic properties. For example, the addition of glass and carbon fibers is interesting because this material requires only a slight increase in the temperature of the printhead extruder. Glass and carbon fibers are widely used in selective laser sintering (SLS) technology, where their addition improves properties, increasing strength and reducing anisotropy of mechanical characteristics (especially in rheological testing) and accuracy, which is especially important in the production of thin-walled elements.

Compared with traditional laminated carbon fibers, carbon fibers made using FDM 3D printing technologies are highly durable and retain all the benefits inherent in carbon fibers due to the nature of 3D printing and the reinforcing power of carbon fibers.

In addition, it is important to note that the surface quality of products created using 3D printing is subject to careful analysis, including the study of the surface roughness and waviness profile parameters. This is of particular importance for FDM technology, where a single layer can be as thick as 0.2 mm. The impact of 3D printing parameters on surface quality, including roughness and waviness, is a subject of research, and analysis methods include both traditional measurement methods and modern approaches based on multi-scale analysis.

The addition of carbon fibers to PLA-based materials is justified and significantly improves their strength characteristics. This article examines not only the strength but also the surface roughness of 3D printed parts from carbon fibers derived from shredded PLA. These fibers are made by two common technologies: FDM and CFF [2].
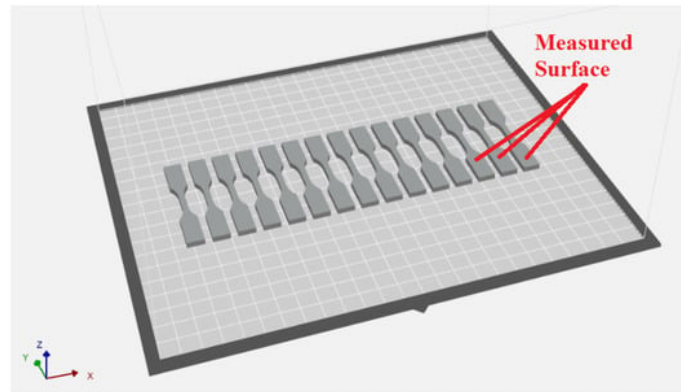
Fused Deposition Modeling (FDM) technology is one of the most common in the field of 3D printing. In this method, the plastic-based material is heated in the printhead to a temperature slightly below the melting point, then extruded through a nozzle and deposited onto the current layer of the model. This technology used materials such as PLA, ABS, nylon, as well as materials with additives such as carbon fiber.

Continuous Filament Fabrication (CFF) is an advanced technology based on the principles of FDM. The printer includes a print head with two independent extrusion nozzles: one nozzle is designed to print plastic filaments and the other to print reinforcing fiber. These two nozzles work in series: first, a plastic filament is printed, then a continuous fiber is applied to the previous layer of plastic, combining them. This makes it possible to create composite materials with using different types of reinforcing fibers, such as fiberglass, carbon fiber, Kevlar and others Fig. 1.

Fig. 1. Structural model: (a) PLA; (b) PA

The 3D models in Fig. 2 were created using SolidWorks software. These models were then approximated using triangles to create STL files [3].

To ensure accurate replication of the 3D sample models, two tolerance parameters were taken into account when saving the files in STL format: a linear tolerance of 0.01 mm and an angular tolerance of 1°. The resulting STL 3D model sample was approximated using 652 triangles, as shown in Fig. 3. Sample models were placed on a construction platform with the metrologically measured surface marked.

Fig. 2. Specimen dimensions

Fig. 3. Sample STL model

Fig. 4. Three-dimensional model on a construction tray

For the production of samples using FDM technology, PLA material with the addition of carbon fibers was used with the following process parameters: layer thickness – 0.2 mm, extruder temperature – 250°C, and filler – 95%.

To prepare samples using CFF technology, the following steps were performed:

Creating a 3D model and converting it to an STL file, similar to FDM models [3,4].

Upload the STL file to the Eiger cloud slicer software provided. Eiger allows you to control the placement of reinforcing fibers and adjust printing parameters such as fill pattern, fill density, roof, floor and wall layers, fiber type, fiber fill type, number of fiber layers, fiber orientation, fiber rings, and more. All of these parameters affect the mechanical properties of the final part.

Preparation of three sets of samples for experimental tests, each containing ten samples.

The first set of samples, marked with the letter "O" (Fig. 5a), had a 100% basic filling, which means a completely solid structure. These specimens were made exclusively of Onyx material with the installation of two wall layers 0.100 mm high and a consumption of 1.56 cm³ of Onyx material [1].

The second set of samples, labeled "ON" (Fig. 5b), had a basic fill of 37% pure Onyx with a triangular pattern, two wall layers, a layer height of 0.100 mm, and an Onyx material flow of 1.09 cm³.

The last set of samples was reinforced with carbon fibers and labeled with the letter "C". The arrangement of the fibres is shown in Fig. 5c. The fibre was laid in layers 5-8 and 19-22, forming 8 layers of reinforced carbon fibre in the sample. In the remaining layers, the filling with plastic material was 37%, and thanks to the reinforcing fiber, the layer height was set at 0.125 mm, which is in line with Eiger's recommendations. Onyx consumed 1.56 cm³ of material and 0.26 cm³ of carbon continuous fiber.
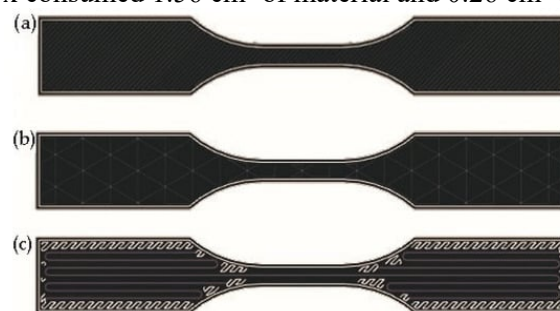

Fig. 5. CFF Sample Preparation – Specimen marked "O" (a);
A sample marked "ON" (b);
Specimen labeled "C" (c) – blue lines indicate reinforcing carbon fiber,
white lines indicate onyx filament.

In the article, samples made using additive technologies on selected machines were subjected to an assessment of the quality of the surface texture. The analysis of the surface topography was carried out on the basis of the analysis of the primary profile, waviness and roughness of the profiles.

Studies show that carbon fiber-reinforced materials have high strength and rigidity. This is confirmed by the results of tensile tests and the analysis of Young's modulus. CFF and FDM materials with carbon fibers are significantly superior to pure materials in terms of mechanical properties.

Fig. 6. An additive installation that uses 3 degrees of freedom

Structure Impact: Visual observation of the structure of materials using SEM images allows for a better understanding of the distribution of carbon fibers within the material. The uniform distribution in FDM and the concentrated presence in CFF account for differences in mechanical properties [3,4].

Advantages of polymer reinforcement. Materials created using CFF technology exhibit higher mechanical properties, which may be due to their special way of constructing internal structures.

It is necessary to pay special attention to the right choice of material and 3D printing technology to obtain models with certain mechanical characteristics. A detailed analysis and overview of the structure of materials also plays a key role in understanding mechanical properties.

Conclusion

The article is devoted to the study of the effect of the addition of carbon fibers on the characteristics of PLA in the process of 3D printing by the Fused Deposition Modeling (FDM) method.

Effects of carbon fibers: The addition of carbon fibers to materials significantly improves their mechanical properties, including tensile strength and modulus of elasticity. Specimens made using CFF technology, which involves the presence of carbon layers, demonstrate the highest strength.

Microstructure: SEM images indicate differences in the structure of materials, especially in the density of carbon fibers. FDM samples are characterized by a uniform distribution of carbon fibers, whereas CFF involves a layered distribution. This has an impact on the mechanical properties.

The study highlights the importance of choosing both material and 3D printing technology to achieve the desired mechanical characteristics and surface texture. CFF technology and carbon fiber-reinforced polyamide-based material are an effective combination to create high-strength parts with good surface roughness.

References

1. Chabanenko A. V., Kurlov A. V. 2021 Control the quality of polymers based on the model of Dzeno Journal of Physics: Conference Series

2. Chabanenko A. V., Kurlov A. V. and Tour A. C. 2020 Model to improve the quality of additive production by forming competencies in training for high-tech industries *J. Phys.: Conf. Ser.* 1515 052065.

3. Chabanenko A. V. and Yastrebov A. P. 2018 Quality Assurance of Hull Elements of Radio-Electronic Equipment by Means of Control System *J. Phys.: Conf. Ser.* 1515 052065.

4. Chabanenko A. V., Kurlov A. V. 2019 Construction of mathematical model of training and professional development of personnel support of additive production of REA IOP Conference Series: Materials Science and Engineering

5. Batkovskiy A. M., Kalachikhin P. A., Semenova E. G., Fomina A. V. and Balashov V. M. 2018 Conficuration of enterprise networks *Entrepreneurship and Sustainability Issues* 6(1) 311–28.

6. Batkovskiy A. M., Nesterov V. A., S emenova E. G., Sudakov V. A. and Fomina A. V. 2017 Developing intelligent decision support systems in multi-criteria problems of administrative-territorial formations infrastructure projects assessment *Journal of Applied Economic Sciences* 12(5) 1301-11.

7. Maiorov E. E., Prokopenko V. T., Mashek A. C., Tsygankova G. A., Kurlov A. V., Khokhlova M. V., Kirik D. I. and Kapralov D. D. 2018 Experimental study of metrological characteristics of the automated interferometric system for measuring the surface shape of diffusely reflecting objects Measurement Techniques 60(10) 1016-21.

8. Nazarevich S.A., Urentsev A.V., Kurlov V.V., Balashov V.M., Rozhkov N.N. Management of development of basic structures of technological systems of machine-building production (2019) IOP Conference Series: Materials Science and Engineering, 537 (4), article No. 042024.

9. Artjuhova M.A., Balashov V.M., Nazarevich S.A., Smirnova M.S. Evaluation of time to failure for radio transmitters under the radiation influence (2019) IOP Conference Series: Materials Science and Engineering, 537 (2), article No. 022016.

10. Nazarevich S., Smirnova M., Tushavin V. Integral criteria for evaluation of scientific and technical research (2015) International Journal for Quality Research, 9 (3), pp. 467-480.

11. Vinnichenko A.V., Nazarevich S.A., Kurlov V.V. Drifting models for evaluating the functional properties of products of innovative value (2021) Journal of Physics: Conference Series, 1889 (4), article No. 042074.

12. Vinnichenko A.V., Nazarevich S.A., Kurlov V.V. Drifting models for evaluating the functional properties of products of innovative value (2021) Journal of Physics: Conference Series, 1889 (4), article No. 042074.

13. Vinnichenko A.V., Nazarevich S.A., Kurlov V.V. Drifting models for evaluating the functional properties of products of innovative value (2021) Journal of Physics: Conference Series, 1889 (4), article No. 042074.

14. Artjuhova M.A., Balashov V.M., Semenova E.G., Nazarevich S.A. The quality of aerospace equipment production analysis (2019) IOP Conference Series: Materials Science and Engineering, 537 (3), article No. 032023.

15. Korshunov G.I., Nazarevich S.A. Parametric Models of the Product Novelty Assessment Through the Basic Structures Approach (2019) IOP Conference Series: Earth and Environmental Science, 272 (3), article No. 032142.

16. Tushavin V.A., Semenova E.G., Smirnova M.S., Frolova E.A. Comparison of qualitative assessments of employees work by randomized indicators ARPN Journal of Engineering and Applied Sciencesthis link is disabled, 2015, 10(16), pp. 7280–7287.

17. Plasticity Estimation with Combined Hardening for the Study of Deformation Processes of Structural Materials under Different Modes of Low-Cycle Loading Chabanenko Alexander Valerievich, Rassykhaeva Maria Dmitrievna Certificate of registration of the computer program 2021619545, 11.06.2021. Application No. 2021618914 dated 11.06.2021.

TRANSITION FROM DIFFERENTIAL EQUATIONS TO PARTIAL EQUATIONS DERIVATIVES TO ORDINARY DIFFERENTIAL EQUATIONS FOR SOLVING THE PROBLEM OF SYNTHESIS OF NONLINEAR CONTINUOUS SYSTEMS WITH DISTRIBUTED PARAMETERS

V. I. Goncharova

Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
goncarovav344@yandex.ru

Abstract. *The paper discusses the implementation of the transition from partial differential equations to ordinary differential equations to solve the problem of synthesizing nonlinear continuous systems with distributed parameters.*
Keywords: *automatic control system (ACS), ACS with distributed parameters, method of separation of variables (Fourier), partial differential equations.*

Let's consider the synthesis of object control with distributed parameters using the example of the plate deformation process. Unlike systems with concentrated parameters, where ordinary differential equations are used, in systems with distributed parameters the task is complicated by taking into account several variables in the state space. This entails a number of difficulties and, as a result, insufficient knowledge, as well as the lack of an effective mathematical apparatus for timely decision-making.

In order to simplify and make it possible to extend the generalized Galerkin method to systems with distributed parameters, let us consider a classic example of the transition from partial differential equations to ordinary differential equations, and, extending the well-known parametric synthesis method, Galerkin, [1-8] to the resulting transition equation.

Since a continuous system is considered in this example, it is necessary and sufficient to use the example from [9], where the control of the plate deformation process is considered. Using data from the source [10], we apply the transition from partial differential equations to ordinary differential equations [11] and extend the resulting equation to the generalized Galerkin method.

It is necessary to find a function satisfying in the domain

$$D = \left\{ (x,y) \in R^2 : \ 0 < x < \pi, 0 < y < \pi \right\}$$

the following equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -(\pi - x)xy,$$

or

$$K_1(x,y)\frac{\partial^2 u}{\partial x^2} + K_2(x,y)\frac{\partial^2 u}{\partial y^2} + K_3(x,y)\frac{\partial u}{\partial x} +$$

$$+K_4(x,y)\frac{\partial u}{\partial y} + K_5(x,y) \cdot u = f(x,y).$$

with boundary conditions

$$u(0,y) = u(\pi,y) = u(x,0) = u(x,\pi) = 10$$

или

$$U(x,y) = U(a,y) = U(x,0) = U(x,b) = d.$$

When $a=b=\pi$, $c=-1$, $d=10$ and $n=9$, let's define the continuous functions of the equation $K_1(x,y)$, $(K_1>0)$, $K_2(x,y)$, $(K_2>0)$, $K_3(x,y)$, $K_4(x,y)$, $K_5(x,y)$ and numeric parameters a,b,c,d.

$$K_1(x,y)=1, K_2(x,y)=1, K_3(x,y)=0, K_4(x,y)=0, K_5(x,y)=0,$$
$$a=\pi, b=\pi, c=-1, d=10.$$
$$f(x,y)=c \cdot (a-x) \cdot x \cdot y$$

We will get the exact solution $U(x,y)$, by decomposing the function into a double Fourier series

$$U(x,y) = d + \sum_{k=1}^{M}\sum_{m=1}^{M} H_{km} \sin(k \cdot \pi x / a)\sin(m \cdot \pi y / b).$$

Let's determine the accuracy of the solution 0.001. This accuracy will be ensured when

$$M=27$$

Let's find the coefficient H_{km}

$$i = 1...M$$
$$j = 1...M$$

$$H_{i-1,j-1} = \frac{-4}{\pi^2 \cdot \left(i^2 + j^2\right)} \times \left(\int_{0}^{a}\int_{0}^{b} f(x,y) \cdot \sin\left(\frac{\pi \cdot i \cdot x}{a}\right) \cdot \sin\left(\frac{\pi \cdot j \cdot x}{b}\right) dy\ dx\right)$$

Thus, we get the exact solution

$$U(x,y) = d + \sum_{k=1}^{M}\sum_{m=1}^{M} H_{k-1,m-1} \sin(\frac{k \cdot \pi x}{a})\sin(\frac{m \cdot \pi y}{b}).$$

and the table U_1 in the field of D

$$i = 1...10$$
$$j = 1...10$$

$$U_{1\,i,j} = U\left(a \cdot \frac{i}{10}, b \cdot \frac{j}{10}\right), \quad U_1 = \begin{matrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n+1,m+1} & ... & a_{m+1,n+1} \end{matrix} .$$

The graph of the exact solution is shown in Fig. 1
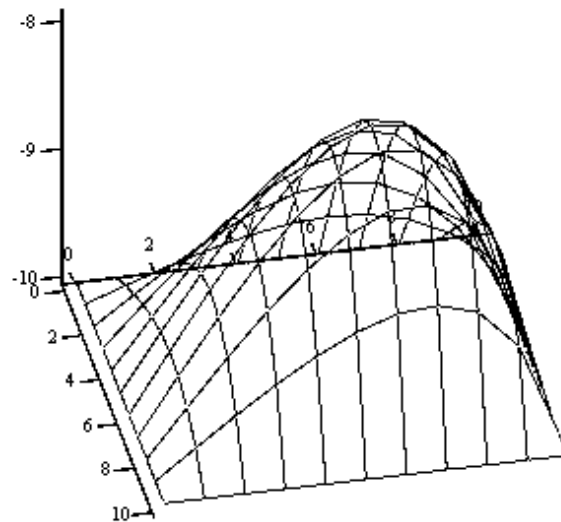
Fig. 1. A graph of the exact solution

To obtain an approximate solution, we introduce $n_1 = \sqrt{n}$ – the order of the trial solution

$$U_n = V(0,0,x,y) + \sum_{k=1}^{n_1}\sum_{m=1}^{n_1} C_{k,m} V(k,m,x,y).$$

$$n_1 = 3$$

Trial functions of the form

$$V_1(k,m,x,y) = x^k \cdot (a-x) \cdot y^m \cdot (b-y)$$

rationing

$$i = 1...n_1$$

$$j = 1...n_1$$

$$VV_{i-1,j-1} = \sqrt{\int_0^a \int_0^b (V_1(i,j,x,y))^2 \, dxdy}$$

We get normalized trial functions of the form

$$V(k,m,x,y) = if\left(k+m \neq 0, \frac{V_1(x,m,x,y)}{VV_{k-1,m-1}}, d\right).$$

To get the verification functions, let's take the trial ones

$$W(k,m,x,y) = V(k,m,x,y)$$

Let's set the operator corresponding to the left part

$$L_1(k,m,x,y,V) = K_1(x,y) \cdot \frac{d^2}{dx^2} V(k,m,x,y) + K_2(x,y) \cdot \frac{d^2}{dy^2} V(k,m,x,y) + K_3(x,y) \cdot \frac{d}{dx} V(k,m,x,y),$$

$$L(k,m,x,y,V) = L_1(k,m,x,y,V) + K_4(x,y) \cdot \frac{d}{dy} V(k,m,x,y) + K_5(x,y) \cdot V(k,m,x,y)$$

and we will find the coefficients of the system of equations $AC=B$ to find the coefficients of trial solutions C_k

$$i = 1...n_1$$

$$j = 1...n_1$$

$$B_{i-1+n_1(j-1)} = \int_0^a \int_0^b (f(x,y) - L(0,0,x,y,V)) \cdot W(i,j,x,y) \, dxdy$$

$$i_1 = 1...n_1$$

$$j_1 = 1...n_1$$

$$i_2 = 1...n_1$$

$$j_2 = 1...n_1$$

$$A_{i_1-1+n_1(j_1-1),i_2-1+n_1(j_2-1)} = \int_0^a \int_0^b L(i_2,j_2,x,y,V) \cdot W(i,j,x,y) \, dxdy$$

In the course of solving a system of equations $AC=B$, we get the vector of coefficients

$$C = A^{-1} \cdot B, \quad C^T = \begin{pmatrix} c_{11} & ... & c_{1n} \end{pmatrix}.$$

Let's define the matrices

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad C = [c_1 \quad c_2],$$

where a_1, a_2, a_{n+1}, b_1, b_2, b_{n+1}, c_1, c_2, c_{n+1} – the elements of the matrices obtained during the implementation of the transition from partial differential equations to ordinary differential equations.

Using the Laplace images, we get

$$\begin{cases} pX(p) = AX(p) + BU(p), \\ Y(p) = CX(p) + DU(p). \end{cases}$$

From the first equation of the system, we get

$$X(p) = (pI - A)^{-1} BU(p).$$

Then

$$Y(p) = C(pI - A)^{-1} BU(p) + DU(p)$$

We get the matrix transfer function

$$W(p) = \frac{Y(s)}{U(p)} = C\left(pI - A\right)^{-1} B + D = C\Phi(p)B + D,$$

where $\Phi(p)$ –the matrix exponent, which has the form

$$\Phi(t) = e^{At}.$$

Using the obtained value of the transfer function, we find the transfer function of the system in a generalized form

$$W(p) = C\left(pI - A\right)^{-1} B = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \left(\begin{bmatrix} p & 0 \\ 0 & p \end{bmatrix} - \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \right)^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 p - r \\ b_2 p - h \end{bmatrix},$$

where

$$r = \frac{b_1 a_4 + b_2 a_2}{a_1 a_4 - a_2 a_3}, \ h = \frac{b_1 a_3 - b_2 a_1}{a_1 a_4 - a_2 a_3}$$

Thus, by implementing the transition from partial differential equations to ordinary differential equations, and using the algorithm for obtaining a matrix transfer function, it is possible to solve the problem of ACS synthesis by the generalized Galerkin method.

References

1. Analysis and optimal synthesis of control systems on a computer / Edited by A.A.Voronov and I.A.Orurka. M.: Nauka, 1984. 340 p.

2. Algorithms of dynamic synthesis of nonlinear automatic systems / Ed. by A.A.Voronov and I.A.Orurka. Saint Petersburg: Energoatomizdat, 1992. 334 p.

3. Shishlakov V. F. Synthesis of nonlinear ACS with various types of modulation: Monograph/Saint Petersburg State University of Economics, 1999.268p.

4. Nikitin A.V., Shishlakov V.F. Parametric synthesis of nonlinear automatic control systems : Monograph / Edited by V.F.Shishlakov / Saint Petersburg State University of Economics, 2003. 358s.

5. Shishlakov V.F. Synthesis of nonlinear pulse control systems in the time domain. Izvestiya VUZov "Instrument Engineering", No. 12, 2003, pp.25-30.

6. Nikitin A.V., Shishlakov V.F. Parametric synthesis of an automatic braking control system for vehicle wheels // Izvestiya VUZOV "Instrument Engineering", No.5, 2004, pp.24-29.

7. *Shishlakov V.F., Gribkov V.N.* Synthesis of discrete self-propelled guns with delay by the method of orthogonal projections // Methods of research and design of automatic systems and devices / LIAP. L., 1990. pp.35-41.

8. Shishlakov V.F. Synthesis of nonlinear self-propelled guns with delay by direct variation method // Methods and means of processing and obtaining data in information management systems / LIAP. L., 1990. pp.30-37.

9. Velmisov, P. A. Mathematical modeling of nonlinear dynamics of the pipeline / P. A. Velmisov, Yu. V. Pokladova, U. D. Mizher // Automatization of control processes. – 2019. – № 3(57). – Pp. 93-101.

10. Aida-Zadeh, K. R. Control of the rod heating process using the current and previous feedback time / K. R. Aida-Zadeh, V. M. Abdullaev // Automation and telemechanics. – 2022. – No. 1. – pp. 130-149

11. Ankilov, A.V. Solving linear problems of mathematical physics based on methods of weighted residuals: a textbook / A.V. Ankilov, P. A. Velmisov, A. S. Semenov; under the general editorship of P. A. Velmisov. – 2nd ed. -Ulyanovsk : UlSTU, 2010. – 179 p.

RADIATED SIGNALS MODELS

A.F. Kryachko

Department of radio and optoelectronic complexes
SUAI
Saint Petersburg, Russian Federation

Abstract. *Emitted radio signals classification on the basis of ability to divide their analytical representation to spatial and temporal was considered. Decomposition options of space-time signals into basic functions was reviewed.*
Keywords: *space-time signal, complex signal, phased array, angular coordinates.*

Emitted radio signals classification on the basis of ability to factorize their analytical representation to spatial and temporal was considered. Four types of emitted radio signals were classified as result of instantaneous antenna diagram equation analysis. Only one of them is able to be representing as the product of antenna diagram and temporal signal function. Other types are not factorized by spatial and temporal arguments. *Emitted radio signals factorization of spatial and temporal arguments was given.* Also an example of space-time *emitted signal forming was given.*

Emitted radio signals classification

Type of emitted signals largely determines the processing algorithms at reception. Therefore, it is advisable to consider the classification of radiated signals on the subject of factorization their spatial and temporal variables and possible methods that give opportunity to create and process such signals at reception.

In most studies of radiolocation locator transmitting part divides on two blocks that are transmitter and transmitting antenna. The Transmitter generates a time function S(t) as a signal that then radiated by antenna. If condition $\Delta\omega L/c \ll 1$, where $\Delta\omega$ – spectrum width of the emitted signal, L – antenna aperture size, c – emitted oscillations propagation speed in the environment, then emitted signals time characteristics aren't related with spatial parameters of radiating them antennas.

Such emitted signals $u(t,\rho) = F(\rho)S(t)$, $|\rho| = 1$ can be written as multiplication of modulating the signal time function *S(t)* and spatial function F(**ρ**) which called radiation pattern:

$$u(t,\mathbf{\rho}) = F(\mathbf{\rho})S(t), \qquad |\mathbf{\rho}| = 1 \ . \tag{1}$$

The appearance of a radar station with phased array antennas and particularly with active electronically scanned array (AESA) gives an ability to construct a radar station transmitting part in a new way. In such radar stations a transmitting antenna and transmitter are combined, that allows to obtain emitted signals with new properties because of their spatial and temporal (S-T) characteristics inseparability. Further, signals that do not satisfy a condition (1) will call S-T emitted signals.

Specific types S-T signals have been considered in the literature on antenna theory [1, 2] and signal processing [3, 4, 5]. However, in the well-known signal theory and signal processing literature such signals are represented as a time argument function only.

Based on the consideration of such signals are conducted their classification [6, 7]. Meanwhile, emitted signals exist in an environment and in time representing S-T functions. Thereby in the basis of emitted signals classification are proposed to use a form of their S-T dependence.

Emitted signals classification with a form of their S-T dependence we will conduct on the basis of transmitting antenna instantaneous beam pattern analysis [2] that can be written:

$$f(t,\mathbf{\rho}) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} d\omega \int\limits_{V_r} \frac{\omega}{\omega_0} i(\omega,\mathbf{r}) \exp j(\omega-\omega_0)t \ \exp j\mathbf{k}(\mathbf{\rho},\mathbf{r})d\mathbf{r}, \ \mathbf{r} \in V_r, \tag{2}$$

where V_r – occupied by the aperture of the transmitting antenna area of space, ω_0 – central emitted signal spectrum frequency, $i(\omega,\mathbf{r})$ – space-frequency current distribution in the aperture of the transmitting antenna, $\mathbf{k}$ – wave vector, $\mathbf{k} = k\,\mathbf{\rho}$, $\mathbf{\rho}$ – unit vector.

From the analysis of (2) follows that emitted signal form can be determined from two conditions:

$$i(w,\mathbf{r}) = i(\mathbf{r})G(w), \mathbf{r} \in V_r , \qquad (3)$$

$$\exp\{j(\mathbf{kr})\} \cong \exp\{j(\mathbf{k_0 r})\}, \omega \in \Delta\omega \ . \qquad (4)$$

Fulfilment of equality in (3) means that the signal with an identical modulations emits from all points of the transmitting antenna aperture. Condition (4) means that for these values r и Δw modulating function signal $S(t)$ delay of the antenna aperture can be neglected. With usage of these conditions, we will provide classification of emitted signals.

Consideration begins with the simplest case when conditions (3) and (4) are carried out. Substituting them in (2) and considering that for most radio signals are performed:

$$\int\limits_{V_r} |i(t',\mathbf{r})|^2 \, d\mathbf{r} < \infty \ \text{ для } \ \omega \in \Delta\omega,$$

we obtain the provision of the emitted signal in the form of antenna transmission pattern $F(\boldsymbol{\rho})$ and time function generated by the transmitter $S(t)$ multiplication. We will call such emitted signals as a narrowband emitted signals or signals with separated spatial and temporal variables.

Now suppose that condition (4) is performed but condition (3) is not. In this case the instantaneous radiation pattern is recorded as:

$$f(t,\boldsymbol{\rho}) = \int\limits_{V_r} i_0(t,\mathbf{r}) \exp\{jk_0(\boldsymbol{\rho},\mathbf{r})\} d\mathbf{r}.$$

Thereby, the instantaneous radiation pattern that characterize the emitted signal is an inseparable S-T function. Such signals can be obtained by changing in time the distribution of current in the antenna aperture when the signal is emitted. Radiation pattern in each moment of time can be found with well-known formulas for a steady state [2] and presents changing in time radiation pattern for signal at a frequency ω_0. We will call signals emitted in this way as a spatio-temporal (S-T) narrow-band emitted signals with modulation of a field in antenna aperture.

Now let a condition (3) be performed and condition (4) is not. In this case, we assume as it has done in [2], that for most radio signals $\omega_0/\omega \approx 1$. Taking this remark into account the instantaneous radiation pattern has a form:

$$f(t,\boldsymbol{\rho}) = \frac{1}{2\pi} \int\limits_{V_r} i(\mathbf{r}) S(t - \frac{(\boldsymbol{\rho},\mathbf{r})}{c}) d\mathbf{r} \exp\{-j\omega_0 t\}. \qquad (5)$$

From (6) follows the instantaneous radiation pattern is an inseparable function of spatial and time variables. Emitted signals of the form (5) we will call S-T wide-band signals.

The difference of this radiated signals type from previous one consist of that for a flat aperture antennas in the direction perpendicular to the aperture the instantaneous radiation pattern can be presented as (1).

In case when either conditions (3) and (4) are not met, we can get S-T wide-band signal with a filed modulation in antenna aperture. The instantaneous radiation pattern for such type of signal determines as (2) and has most general form.

Thus, all set of emitted signals can be divided on four groups [8]. The first group includes the most commonly used in present time signals with separated spatial and temporal variables. Analysis and synthesis of such emitted signals is produced separately for radiation pattern and time function $S(t)$.

Three other groups of emitted signals cannot be represented in (1) form, and their synthesis and analysis are necessary to conduct with combined positions of signals and antennas theory.

Development of space-time emitted signals models

For synthesis and analysis convenience as well as generation and processing devices construction it is advisable to represent S-T signals in the form of an expansion in the basis functions. The specific character of such expansions is that S-T signals are not separable functions of spatial and temporal variables. Therefore, before we proceed directly to expansions, let's consider some requirements to the form of the expansions and the basis functions used in them.

On the basis of physical considerations, power and energy of emitted S-T signals are limited that is, the signals belong to the space L_2 where performed:

$$\int_{V_r} \int_{-\infty}^{+\infty} |i(t,\mathbf{r})|^2 d\mathbf{r} dt < \infty.$$

From the representation simplicity, analysis, generation and processing of S-T signals point of view, a natural choice of those expansions where the spatial and temporal variables are separated. In this case, the basis functions of the spatial argument should belong to $L_2(V_r)$, and time argument to space of $L_2(-\infty,\infty+)$. Consider such expansions for narrow-band S-T emitted signals with modulation of the field in antenna aperture.

One of the possible signals expansions in the form of a double row can be written with usage of the decomposition presented in [4]. It is written:

$$i(\mathbf{r},t) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} c_{ij}\eta_j(\mathbf{r})\psi_i(t) , \tag{6}$$

where $\eta_j(\mathbf{r})$, $\psi_i(t)$ are orthogonal functions of space and time variables, and c_{ij} are elements of coding matrix [3, 4].

As with increasing numbers of i and j coefficients c_{ij} tend to zero, then the number of row members (10) can be limited by values of K and N and rewritten in vector form:

$$i(t,\mathbf{r}) = \tilde{\boldsymbol{\eta}}(\mathbf{r})\mathbf{C}\boldsymbol{\psi}(t),$$

where $\boldsymbol{\eta}(\mathbf{r})$, $\boldsymbol{\psi}(t)$ are vectors of size $K \times 1$ and $N \times 1$ respectively, and $\mathbf{C}$ – coding matrix.

The instantaneous radiation pattern is also written in this form:

$$f(t,\boldsymbol{\rho}) = \tilde{\mathbf{F}}(\boldsymbol{\rho})\mathbf{C}\boldsymbol{\psi}(t), \tag{7}$$

where orthogonal functions $F_i(\boldsymbol{\rho})$ are defined:

$$F_i(\boldsymbol{\rho}) = \int_{V_r} \eta_i(\mathbf{r})exp\{jk(\boldsymbol{\rho},\mathbf{r})\}d\mathbf{r} .$$

The advantage of decomposition (6) is that it can use standard basic functions, but its drawback is double view row representation.

The current distribution in the antenna aperture and consequently emitted S-T signal, can be represented as decomposition into a single row with the basis of orthogonal functions spatial and time variables using the decomposition S-T functions $i(\mathbf{r},t)$ similar to Korunov-Loeve decomposition for random processes. In this case we obtain the representation of S-T signal in the form of sum of orthogonal time signals emitted through orthogonal radiation patterns.

Kotel'nikov expansion is one of the most used expansions of S-T signals. The possibility of such a decomposition is that real signals have a spectrum limited by upper ω_b frequency and emitted by antennas with the finite dimensions apertures, those with a limited spectrum of spatial frequencies.

The expansion of S-T signal (the field) in the Kotelnikov series by angular coordinates and time is recorded:

$$e(\alpha,\beta,t) = \sum_{k=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} f(\frac{k\pi}{\Omega_{b_\alpha}},\frac{n\pi}{\Omega_{b_\beta}},\frac{l\pi}{\omega_b}) \times \frac{\sin\Omega_{b_\alpha}(\alpha - k\Delta\alpha)}{\Omega_{b_\alpha}(\alpha - k\Delta\alpha)} \times$$
$$\times \frac{\sin\Omega_{b_\beta}(\beta - n\Delta\beta)}{\Omega_{b_\beta}(\beta - n\Delta\beta)} \times \frac{\sin\omega_b(t - l\Delta t)}{\omega_b(t - l\Delta t)},$$

where $\omega_{b\alpha}$ and $\omega_{b\beta}$ are the maximum value of spatial frequencies with respect to the angular coordinates α and β, $\Omega_{b\alpha} = x_m/\lambda$, $\Omega_{b\beta} = z_m/\lambda$, x_m and z_m are the maximum value of antenna dimensions by angular coordinates α and β , $\Delta\alpha=\pi/\Omega_{b\alpha}$, $\Delta\beta=\pi/\Omega_{b\beta}$, $\Delta t = \pi/\omega_b$ – intervals of discretization by angular coordinates and time.

In addition to the above expansions, another expansions also can be used. For example, the expansions of several variables function on the basis of several variables function. However, such representations do not allow to separate spatial and temporal variables, which makes difficult the analysis of S-T signals their generation and processing.

Space-time signal formation

Consider an example of S-T signal formation in case of one angular coordinate coding with usage of (7), where vector $F(\rho)$ has the size 2×1, consist of radiation patterns of half of the elements linear phased array antenna (picture 1) matrix C – unit size 2×2, and vector $\Psi(e)$ size of 2×1 in the first line has pulses of the seven-element Barker code with initial phases "0", and in the second line has components with initial phases "π" – picture 1a.

As a result, a complex S-T FM signal is emitted whose law intra-phase modulation will depend on the direction of radiation – "0+$\Delta\varphi(\alpha)$" and "π-$\Delta\varphi(\alpha)$" – picture 1b, where $\Delta\varphi$ is determined by the formula:

$$\Delta\varphi = \varphi(\alpha) = \frac{\pi D \sin(\alpha)}{\lambda},$$

where D – base between phase centers of transmitting radiation patterns, λ – wave length.
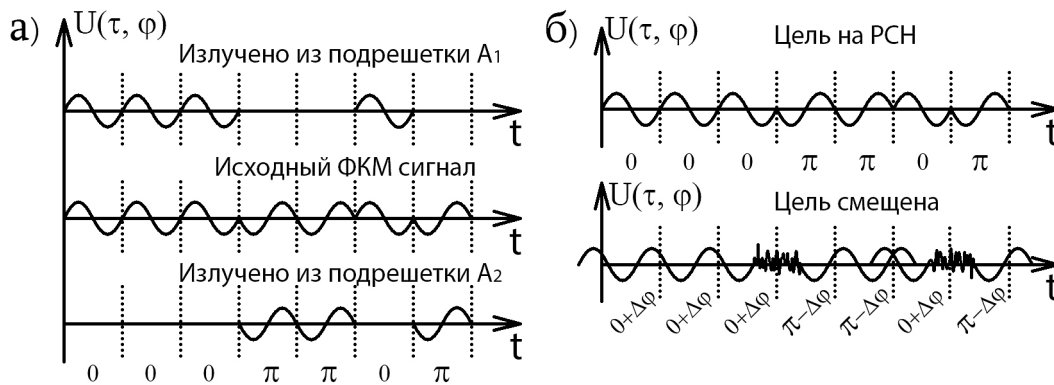
Picture 1. S-T signal obtained by phase modulation
of the center radiation position of the transmitting antenna

Dependence of the emitted signal phase structure from the distribution direction allows to evaluate the angular coordinates of objects without receiving antennas directional properties usage.

Literature

1.   Сканирующие антенные системы СВЧ т. 3 пер. с англ. Под редак. Г.Г.Мариова и А.Ф.Чаплина – М.: Советское радио, 1971, с.496

2.   Минкович Б.М., Яковлев В.П. Теория синтеза антенн – М.: Советское радио, 1969, С. 298.

3.   Дробович. Орби, Боннасье. Сжатие диаграммы направленности антенной решетки методом пространственно-временного кодирования. – Зарубежная радиоэлектроника 1971. № 3, С. 18-30.

4.   Погорелов А.И., Использование сложных диаграмм направленности в системах пеленгования для разрешения сигналов в широком секторе обзора – Радиоэлектроника. Респ. межвуз. науч. техн. сб. 1978 г.

5.   Радиотехника т. 10 кн. 1 под ред. Р.Г. Мириманова – М.: 1976.

6.   Тузов Г.И. Статистическая теория приема сложных сигналов – М: Советское радио, 1977, с.400.

7.   Рихачек., Упрощенный способ выбора радиолокационных сигналов Зарубежная радиоэлектроника, 1973, №2 с. 20-37.

8.   Лексаченко В.А., Поддубный С.С. О классификации излучаемых сигналов.- Теория и техника обработки сигналов в многоканальных локационных системах : Межвуз. сб.-М.: 1980, С. 3-8.

EXPERIMENTS ON THE RECOGNITION OF OBJECTS ON THE SURFACE BASED ON THE PROCESSING OF LASER DATA FROM SMALL AIRCRAFT

Victoria Afanaseva

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: victoria_afanaseva@mail.ru

Abstract. *Developing technologies in the field of computer vision and data processing are attracting increasing attention to the problems of segmentation of the Earth's surface. One of the key aspects of this area is the development of object classification algorithms.*
In this study, the focus is on the task of segmenting objects in three-dimensional space. The algorithm can be applied, for example, to control the environmental situation of a certain area of the earth's surface, and has been successfully tested on real data.
Keywords: *point cloud, segmentation of the Earth's surface, lidar.*

Introduction

Computer vision and data processing technologies are actively developing in the modern world, which attracts the attention of researchers to the complex tasks of segmentation of the earth's surface. One of the main directions of this development is the creation of effective algorithms for classifying objects in three-dimensional space based on data obtained from laser systems [1-2] installed on unmanned aerial vehicles.

The study covers algorithms for classifying objects in three-dimensional space with an emphasis on their application in monitoring the environmental situation on specific areas of the earth's surface [3]. The results of successful testing of the developed algorithms on real data are also presented, which highlights their potential in modern challenges of data processing and automated object recognition.

Segmentation of the Earth's surface

An experiment was performed in the study of this algorithm. To conduct an experiment to control the number of trees, data from a point cloud [4] was used from a "LAZ" file created as a result of flights from a small-sized onboard laser system [5]. First, forest indicators were extracted by classifying the point cloud by terrestrial and plant features, and then individual tree attributes were extracted by segmenting vegetation on each tree (Fig. 1).
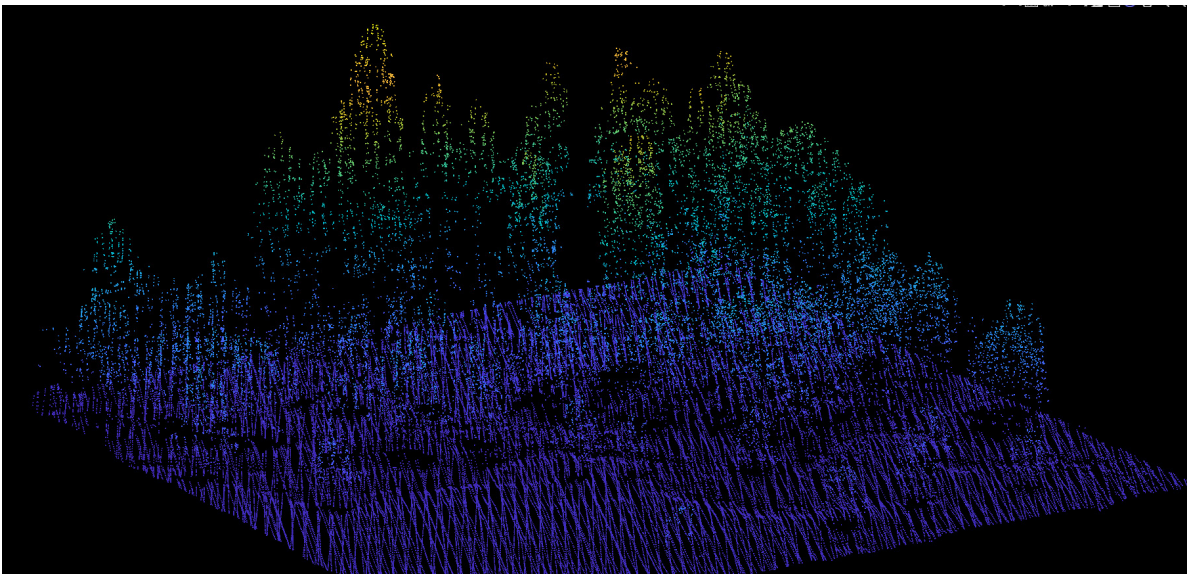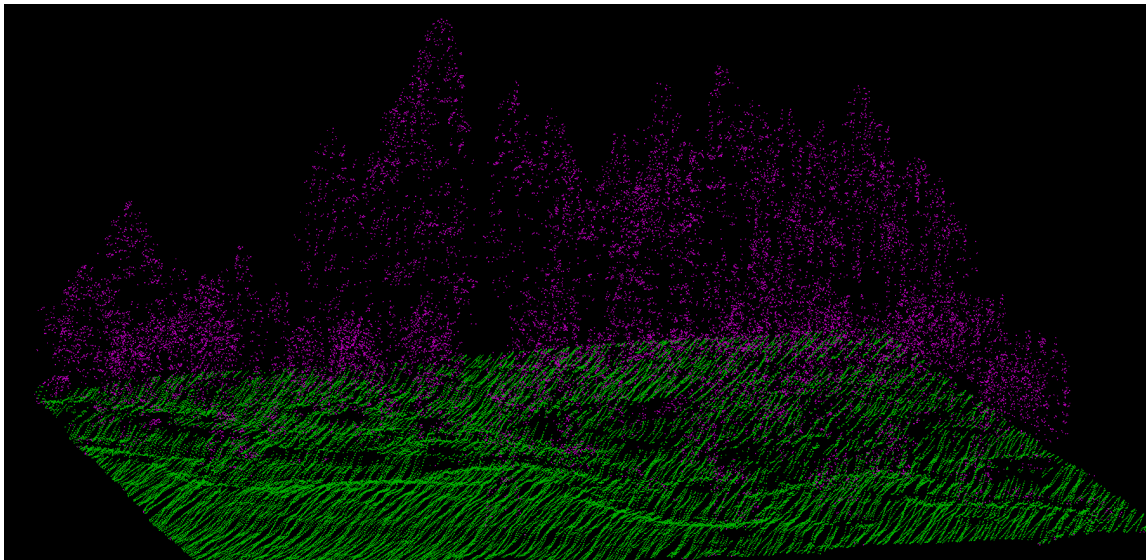

Fig. 1. Point Cloud Visualization

Fig. 2. Segmentation of the Earth's surface

Fig. 2 shows the results of processing a point cloud, where vegetation data is highlighted based on segmentation methods and information about forest indicators is extracted [6].

Relief correction

To eliminate the influence of the relief, an approach was applied as a result of which the normalization of heights was carried out.
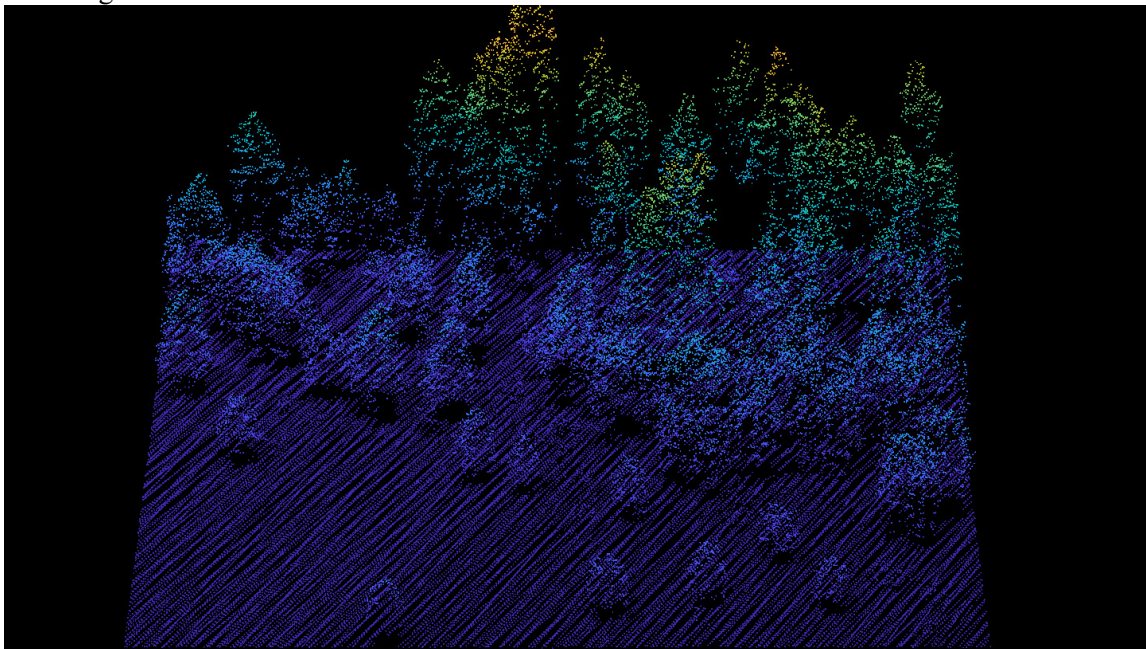

Fig. 3. Height normalization

To calculate forest indicators and tree parameters, points with normalized height were used as input data. Next, duplicate points along the x and y axes are eliminated. Interpolation has been applied to estimate the location of each individual tree in the point cloud, as shown in Fig. 3.

Search for forest and individual tree parameters

The extraction of forest characteristics itself occurs as follows. The point cloud is divided into grids according to a given size and its own forest indicators are calculated in each sector. It is assumed that the points with a normalized height below the specified cut height are the ground, and the rest are vegetation.

Fig. 4 shows the definitions of several forest indicators that were calculated based on the extracted point cloud data. The flat cover is the proportion of the forest covered by the vertical projection of the tree crowns. The gap fraction is the probability that a laser beam will pass through the crown without colliding with foliage or other vegetation elements. It can be calculated as the ratio of the number of irradiations of the earth to the total number of their returns. The leaf area index (LAI) is the amount of one—sided leaf area per unit of land area. The value of "LAI" is calculated using an equation using the following parameters: ang is the average scanning angle, GF is the gap fraction and k is the attenuation coefficient, which is closely related to the distribution of the angle of inclination of the sheet.
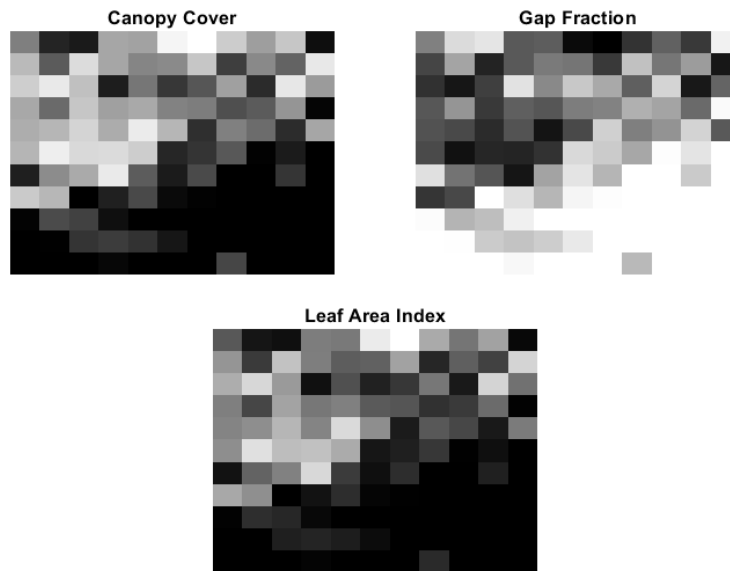
Fig. 4. Extracting metrics

Tree crown height models are raster representations of the height of trees, buildings, and other structures above the terrain. These models were used as input data for tree detection and segmentation (Fig. 5).
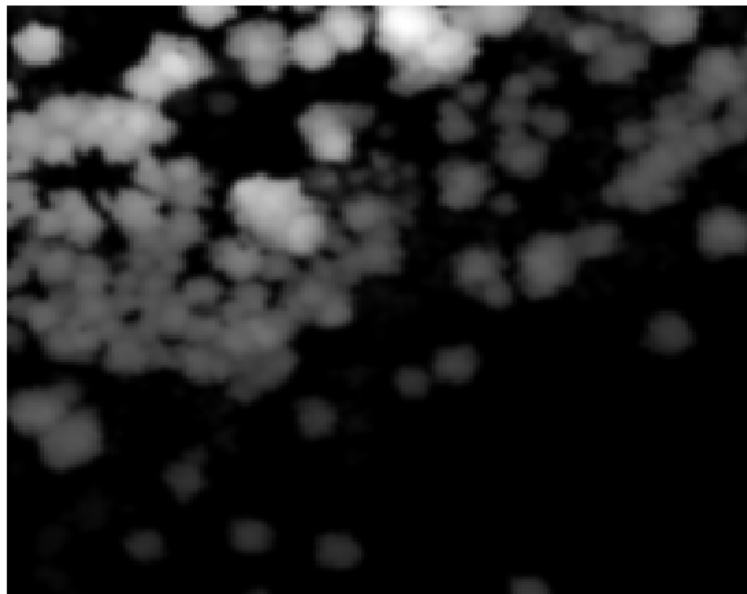
Fig. 5. Model of tree crowns

The detection of tree tops was performed through local maxima within the size of the models. To determine the top of the tree, an auxiliary function was used, which considers only points with a normalized height exceeding the minimum specified parameter (see Fig. 6).
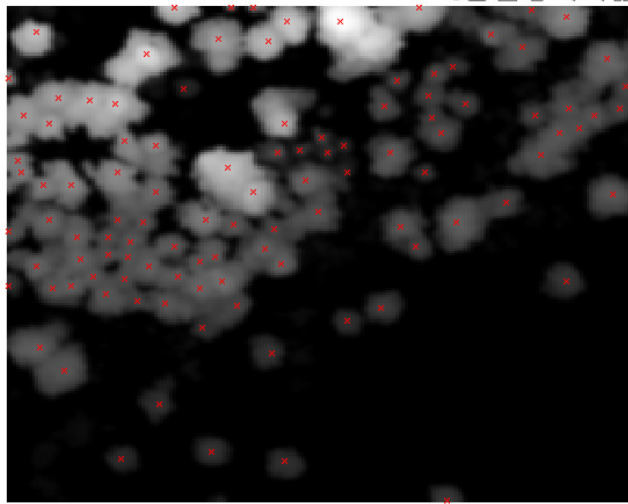
Fig. 6. Detecting tree tops

Marker-driven segmentation is used to highlight all trees, which allows you to highlight each individual tree. First, a binary marker image is created, where the vertices of the trees are marked with the value 1. Then filtering is applied by applying minima to the model complement to remove minima that are not the vertices of the tree. After that, segmentation is performed on the filtered model complement to separate the individual trees. At the end, individual segments of the tree are visualized (Fig. 7).
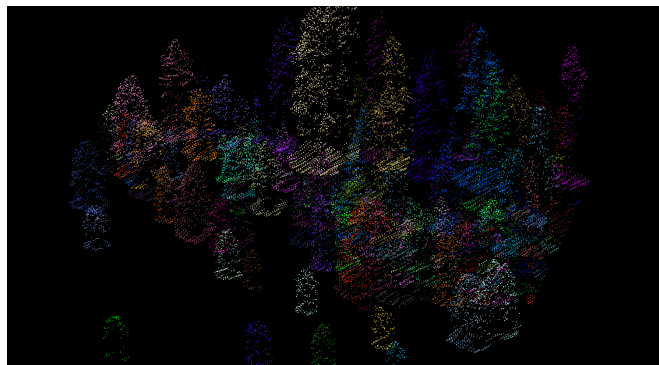
a) Side view
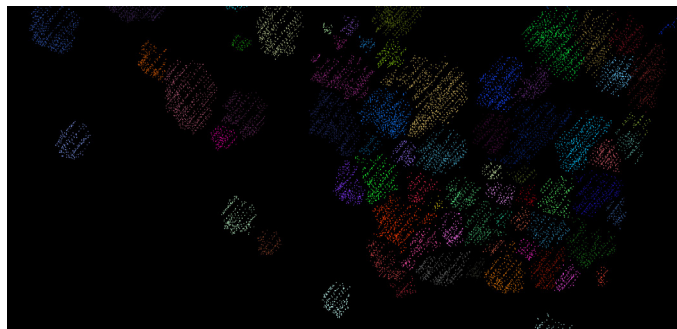
b) View from above

Fig. 7. Detection of individual trees

To determine individual attributes of a tree, points belonging to individual trees should be identified by labels [7]. Then the attributes of the tree are extracted, such as: the location of the top of the tree along the x and y axes, height, crown diameter and area of the tree, etc. Table 1 shows a part of the table with tree attributes.

Table 1 – Attributes of trees

Tree ID	Number of points	The X coordinate of the top of the tree	The coordinates of the top of the tree in Y	Tree height, m	Crown diameter	Crown area, m2
1	374	268671.550000000	4171857.95000000	29.5124272827084	7.37799500249552	42.7530000006176
2	22	268667.250000000	4171868.09000000	21.4752094430787	0.990502262071719	0.770549999979197
3	350	268669.770000000	4171873.89000000	24.2164764458571	6.98224215758542	38.2894999992351
4	53	268668.940000000	4171883.90000000	19.5086963270965	3.19561482441011	8.02044999954333
5	117	268669.140000000	4171886.82000000	19.3986687364654	4.02650370430216	12.7334499998830
6	101	268667.110000000	4171897.94000000	13.2987935773649	3.27535256602415	8.42569999975334
7	266	268671.540000000	4171906.01000000	14.9785056705464	7.40896716957415	43.1126999996261
8	40	268667.300000000	4171909.82000000	8.42299311099941	2.05450239154687	3.31515000011233

Thus, we found out the number of trees – 110 (it is enough to open the treeMetrics array), and also received information about each of the trees:
- Number of points in each tree,
- The x coordinate of the top of the tree,
- The y coordinate of the tree vertex,
- Height of the tree,
- Diameter of the crown,
- Crown Square.

ConclusionS

In the course of this study, an algorithm for segmentation and analysis of lidar data was developed and successfully tested. The results obtained emphasize the effectiveness of the proposed approach to lidar data processing for extracting information about forest areas and individual trees.

Therefore, it can be argued that this area is promising and in demand today in many areas where monitoring of the Earth's surface with high detail of the observed territories is required.

REFERENCE

1. Matrice 300 RTK platform. URL: https://www.dji.com/ru/matrice-300 (accessed 13.11.2023);

2. Zenmuse L1 module. URL: https://www.dji.com/ru/zenmuse-l1 (accessed 11/14/2023);

3. Formation of three–dimensional terrain models based on lidar surveys to identify structural changes in the Earth's surface / V. A. Nenashev, V. I. Afanasyeva, A. A. Zalishchuk [et al.] // Proceedings of MAI. – 2023. – No. 131. – DOI 10.34759/trd-2023-131-15 . – EDN XZHFRA.

4. Nenashev, V. A. Integration of heterogeneous information from lidar and camera in an on-board aviation system for observing the Earth's surface / V. A. Nenashev // Processing, transmission and protection of information in computer systems 23 : Collection of reports of the Third International Scientific Conference, Saint Petersburg, April 10-17, 2023. – Saint Petersburg: Saint Petersburg State University of Aerospace Instrumentation, 2023. – pp. 188-195. – DOI 10.31799/978-5-8088-1824-8-2023-3-188-195. – EDN UOCDED.

5. Afanaseva V.I. Creation of 3d terrain models based on lidar survey. Bulletin of the UNESCO department "Distance education in engineering" of the SUAI: Collection of the papers. Saint Petersburg, Issue 8. – SPb.: SUAI, 2023. – 23 p. ISBN 978-5-8088-1825-5;

6. Nenashev, V. A. Classification and recognition of ground objects in the flow of radar frames based on a neural network approach in the forward viewing area of airborne radars of a multi-position system / V. A. Nenashev, S. A. Nenashev // Proceedings of the International Conference on Computer Graphics and Vision "Graphicon". – 2023. – No. 33. – pp. 572-580. – DOI 10.20948/graphicon-2023-572-580 . – EDN TMUCPN.

7. Afanasyeva V.I. Classification of the observed zones on the generated images of the video frame stream based on segmentation in on-board monitoring systems of the Earth's surface //Collection of reports of the 74th International Student Scientific Conference of GUAP (April 19-23, 2021) – Saint Petersburg: GUAP, 2021. – pp. 126-133.

SMART LIGHTING

Emily Alaimo

University of Enna "Kore" – Italy
emily.alaimo001@unikorestudent.it

The paper simulates a network scenery using TrueTime for a smart lighting system. The entities involved are two sensors and an actuator located within a single wireless network. Specifically, there is an ambient light sensor managed by a fuzzy logic controller to adjust the brightness of the light source based on the surrounding brightness, and a PIR motion sensor modeled with MBSD to detect movement and identify whether it is 'presence of a person' based on the heat emitted by the body itself. The data collected by the sensors, such as brightness and presence detection, are transferred wirelessly to an actuator responsible for managing the lighting based on the received information. In particular, the actuator is tasked with turning on (with appropriate brightness) the light source in the presence of human presence, or turning it off/keeping it off in the absence of such detection.

Introduction

In the context of the evolution of intelligent technologies, lighting has assumed a central role in enhancing user experience within environments. In recent years, numerous studies and research have been conducted both in academic and industrial settings to develop and improve smart lighting systems. These studies have led to the introduction of a wide range of commercial solutions designed to meet a variety of needs, both in residential and commercial settings. Smart lighting systems offer a wide range of functionalities ranging from simple adjustment of light intensity to the creation of customized lighting sceneries.

The paper presented proposes a smart lighting system that leverages cutting-edge technologies to enhance user experience. Specifically, it utilizes an ambient light sensor and a PIR motion sensor which, through appropriate control methodologies, communicate with an actuator to dynamically adjust lighting based on the surrounding environment conditions.

The following sections will provide a more detailed illustration of the work, specifically:

- Section 1: Proposed Approach, in terms of implemented (wireless) architecture and any algorithms or control methodologies that may be present

- Section 2: Related Works, aims to examine relevant literature, critically evaluating them to highlight innovations and gaps in the field.

- Section 3: The simulated scenery is presented, essentially illustrating how the proposed approach was simulated using Matlab/Simulink/TrueTime.

- Section 4: The results obtained from the simulations must be shown, including the evaluation of network performance.

- Section 5: Conclusion.
- Section 6: References.

1. PROPOSED APPROACH

The smart lighting system has been designed to operate within a wireless context, offering flexibility and ease of installation. The system architecture comprises three main components: sensors, actuators, and the wireless communication network (IEEE 802.15.4 (ZigBee)). The sensors include an ambient light sensor and a PIR motion sensor, both placed within the environment to be illuminated. The actuator is responsible for adjusting the lighting based on the information gathered by the sensors.

The ambient light sensor is the first node of this network, it periodically detects the brightness of the surrounding environment, which can range from 0 to 100. In order to implement the sensor, I have used two functions: one for initialization ('sensore_luceAmbientale_init.m') and the function that the sensor must execute ('sensore_luceAmbientale_fnctn.m'). It is precisely thanks to the latter that the sensor sends, with wireless network, to the actuator. However, it does not simply send what it detects because the sensor integrates a fuzzy logic management (RegolatoreLuminosita.fis), whose input is represented by the external brightness of the surrounding environment detected by the sensor.
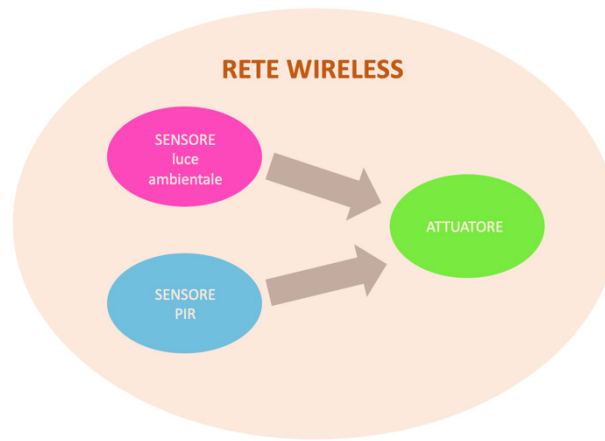
The initial logical model

Through the membership functions, it adjusts the brightness to be assigned to the light source in question. Both input and output values of the two variables range from 0 to 100, specifically:

- input 'luceAmbientale' is: Buio[<30]; Moderato [20:70]; Luminoso [>60].
- output luminosita is: Bassa[<30]; Media[30:70]; Alta[>70].

The purpose of the fuzzy is to automate the regulation of the brightness of the light source, as it will be higher in darkness and lower in a bright environment.
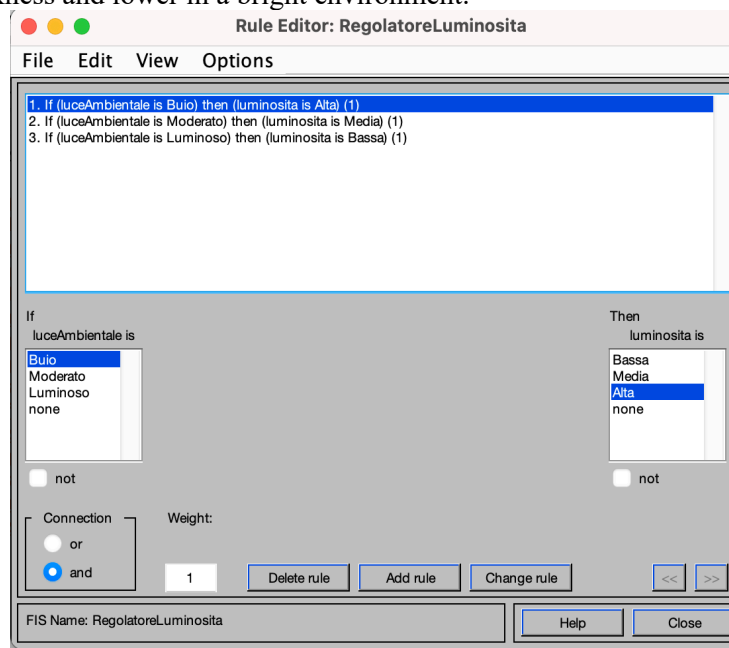
So, returning to the function related to this sensor, it will be tasked with sending to the fuzzy logic controller what it is detecting from the external environment. This regulates the brightness (output), which will then return to the sensor (input) so that it can send the processed data to the actuator.

The PIR sensor is the second node of the wireless network, it periodically detects any movement in the environment in which it is located and the respective body temperature of the moving object. In particular, thanks to the use of MBSD, it can understand whether the moving object is a person or not: after detecting movement [0-1] and temperature [0:50], the values are inputted into the block that allows state-based modeling of 'person detection'. In particular, the 'Rilevamento' output of the block can take on a value of 0 or 1 following this logic: if the received movement is 0, it remains in the 'inactive' state with 'Rilevamento=0', if the detected movement is 1, it is necessary to verify that the detected moving object is a person, therefore if the detected body temperature is between 35 and 40 degrees, consequently set 'Rilevamento=1', otherwise, if the temperature is not within that range, 'Rilevamento=0'. The output of the state model is the input of the PIR wireless sensor implemented using two functions: an initialization

one ('sensore_PIR_init.m') and the function that the sensor must execute ('sensore_PIR_fnctn.m') in which the procedure for sending 'Rilevamento' to node 3 of the network, the actuator, is present.

The actuator is node 3 of the wireless network and is responsible for managing the lighting of the reference light fixture based on what it receives from the two sensors listed above. This node requires 3 functions for implementation: one for initialization ('attuatore_init.m') which involves configuring the mailbox to manage the received data (from the ambient light sensor or from the PIR sensor), one necessary for network management ('nethand_attuatore_fnctn.m') as the actuator must listen on the reference network to receive data from the ambient light sensor and the PIR sensor, and a function that characterizes the behavior of the actuator (attuatore_fnctn). Specifically, the 'attuatore_fnctn' function is structured in such a way that, if a detection signal has been received from the PIR sensor equal to 1, the light fixture must be turned on with brightness equal to that received from the ambient light sensor, otherwise, if 'Detection = 0', it turns it off or keeps it off.

2. RELATED WORKS

Smart lighting represents a promising evolution in the lighting technology sector, offering a wide range of applications in various fields. In residential settings, smart lighting solutions allow users to adjust brightness, color, and ambiance according to their preferences, enhancing living comfort and optimizing energy efficiency. In commercial and industrial contexts, smart lighting systems can significantly contribute to reducing energy costs and creating more productive and comfortable work environments through centralized control and intelligent automation. Beyond practical aspects, smart lighting also plays a crucial role in smart cities. Intelligent lighting networks enable more efficient management of public lighting, dynamically adapting to environmental conditions and usage patterns, thus improving safety and reducing light pollution. For example, in "Smart hybrid wind-solar street lighting system fuzzy based approach: Case study Istanbul-Turkey", fuzzy-based control approach is proposed to control the street lighting systems depending on solar and wind renewable energy sources. The proposed street lighting controller depending on the utility, making the efficient utilization of renewable energy such as solar and wind energy which is very necessary in this energy crisis world and the arising of the environmental pollution concerns. The authors of "Universal Smart Lighting System" instead propose an innovative universal device for energy-efficient and intelligent lighting control. The device offers advanced features such as automatic human detection, smart turn on/off, and dynamic lighting adjustments. It can detect natural factors and respond based on ambient light. The proposal explores the device's features, applications, and versatility in managing LED lights, emphasizing energy savings and cost reduction.

Despite the numerous advantages offered by smart lighting, it's important to consider some critical issues. Firstly, widespread adoption of these technologies may raise concerns about privacy and data security, especially when dealing with systems that collect information through integrated sensors. Additionally, interoperability and standardization of devices and communication protocols pose significant challenges to ensuring effective integration and management of smart lighting systems within existing infrastructures.

Moreover, the aspect of economic and technological accessibility of smart lighting may limit its dissemination in certain communities or socio-economic contexts. It is therefore essential to adopt inclusive and sustainable approaches in implementing these technologies, ensuring they are accessible to all and contribute to reducing social and economic disparities.

3. SCENARIO

The proposed approach has been simulated using Matlab/Simulink/TrueTime:
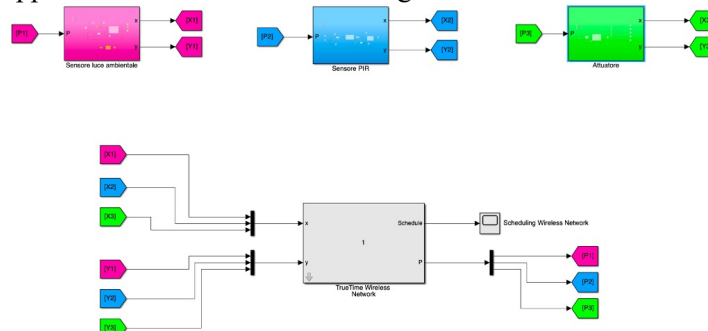
As already anticipated by the logical model, there is a wireless network (802.15.4 (ZigBee)) with 3 nodes enclosed in 3 subsystems, whose outputs are the coordinates (X,Y) of the respective node, and the input is the relative power to the battery of the various sensors. This representation is made possible through the use of goto (transmission) and 'from' (reception).

SENSORE DI LUCE AMBIENTALE:
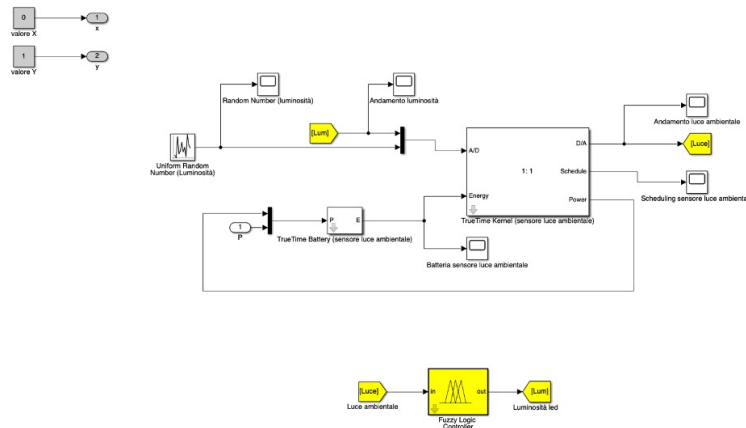
The ambient light sensor is node 1 of network 1, characterized by a TrueTime kernel with an initialization function 'sensore_luceAmbientale_init' with:

- coordinates (x, y);
- two inputs: one consists of the Uniform Random Number block present in the Simulink library with values ranging from 0 to 100, to simulate data related to external brightness; the other represents the brightness to be attributed to the light source;
- one output: the ambient light it is detecting;
- battery (TrueTime battery) with 'initial energy' = 3330 mA;
- fuzzy logic controller (also present in the Simulink library) which adjusts, as explained in section 1, the brightness to be attributed to the light source based on the ambient light it detects. The input of this block corresponds to the output of the sensor while the output of this block corresponds to the input of the sensor thanks to the use of 'goto' and 'from'.
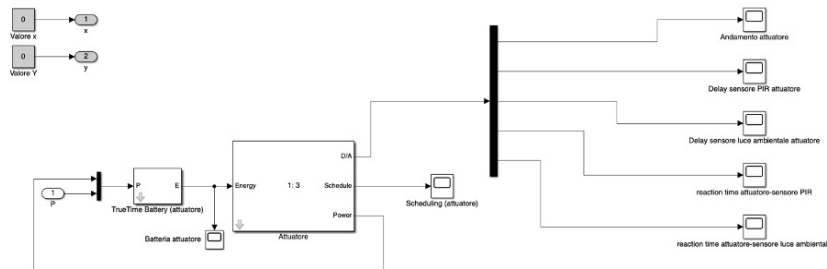
SENSORE PIR:
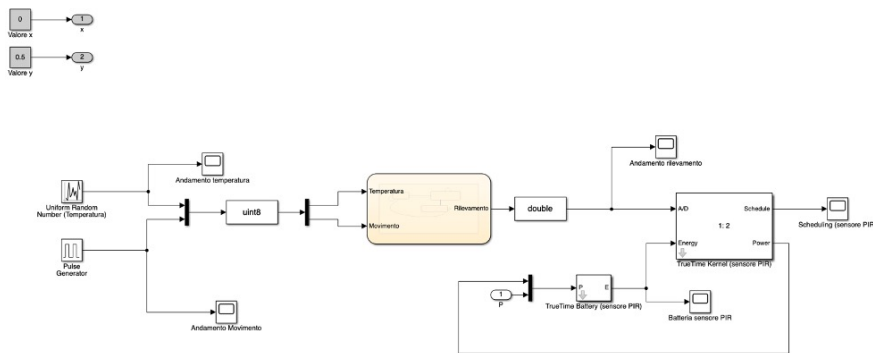
The PIR sensor is the second node of network number 1. It is characterized by a TrueTime kernel with an initialization function 'sensore_PIR_init', featuring:

- coordinates (x,y);
- no output;
- battery (TrueTime battery) with 'initial energy' = 3330 mA;
- an input named 'Rilevamento' derived from the 'chart' block present in the Simulink library allowing me to create the state model presented in 'section 1'. Specifically, to simulate the proposed approach, i used two blocks from the Simulink library to capture any movement (pulse generator [0-1]) and temperature (Uniform Random Number [0-50]) from the outside. Thanks to the 'mux' and 'demux' provided by Simulink, after conversion to 'uint8' to ensure better exportability, the signals taken from the outside reach the 'chart' block:
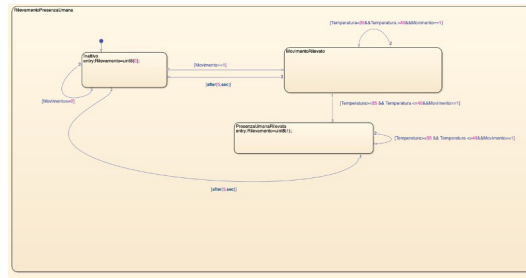
The block, as already mentioned, aims to simulate the possible detection of human presence through MBSD. The representation must occur through the use of states, transitions, and permanences in a state. As shown in the figure, as long as the input movement is zero, it remains in a state called 'Inattivo' where the output variable 'Rilevamento' is zero. A transition occurs if it receives a movement equal to 1, leading to the 'MovimentoRilevato' state where it remains if the detected temperature is not between 35 and 40 degrees and the movement continues to remain equal to one; it returns to the 'Inattivo' state after 5 seconds. From the 'MovimentoRilevato' state, there is a transition to the 'PresenzaUmanaRilevata' state where the output 'Rilevamento' is set to 1, only if, in addition to detecting movement, the temperature of the moving body is between 35 and 40 degrees, and it remains in these circumstances, otherwise after 5 seconds, it returns to the 'Inattivo' state.

ATTUATORE:

The actuator is the third node of network number 1. It is characterized by a TrueTime kernel with the initialization function 'attuatore_init', with:

- Coordinates (x, y).
- No inputs (data is received via the network for processing).
- 5 outputs: the first one is used to display the behavior of the actuator: how it adjusts the reference light source according to the received signals (brightness and person detection); the second and third represent the delay between the transmission of the signal by the PIR sensor and its reception by the actuator, and the transmission of the signal by the ambient light sensor and its reception by the actuator, respectively. The calculation of the delay is made possible through the implemented functions 'sensore_PIR_fnctn', 'sensore_luceAmbientale_fnctn', where the current time (ttCurrentTime) is added (timestamp) to the message to be sent to the actuator when the sensor sends the useful signal to the actuator, and 'attuatore_fnctn' where, processing the received data, it is possible to calculate the difference between the time when i actually received (ttCurrentTime) and the time when the sensor sent me the message: the result of this difference is the delay. Furthermore, in 'attuatore_fnctn', the calculation of the reaction time is also present, that is the time between receiving the signal by the actuator (whether it is brightness or detection) and its execution, which in my case corresponds to adjusting the light source. Outputs 4 and 5 indeed represent the trend of the reaction time from the two signals of the two sensors.

4. PERFORMANCE

After viewing the approach and the respective scenery used, it is useful to understand how the system behaves by proceeding with a simulation (3600 seconds).

If no detection is received:
```
SENSORE PIR: mando segnale rilevato:0 T=0.4
SENSORE DI LUCE AMBIENTALE: mando il segnale rilevato:90.3333 T=0.4
ATTUATORE: rilevamento ricevuto:0 t=0.40034 delay=0.00034
ATTUATORE: luminosità da attribuire ricevuta:90.3333 t=0.46034 delay=0.06034
ATTUATORE: spento
reaction time dalla ricezione del segnale del sensore PIR=0.18 current time =0.58034
reaction time dalla ricezione del segnale del sensore luce ambientale=0.12 current time =0.58034
```

If detection is received:
```
SENSORE PIR: mando segnale rilevato:1 T=31.6
SENSORE DI LUCE AMBIENTALE: mando il segnale rilevato:90.3333 T=31.6
ATTUATORE: rilevamento ricevuto:1 t=31.6003 delay=0.00034
ATTUATORE: luminosità da attribuire ricevuta:90.3333 t=31.6603 delay=0.06034
ATTUATORE: accendo con luminosita=90.3333
reaction time dalla ricezione del segnale del sensore PIR=0.18 current time =31.7803
reaction time dalla ricezione del segnale dal sensore di luce ambientale=0.12 current time =31.7803
```

The performance of the actuator described above as a function of the received signals is as follows:
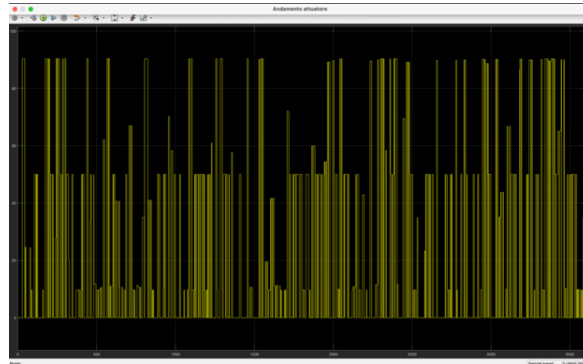
An analysis consists of two graphs related to the ambient light sensor to better illustrate the functioning of the fuzzy controller. In the first image, the values captured by the sensor in terms of ambient brightness are shown, while in the second image, the brightness that the ambient light sensor must send to the actuator to regulate the light fixture is depicted. It is observed how a high ambient light value corresponds to a low brightness value to be assigned, and conversely.

1) 2)
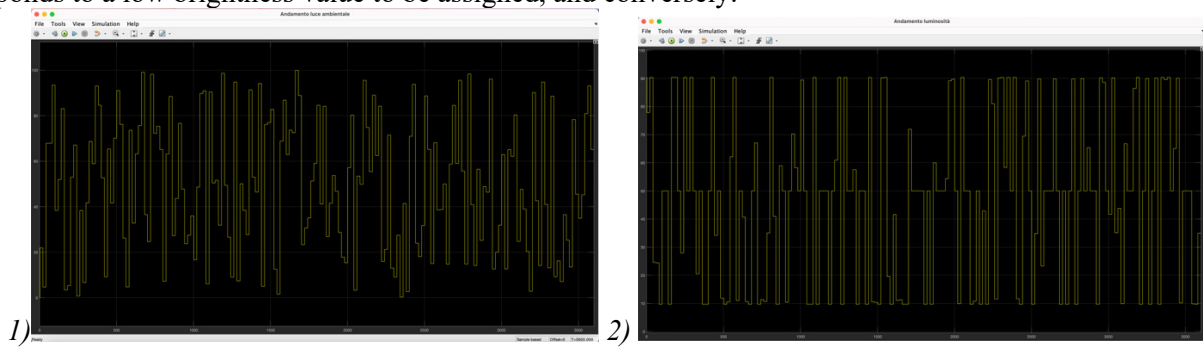
Regarding performance metrics, namely performance indices to refer to, both the delay and reaction time (explained in Section 3) have been calculated. The graphs below represent the trend of the calculated delay and reaction time, reflecting the values they assume during the simulation. Particularly in both cases, as can be observed, they remain constant around zero.

5. Conclusion

In conclusion, this paper has proposed and simulated a smart lighting system that integrates two sensors and an actuator within a wireless network, using TrueTime as the simulation environment. The analysis of the simulations has highlighted the success of the proposed approach in improving energy efficiency and user experience by providing optimal lighting according to circumstances. Smart lighting solutions represent an effective way to enhance user experience, extending beyond the scope of domestic lighting to a wide range of applications, including offices, shops, and public spaces. The combination of advanced technologies, automation, and monitoring enables the creation of a more efficient and eco-friendly illuminated environment.

For possible future developments of the proposed smart lighting system, some areas of improvement that could be considered are:

- personalization of the user experience: The system could be enhanced to offer an even more personalized user experience. For example, users could be given the ability to define their preferred lighting sceneries or to schedule specific lighting routines based on their habits and needs.

- integration with other smart technologies: the system could be integrated with other smart technologies in the home or workplace environment, such as home automation systems or voice control devices.

- optimization of control algorithms: control algorithms can be optimized to ensure even more effective lighting management. This could include implementing more advanced fuzzy logic algorithms or using machine learning techniques to dynamically adapt control parameters based on data collected from sensors.

References

1. Reti di calcolatori e Internet. Un approccio top-down; 7° edizione; James F. Kurose, Keith W. Ross, ISBN 13: 9788891902542.

2. Matlab Simulink documentation MathWorks: https://it.mathworks.com/help/matlab/

3. Wadi M., Shobole A., Tur M. R.and Baysal M. "Smart hybrid wind-solar street lighting system fuzzy based approach: Case study Istanbul-Turkey," *2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG).*

4. Saranya G., Duraiarasu E. and Manoj S. "Universal Smart Lighting System," 2023 IEEE Sustainable Smart Lighting World Conference & Expo (LS18).

RYDBERG ATOMIC SPECTRUM SENSOR WITH SENSITIVITY UP TO 20 GHZ

Alexey Anisimov
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
Anisimov_junior@mail.ru

Abstract. *The method of a non-resonant high-frequency local oscillator is considered when operating at frequencies from zero to 20 GHz; promising sensors that surpass the characteristics of existing devices are considered.*
Keywords: *Rydberg atoms, sensors, accuracy.*

The discussed atomic radio frequency (RF) receiver and spectrum analyzer utilize thermal Rydberg atoms bound to a planar microwave waveguide. Such a system provides an internal sensitivity of up to -120(2) dBm/Hz, a bandwidth of 4 MHz, and a linear dynamic range of over 80 dB. With the use of a preamplifier, high sensitivity spectrum analysis of more than -145 dBm/Hz is achieved. The "rabbit ear" antenna allows for the detection of weak radio signals such as FM, AM, Wi-Fi, and Bluetooth. Also considered are the interactions of waves with Rydberg atoms and new opportunities for creating small sensors that surpass standard radio frequency devices in sensitivity, bandwidth, and accuracy. Quantum sensors possess unique properties that distinguish them from traditional technologies. Their absolute uniformity and precision, associated with fundamental principles, make them particularly efficient.

A radio frequency quantum sensor based on thermal Rydberg atoms is considered, operating in the range of 0 to 20 GHz and outperforming known spectrum analyzers in terms of performance. The discussed sensor features enhanced sensitivity due to the confinement of the RF field in a small mode volume, as well as the use of off-resonance heterodyne techniques to increase sensitivity at selected frequencies. This system can operate with or without a low-noise preamplifier for input radio frequency signals, opening up new possibilities for miniaturization and performance enhancement.

Rydberg radio frequency sensors are important for several reasons. As quantum sensors, they must overcome the limitations of traditional receivers. Firstly, their response is related to fundamental constants, ensuring high calibration accuracy over a wide frequency range [1]. Secondly, they have a high bandwidth that is independent of frequency, unlike passive receivers [2]. Thirdly, they provide the opportunity to avoid internal thermal noise even at room temperature, thanks to the use of low-entropy laser beams to measure the quantized internal states of atoms.

Rydberg sensors require further performance improvement before they become useful as real-time radio frequency receivers. Firstly, previous efforts have been limited by the frequencies of resonant transitions between Rydberg states. Utilizing all available transitions requires a powerful and fast laser system. Additionally, existing sensors have limited sensitivity to radio frequency fields, making their application in real-world conditions challenging. These realities complicate the creation of a wideband Rydberg sensor for weak radio frequency fields.

The discussed system addresses the issue of poor device connectivity in free space by directing the electric field into a planar microwave waveguide, which concentrates it in the subwavelength region. Rydberg atoms are created and probed above the gap between the signal trace and the ground plane, where the evanescent electric field is concentrated. Such a sensor significantly improves sensitivity, which is critically important for radio frequency applications in communication, probing, and spectrum monitoring.

Resonance heterodyne technique was also applied, combining the input radio frequency signal with a powerful local oscillator. This method allows operation over a wide range without tuning the Rydberg laser. Heterodyning enhances the response at any arbitrarily chosen frequency of the radio frequency input signal and linearizes the sensor. In known experiments, resonance local oscillators were also used for sensing electric fields in free space.

Continuous operation with good sensitivity was achieved in the range from DC to 20 GHz. The high-frequency limit of our system is mainly determined by the limited bandwidth of the vacuum feedthroughs and microwave circuitry, which can be improved in future projects.

In the experiment discussed in [3], the coupling between atoms and the circuitry is significant enough to directly detect the presence of Rydberg atoms with minimal probing of the microwave waveguide. The explored experience serves as an initial step towards expanding fundamental work on micro-

wave Rydberg quantum electrodynamics into room-temperature regimes. Additionally, reading data from the Rydberg waveguide will be a useful tool for further iterations of sensors. Firstly, direct microwave detection of Rydberg populations is not sensitive to Doppler shifts, which are characteristic of modern optical readout schemes. Secondly, strong collective coupling between the Rydberg region and the microwave circuitry can lead to a range of impressive experimental possibilities, such as quantum frequency conversion between optical and radio frequency signals, generation of tangents and/or Rydberg masers.

The discussed experimental configuration also allows for the advantages of standard antennas to be utilized when measuring fields of signals in the surrounding environment. In most previous works with Rydberg atom sensors, the main focus has been on using atoms in free space [4]. While such an approach has several advantages (such as absolute accuracy, detection of THz frequencies, and low field perturbation), these sensors suffer from poor coupling to any specific RF mode, and the spectral response cannot be easily tuned or narrowed. By interacting with an external antenna, it becomes possible to select a precisely defined radio frequency mode using spectral selectivity and antenna gain. In this paradigm, the Rydberg sensor replaces the back end of the receiving system, providing data readout that can surpass the limits of Johnson noise and alleviate some impedance matching issues of traditional electronic sensors.

The investigated experimental configuration enables the use of standard antennas for measuring fields of signals in the surrounding environment. In contrast to existing works, which have focused on using atoms in free space, such an approach allows for specific RF modes to be selected and the spectral response to be easily tuned. The Rydberg sensor replaces the back end of the receiving system, providing data readout with high sensitivity and aiding in overcoming impedance matching issues of traditional electronic sensors.

In addition to increasing the coupling of RF atoms, an important research direction is the study of alternative probing schemes for thermal Rydberg atoms. Existing methods, such as Electromagnetically Induced Transparency (EIT), have their limitations compared to the quantum limit of sensitivity. Employing more sophisticated probing schemes can significantly enhance sensitivity and throughput. Additionally, integrating a resonator for energy recirculation in a laser currently at 480 nm could lead to improved sensor performance and reduce its cost and size.

Rydberg radio frequency sensors represent a promising technology for the future. In the long term, Rydberg quantum sensors may play a crucial role in integrating classical and quantum communications. Current experiments on quantum frequency conversion and the interaction of Rydberg atoms with superconducting resonators are also aimed at this goal. However, to fully understand the potential of these quantum tools and their application in real-world tasks, further fundamental research is required.

References

1. D. H. Meyer, Z. A. Castillo, K. C. Cox, and P. D. Kunz, Assessment of Rydberg atoms for wideband electric field sensing, Journal of Physics B: Atomic, Molecular and Optical Physics 53, 034001 (2020).

2. K. C. Cox, D. H. Meyer, F. K. Fatemi, and P. D. Kunz, Quantum-Limited Atomic Receiver in Electri- cally Small Regime, Physical Review Letters 121, 110502 (2018).

3. J. A. Gordon, M. T. Simons, A. H. Haddab, and C. L. Holloway, Weak electric-field detection with sub-1 Hz res- olution at radio frequencies using a Rydberg atom-based mixer, AIP Advances 9, 045030 (2019).

4. M. Jing, Y. Hu, J. Ma, H. Zhang, L. Zhang, L. Xiao, and S. Jia, Atomic superheterodyne receiver based on microwave-dressed Rydberg spectroscopy, Nature Physics 16, 911 (2020), arXiv:1902.11063.

AUTOMATION AS A METHOD OF RISK MANAGEMENT

Maria Belova
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: *marebel13@mail.ru*

Abstract. *This study considers risks and how to manage them. A methodology of management by automation is developed and proposed, which will contribute to the optimisation of this process. As a result of the study, a general picture of how modern risk management may look like is formed.*

Production in the modern world is accompanied by the introduction of new technologies and developments that are relevant for a certain period of time. Due to the fact that enterprises introduce new untried technologies, there is a high probability of occurrence of risks, which must be monitored and eliminated, otherwise the output will be losses, which will be a loss to the enterprise.

In order to create the possibility of rapid adaptation of new technologies, it is necessary to monitor and eliminate all risks as soon as possible to prevent losses of the first and second kind. Thus, **the purpose of the study** is to develop recommendations for the formation of the risk management process and to propose possible tools for automating the process at this stage of development.

Risk management is not a new process in quality assurance and control in domestic enterprises for a long time. Companies implementing quality management systems and other management systems use the principles of risk-oriented thinking, which is required by GOST R ISO 9001-2015 [1], and in order to create a risk management that meets the development of the modern Russian market, GOST R ISO 31000-2010 "Risk Management. Principles and Guidelines". Risk management is coordinated actions to manage the organisation taking into account the risk [2]. The organisational structure of risk management, financing of risk management system maintenance and the level of risk management automation are the least developed elements of corporate risk management culture in companies [3], therefore, increasing risk-oriented thinking and automation of processes that ensure the operability of risk management is one of the priority areas of quality management development.

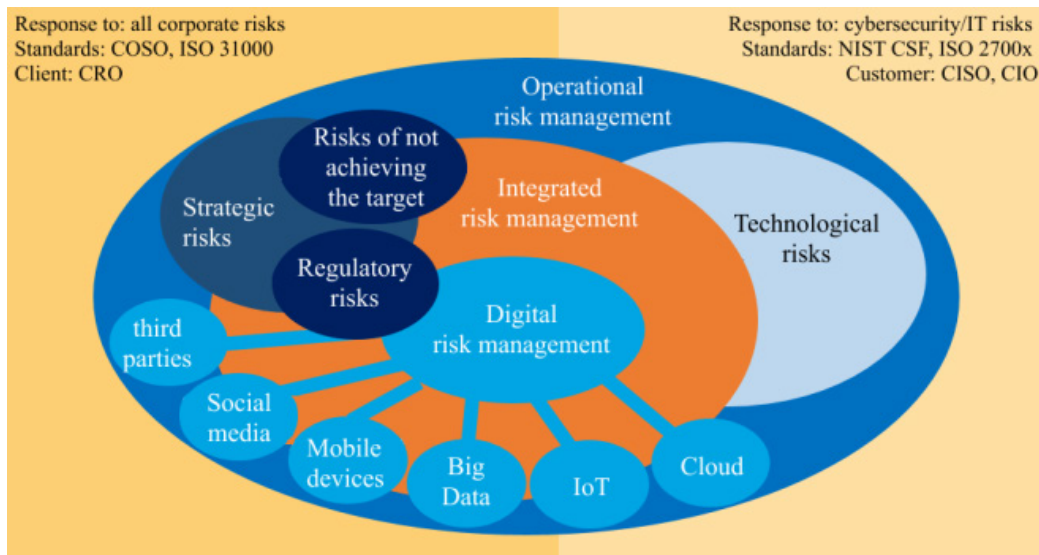Risks are divided into the following segments (Fig.1).

Fig. 1. Risk clustering

Often the measures taken at the enterprise to form a risk management process are not enough. This is due to the perception of Russian companies not only of risk management, but also of quality management in general. Most companies take the quality management system (QMS) at the enterprise for granted, as a necessary implemented standard that must be complied with, and constantly confirm this compliance in order not to lose the certificate. Because of this, the majority of QMS processes are brought to a proper state just before the audit, internal regulatory documentation is not updated for long periods of

time, and the system itself does not work, but simply exists on paper. Such a formed attitude to QMS is certainly not observed in all companies, but the problem exists, and companies often do not look into the future for the development of quality management system and management of the organisation as a whole. According to the latest data of the International Organisation for Standardisation, ISO 9001:2008 certificate was issued more times than ISO 9001:2015 (Fig.2) [4], which indicates that the popularity of quality management system development is decreasing.
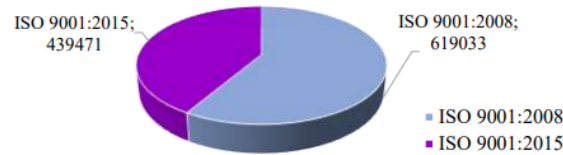
Fig. 2. Comparison of the number of issued ISO 9001:2008 and ISO 9001:2015 certificates

Management is constantly changing. In the era of Industry 4.0, smoothly transitioning to Industry 5.0, these changes are much faster. And to be able to control risks it is suggested to use one of the methods of process optimisation – automation.

Process automation is the use of technology and software to optimise and improve the efficiency of a company. It allows to automate routine tasks, speed up processes, reduce the probability of errors and improve the quality of products or services [5].

Currently, the automation of risk management processes can be presented as follows:
1) Automation of risk management processes;
2) Centralised documentation of risk management processes;
3) Automated continuous monitoring of key risk indicators;
4) Automation of processes for agreeing changes to the risk management system;
5) Operational reporting.

There are two other options for deploying an automated risk management system: either a specialised module is implemented as part of a large enterprise management information system, or a separate system is implemented. In the first case, the project timeline is determined by the timeline of the ERP system itself, but as a result the company gets an integrated solution. In the second case, there is no dependence on which ERP system is used at the enterprise, but, as a rule, additional integration work is required.

That is why companies with a holding structure, numerous subdivisions of which have their own IT infrastructure and different ERP-systems, can be recommended to implement an independent risk management system [6].

Implementation of automated risk management

The automation of the risk management process will be based on incident management and risk management itself. Incident management will automate the collection and recording of incident information and identify links between risks and incidents. The collection and recording of incidents can be done either manually or automatically through the use of sensors in production that will transmit signals to information systems [7]. Then automatically in the system will be created risk maps, identification, classification and evaluation of risks. It can be said that all risk management processes will be automated.

Automation on the basis of software ready solution is realised with the help of the following technologies:

– building a process model of risk management with the help of Workflow task assurance systems – includes the construction of all enterprise process maps and the formation of links between them in the form of data flows;

– data flows will reside in data warehouses;

– data will be processed and analysed automatically using machine learning methods with the ability to build models and forecasts;

– use of dashboards and infographics for reporting and visualisation of operational, analytical and statistical data;

– automated document management systems will ensure integration of regulatory documents from different management systems.

Automation by one's own efforts implies the creation of a quality service or risk management group in the enterprise. This group of employees within the framework of risk management activities – is responsible for the register, map and risk passports / cards; ensures timely fulfilment of responsibilities in the risk management system; proposes measures for the development of the risk management structure; is responsible for the creation of the risk management system; is responsible for the selection of key risk indicators; establishes risk owners; controls the development and implementation of risk management measures [8]. Reporting should be present in all departments and at all levels of management, reporting can be collected in the process and as a result of internal audits.

An opportunity that should be considered by all enterprises in the automation phase is automated document management systems. Documentation can be electronic, and in order to ensure that it is compiled correctly, there are ready developed templates that can be adapted to the specific organisation. Electronic document management systems will ensure that information is shared more quickly between departments and that information is collected and processed. For example, to collect up-to-date information on incidents that have occurred in the enterprise, it is possible to have a single register stored on a file balloon publicly accessible to risk owners, which will be filled in directly by the owners when an incident occurs, or a single incident registration request form can be developed and these requests can be sent to the responsible person for checking and registration. Storing documentation in a single system with the ability to configure user access will reduce the time it takes for a document to be approved, allow you to view past versions of the document, and when integrated with other information systems, automatically fill in some fields and update the information. Such software solutions as Microsoft SharePoint Server, DocsVision, DocLogix, DELO, DocSpace can optimise the electronic document management system of the whole enterprise not only in terms of QMS and risk management [9].

As for electronic forms, it can be both elementary use of Microsoft Excel tables with shared access of employees to the file and Google services – forms and tables (provided that public information is processed), and the use of Adobe and Microsoft software products with support for a wider functionality in the field of electronic forms.

In electronic forms can be created: a table of responsibility for the risk management process, a matrix of possible types of risk for key areas of the enterprise, a balanced classification of risks of the enterprise. Further on the basis of joint work of risk management group in electronic forms and electronic document flow system will be created risk register, risk maps, risk impact plan and risk management report. Risk register for confectionery enterprise can be more than one, it is possible to divide the register into processes [10].

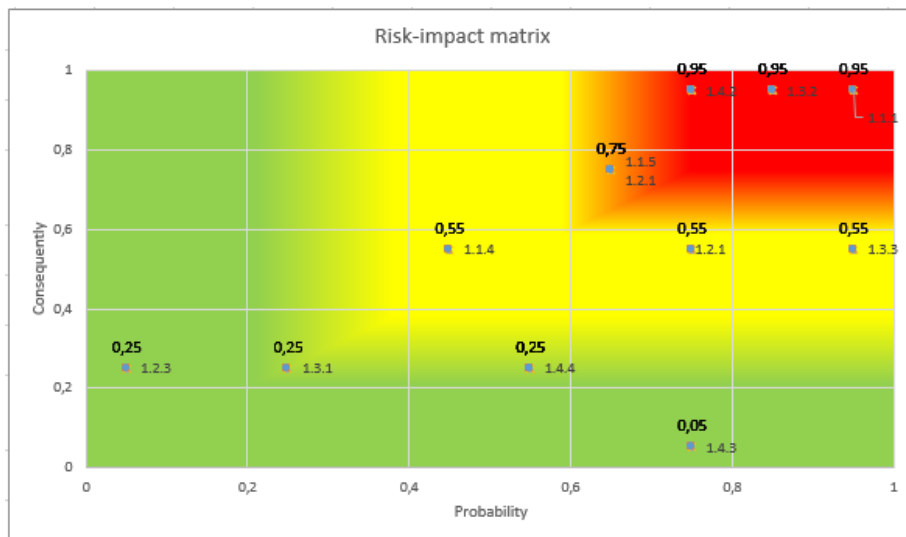The distribution of the risk map can be represented as follows (Fig. 3).

Рисунок 3. Карта рисков процессов

In automated risk management systems these documents are prepared automatically and promptly, on the basis of these documents can be built a map of guarantees – a document that defines the areas of responsibility of the subjects of the quality management system of the enterprise and other stakeholders in the Scientific Journal of NIU ITMO. Economics and Environmental Management Series No. 4, 2020

84 with regard to individual risks of the company. The assurance map reflects the coverage of risks by control functions of the organisation by levels in accordance with the model of "Three Lines of Defence", where the first line performs business functions, the second line – monitoring functions, and the third line of defence conducts independent assessment and control. The assurance map can also be compiled manually, using the company's risk and process classifiers, risk map and other internal documentation [11].

To implement partial automation of the risk management process can also be used software tools aimed at solving enterprise problems and focused on the quality management system and requirements of GOST R ISO 9001. One of the most accessible is Microsoft Project, which although not focused on the standard and QMS purposefully, but contains a lot of tools that allow to optimise the work on QMS tasks.

After drawing up the planned risk management documents it is necessary to carry out prevention measures according to the risk management plan. This will reduce the risk level and bring it into the permissible "green" zone. For each of the identified most dangerous risks in the process, a detailed plan of measures to prevent or reduce the threat is prepared. It is also necessary to reflect in the plan of measures to realise opportunities, which can help not only the development of QMS, but also the company as a whole [12]. For example, for the possible risk "non-compliance with the recipe", where the opportunity would be "improving the recipe", such prevention activities are carried out as:

– training and certification of personnel on process, QMS aspects;

– introduction of thorough parameter control at each production site (not just incoming inspection and finished product inspection);

– daily verification of measuring equipment;

– creation of a product route map;

– use of natural additives and preservatives;

– use of software designed for food production (e.g. "Recipe Manager") (step-by-step display of components according to a given recipe with the display of their parameters) [13];

– introduction of an automated timing system with sound and visual effect (the worker can see and hear by means of automatic signals at what time it is necessary to proceed to the next stage of the process);

– modernisation of recipes, taking into account the experience of other companies (benchmarking).

Thus, from the very initial stage to prevention and monitoring measures in the enterprise risk management process, there are more and more elements of automation, which are often essential in the digitalisation of the economic space, so every company that does not make automation a priority should think about investing in the IT infrastructure of the enterprise.

Conclusion

This study analyses risk management. During the study of the current situation of risk management and quality management development, the tendency and need for automation of enterprise processes were identified. Possible tools for automation and measures for the implementation of risk management were proposed, the scheme of the risk management process, risk register and risk map of one of the main processes of the enterprise were drawn up.

References

1. GOST R ISO 9001-2015. Quality management systems. Requirements. – Moscow: Standardinform, 2015. 32 c.

2. GOST R ISO 31000-2010. Risk management. Principles and guidelines. Moscow: Standardinform, 2011. 24 c.

3. Kuznetsova M.O.. Practices of risk management implementation in Russian industrial companies: results of empirical research // Strategic decisions and risk management. T. 10. №4. 2019. C. 410-423.

4. ISO Survey of certifications to management system standards. 2018. URL: https://isotc.iso.org (accessed 20.01.2024).

5. Just about complicated: What is business process automation, 2017. URL: https://vc.ru/ (access date: 02.02.2024).

6. Automation of risk management and internal control: approaches, methodology, features, 2021. URL: https://www.gmcs.ru (date of reference: 02.02.2024).

7. Galinovskiy A. L. Information systems of quality management in automated and automatic productions: a textbook / A. L. Galinovskiy, S. V. Bochkarev, I. N. Kravchenko [et al] ; edited by A. L. Galinovskiy. – Moscow : INFRA-M, 2020. – 284 с.

8. Dolzhenkova A. V. Formation of the enterprise risk management system // XII Annual scientific session of postgraduates and young scientists: materials of the interregional scientific conference: in 2 vol. / Ministry of Science of Higher Education. Russian Federation, Vologda State Univ. Vol. 2: Social and Humanitarian Direction. Vologda: Vologda State University, 2018. С. 113-117.

9. Moiseeva, A.V. Software of the quality management system / A.V. Moiseeva. – Text : direct, electronic // Young scientist. – 2017. – № (144). – С. 259-261. – URL: https://moluch.ru/archive/144/40269/ (date of reference: 10.02.2024).

10. Kasyanenko, N. N. Implementation of risk management in the organisation // Saint Petersburg Journal of Electronics. 2016. № 1 (84). С. 129-132

11. Tokun M., Sadova N. Lines of defence of the company. Map of guarantees. Audit-it.ru: Accounting. Taxes. Audit. 2019. URL: https://www.audit-it.ru/articles/audit/a104/979017.html/ (accessed 13.02.2024).

12. Rakhimova, N. N. Risk management, system analysis and modelling: study guide / N. N. Rakhimova. Orenburg: OGU, 2016. 190 с.

13. Vershinina N. Secrets of increasing the efficiency of production. Equipnet.ru: Buying and selling equipment for business. 2017. URL: https://www.equipnet.ru/ / (date of reference: 13.02.2024).

ALGORITHMS FOR PROCESSING AND RECORDING READINGS OF TECHNICAL OBJECTS OF RESEARCH

Danil Bobryshov

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: danil.bobryshov@mail.ru

Abstract. *The operation and state of technical objects is described by input and output data describing the nature of the activity. For example, power supply systems are described by electrical parameters of power, current, voltage, power quality, power factor and many other data. For many systems of mechanical kind of action, the work is described by dependencies and data of power, pressure, flow, consumption and so on. It is worth noting that today, given the active development of digital technologies and their implementation in various kinds of systems, monitoring of parameters is one of the important tasks, because on the basis of recorded data analysis and prediction of the behavior of the evaluated apparatuses is made [1]. Considering the specifics of the construction of the system of many objects that do not have the possibility of data exchange with the system of processing of fixed parameters, there is an urgent need to build ways to register the readings of technical objects from such apparatuses and the formation of certain algorithms for processing values.*

Technologies and methods of recording readings of technical objects

Before carrying out the processing of results, it is necessary to ensure the collection of data describing the behavior and operation of the object or system under study. For particular cases that allow the implementation and use of digital technologies, data collection is carried out by building logical systems using apparatuses that record values and devices that store and process the data. As an example, a power supply system using measuring transformers aimed at recording high values of electrical energy, as well as digital voltmeters, ammeters, wattmeters, varmeters, ohmmeters, frequency meters, etc. can be used. The logical processing system may be a server that stores, processes the measurement data, and controls the entire system. The structure of such control will consist of three main levels (Fig.1):
- storage, processing and control of the digital system apparatuses;
- physical control of the system;
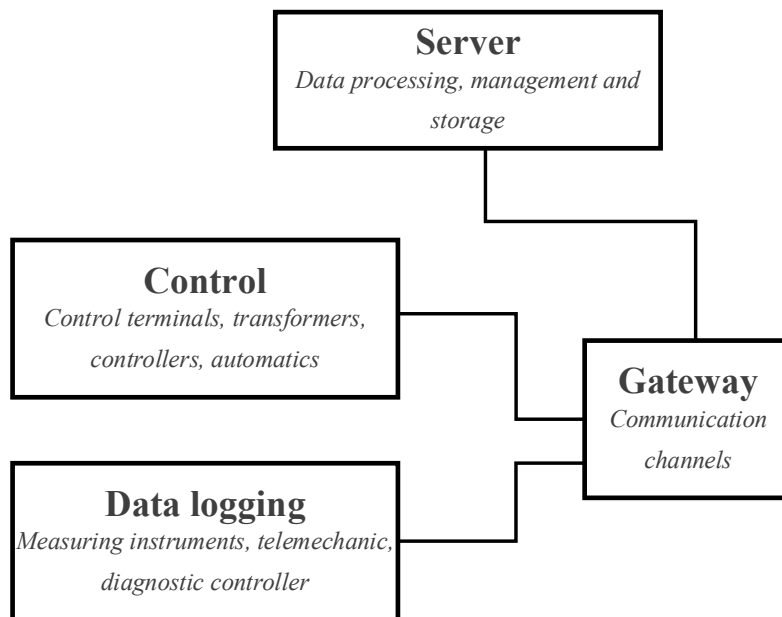- monitoring and fixation of values of the system under study.

Fig. 1. Structure of a digital system for monitoring and control of a technical object

The first unit will mainly consist of a server and a logical decision-making system, based on a pre-defined control algorithm or a model based on the characteristics to be recorded. The second level may also contain elements of logical control, limited to accepting commands and projecting their execution on the physical elements of the system, for example, organizing the switching of an electrical object or hydraulic system, considering the fixation of excessive pressure readings in individual elements of the system. The third block is aimed at registering the values of pierced processes, and sending the readings to the server for processing and selecting a behavior scenario. All these levels are united by means of communication channels called gateways. The main advantage of such a system is monitoring and registration of emergency situations, as well as providing an instant response of the logical system and preventing damage to the system. Such structures and systems are characterized by a high degree of reliability, because the reaction to abnormal situations occurs almost instantaneously, while, of course, it should be understood that there is a probability of failure at each level [2]. To avoid this error, systems can be duplicated, for example, connecting an additional server, or made in such a way as to minimize damage.

For systems that use in their design panelized measuring systems without the possibility of providing a communication protocol between logical elements of processing or control. A prime example of the technology and method of recording such values is technical vision and machine learning in automating the process of performing verification and certification of electrical control and measuring instruments. Fixation of values from measuring devices without communication protocols, illustrating parameters by means of outputting values on a dial and fixation of these parameters by technical vision, is caused by various features. For example, for digital measuring devices displaying the recorded data on the dial of the instrument it is possible to use a vision system with the use of optical character recognition libraries, such as Tesseract. OCR, PyOCR, OCRopus, EasyOCR, etc. [3]. [3]. The logic of work is to determine the boundaries of the device screen by detection method and read the values from the screen using the library. Intermediately, the image is converted into black-and-white format, which is mostly a gradient and sorting of hue values by threshold value for each hue (Fig. 2).
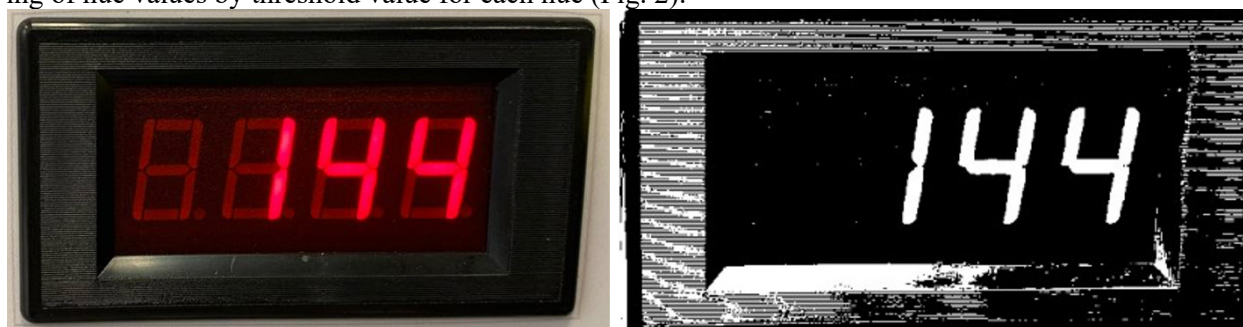

Fig. 2. Example of image processing by set RGB boundary

Conversion to black and white image format allows to exclude small details and to emphasize the main values required for registration. For pointer measuring devices, the task of recording a value is characterized by its degree of processing. For this type of devices, it is necessary to match the arrow and the dial of the device to set the recorded value. To solve these problems, various algorithms are used for fixing reference values and matching them through mathematical processing algorithms, such as setting the range of arrow hit, pointer deflection angle and pointer magnitude [4].

Also, in the field of unmanned aerial vehicles, which is gaining active development, vision technology is used for territory scanning, terrain mapping, site identification and control of device landing, orientation in space, etc. In such cases detection, tracking, depth estimation and segmentation techniques are applied, most often such algorithms are used together. In industry, this technology can be applied to the production of boards to detect defects. For the operation of such systems it is necessary to realize the construction of algorithms for processing of fixed data.

Algorithms for processing parameters of technical objects

There is a large number of algorithms for processing parameters of technical objects. These systems depend on the type of data, processing tasks, volume of input values and allocated resources.

Pattern and pattern recognition algorithm is a computer vision method used for automatic detection of objects and individual images in the studied image. The main purpose of this method is to search and classify objects based on their characteristics and features. Such a system is based on pattern and pat-

tern recognition using mathematical and statistical methods [5]. Such methods include: support vectors, neural networks, clustering algorithms, etc. The process of image and pattern recognition involves several steps:

- image preprocessing, with noise reduction and image preparation for processing;
- extraction of characteristics and features from images and templates using various methods;
- comparison of extracted features with predefined samples and templates;
- classification and identification of objects in the image based on the comparison results.

Pattern and template recognition algorithms are widely used in various fields such as medicine, robotics, industrial automation, biometrics, etc. These methods allow automating the processes of object recognition and identification, which increases the efficiency and accuracy of the system.

Another example of a system for processing the parameters of technical objects is the algorithm for predicting failures and failures, which is a method of data analysis aimed at identifying the preconditions of possible failures or malfunctions in systems, equipment, machinery or processes before their actual occurrence [6]. This type of algorithms is applied in various fields such as industry, transportation, medicine, information technology and many others, where it is necessary to fix and prevent possible problems. This is necessary to minimize the risks of failure occurrence. The main purpose of this algorithm is that prediction of a possible failure or problem before it occurs. To do this, algorithms analyze large amounts of data using machine learning, statistics, artificial intelligence and other techniques. The process of a failure and failure prediction algorithm typically involves the following steps:

- data collection and preprocessing;
- feature extraction;
- construction of the prediction algorithm;
- training and testing the model;
- prediction and acting.

Failure and failure prediction algorithms ensure the reliability and efficiency of various systems and processes, helping companies prevent potential problems and minimize losses.

Another example is the optimal maintenance scheduling algorithm, which is a method or process for optimizing the scheduling of maintenance work and equipment maintenance [7]. The objective of this algorithm is to minimize the downtime of the equipment, to ensure its proper functioning and to increase the service life of the apparatus or system. This algorithm takes into account various factors such as type and condition of equipment, regulated maintenance schedules, availability of spare parts and personnel, and budget constraints. Optimal scheduling helps in determining the timing and sequence of maintenance activities to ensure smooth operation of the production, equipment or system. The optimal maintenance scheduling algorithm is often based on mathematical models or optimization techniques to find the best solution considering all the constraints and objectives of the maintenance process. This algorithm plays an important role in equipment maintenance management in industrial plants, transportation systems, aviation and other industries where it is important to maintain high production efficiency and safety of operations. It helps to reduce maintenance costs, increase equipment reliability and service life, and reduce the risk of accidents and downtime.

A separate algorithm for monitoring and controlling the state of an object, which is a set of instructions, programs or procedures that allow to observe and control the state of certain objects or systems [8]. This algorithm usually consists of such stages as:

- determination of object parameters;
- establishment of threshold values;
- data collection;
- data analysis;
- decision making;
- control and monitoring.

The important elements of this algorithm are parameter determination, threshold setting, data collection, analysis and decision making based on the obtained information. The algorithm of monitoring and control of the object state can be implemented with the help of special software systems, sensors, data collection devices and other technologies that help to automate the control process and ensure the reliability and efficiency of objects.

Conclusion

To summarize, it can be noted that algorithms of processing and registration of readings of technical objects of research are important aspects of correct functioning of the system today. Thanks to the

active development of digital technologies and instrumentation, it is possible to build a system with a special degree of registration and response to disturbing influences, as well as considering different scenarios of system operation. A prime example of digital technology is technical vision, which provides the interaction of a largely analog system with a superimposed logical system.

Results processing algorithms are also an important part of new digital systems. Collecting and analyzing data to ensure efficient functioning of technical systems is among the primary tasks. The parameters are further processed by algorithms and methods for processing the parameters of technical objects. In this paper, as one example, attention was paid to the need for data collection for analytical prediction of object behavior in order to prevent accidents and ensure prompt system response. The structure of control and monitoring of technical systems was discussed, including the levels of storage, data processing and physical control.

The paper presented an analysis of processing algorithms including pattern and pattern recognition, failure and failure prediction, optimal maintenance planning and object condition monitoring. These algorithms are important tools for improving the efficiency, reliability and safety of technical systems.

References

1. Privalova T.V., Karpov M.V., Possibility of transition to digital technologies in the development of complex technical systems. Izvestiya Tula State University 2019, P. 222- 228. Date of circulation (07.01.2024);

2. Kuryanov V.N., Kushch L.R., Gorbunova N.R., Bondarev I.V., Tsypik V.V., Digital substations. Experience of realization. Science, Education and Culture 2018, Date of circulation (15.01.2024);

3. Bulatnikov E.V., Goeva A.A., Comparison of computer vision libraries for use in an application using flat image recognition technology. Bulletin of Moscow State University 2015. Date of circulation (15.01.2024);

4. Bobryshov A.P., Tretyakov N.K., Rysin A.V., Kuzmenko V.P., Solyony S.V., Mathematical model of data processing from analog arrow control and measuring devices. Instrumentation 2023. Date of circulation (17.01.2024);

5. Goryachkin B.S., Kitov M.A., Computer vision. E-Scio 2020. Date of circulation (24.01.2024);

6. Avdeev A.S., Barysheva N.N., Algorithm and systems of forecasting of failures and failures of the equipment. Bulletin of Altai State Agrarian University 2019, P.134 – 140. Date of circulation (05.02.2024);

7. Okladnikova E.N., Sugak E.V., Planning of the maintenance system. Siberian Aerospace Journal 2006 P. 66 – 70. Date of circulation (12.02.2024);

8. Pankin A.M. Methodology of development of algorithms for control of technical state of continuous objects. Proceedings of the International Symposium "Reliability and Quality" 2016. Date of circulation (17.02.2024).

ANALYSIS OF THE INFLUENCE OF MATHEMATICAL ALGORITHMS OF MEASUREMENT PROCESSING OF ELECTRICAL CONTROL AND MEASURING DEVICES ON THE RESULTS OF VERIFICATION

Aleksey Bobryshov

***Saint Petersburg State University** of Aerospace Instrumentation,*
Saint Petersburg, Russia
E-mail: ap.bobryshov@mail.ru

Abstract. *Evaluation of verification of electrical test and measurement devices is carried out by determining the compliance of the established measurement errors of the device and the declared accuracy of the device. Correct calculation of this parameter directly depends on the investigated measurement results of the device under test. Verification is carried out in accordance with the regulatory rules that establish requirements for the conditions and equipment of certification, such as ambient temperature, atmospheric pressure, relative humidity, vibration, magnetic field, electrical network pulsations, etc. [1]. Non-compliance with the conditions described in the standards threatens to distort the measurement results. It is worth noting that it is not always possible to meet the established requirements, and in many cases, for example, in case of a voltage surge, it is possible to obtain a falsified measurement value in time. The probability of such a result is rather small and is excluded by the certification methodology, but it is not equal to zero.*

Deviations of electrical control and measuring instruments

Deviations of measured values from the actual values are measurement errors. Detection, analysis and elimination of these values is the key task of verification of electrical control and measuring instruments (CMI). These deviations are divided into several types, depending on the influencing parameters and phenomena. The most important division is into computational and manifestation errors. Computational deviations include absolute, relative and reduced deviations. Using these errors are analyzed the results of verification and determination of the suitability of CMI. By the nature of manifestation of errors can be divided into: random systematic and gross. Systematic deviations include errors due to the nature of manifestation or source. It is worth noting that there are quite a large number of deviations, but the main feature of these values is in the manifestation. These outliers can be detected either separately or together, forming a mixed form. An example would be an operator error in a single measurement that caused the measured value to deviate significantly from the original set value. Such an error can be referred to gross random or, more precisely, gross probabilistic error. In this case, the kind of error manifestation is subjective.

All electrical CMI today, globally can be divided into two main types, these are digital measuring devices and devices of mechanical action magneto-electric, electromagnetic or analog pointer. Both the method of verification of these devices, and certain errors in each type of device are different. For example, when carrying out the certification of arrow CMI, in order to correctly determine the deviation of the arrow of the device, a methodology that implies the use of a mirror is used. This is necessary to reduce the possible manifestation of deviation of measurement results in the case of incorrect determination of the deviation of the arrow of the device. Such an error occurs when verifying only analog pointer CMI and also bears the name of readout error, which, in turn, consists of interpolation and parallax errors [2]. These deviations, given the nature of manifestation, occur inevitably, differing only in the degree of influence, which depends on the subject. The interpolation error is always manifested when determining the position of the CMI indication arrow between two neighboring scale divisions. This error has a value of ± 0.2-0.1% of the division value, within which the metrological readings of the measuring instrument are determined [3]. As mentioned above, digital instruments also have a deviation inherent in their functionality. Such error is called quantization, at that the character of manifestation is not subjective and most often such deviation is considered when carrying out calibration measures, because they cannot be excluded [4]. The error of parallax, as well as interpolation, is always manifested at perpendicular view on the measurement scale. The error depends on y – distance between the scale and the measuring device arrow, z – displacement of the observer's gaze in the plane parallel to the scale and l – distance from the observer to the CMI arrow. The error is determined by the following formula:

$$\Delta_{par} = \frac{y \cdot z}{l}$$

Subjective errors depend on the operator, his experience and degree of fatigue. This manifestation can be both systematic, in case of violation of the rules of verification methodology, and random in manifestation, due to fatigue, inattention, etc. The latter indicates the possibility of distortion of only a certain number of measurements from the entire sample of verification results. It should be noted that these errors can falsify the measurement results up to the value of gross error, i.e. by a sufficiently large value. Taking into account such data is necessarily reflected in the final certification and reduces the degree of quality, not only the verification itself, but also the determination of the final accuracy of the electrical control and measuring device. Compliance with the requirements of verification reduces the manifestation of such errors, but even in the standards for the conditions there is a calculation of coefficients to account for changes in the conditions of certification, to level the influence of distortions on the results of measurements.

Verification is accompanied by determination of compliance with the accuracy and measurement results established by the instrument. Compliance with the established standards for testing: environmental conditions, methods of fixing values, methods of calculations, determines the purity of the experiment its accuracy and precision. It is also worth highlighting the reproducibility of measurements, which determines the compliance of all conditions and techniques when repeating the experiment. All these parameters determine the quality of verification activities. In support of this conclusion it is necessary to compare accuracy with reliability, which is a qualitative characteristic [5]. Such parameters as probability of failure-free operation, failure rate, failure-free rate, etc. characterize the value of reliability and can be expressed quantitatively [6]. In turn, the accuracy parameter is described by such quantities as error, the probability of accepting the device as fit or unfit, the conditions of verification, as well as errors of the first and second kind. By analogy with reliability, we can conclude that the complex accounting of the described parameters determines the quality of CMI.

Separately, it is possible to identify certain errors that are random, coarse or mixed. The results of measurements can be considered from the point of view of probability, determining the errors of the first and second kind, as well as the probability of recognizing the investigated device as valid and vice versa. The indicators of these calculations can be considered from the point of view of the quality of the tested CMI.

Influence of coarse manifestations of measurement deviations on the results of verification

The errors of electrical measurements recorded by CMI can be considered from the point of view of statistics, according to a certain law of distribution of the value. Devices are designed to display the real physical quantity with the highest possible approximation, i.e. accuracy, but due to various kinds of circumstances, it is possible to manifest measurement errors. Hence we can conclude that the calculated value of the error of the device tends to a certain value. It is worth considering that in the process of operation the accuracy of the device decreases, for various reasons, and the readings of the device will also tend to the actual measured value of the electrical quantity, but with large deviations. This conclusion confirms the theory put forward earlier about the aspiration of the device readings to a certain value. Clearly distinguished among all deviations are gross errors, the manifestation of which has a probabilistic nature, for a variety of reasons and are statistical outliers that distort the results of the evaluation. It should be noted that the accuracy parameter refers to a qualitative characteristic describing the compliance of the device to perform a direct function according to the established norms and requirements. At the same time, compliance with accuracy determines correctness, which also refers to quality parameters in the context of measurements by electrical CMI. Considering measurement errors from the point of view of statistics, gross deviations will be the most prominent among other values.

Figures 1 and 2 show the results of compliance with the normal law of distribution of the reduced errors of electrical measurements of two different electrical CMI. Analysis of deviations for compliance with the distribution law was carried out using the Shapiro-Wilk, Kolmagorov-Smrnov and Lilliefors criteria. These criteria were chosen because of the adequacy of working with small samples. To solve the same problem Pearson's criterion is not suitable, because for a more accurate analysis of the results, there is a need to increase the sample to 30 or more values. According to the analysis of the results in Statistica program it was determined that the compliance of measurement errors with the normal distribution law. This can be seen from the correspondence of the graphs of Figures 1 A) and 2 A) to the normal distribution diagram. The Kolmagorov-Smirnov and Liliefors criterion in both cases shows exceeding the significance level greater than 0.2, which corresponds to the normality of the distribution law. The Shapiro-Wilk criterion also confirms the hypothesis of normality of distribution as in both cases the significance level is greater than 0.05.

In Fig. 1 Б) and 2 Б), gross error outliers were artificially introduced at random. The Kolmagorov-Smirnov criterion indicates that the results of deviations with gross errors conform to the normal distribution law, while the Lilliefors and Shapiro-Wilke criteria reject this hypothesis. Visually, the graphs have changed a lot in comparison with the initial ones, Fig. 2 A) has acquired a distribution visually more similar to the uniform one.

А)   Б)
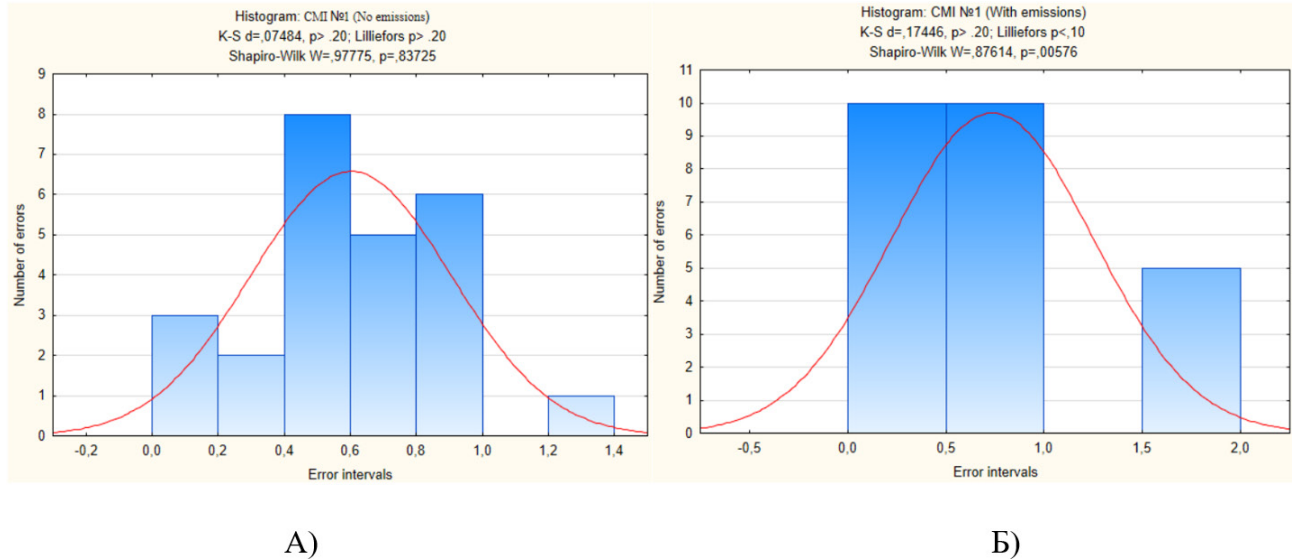
Fig. 1. Histogram of evaluation of compliance deviations of electrical CMI № 1
А) without emissions, Б) with emissions

А)   Б)
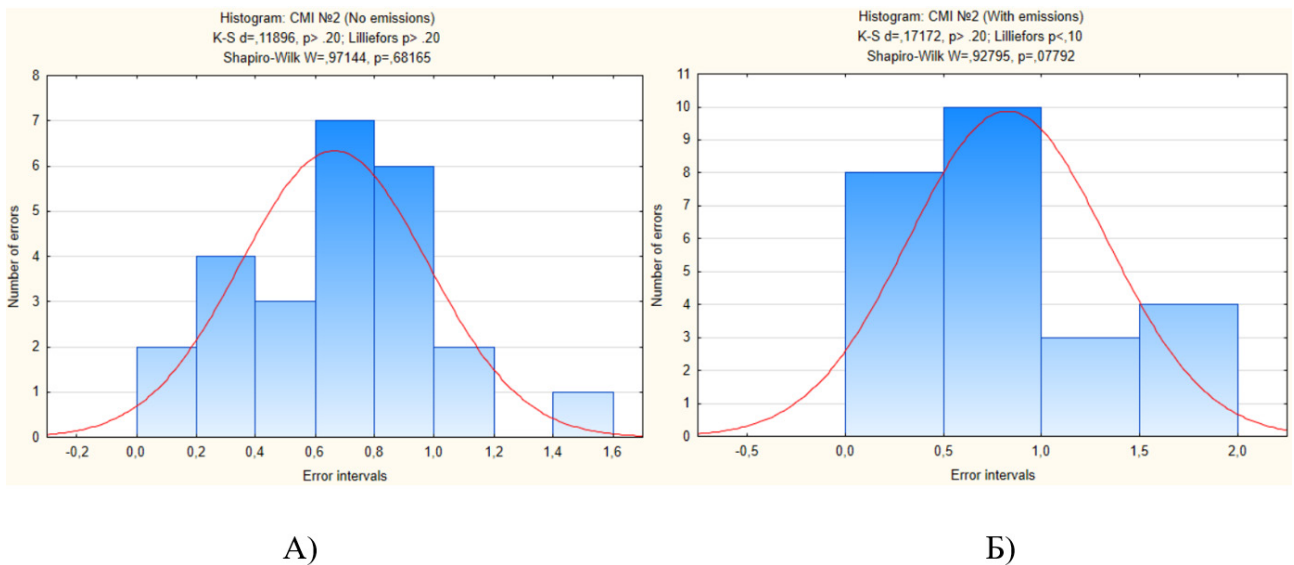
Fig. 2. Histogram of evaluation of compliance deviations of electrical CMI № 2
А) without outliers, Б) with outliers

These estimates show the extent to which the introduced changes affect the magnitude distribution. Prior to the introduction of the errors, none of the reported measurement errors in CMI 1 and CMI 2 exceeded the ± 1.5 % stated in the device data sheet, with the average accuracy values for these devices being 0.6 % and 0.665 %. After entering the deviation values 3, the values for CMI 1 and CMI 2 exceeded the accuracy limits, with the average for CMI 1 being 0.737 % and for CMI 2 being 0.833 %. According to the requirements.

If during the verification of an electrical instrument at least one measurement falls outside the accuracy limits of the instrument, it is common practice that the instrument is unusable, which is confirmed by the requirements for carrying out measurements [7]. This is due to the fact that the accuracy of the device is one of its main qualitative indicators and serves to determine the compliance of measured values with reference or standard values. Of course, it is possible to notice the influence of coarse deviations on

the results of average accuracy indices of electrical CMI. For average accuracy values close to the boundary value, coarse deviations will have a much stronger influence, because in this case the accuracy of the instrument may go beyond the nameplate values. In attestation and evaluation of parameters of deviations of electrical measurements one should not forget about the quality of evaluation. According to the standard "Reliability and requirements for methods of verification of measuring instruments" it is possible to evaluate the attestation of electrical instrumentation in terms of quality [8]. The basis of mathematical apparatus in these requirements is the probability of correct attestation of CMI and possible occurrence of false attestation errors of the 1st kind or false non-attestation of the 2nd kind. These criteria describe not only the compliance of the CMI accuracy with the declared passport values, but also establish the quantitative value of not only possible errors, but also the difference of change or influence of parameters on the certification as a whole. In this way it is possible to additionally or more accurately assess the impact of the proposed implemented techniques.

Figures 3 and 4 show the constructed operational characteristics of the verification of the deviation results of the two electrical CMI investigated earlier for the change of the distribution law considering the influence of gross errors.
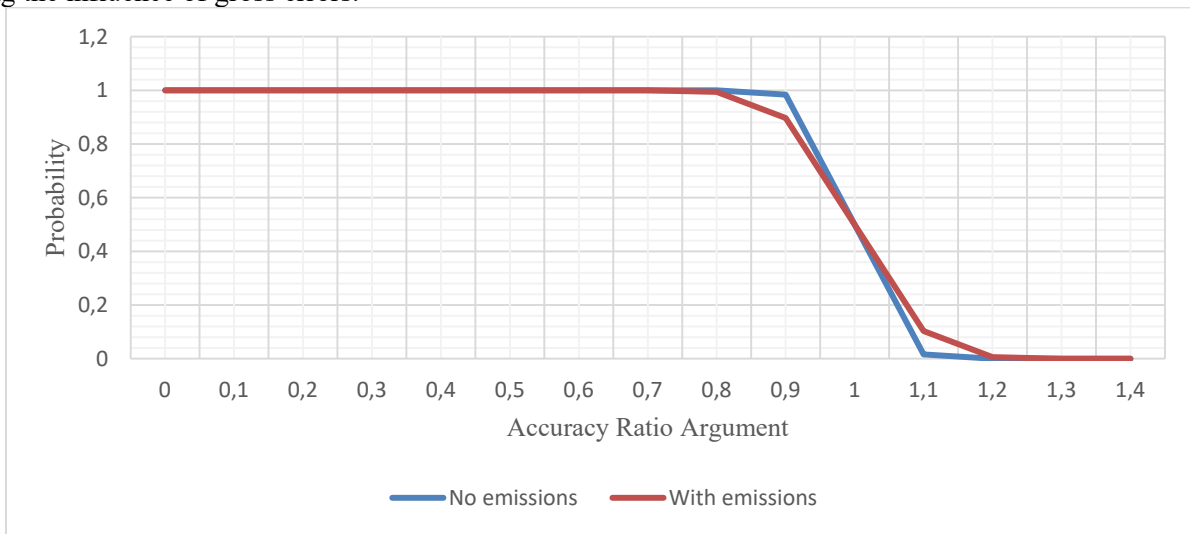
Fig. 3. Graphs of operational characteristic of verification of electrical CMI №1
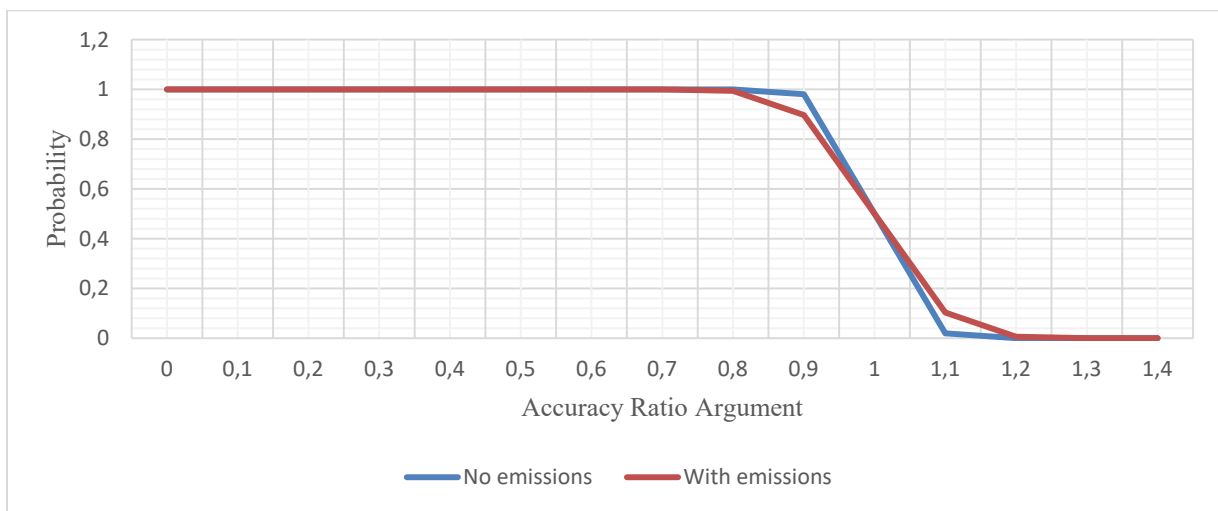
Fig. 4. Graphs of operational characteristic of verification of electrical CMI № 2

Values below the curve up to 1 on the X-axis of the accuracy ratio argument describe the area of correct estimation. For example, for the point 0.5 on the X axis, the probability of correct estimation is 100% in both cases. For the point 0.9, the difference of the correct attestation values between the values with and without outliers is clearly visible, it reaches about 5% at the point, for the difference between the

average accuracy ratios. At the same time, an increase in false attestation can be seen in the area under the curve for arguments of the accuracy ratio greater than 1.

Conclusion

The main task of the electrical CMI certification is to set the accuracy index of the investigated measuring device and assess its compliance with the established requirements of the device functioning. Errors appearing as a result of verification can sufficiently distort the results of the verification event. Such errors can have a random manifestation, systematic, methodological, etc. Coarse outliers are most clearly distinguished among the whole sample of measurement results. Such phenomenon is clearly visible when considering the results of attestation from the point of view of statistics. According to numerous researches of scientists it has been revealed that the results of deviations of electrical CMI measurements correspond to the normal distribution law, which was considered in the study and the given graphs and evaluation results confirmed this theory. The Kolmagorov-Smirnov, Lilliefors and Shapiro-Wilke criteria indicated the conformity of the presented samples, before the introduction of coarse outliers, to the normal distribution law. At the same time, the estimation of the average accuracy and indicated that the results fully satisfied the accuracy of the device. In addition, these results were considered from the point of view of probability and reliability, where the construction of operational characterization revealed a clear influence of coarse outliers on the certification of measuring instruments.

Summarizing the conducted research, we can confidently establish that the expediency of using different methods of exclusion of gross errors is justified. This conclusion is based on the average estimate of the accuracy parameter in terms of statistics and the probability of verification of electrical CMI.

References

1. GOST 8.395-80. State system for ensuring the uniformity of measurements. Normal conditions of measurements during verification. General requirements;

2. Bespalov B.B., Tarasova N.A., Burmistrov V.V. Brief error theory. 2008. №3. C. 11-12. Date of circulation (05.12.23);

3. Kirichenko I., Morneva M., Kashura A., Popov S. Metrological support of production in mechanical engineering 2017. C. 76-79. Date of circulation (25.12.23);

4. Bobryshov A.P., Research of peculiarities of maintenance of automation of verification of electric control and measuring devices. Collection of articles of the II International Research Contest 2023. C. 57-67. Date of circulation (07.01.24);

5. Lysyutina A.I., Quality of production: The concept and characteristics of quality. Izvestiya TulSU. Technical Sciences 2020. №3 C. 282-285. Date of circulation (14.01.24);

6. Vasin S.G. Quality management. Universal approach 2014. Date of circulation (15.01.24)

7. GOST R 8.563-96. State system for ensuring the uniformity of measurements. Methods for performing measurements;

8. MI 187-86, MI 188-86 Reliability and requirements for methods of verification of measuring instruments.

APPLYING THE GRNN MODEL TO FILL THE MISSING VALUES

Viktoriya Bozhenko

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: vibozhenko@yandex.ru

Abstract. *The study describes the application of a generalized regression neural network (GRNN) to solve the problem of filling the missing values in a dataset. The basic concepts of network operation, the features of this architecture, its advantages and disadvantages are considered. A technique for using GRNN to restore passes is also provided. In addition, the article outlines possible prospects for further research in this area and demonstrates the results of training a model on the dataset.*
Keywords: *GRNN, missing values, neural networks, regression, data preprocessing.*

Filling missing values in the dataset is an important step of preprocessing and can have a significant impact on the results of the analysis. Modern researchers often face the problem of missing data that needs to be processed, because the absence of some values or their incorrect filling leads to the impossibility of using machine learning methods and making decisions based on available data, making the analysis of such data useless [1]. Thus, finding effective methods in order to properly process the gaps is an actual problem. There are various methods for filling such values, the simplest ones include filling with the average or median value for a column [1], however, such methods do not allow to see complex dependencies in the data, but only describe the general picture [2]. The task of restoring numerical data gaps can be reduced to a forecasting task. One of the promising approaches to solving this problem is the use of neural networks. When starting to develop a neural network model, usually we face the problem of choosing the optimal neural network architecture. Sometimes simple fully connected neural networks are used, but a more effective solution is GRNN models that can find hidden nonlinear patterns in the data [3]. These are the models that will be considered in this paper.
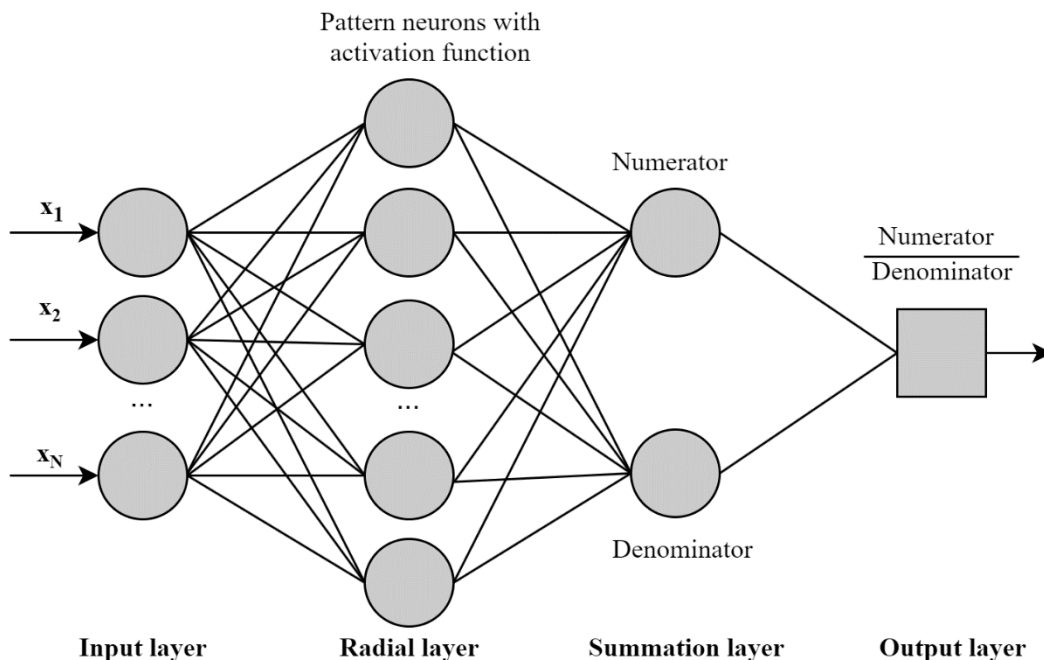
Fig. 1. GRNN Architecture

A generalized regression neural network (GRNN) is a type of neural network with a radial basis. This network can be used for regression, forecasting and classification. Based on the existing training data X and Y, the network calculates the predicted $\hat{Y}$ [4]. GRNN consists of several layers: the input, the hid-

den (pattern), the summation and the output (division). Observations are received on the input layer, the activation function of this layer is linear. The second layer of such a network is called a radial-basis network with a radial-basis activation function, the summation layer transmits the numerator (the sum of multiplication of the training and activation outputs) and the denominator (the sum of all activation functions), and the output layer contains one neuron that calculates the output data by dividing part of the numerator by part of the denominator and is designed to evaluate the weighted average [5, 6].

The probability density function used in GRNN is a normal distribution and is based on the following equation:

$$\hat{Y}(X) = \frac{\sum\limits_{i=1}^{n} Y_i e^{\left(-\frac{D_i^2}{2\sigma^2}\right)}}{\sum\limits_{i=1}^{n} e^{\left(-\frac{D_i^2}{2\sigma^2}\right)}}, \quad D_i^2 = (X - X_i)^T (X - X_i),$$

where X – is the input sample, X_i – training sample, σ – speed parameter, Y_i – is the output, D_i^2 – is the Euclidean distance from X.

The following equation is used as the activation function:

$$activation = e^{\left(-\frac{D_i^2}{2\sigma^2}\right)}.$$

D is used as a measure of how well each training sample can represent predictions. σ – the only important parameter that needs to be determined, while there is no specific rule on how to select it, as a rule, a value from 0 to 1 is selected.

GRRN is a one–pass algorithm, input weights are used for GRNN training – these are the transposed training inputs, and the output weights are the learning targets. During training, weights and parameters are tuned to minimize prediction error. The best approach is to find where the mean squared error (MSE) is minimal:

$$MSE = \frac{1}{n} \sum\limits_{i=1}^{n} (\hat{Y}_i - Y_i)^2.$$

The advantage of the GRNN network can be considered the definiteness of the structure, because the network actually accommodates all the training data. On the other hand, such a neural network structure is its biggest drawback: with a large amount of training data, the network speed decreases, sometimes very significantly, due to a noticeable increase in the complexity of the architecture. At the same time, with a small amount of input data, the network learns very quickly. GRNN allows you to count nonlinear dependencies in the data, which makes it an effective tool for approximating complex functions and predicting missing values. This method works on the radial basis function, which allows you to consider the data structure and allows you to model complex dependencies between variables [7].

The general algorithm for filling missing values can be represented as follows:

1. Check for data gaps at the preprocessing step,

2. Identify the feature or features in which you need to fill the gaps – Y,

3. Select from the complete data sample the part in which the values in the rows for this feature are completely filled, this particular sample will be used to train the neural network,

4. Train the GRNN model on a dataset without omissions, evaluate the accuracy of the resulting model on validation data,

5. Use a sample in which there are missing values in the target column Y, predict the values of this feature using a trained GRNN model, obtaining the restored values based on the available features X,

6. Replace the empty values in the original dataset with the restored values obtained in the previous paragraph.

There are various libraries for implementing neural networks. In this paper, the Python programming language, keras and pygrnn libraries were used to develop the GRNN model. The mean squared error (mse) was chosen as the loss function. An open dataset that contains information about diabetes has been selected as a dataset to test the network. To create the model, the glucose level was selected as the

target column, in which 15% of the missing values were found. The network is created and trained based on the prepared data, before training the network was performed data processing, which included data normalization and removal of outliers for a more accurate forecast. Rows in which the value of the target column is not filled in did not participate in the training. The remaining data were divided into training and validation samples, with 1/4 of the validation. Fig. 2 shows the results of the MSE for 125 epochs of learning. On the last epoch, the MSE is 0.0144 on training, and 0.0377 on validation. Such results allow us to apply the model to fill in the missing values based on the available data, since the spread relative to the initial values is low. However, in the course of this work, a dataset was also analyzed in which the number of completed training data was 150 rows, while it was not possible to train the GRNN model to effectively identify patterns in the data, which may be due to a small amount of training data. In addition, in order to obtain the highest quality metrics, it is necessary to apply not only data preprocessing methods, but also methods to identify the most significant features that affect the target value of the predicted parameter.
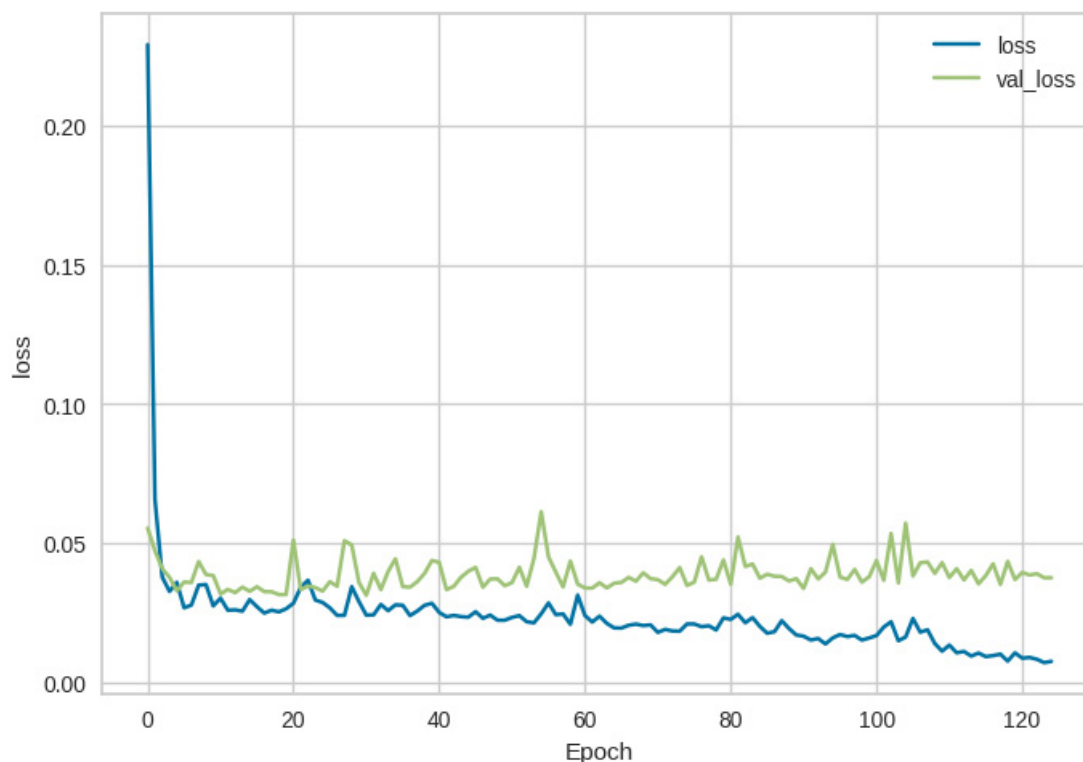
Fig. 2. Results of model training

Finally, the paper presents a technique for filling in missing values using GRNN. It can be concluded that GRNN is a good tool that can be effectively used to solve this problem, as it allows you to automatically extract complex dependencies from data and make high-accuracy predictions. However, it is worth noting that it is necessary to prepare the data correctly, as well as take into account the quality of the resulting model. If the quality metrics of the model are too low, then it is impractical to use this method. It is of interest to further study the dependence of the effectiveness of this method on the number of objects for training and compare this approach with other methods.

References

1. *Bozhenko V.V. Data preprocessing in machine learning.* Bulletin of the UNESCO department "Distance education in engineering" of the SUAI: Collection of the papers. Saint Petersburg, Issue 8. SPb.: SUAI, 2023.

2. Bozhenko V.V., Tatarnikova T.M. Application of Data Preprocessing in Medical Research. // Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), Saint Petersburg, 2023. P. 1-4.

3. Stashkova O.V., Shestopal O.V. The use of artificial neural networks to restore gaps in the source data array // Bulletin of Higher Educational Institutions. North Caucasus Region. Technical Sciences. 2017. № 1. P. 37–42.

4. Applications of General Regression Neural Networks in Dynamic Systems, URL: https://www.intechopen.com/chapters/63920

5. Andreev, P. G. The use of an artificial neural network of the GRNN type in forecasting problems / P. G. Andreev, T. V. Andreeva, N. K. Yurkov // International Conference on Soft Computing and Measurements. 2017. T. 2. P. 63-66.

6. A GRNN-based Approach towards Prediction from Small Datasets in Medical Application, URL: https://www.sciencedirect.com/science/article/pii/S1877050921006621

7. Shpakov, A.V. The use of neural networks for the approximation of experimental data / A.V. Shpakov, T. A. Lavina // Trends in the development of science and education. 2022. № 84-2. P. 60-64.

AUTOMATION WITH LASER GEOMETRY SCANNING

Daniele Casadio

Saint Petersburg State University of Aerospace Instrumentation
E-mail: *kazdanila@gmail.com*

Abstract. *Automation using laser geometry scanning is an advanced technology that is increasingly being used in a wide variety of industries, including construction, engineering, archaeology and manufacturing. This technology makes it possible to collect three-dimensional data about objects or territories with high accuracy and speed, which opens up new opportunities for analysis, design and quality control. Fig. 1, 2.*

Point clouds provide a quick visualization of a real-world object. At the same time, they are successfully used for measuring and controlling objects, 3D printing, visual visualization of hard-to-reach places or large extended objects, creation of three-dimensional and mathematical models, pattern recognition, in automated analysis, reconstruction and operation, and are also the basis for reverse engineering of real objects. In the case of reverse engineering, the first step in analyzing this type of data is usually to repair the surface and triangulate it to obtain a basic approximation in the form of a polygon mesh. Triangulation can then be used to approximate such a surface with higher-order functions, such as parametric or Nurbs surfaces. [1]
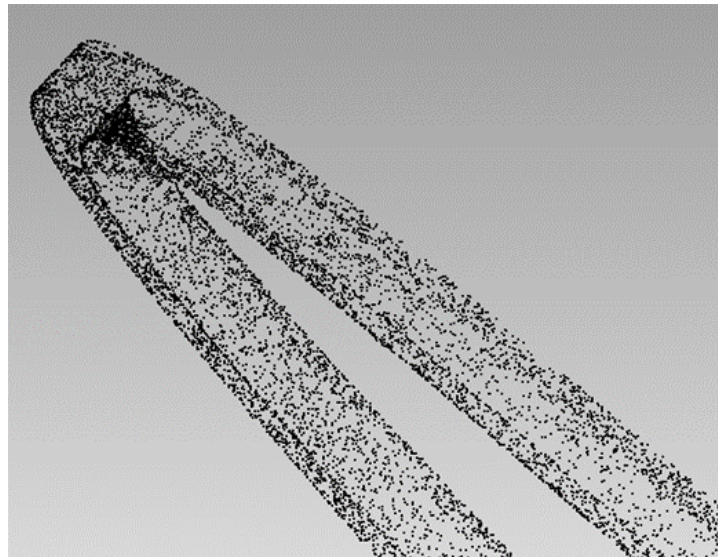
Fig. 1. Object Surface Point Cloud
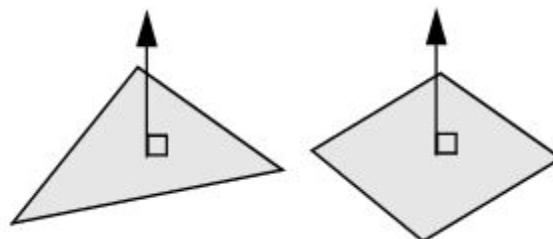
Fig. 2. Face normals.

Laser scanning is the process of capturing digital information about physical objects through the use of a laser beam. The scanner projects laser beams onto an object and measures the time it takes for the beam to bounce back to the sensor. This data is then converted into accurate three-dimensional models of the object.

Advantages of the technology

High accuracy and speed: Laser scanning allows you to obtain a detailed three-dimensional model of an object with a high degree of accuracy in a short period of time.

Versatility: The technology can be used to scan objects of varying size and complexity, from small parts to entire metropolitan areas.

Improved design and planning quality: With 3D models obtained from laser scanning, engineers and designers can analyze designs more accurately, identify potential problems, and optimize processes.

Reduce costs and time: Automating data collection processes reduces the time and cost of data processing, as well as minimizes the likelihood of errors [2,3].

A polygon mesh (model) is a set of topologically related simple geometric two-dimensional primitives that describe the surface of an object. These primitives are polygons, which are shapes with straight sides (3 or more sides) defined by points in three-dimensional space (vertices) and lines (edges) connecting these points. Using a 3D scanner, Fig. 2,3.

Fig. 2. 3D scanner

Fig. 3. 3D Scanning Engine

The inner area of the polygon is called the face. Typically, in a polygon mesh, vertices and edges are shared by different faces to implement their topology. In this case, they are referred to as shared vertices or shared faces. The outer edges of the mesh are called boundary edges. The front side of a polygon face is graphically represented by a vector perpendicular to it, called the face normal. The order in which

the vertices surrounding the face are listed determines the direction of the face (which way the face is facing and which way the face is facing the wrong side). This fact can be important because polygons are only visible from their face. The polygon model is more accurate and closer to the original feature the smaller the size of the polygons and the greater the number of polygons. Therefore, polygonal models are divided into high poly and low poly, and since the amount of data in polygonal models greatly affects the performance of technical equipment, these two types of polygonal models have different applications. Low-poly models are mainly used where you need high performance and speed to render the model in a program, such as in games. However, it is possible to increase the detail of the model by using tessellation. Tessellation is a method by which it is possible to increase the number of polygons in a three-dimensional model using Bézier curves. In this case, each polygon of the model is divided into a specified number of related polygons, which are lined up according to the general direction of the model's surface. In this way, you can first create a simple model and then quickly and easily increase its detail. Polygonal modeling is modeling only the surface of an object, so it belongs to hollow modeling [4,5]. This manifests itself in the fact that when editing such a model, if you remove some of the polygons, a hole will appear in the surface through which the entire interior of the model will be visible. That is, such a model, unlike solid models, does not have information about its volume. Therefore, the disadvantages of polygonal models include the fact that it is impossible to obtain information about the physical properties of an object, such as mass, volume, center of gravity, etc., from such models.

Conclusion

Automation using laser scanning of geometry opens up new horizons in design, manufacturing and research. Thanks to its high precision, speed and flexibility, this technology is becoming an integral part of modern industry, contributing to its efficiency and innovation.

The use of laser scanning to automate processes is making significant changes in the approaches to the development and implementation of innovative solutions. This technology makes it possible to obtain three-dimensional models of objects with incredible detail and accuracy, which in turn provides a deeper understanding of their structure and functioning. In this way, engineers and designers can not only improve existing products, but also create entirely new ones that would not be possible without this level of visualization and analysis.

The use of laser scanning in automation also helps to reduce the time and cost of design and production. Automatic data collection reduces the need for manual labor and minimizes the likelihood of human errors. This is especially important in industries such as aerospace and automotive, where high demands on precision and reliability make every detail critical.

In addition, laser scanning plays a key role in supporting sustainable development and environmentally friendly production. Accurate modeling and analysis help optimize the use of materials, reducing waste and improving resource efficiency. In construction, for example, this makes it possible to accurately assess the amount of materials needed, avoiding their excessive use and contributing to a reduced environmental impact.

In the field of cultural heritage, laser scanning opens up new possibilities for documenting and preserving historical monuments. The three-dimensional models created with it can serve to virtually reconstruct objects damaged by time or disasters, providing an invaluable resource for researchers and future generations.

Finally, advances in laser scanning and related automated technologies are contributing to the advancement of science and technology. Researchers have the opportunity to study objects and processes with an unprecedented level of detail, discovering new patterns and mechanisms in a variety of fields, from biology and medicine to archaeology and planetary science. In this way, automation using laser scanning not only increases the efficiency of existing processes, but is also a powerful stimulus for innovation that contributes to the development of knowledge and technologies of the future.

References

1. Chabanenko A V, Kurlov A V 2021 Control the quality of polymers based on the model of Dzeno Journal of Physics: Conference Series
2. Pavel Kosushkin Edition: Vector High Technologies No. 2 (23) 2016
3. Batkovskiy A M, Nesterov V A, S emenova E G, Sudakov V A and Fomina A V 2017 Developing intelligent decision support systems in multi-criteria problems of administrative-territorial formations infrastructure projects assessment *Journal of Applied Economic Sciences* **12(5)** 1301-11.

4. Maiorov E E, Prokopenko V T, Mashek A C, Tsygankova G A, Kurlov A V, Khokhlova M V, Kirik D I and Kapralov D D 2018 Experimental study of metrological characteristics of the automated interferometric system for measuring the surface shape of diffusely reflecting objects Measurement Techniques 60(10) 1016-21.

5. A V 2018 Quality management of radioactive electronic equipment *RIA: Journal: "Standards and Quality"* **2** 90-94.

6. V. L. Minaev, G. N. Vishnyakov, A. D. Ivanov, G. G. Levin Methods of control of geometrical parameters and internal stresses of products of additive technologies 2020.14.1.42.54

LIDAR-BASED AUTONOMOUS 3D SCANNING OF ROOMS

Egor Dolgov

Saint Petersburg State University of Aerospace Instrumentation, Ivangorod Branch,
Saint Petersburg, Russia,
E-mail: egorka.dolgov.2000@mail.ru

Abstract. *The work presents the issue of forming a point cloud during autonomous 3D scanning of premises by a robotic system. The tool used is LiDAR placed on a mobile platform with a manipulator or a moving bracket. The work solves the problem of precise positioning of all elements of the system. The direct kinematics problem is solved at the first stage of constructing a point cloud. A simplified technique is defined for automation instead of solving the inverse kinematics problem when there are voids in the point cloud. Mathematical models were built for a system with a certain set of elements to solve both problems. The question of adapting the resulting models to arbitrary sets of elements is discussed at the end.*

Keywords: *Autonomous 3D Scanning, Automation, Kinematics Problem, Mathematical Model.*

3D scanning of rooms is used in various tasks. This is the task of obtaining a digital copy of the premises, preserving historical values, archeology, speleology, searching for victims under the rubble, and others [1] [2]. The study of premises is carried out using robotic platforms controlled by humans. The technical means used, the method of communication with the robotic platform and the situation require the abandonment of direct control and automation of the process.

The result of the 3D scan will be a cloud of points. It is used to assess a situation or obtain a set of surfaces for analysis [3]. A point cloud consists of sets of coordinates of individual points. Most modern 3D scanners are time-of-flight and determine the distance to the nearest obstacle. These include acoustic and light scanners. LiDAR has the highest resolution and the smallest error.

LiDAR on a mobile robotic platform can obtain the distance to an obstacle – the basis for the coordinates of a point in space. The direct kinematics problem converts this value into point coordinates. The direct problem of kinematics is simple. But solving it automatically requires description and adaptation.

The solution to the direct kinematics problem should be divided: the first task is the location and orientation of the mobile platform, вторая – вопрос позиционирования места крепления LiDAR, the third task is to determine the coordinates of a point based on previous tasks and sensor readings.

In the first task, you need to determine the position of the mobile platform *B* (the reference point of the manipulator on it) and its rotation angle around the *Oz* axis relative to the base β – fig. 1. The position can be determined using:

- sensors (magnetometer and gyroscope, satnav, etc.);
- LiDAR measurements of distances to obstacles using a predefined algorithm and previous measurements;
- mathematical models [4] [5].

The positioning method is based on the hardware characteristics of the platform and scanning task. In what follows, we will assume that the coordinates (x_0, y_0) and the angle β are already known.
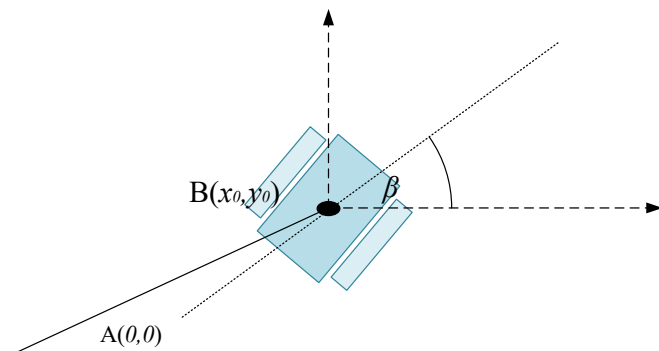
Fig. 1. Mobile platform: horizontal displacement from the reference point
and rotation angle from the reference direction

In the second problem, we neglect the rotation of the manipulator around the base. Most manipulators perform movements in one plane. In this case, the solution to this problem is to determine the coordinates of the position of the last joint (the place where the LiDAR is attached). The complexity of the task depends on the number of links in the manipulator (fig. 2). To solve the problem, we can use the rotational matrix [6] or use a right triangle (the manipulator links are vectors, the point is the sum of vectors).

We will consider the manipulator in coordinates (x,z). To find a point (fig. 2, left), each link of length l_i must be rotated around its origin by the sum of all previous angles θ_j:

$$x_i = l_i \cdot \cos\left(\sum_{j=1}^{i} \theta_j\right),$$
(1)

$$z_i = l_i \cdot \sin\left(\sum_{j=1}^{i} \theta_j\right).$$
(2)

By adding separately the values along the axes Ox – (1) and Oz – (2) we obtain the required values:

$$x = \sum_{i=1}^{n}\left[l_i \cdot \cos\left(\sum_{j=1}^{i} \theta_j\right)\right],$$
(3)

$$z = \sum_{i=1}^{n}\left[l_i \cdot \sin\left(\sum_{j=1}^{i} \theta_j\right)\right],$$
(4)

where n is the number of manipulator links.
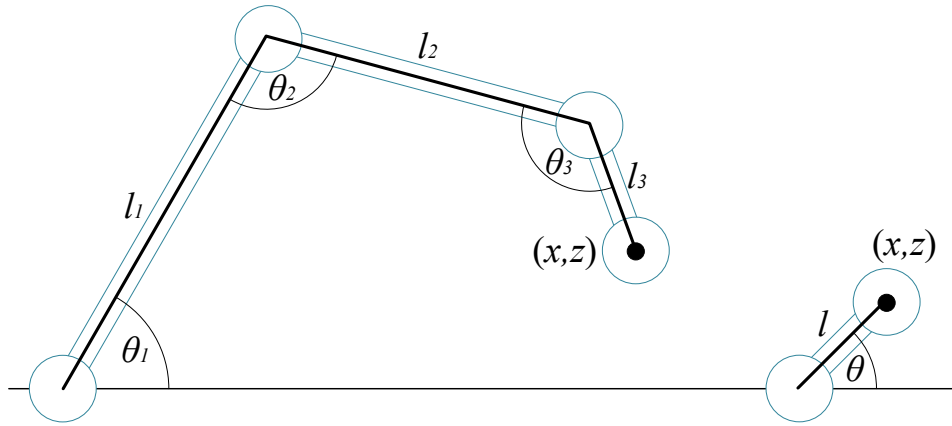
Fig. 2. Manipulator: multi-link (left) and single-link (right)

When solving the problem of automatically collecting a point cloud indoors using LiDAR placed on a mobile platform, the manipulator deflects it relative to the floor plane. This simplified version (fig. 2, right) is solved by (3) and (4).

The lengths of all links l are known in advance. If the set of angles θ for each link is known in advance, it is necessary to include pre-calculated values in the firmware of the mobile scanner controller. This will speed up calculations and reduce the time it takes to generate a point cloud. If the algorithm of the mobile scanner includes dynamic determination of angles θ, then it is worth using memoization.

In the third task, for most LiDAR installations, the determination of relative coordinates (x,z) is performed similarly, therefore (3) and (4) can be supplemented:

$$x' = \sum_{i=1}^{n+1}\left[l_i \cdot \cos\left(\sum_{j=1}^{i} \theta_j\right)\right],$$
(5)

$$z' = \sum_{i=1}^{n+1}\left[l_i \cdot \sin\left(\sum_{j=1}^{i} \theta_j \right) \right], \tag{6}$$

where $l_1..l_n$ are the known lengths of the manipulator links, l_{n+1} is the distance to the point obtained by LiDAR (fig. 3, left), x' and z' – are the values of the known coordinates of the obtained point without taking into account the positioning of the mobile platform.. The result will be a point with coordinates $(x',0,z')$, which must be rotated around the Oz axis:

$$x'' = x' \cdot \cos\beta, \tag{7}$$

$$y'' = x' \cdot \sin\beta, \tag{8}$$

$$z'' = z'. \tag{9}$$
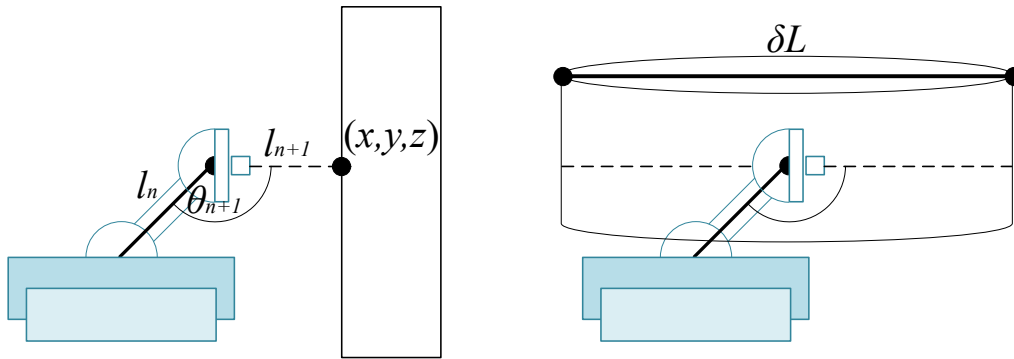
Fig. 3. Operating ranges of the manipulator: direct kinematics (left) and inverse kinematics (right)

Let us shift the point obtained in (7), (8), (9) to (x_0,y_0,z_0) and express it through (5) and (6):

$$x = x_0 + \sum_{i=1}^{n+1}\left[l_i \cdot \cos\left(\sum_{j=1}^{i} \theta_j \right) \right] \cdot \cos\beta, \tag{10}$$

$$y = y_0 + \sum_{i=1}^{n+1}\left[l_i \cdot \cos\left(\sum_{j=1}^{i} \theta_j \right) \right] \cdot \sin\beta, \tag{11}$$

$$z = z_0 + \sum_{i=1}^{n+1}\left[l_i \cdot \sin\left(\sum_{j=1}^{i} \theta_j \right) \right]. \tag{12}$$

Constants and memoization speed up (x,y,z) calculations. A point cloud is obtained without reducing the speed of point collection.

The solution to the inverse kinematics problem is to determine the position (x_0,y_0,z_0), angle β and all angles θ at which the analyzed area is within the optimal range of the LiDAR. If any area needs to be explored in more detail, then this approach is used. The permissible range of operation is determined (fig. 3, right) taking into account the deviation within the documented limits of all moving parts. The mobile platform is moved so that the resulting area coincides with the analyzed area.

The mathematical model obtained in (10), (11), (12) allows you to automate the process of obtaining a point cloud during 3D scanning using LiDAR of rooms or enclosed spaces. In this work, we considered an autonomous mobile platform. This platform can be implemented by wheeled and tracked ground robots (Fig. 1-3) and unmanned aerial vehicles [7]. The selected type of mobile platform only affects the method of obtaining and the meaning of the basic values (x_0,y_0,z_0). This paper does not address the issue of 3D scanning in motion. Each measurement is carried out by a standing mobile platform. The task of further research is to adapt the mathematical model to continuous motion.

References

1. Jaboyedoff, M., Oppikofer, T., Abellán, A. et al. "Use of LIDAR in landslide investigations: a review". Nat Hazards 61, 5–28. 2012. 10.1007/s11069-010-9634-2.

2. Collaro C., Herkommer M. Research, "Application, and Innovation of LiDAR" Technology in Spatial Archeology. In Encyclopedia of Information Science and Technology, Sixth Edition. 2023. 10.4018/978-1-6684-7366-5.ch054.

3. Sorokin A.A., Yakovleva E.A., Kovalenko R.A. "3D Machine Vision Systems on Microcontrollers and Microprocessors". 2021. 136 p.

4. Bartenev V.V., Yatsun S.F., Al-Ezzi A.S. "A mathematical model of motion of mobile robot with two independent driving wheels on horizontal plane". Izvestiya of Samara scientific center of the Russian academy of sciences. 2011. T. 13. No 4. Pp. 288-293.

5. Sorokin A.A., Kovalenko R.A., Yakovleva E.A. "Problems of adapting simple motion models of wheeled robots to available components". Innovative approaches in modern science. 2020. T. 21(81). Pp. 112-118.

6. Ayiz C., Kucuk S. "The kinematics of industrial robot manipulators based on the exponential rotational matrices". 2009 IEEE International Symposium on Industrial Electronics. 2009. pp. 977-982. 10.1109/ISIE.2009.5222601.

7. Kovalenko R.A., Sorokin A.A., Yakovleva E.A. "LiDAR as a spatial orientation tool" XXXIII international scientific and practical conference Issues of technical and physical and mathematical sciences in the light of modern research. 2020. T. 11(26). Pp. 20-25.

DEFENSE AGAINST ARTIFICIAL INTELLIGENCE

Evgenia Galeeva
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: zenagaleeva1315@gmail.com

Abstract. *The article deals with the urgent problem of ensuring security and protection from threats associated with the use of artificial intelligence (AI). The author notes the importance of developing effective mechanisms of protection against potential threats from AI and their negative consequences for society. In the development of AI there are dangers associated with its inordinate use in various areas of life.*
Keywords: *machine learning, artificial intelligence, law, protection, threat, society, development.*

Artificial Intelligence (AI) is a technology that allows a system, machine, or computer to perform tasks that require intelligent thinking, that is, to mimic human behavior for incremental learning using the information gained and solving specific questions. [1]

Machine Learning (ML) has a great impact on people's lives and work. The demand for ML-based products is growing every day. Better data generation, more accurate predictions and calculations are required. However, the stronger the AI industry develops, the more questions arise about protecting humans from the impact of AI on them. [2]

Machine learning is already being actively used in the sphere of information technologies, implemented in medical and military, transportation, entertainment, as well as in many other areas of human activity. However, in many countries there are still no laws regulating the activities of artificial intelligence and everything related to it. This is a big problem, as it is not regulated who is responsible for the actions of neural networks and AI.

Artificial intelligence has no concept of morality, decency, it creates what the user orders. Therefore, AI performs the role of a tool for work and entertainment, on a par with other programs and algorithms. Therefore, only the issuing company that owns the product places restrictions on the generated data. Because of this, attackers can take advantage of artificial intelligence for selfish purposes. They can create fake data and pass it off as true data, because the level of generated data is so high that it is impossible to distinguish it from real data.[3]

Also, one of the main concerns associated with the development of artificial intelligence is the possibility of losing control over it. An AI that is out of control can direct its actions against a human being, which can lead to serious consequences. To avoid these situations, AI should be tested thoroughly, modeling its work in laboratory conditions. Study and eliminate situations in which AI gets out of control. Depending on the environment of use, the consequences vary, but in all cases, AI out of control leads to dangerous situations.

Let's consider some situations:

1. Malpractice:

If the AI was designed to perform certain tasks, then without control it could continue to operate uncontrollably and accomplish its task in a way that causes damage, harm, or loss of life;

2. Error propagation:

AI can begin to exhibit undesirable behavior or make errors in conclusions if its control is not exercised. Errors can occur in a variety of areas, including finance, security, healthcare, and more;

3. Threat to privacy:

In some cases, AI may contain sensitive data or have access to vulnerable systems such as video surveillance systems or company networks. If control over the AI is lost, it could lead to the leakage of sensitive information or abuse of access.

4. Religious or ethical issues:

If AI has advanced intelligence and self-learning capabilities, its loss of control may raise questions of religion and ethics. For example, there would be the question of what status and rights AI should have and how to protect them;

5. Social implications:

AI can affect social structures, jobs and the economy as a whole. Losing control over AI can lead to increased inequality, job losses and imbalances in society.

As they work, neural networks get better and better each time, they do not lose their performance over time, unlike humans. Artificial intelligence can automate workplace processes, thereby reducing the number of jobs an employer could hire people for. It can analyze data, perform different tasks and make complex decisions. With proper training and supervision of a single neural network, it can replace a huge number of workers. [4]

In the context of machine learning models, the problem of distinguishing malicious from harmless information arises and cannot always be solved with a high degree of reliability. With such models, there is a risk that malicious data may eventually be classified as harmless, as machine learning models are trained on the data provided and seek to optimize their predictive abilities. [5]

This phenomenon, called overlearning, occurs when machine learning models contain "incomplete" knowledge about the data and are unable to generalize beyond the available training dataset. In this case, the models may distort classification by not sufficiently considering the context and characteristics of the malicious data.

The overfitting of machine learning models, which can lead to the classification of unsafe data as safe, poses a significant threat to the safety and reliability of systems. Therefore, mechanisms and methodologies are needed to combat this problem, including continuous updating and retraining of models using relevant and representative data, as well as the use of additional methods to validate and control results, including manual moderation of incoming queries.

It is not always easy to understand the true logic that machine learning models use to make decisions. This is due to the many hidden layers of classifiers or neurons that may be included in such models. As a result, it is almost impossible to discern what specific logic was used to make a decision.

Perception and interpretation of decisions made by machine learning models is complicated due to their vagueness and lack of transparency. However, it is essential for trust and understanding of the decision-making process, especially in cases involving important applications such as medicine, defense enterprises.

This complexity is due to the fact that the neural network is trained on a huge amount of different data, which may include redundant information, knowingly false data. At the same time, the machine itself pulls out of this data the features that are inaccessible for direct human perception.

Traditional security methods do not provide adequate protection. New aspects of threat modeling must also be explored to prevent new forms of AI abuse. Information security professionals must always think ahead and take into account that the threshold for entry into cybercrime has decreased significantly and the tools that help hackers steal data and break into systems have improved dramatically.

Machines must be able to analyze and identify abnormal behavior for their system. Such systems need to be modeled so that they can confront inputs when strange behavior is detected. This means that AI must be given the ability to detect when user interactions fall outside of an acceptable scenario.

References

1. Аксенова Е.И. Экспертный обзор развития технологий искусственного интеллекта в России и мире. Выбор приоритетных направлений развития искусственного интеллекта в России. М.: ГБУ «НИИОЗММ ДЗМ», 2019. – 38 с

2. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. Питер, 2018. – 480 с

3. Татарникова Т.М., Бимбетов Ф., Богданов П.Ю. Выявление аномалий сетевого трафика методом глубокого обучения // Известия СПбГЭТУ ЛЭТИ. 2021. № 4. С. 36-41.

4. Татарникова Т.М., Богданов П.Ю. Обнаружение атак в сетях интернета вещей методами машинного обучения // Информационно-управляющие системы. 2021. № 6 (115). С. 42-52.

5. Татарникова Т.М., Журавлев А.М. Нейросетевой метод обнаружения вредоносных программ на платформе ANDROID // Программные продукты и системы. 2018. № 3. С. 543-547.

RECOGNIZING TRAFFIC ACCIDENTS USING ARTIFICIAL INTELLIGENCE

Matvey Golovkin

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: matvey.golovkin@mail.ru

Annotation. *The process of creating a neural network that solves the problem of computer vision in detecting and localizing traffic accidents on public roads is described. The stages of data preparation and processing, selection of the correct architecture of the neural network and justification of the accuracy indicators of its training are given.*
Keywords: *machine learning, artificial intelligence, data analysis, data partitioning, convolutional neural network, learning error, traffic accident, road safety.*

Neural network (neural net) is one of the methods of machine learning, in which a computer program imitates the work of the brain. A distinctive feature of neural networks is that knowledge about the task it solves is distributed over the network itself, rather than written explicitly into the program. The knowledge is modeled as connections between processing elements (artificial neurons) and the adaptive weights of these connections. Teaching a neural network to solve the classification problem means adjusting the weights so that the output results are within the acceptable error [1].

Convolutional Neural Networks (CNN), which is a special class of neural networks specialized in image and video processing [2], was used to implement the model.

In CNNs, the outputs of intermediate layers form a matrix (image) or a set of matrices (multiple image layers). For example, three image layers (R-, G-, B-channels of an image) can be fed to the input of a convolutional neural network. The main types of layers in CNNs are convolutional layers, pooling layers and full-connected layers. Fig. 1 shows a CNN with an image of the digit 7 as input.
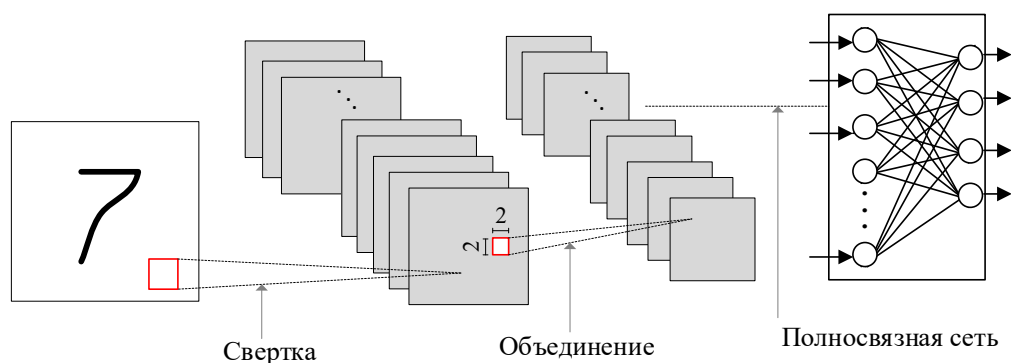
Fig. 1. Structure of convolutional neural network

For recognition of road traffic accidents (RTA) 3 models of neural network are implemented: two classifiers, the first of which determines whether there is a car in the photo, and the second whether there is an accident in the image with a car, and a localizer that highlights the car or accident in the frame of the appropriate color: if the accident, then red, if the car, then green.

All three models have the same architecture except for the output neurons. There are 5 convolutional layers in total, after each layer Max – Pooling is applied to resize the image by adding pixel block values, also one of the dropout regularization methods was used to prevent overtraining. In all layers except the last one, the ReLu neuron activation function is used. After using the Flatten function, the results of the convolutional layer computation are transferred to the full-link neural network [3-5]. The first full-link layer has 512 neurons, the same number as the number of channels at the output of the last convolutional layer. With each layer, the number of neurons decreases. In the full-link layers, the ReLu activation function is also used. If the neural network is trained for classification, the last layer has Sigmoid activation function, for localization a Linear activation function is used.

A dataset of 20000 car photos and 1200 accident photos was collected to train the neural network. 13000 other images were also collected.

All images from the dataset are converted to an array and their size and color palette are changed. The size is reduced to 208 pixels in height and width. The palette is changed from BRG to RGB. Images for training and validation samples are divided by Pareto law in the ratio of 20/80.

For the dataset of accidents and cars, the area with the desired object is cut out from the initial image using the marked coordinates.

The neural network was trained several times to identify the best model and its architecture. For good car recognition, the neural network was trained on 60 epochs, due to which a training accuracy rate of 98% was achieved, this can be seen in figure 2.
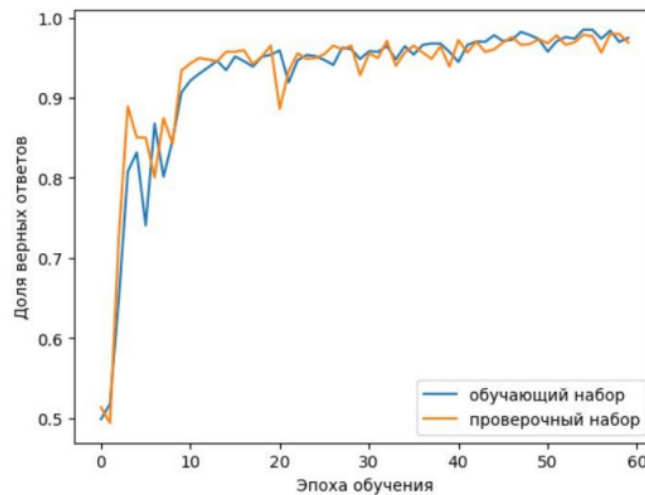
Fig. 2. Learning accuracy index

To distinguish between a car and an accident involving a car, the neural network was trained on 1300 car images and 1500 accident images, resulting in an accuracy rate of 93%.

The ready-made models are combined into one algorithm to recognize accidents and car images. The predictions of the models are passed to a function that evaluates the content of the image as a percentage. If the probability that the image depicts a car is greater than 30% and the probability of detecting an accident is greater than 25%, the function returns a tuple with information about the recognized objects. If the second condition is not met, the function returns a tuple with information that only a car is in the image. If the first condition is not met, the tuple will retain the information that nothing is detected in the image.

The results of model testing are presented in Figs. 3, 4 и 5. Fig. 3 shows that the neural network detected neither the car nor the accident, Fig. 4 shows that the neural network detected a car in the image, Fig. 5 shows that the neural network detected an accident in the image.

Fig. 3. Nothing detected

```
In [8]: #Машина
        predict_bounding_box('C:\\Users\\Matvey\\Downloads\\проверка\\18015-MC20BluInfinito-scaled-e1666008987698.jpg')
```
Обнаружена машина

Fig. 4. Machine detected

```
In [9]: #Авария
        predict_bounding_box('C:\\Users\\Matvey\\Downloads\\проверка\\Без названия11.jpg')
```
Обнаружена авария
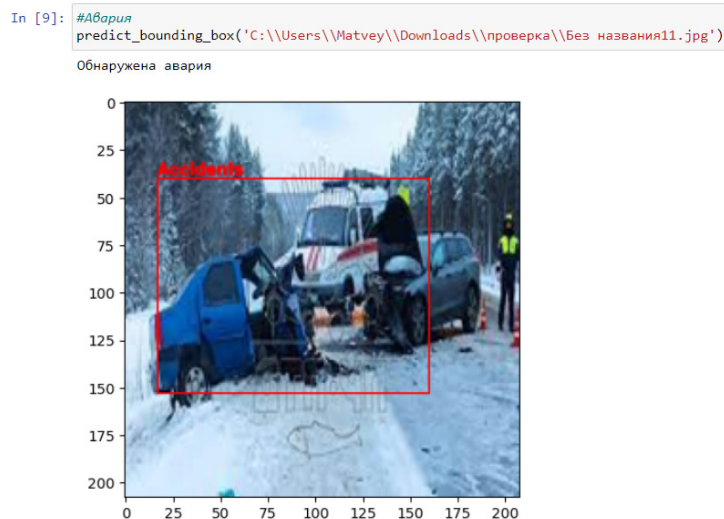
Fig. 5. Accident detected

References

1. Аксенова Е.И. Экспертный обзор развития технологий искусственного интеллекта в России и мире. Выбор приоритетных направлений развития искусственного интеллекта в России. М.: ГБУ «НИИОЗММ ДЗМ», 2019. – 38 с

2. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. Питер, 2018. – 480 с

3. Татарникова Т.М., Бимбетов Ф., Богданов П.Ю. Выявление аномалий сетевого трафика методом глубокого обучения // Известия СПбГЭТУ ЛЭТИ. 2021. № 4. С. 36-41.

4. Татарникова Т.М., Богданов П.Ю. Обнаружение атак в сетях интернета вещей методами машинного обучения // Информационно-управляющие системы. 2021. № 6 (115). С. 42-52.

5. Татарникова Т.М., Журавлев А.М. Нейросетевой метод обнаружения вредоносных программ на платформе ANDROID // Программные продукты и системы. 2018. № 3. С. 543-547.

COMPUTER VISION AND MACHINE LEARNING IN AUTOMATING THE VERIFICATION OF SCHOOL TEST PAPERS

Mikhail Gordeev *, Dmitry Doronin **

* *Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia*
**University ITMO, Russia
E-mail: nadgeka2002@gmail.com

Abstract. *The issue of automating the assessment of students' test papers to reduce the workload on teachers is being considered. The proposed solution involves implementing a software system based on computer vision and machine learning. The system operates with images of completed test sheets obtained through scanning or capturing with a smartphone. The outcomes of the system's training and its performance in assessing school tests are presented.*
Keywords: *machine learning, text recognition, text verification automation, test work verification.*

Introduction

Using testing as one of the most important components of the learning process helps educators determine, based on its results, the level of material comprehension by students and enables the formation of strategies to adjust the educational process. However, test grading is a time-consuming task, demanding a significant amount of a teacher's time. Today, its reduction is possible through the use of automated testing processes, leveraging elements rooted in machine learning.

Existing systems for grading papers using a general visual assessment of images without extracting specific sections or processing text [1,2] do not provide high accuracy. They can be used at an initial stage to eliminate corrupted test sheets. Solutions are known that rely on automatic grading using optical scanning of marks and subsequent comparison of the extracted image with a pre-prepared template. However, these solutions can only be applied to multiple-choice tasks and require users to have high-precision image scanning equipment [3].

Solutions are known that process images obtained from non-professional equipment, but they can only be used with templated task sheets [4].

The majority of these and similar solutions listed do not have open-source code, making it impossible to integrate them into software systems used in educational institutions.

The goal of this work is to propose a solution in the form of a software system based on digital processing of test sheets as images, obtained through smartphone capture or scanning, analyzing them, and recognizing printed and handwritten texts.

Stages of recognition system operation

Based on the general principles of building recognition systems [5], working with test texts consists of several stages, taking into account the specifics of the task, as depicted in the blocks in Fig. 1.

The process starts with generating a test using a question database. The test is presented in a specific standardized format to ensure its proper handling during assessment. The test is printed on paper in a defined format, usually A4 size.

Students fill in the printed test sheets with their respective answers in the designated spaces. The completed test sheets are then scanned or photographed. Examples of photographed and scanned test sheets for the subject "Literature" for 7th-grade students are shown in Fig. 2.

The obtained images of completed test sheets have their text areas identified and then recognized. Finally, the system checks the answers (compares them with the answer key) and automatically generates test results based on this comparison.

For the efficient operation of the test grading system using machine learning, a standardized format for tasks is necessary. This involves multiple-choice questions as well as short-answer tasks. Test papers are generated from a pre-existing bank of questions. This approach enables quick and efficient test creation while offering a wide array of options for students.

The original dataset of completed test papers was obtained based on an agreement with the administration of the "Learning Together" school in Saint Petersburg for conducting testing. The agreement ensures the confidentiality of the participating students involved in automated testing. Hence, the obtained test sheets were anonymized to maintain confidentiality.
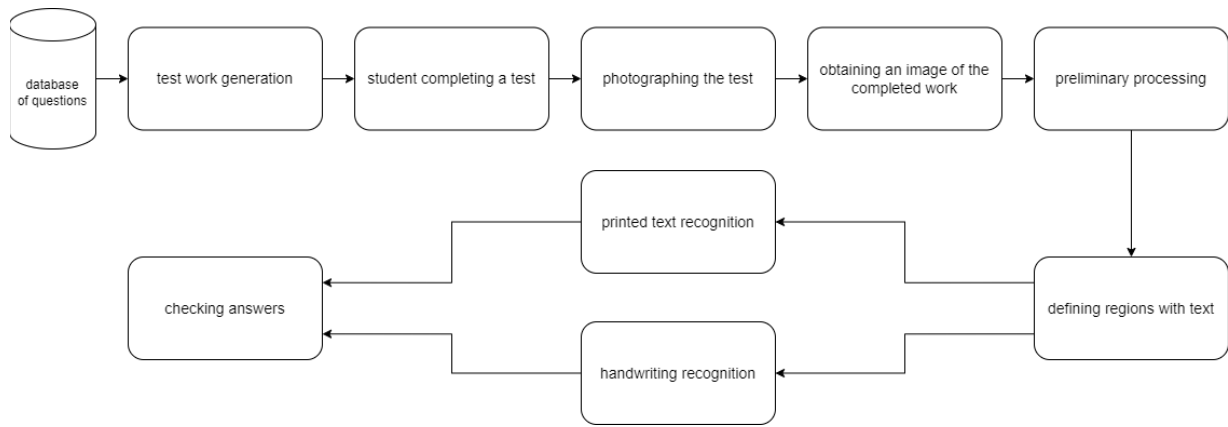
Fig. 1. Stages of System Operation

Image Acquisition. The test sheets were converted into images using both a scanner and a Sony Xperia 1 III smartphone. This enabled the experiment to have images of varying quality (see Fig. 2).
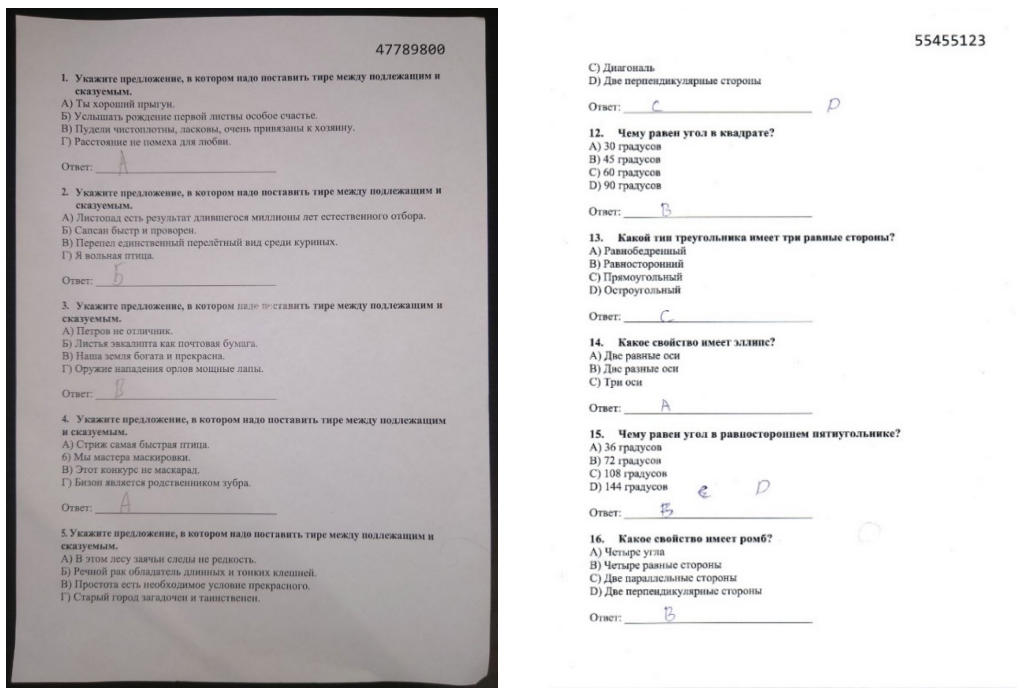

Fig. 2. Photographed and Scanned Completed Test Sheets for 'Literature'

Preprocessing of the Image. The image of a completed test sheet may have various defects that affect the quality of recognition and the overall assessment result for the student. For instance, when scanned or photographed at an angle, the image requires initial perspective correction and distortion removal.

Given the system's orientation towards school use, it's essential to consider various physical document distortions. These might include folds, tears, creases, among others, rendering basic distortion correction methods ineffective [6].

Consider the following as part of the initial image processing:
– Perspective correction with prior detection of the document contour in the image [7];
– Image enhancement by noise and interference reduction through averaging filters [8];
– Contrast and sharpness enhancement;
– Image binarization, assigning pixel values of 0 or 255 to reduce the volume of processed information;
– Removal of lines delineating the answer area to enhance the readability of handwritten text in the test [9].

Segmentation. "After processing and binarization of the image, segmentation of areas containing printed and handwritten text is performed using the YOLOv8 neural network [10], which has shown promising results in document processing. YOLOv8 employs object detection for image segmentation, determining the boundaries of text regions, rather than pixel classification."

After the completion of segmenting areas containing printed and handwritten text, their values and positions relative to each other are determined. This allows for the differentiation of handwritten responses and student notes.

For model training, manual labeling of regions on images related to handwritten and printed text was performed (see Fig. 3). The LabelImg utility [11] was used for image annotation during segmentation, allowing the export of image regions in a format suitable for training the YOLOv8 neural network.
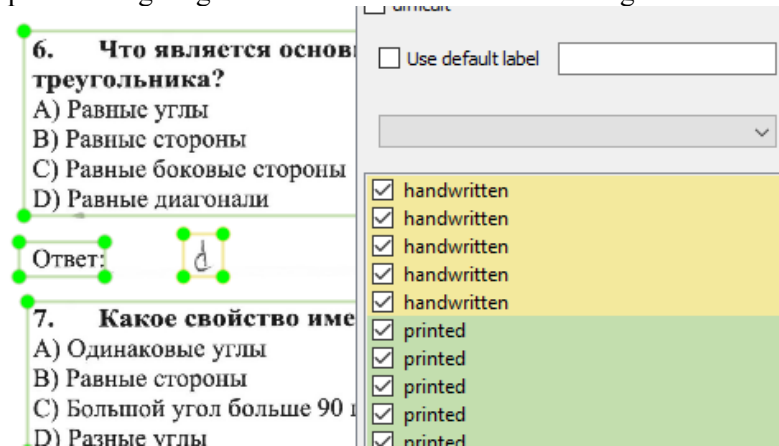
Fig. 3. Example of image annotation using the LabelImg utility

Text recognition of the test and answer verification

Tesseract [12] was used for recognizing printed text. However, for handwritten text recognition, neural network algorithms like Transformer [13], belonging to the deep learning algorithm class, proved to be more effective. They can handle variable-length sequences and are suitable for recognizing handwritten text.

After recognizing the key printed words indicating the position of the answer area, the handwritten response recognized by the system was compared against the required result to determine its accuracy.

The dataset used for training the model used in recognizing handwritten text was the "Cyrillic Handwriting Dataset" [14], which comprises 73,830 segments of handwritten Cyrillic text. Fig. 4 shows an example from the utilized dataset.
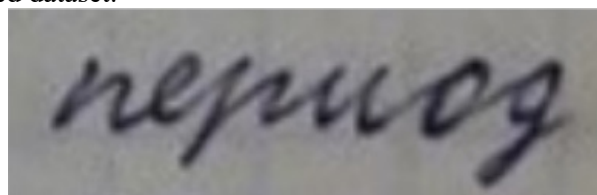
Fig. 4. Example image from the 'Cyrillic Handwriting Dataset'

Practical results of the experiment

The described solutions were implemented into a software system for conducting experiments, utilizing a computer equipped with an Intel® i7-8700k 4.70 GHz processor and an NVIDIA GeForce RTX 3080 graphics card.

The assessment of experiment results was based on the accuracy of segmenting handwritten and printed text regions and determining the content within these regions.

Two models were trained to evaluate the effectiveness of pre-processing images for the YOLOv8 neural network. The first model was trained on a pre-processed dataset containing normalized grayscale images with removed lines. The second model was trained on the original dataset containing color images. In both cases, the training effectiveness decreased after 80 iterations. It was determined that the model trained on the pre-processed dataset achieved a high average precision (mAP) of 98%. The output of its performance is presented in Fig. 5.
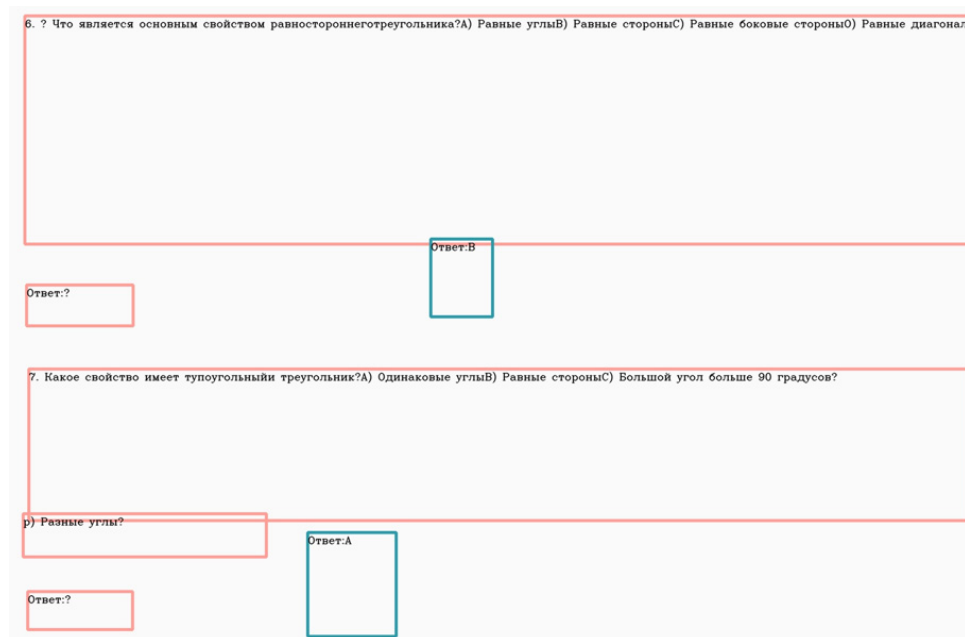
Fig. 5. Example of correct segmentation by YOLOv8

The Transformer model trained on the "Cyrillic Handwriting Dataset" achieved a high Word Error Rate (WER) of 33% and Character Error Rate (CER) of 8%. These results allow for precise identification of answers to multiple-choice test questions, given that the area containing the student's written answer is correctly identified, with high accuracy.

The experiments revealed that the utilized "Cyrillic Handwriting Dataset" lacks a sufficient variety of individual characters written in different handwriting styles, hence resulting in a CER of only 33%.

Conclusion

The developed software system based on the chosen solutions demonstrates the potential for automating test evaluations in schools and other educational institutions. It has already made it possible to relieve teachers from routine work in a test mode.

The experiments conducted with the developed system revealed that despite the high accuracy in identifying areas where answers are written, the precision in evaluating tests decreases due to relatively high errors in recognizing the answers themselves, attributed to inaccuracies in recognizing handwritten characters.

The selected dataset for training, the "Cyrillic Handwriting Dataset," is deemed insufficient. For future experiments, it's recommended to expand this dataset to improve the recognition of individual characters.

The further development of the test evaluation software system involves exploring opportunities to implement the assessment of test sheets with multiple-choice answers.

Gratitude

The authors thank A.M. Sergeev for his guidance and valuable advice in preparing this work**.**

References

1. Barskaya G. B., Krivodanova A. A., Sbrodov S. O., Shestakova A. O. Information system for recognizing and verifying test tasks / In the collection: Mathematical and information modeling. Materials of the All-Russian Conference of Young Scientists. Tyumen, 2022. P. 122-131.

2. Fisun N. T., YAblonskaya E. Yu.   Analysis of algorithms for processing text information for automated verification of open-type tests // Bulletin of the Kherson National Technical University. 2012. № 1 (44). P. 382-388.

3. Why do scantron-type tests only read #2 pencils? Can other pencils work? URL: howeverythingworks.org/2006/05/29/question-1529 (accessed 21.11.2023)

4. Xiaoshuo L., Tiezhu Y., Xuanping H., Zhe Y., Gang X. BAGS: An automatic homework grading system using the pictures taken by smart phones // *ArXiv* abs/1906.03767 (2019).

5. Tou J., Gonzalez R. Pattern recognition principles. Addison-Wesley Publishing Company, 1977. 377 p.

6. Borisenko D. I. Methods of searching for angular features in images // Young Scientist. 2011. № 5(28). Vol.1. P. 120-123.

7. Vostrikov A. A., Sergeev A. M., Solov'ev N. V., Solov'eva T. N. Introduction to Bitmap Segmentation. Saint Petersburg: SUAI, 2017. – 34 p.

8. Sergeev M. B., Solov'ev N. V., Stadnik A. I. Methods of increasing the contrast of bitmap images for digital video processing systems // Information and control systems. 2007. № 1. P. 2-7.

9. Erosh I. L., Sergeev M. B., Solov'ev N. V. Image processing and recognition in preventive security systems. Saint Petersburg: SUAI, 2005. – 152 p.

10. YOLOv8. URL: ultralytics.com/yolov8 (accessed 26.01.2023)

11. LabelImg. URL: github.com/heartexlabs/labelImg (accessed 20.12.2022)

12. Tesseract. URL: github.com/tesseract-ocr/tessdoc (accessed 01.02.2023)

13. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser Ł., Polosukhin I. Attention is all you need // Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017). P. 1 – 15.

14. Cyrillic Handwriting Dataset. URL: kaggle.com/datasets/constantinwerner/cyrillic-handwriting-dataset (accessed 17.01.2023)

DESIGN OF A MICROPROCESSOR INDOOR CLIMATE CONTROL SYSTEM WITH TIME FUNCTIONS

Igor Grigoriev

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: grigoryev_igor@mail.ru

Abstract. *This work discusses the main stages of designing a microprocessor system (MS) for controlling climatic parameters in a room with time functions, argues for the choice of certain tools, modules and development components, and tests the developed MS.*

Keywords: *designing a microprocessor system, programming the microcontroller (MCU), ATmega128, CodeVisionAVR, C programming language.*

Introduction

Microclimate is a complex of physical factors of the indoor environment of a room. The government standard, which establishes construction requirements for the microclimate of public and residential buildings, defines it as "the state of the indoor environment of a room that affects a person, characterized by air temperature and enclosing structures, humidity and air mobility" [1].

The microclimate has a direct impact on human health and well-being. In an optimal microclimate, a person feels a sense of comfort, and his body does not waste energy adapting to external conditions.

Currently, a huge number of people live in large cities where there are problems of environmental pollution and, consequently, indoor microclimate, therefore, indoor climate control devices are an important and relevant component of a person's daily life, as well as his health.

Thus, the purpose of this work is to review the stages of designing an MS for controlling climatic parameters in a room with time functions.

Analysis of the method for constructing MS for controlling climatic parameters in a room with time functions

The main function of the designed device is to control the climate parameters in the room. The purpose of the device is to turn on or off actuators when the temperature or humidity is lower or higher than specified, respectively.

The main electronic components that make up the designed MS include:

1) MCU – the device that controls the entire system. Sets the required temperature and humidity level, and also provides control signals to turn on/off the actuators.

2) Sensors – devices necessary for reading indicators of various physical quantities. Based on the data received from the sensors, the operation of the MCU will depend.

3) Real–time clock – is a device necessary for the implementation of calendar functions (turning on/off the system at user-defined time intervals).

The operation of most currently used digital temperature and humidity control systems is based on the following principle: the device has one or more inputs to which the measured values of temperature and humidity are received from the sensors. At the output of the device there are output relays to which actuators (humidifier, conditioner, heater, fan) are connected. When the set temperature or humidity parameters are reached, a voltage is applied to the selected output and the device starts working if the current time corresponds to the time interval in which the device is allowed to operate. If it is necessary to shut down the device, the voltage is removed from its output and the device turns off.

The block diagram design

Based on the analysis of the method of constructing an MS for controlling climatic parameters in a room with time functions, a block diagram was designed.

The block diagram defines the main functional parts of the product, their purpose and interrelationships [2].
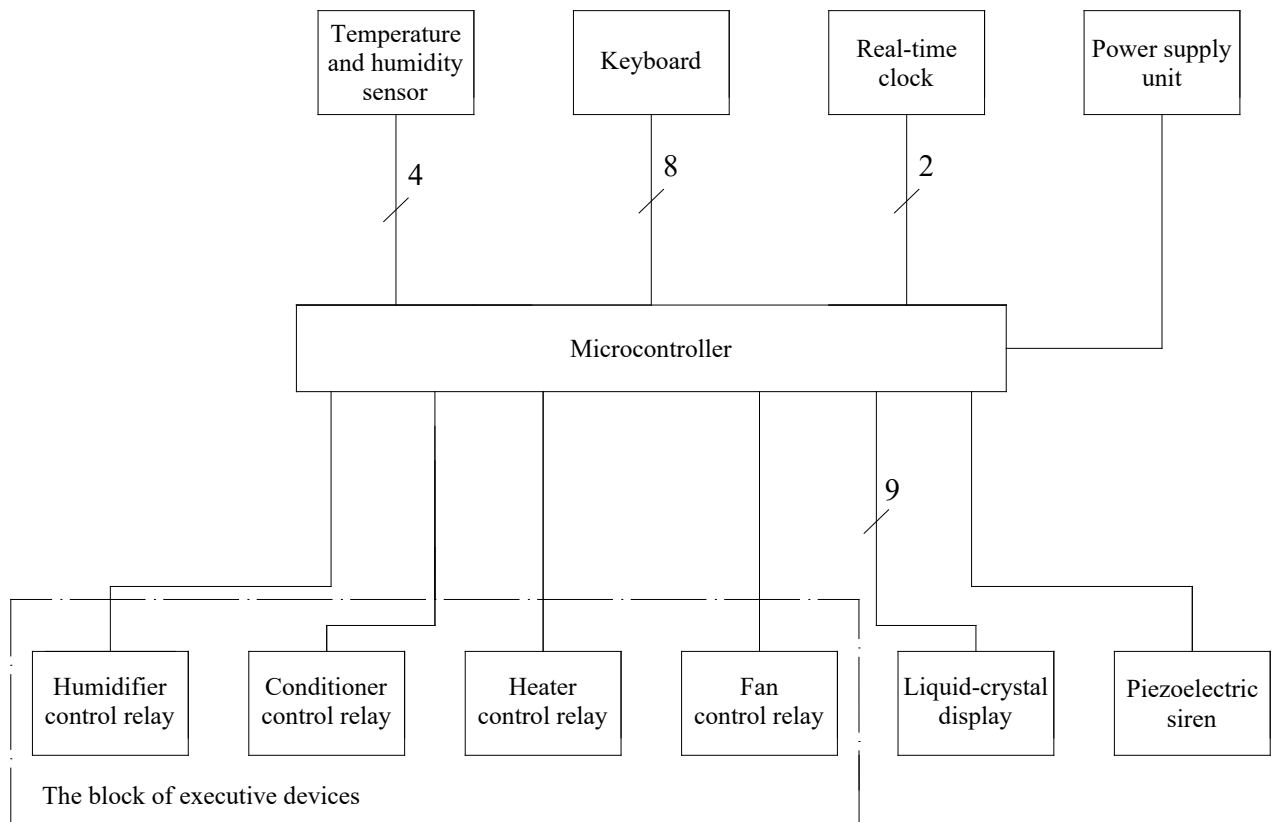
Fig. 1. The block diagram

The block diagram consists of the following blocks:

1) MCU, temperature and humidity sensor, real–time clock – the purpose of these blocks is given in the previous section;

2) Keyboard – is intended for entering various data into MCU;

3) The block of executive devices – is designed to control actuators (fan, humidifier, etc.);

4) Liquid crystal display (LCD) – is designed to display various symbolic information;

5) Piezoelectric siren – is designed to confirm the keystroke by sound accompaniment;

6) Power unit – is designed to power the device and all its logical elements, as well as to maintain its operation.

The block diagram scheme will be used to select the element base and develop a circuit diagram.

Element base selection

In the designed MS, the coordinating and managing component is ATmega128 8-bit MCU. Below are the main characteristics of this MCU [3]:

1) RISC architecture:

– 133 instructions, most of which are executed in one machine cycle;

– 32 8-bit general-purpose registers, integrated peripheral control registers;

– Completely static work;

– Performance up to 16 million operations per second at a clock frequency of 16 MHz;

2) Non-volatile program and data memory:

– Wear resistance of 128 kB of in-system reprogrammable flash memory: 1000 write/erase cycles;

– In-system programming with the built-in boot program;

– Guaranteed two-operation capability: the ability to read while writing;

– Gear resistance of 4 kB of EEPROM: 100000 write/erase cycles;

– Built-in 4 kB static RAM;

– Optional ability to address external memory up to 64 kB in size;

3) Features of peripheral devices:

– Two 8-bit timer-counters with separate prescalers and comparison modes;

– Two advanced 16-bit timer-counters with separate prescalers, comparison modes and capture modes;

– Real-time counter with separate generator;

– Two 8-bit pulse-width modulation (PDM) channels;

– 6 PDM channels with programmable resolution from 2 to 16 digits;

4) Operating voltage: 2.7 to 5.5 V.

A real-time clock is used to get information about the current date and time FM30C256-S/S020 [4].

The main characteristics of the watch include:

1) Current consumption from a backup source up to 1 μA;

2) Real-time representation in the range from seconds to centuries (binary-decimal format);

3) A standard quartz resonator with a frequency of 32.768 kHz is used.

A lithium CR2032 battery is used for backup power supply of the clock chip, the main features of which include:

1) Long charge storage period (self-discharge of about 2% at 25°C);

2) Nickel-plated sealed framework;

4) Operating temperature range: -20 to 60°C.

To set the clock frequency of the clock chip, a quartz resonator DT-38 is used, the main characteristics of which include:

1) The resonant frequency: 32.768 kHz;

2) Operating temperature range: -10 to 60°C.

MOCD213 optocouplers are used to match the levels of external control signals with the MK input signals. In this MS, they can be used to expand the functionality of the system. The maximum input current of this optocoupler corresponds to 60 mA, and the output voltage is 70 V.

To lower and stabilize the input voltage to the required value, a step-down stabilizer MAX727-5.0CCK is used.

Its main characteristics include:

1) Input voltage range: 9 to 40 V;

2) Fixed output voltage: 5 V;

3) Maximum operating current: 2.5 A.

There are seven connectors on the circuit diagram, which have the following functional purpose:

1) MSTB-8 connectors can be used to supply external control signals;

2) The 5-EHDR-4 connector is used to connect an external power supply;

3) The PLD-16 connector is used to connect a 4x4 matrix keyboard, as well as a temperature and humidity sensor;

4) The MSTB24 connector is used to connect external actuators;

5) The PLS-16 connector is used to connect a LCD;

6) The PLD-10 connector is used to connect the programmer;

The SMA-24P piezoelectric siren with a sound intensity of 92 dB is used to sound the keystroke.

Relays are used to control external actuators TTI TRD–5V, the main characteristics of which include:

1) Switching voltage: 250 V;

2) The maximum switching current: 12 A;

3) The control voltage of the coil: 5 V;

4) The rated current of the coil: 10 mA.

The circuit diagram design

Based on the above element base and the block diagram, a circuit diagram was designed.

Circuit diagram defines the full composition of the elements and the relationship between them and gives a complete (detailed) idea of the principles of operation of the product [2]. It is used to study the operation of the product, as well as during its adjustment, control or repair. The schematic diagram also serves to develop the wiring diagram and drawings.

The main function of the device is performed by MCU D1. The clocking of the MCU is carried out from the quartz resonator ZQ2.

DB1 – DB8 optocouplers are used to match the levels of external control signals with the MCU input signals.

Resistors R1 – R16 are necessary to limit the current through the input light-emitting diode (LED) of the optocoupler.

The real-time clock function is performed by the D2 chip connected to the MCU through the I²C bus. A quartz resonator ZQ1 is connected to the terminals X1 and X2 of the clock to set the clock frequency. The uninterruptible power supply for the clock chip is a lithium battery BT1.

The functionality of the control unit can be expanded by connecting external devices to connectors X1 and X2.

The MCU receives and processes data coming from the temperature and humidity sensor through the X4 connector.

Depending on the sensor readings, the MCU controls the actuators using relays K1 – K8 through transistors VT1 – VT8, switched on in the emitter follower mode. The actuators are connected to the X5 connector. Diodes VD1 – VD8 are connected in parallel to the relay coils to protect against self-induction electromotive force (EMF).

All necessary parameters are set by the user using a 4x4 matrix keyboard connected through a X4 connector.

A LCD connected to the X6 connector is used to display information. The display operates in 4-bit data bus mode. A tuning resistor R3 connected to the X6 connector is necessary to adjust the contrast level of the display.

The VT1 transistor assembly consists of two MOSFET-type transistors, one of which is designed to control the backlight of the indicator, and the second to control the piezoelectric siren BA1.

The RN1 and RN2 resistor assemblies perform the function of protecting the MCU D1 I/O ports from static electricity.

The RN3 resistor assembly is necessary to ensure bus operation on I²C lines.

The RN4 resistor assembly is necessary to provide the required signal levels in the MCU reset mode.
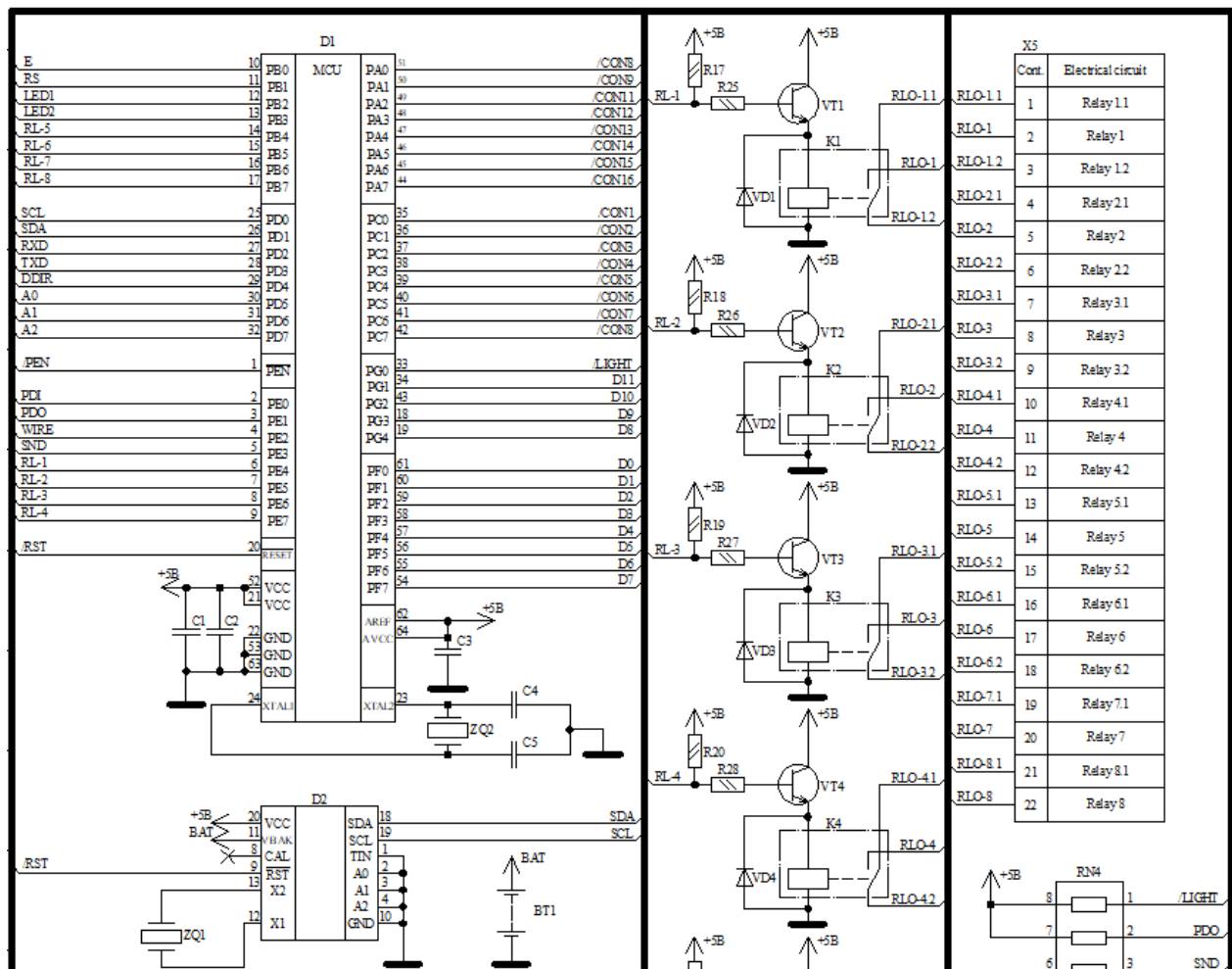
Fig. 2. A circuit diagram. The fragment

The device is powered through the X3 connector, which is supplied with an external constant voltage from 9 to 40 V. The supply voltage is supplied to the step-down stabilizer DA1, limiting it to 5 V.

The software is loaded into the MCU by the programmer through the X7 connector.

The HL1 LED indicates the presence of an input voltage, and the HL2 LED indicates the presence of a 5 V voltage at the output of the stabilizer.
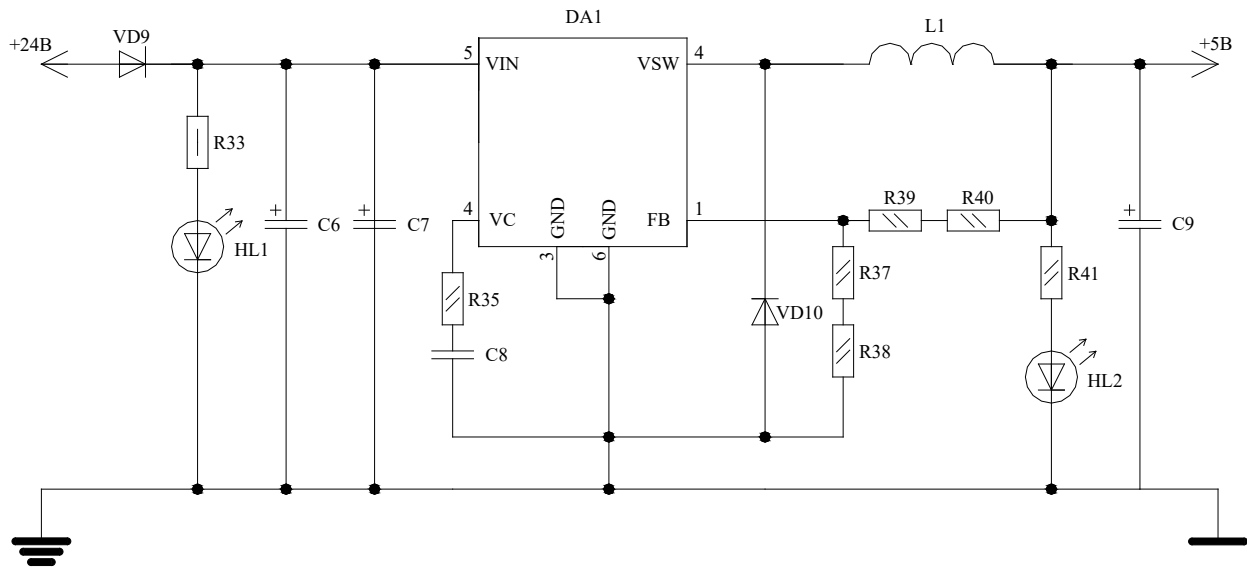
Fig. 3. A circuit diagram. The fragment. Connecting a step-down stabilizer

Based on the circuit diagram of the of the projected MS, a drawing of the printed circuit board of the device under development is performed, and the reliability of the system is calculated under specified operating conditions.

Development of a printed circuit board

A printed circuit board (PCB) is a compound made of an insulating base and metal layers, which serves for the electrical installation of the electronic components and assemblies, as well as their mechanical fastening.

A two-sided PCB type was chosen for the construction of the projected MS, since such a PCB allows for more complex circuits than one-sided PCB, and also, in comparison with multilayer PCB, a two-sided board is more economical.

Accuracy class 4. The dimensions of the PCB are 137.5 x 72 mm. PCB thickness is 1.5 mm. The type of board according to its properties is hard.

For the manufacture of a PCB, the positive combined method was chosen, since it combines all the techniques of manufacturing PCB.

When choosing the PCB base material, it is necessary to pay attention to the following factors: the expected mechanical effects (increased humidity and ambient temperature), PCB accuracy class, implemented electrical functions, performance, operating conditions and cost.

Foiled fiberglass FR-4 government standard 26246.5-89 was chosen as the PCB base material [5].

The placement of the electronic components on PCB is carried out in accordance with industry standard 4.010.030-81 [6].

Soldering of integrated circuits and electronic components is carried out using tin-lead solder with 61% tin content, government standard 21931-76 [7].

After installing the components, the board must be treated with a Flux-Off cleaner, rinsed with water, dried and varnished with Urethane Clear, except for the surfaces of the connectors and piezoelectric siren.

In everyday conditions, for the convenience of using the device, the board can be placed in a plastic rectangular case. There are four mounting holes on the board for attaching the board to the housing and for attaching the LCD to the board.
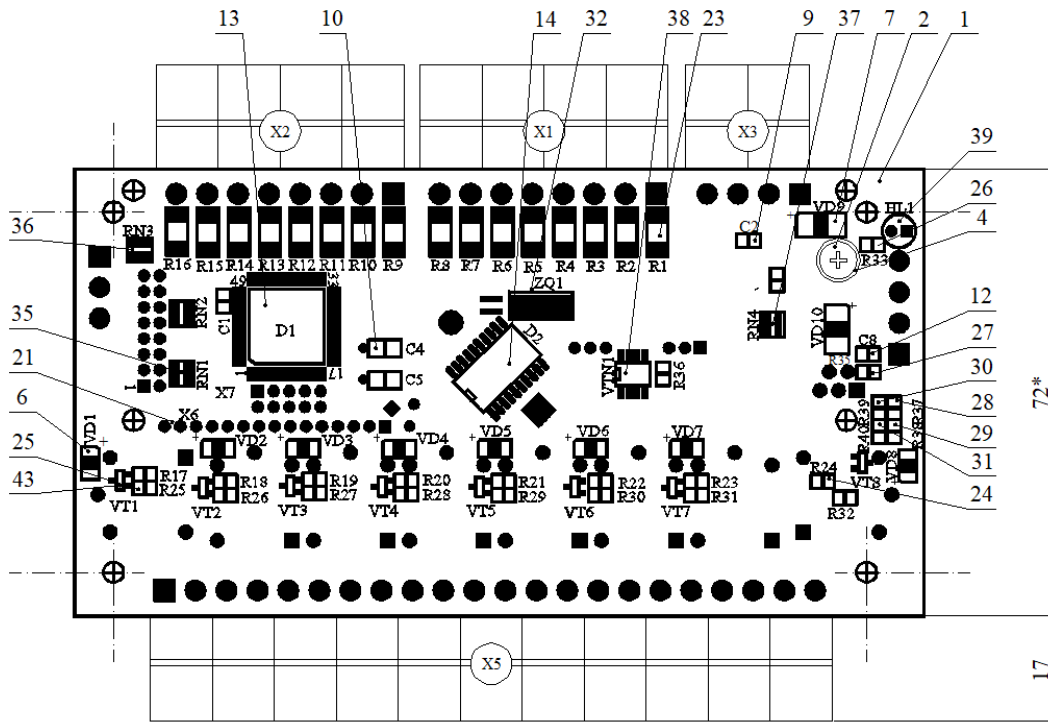
Fig. 4. An assembly drawing. The fragment. Bottom view

Wire routing

The topology design stage is a transition from circuit information (logic or electrical circuits) to geometric information – placing circuit elements in the field of the PCB drawing and creating a drawing of conductors connecting these elements.

Only at this stage will the real characteristics of the conductors become known, their length, width, area and, consequently, their capacitance, resistance and inductance, which will ultimately determine a number of important characteristics of the product, for example, its performance.

The main task of tracing electrical connections is the optimal arrangement of paths for the passage of conductors on printed circuit boards. This task is reduced to sorting through all possible options for the placement of conductors, in order to identify the optimal one with the minimum sum of the lengths of all placed conductors.
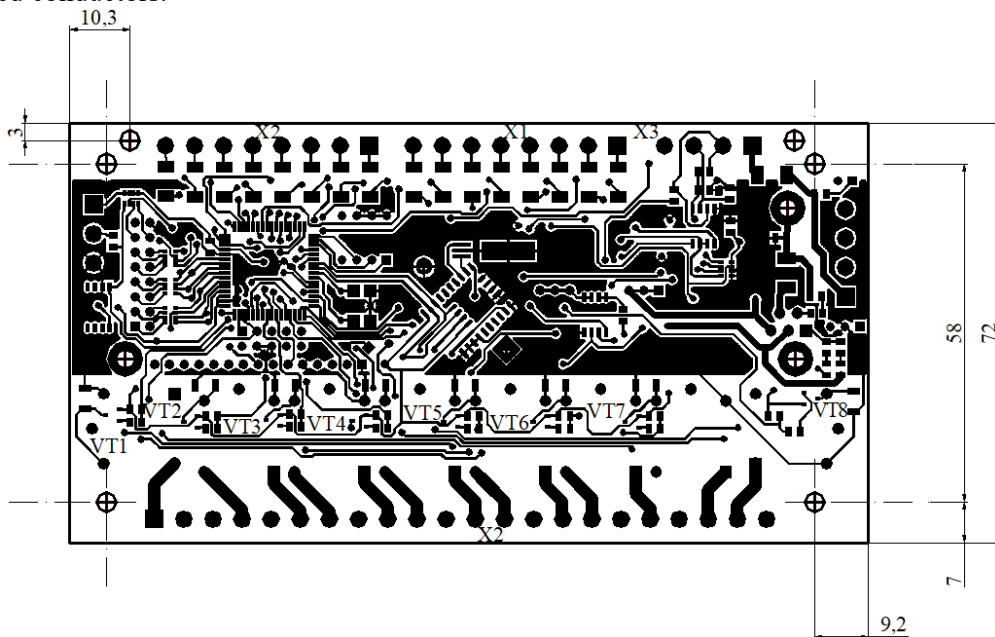
Fig. 5. A wire routing. The fragment. Bottom view
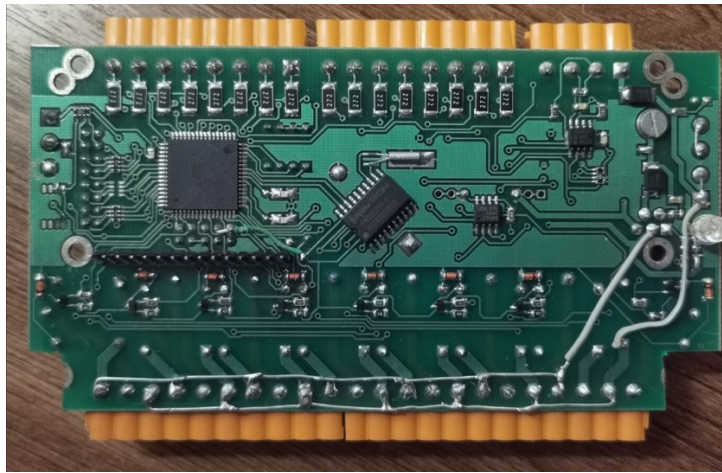
Fig. 6. PCB. Bottom view. The finished product

Development of a software model
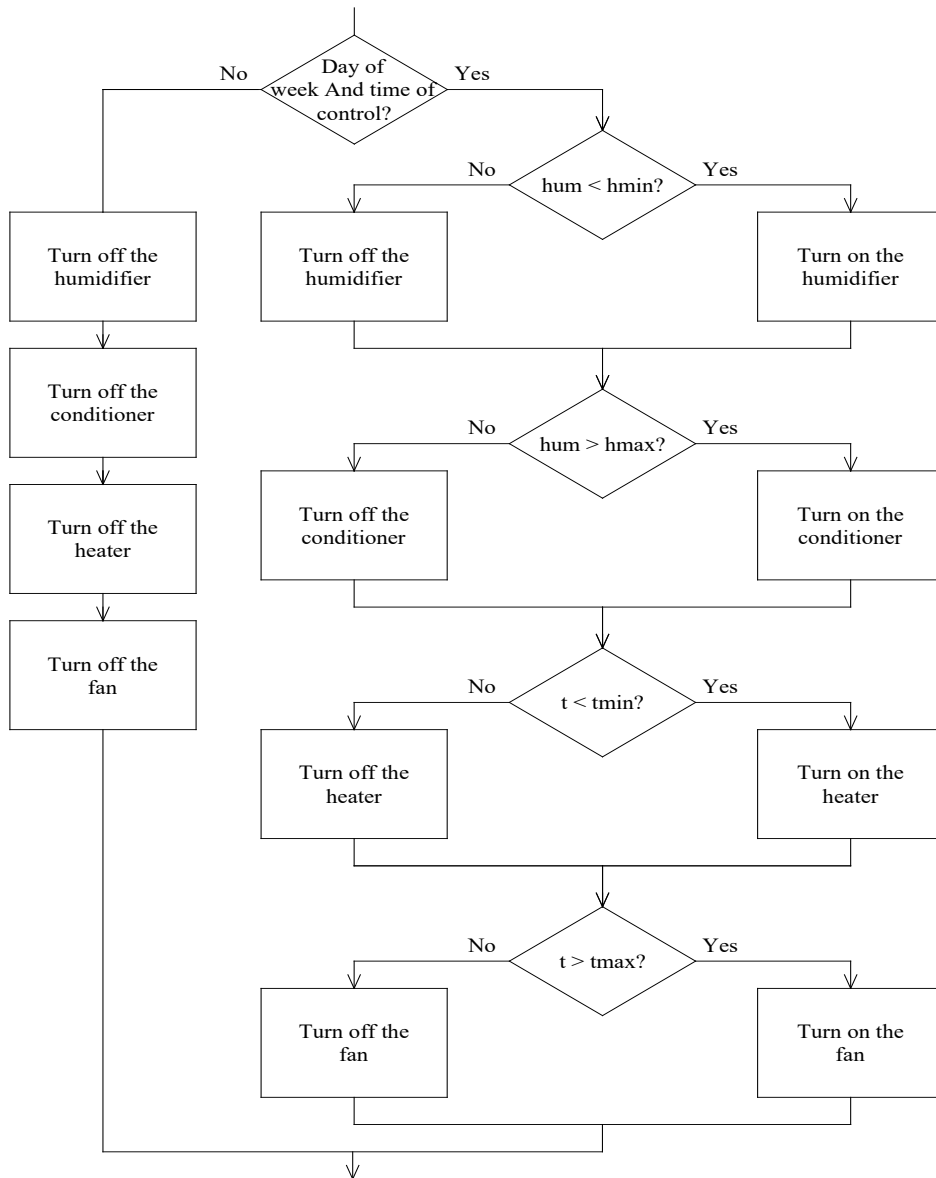
Fig. 7. A flowchart of the algorithm for the operation of the actuation device. The fragment

```
fx  main

204  if ((fonweekday[weekday-1]==1) && (fpoweron)) {
205
206      if (hum<hmin) {humider_on} else {humider_off}
207      if (hum>hmax) {conder_on} else {conder_off}
208      if (temper<tmin) {heater_on} else {heater_off}
209      if (temper>tmax) {cooler_on} else {cooler_off}
210
211      }   else {
212  cooler_off
213  humider_off
214  conder_off
215  heater_off
216  }
217
218                  };
219
220  //-------------------------------------------меню
221      switch (menu)  {
222  #include "menu.c"
223
224  //=================================================   //включение системы по дням недели
225
226  #include "menu_1.c"
227  #include "menu_2.c"
228  #include "menu_3.c"
229  #include "menu_4.c"
230  #include "menu_5.c"
231  #include "menu_6.c"
232  #include "menu_7.c"
233      //=================================================  //установка даты
234  #include "menu_8.c"
235      //=================================================  //установка времени
236  #include "menu_9.c"
237      //=================================================  //установка максимумов и минимумов
238  #include "menu_10.c"
239      //=================================================  //контроль реле
240  #include "menu_20.c"
```

Fig. 8. CodeVisionAVR development environment. Listing of the program code. The fragment

Based on the circuit diagram design of the projected MS, a flowchart of the device is being developed.

To develop a software model, an algorithm for the operation of the device is compiled for further writing the program.

First, the current day of the week and time are determined. If the control is disabled at this time, all actuators will be disabled. If the control is turned on, the current climatic parameters in the room are compared and when they are not within the previously set limits, the corresponding actuator is switched on. If the parameter is within acceptable limits, the actuator is turned off.

The ATmega128 MCU is supported by a full set of software and hardware for design, including: C compilers, macro assemblers, software debuggers/simulators, in-system emulators and evaluation kits [3].

To write a program executed by MCU, the C programming language is used [8], since according to the design of the language, its constructions are closely compared to typical machine instructions and assembly language.

After developing a flowchart of the device's operation algorithm, a program is written.

Atmel's CodeVisionAVR environment is used as a program writing environment [9], since it supports the entire family of AVR microcontrollers, allows you to write a written program to the MCU memory, and also interacts with many popular programmers.

The CodeVisionAVR environment does not have the ability to compile and debug written code. The Olimex AVR-ISP-MK2 in-circuit programmer is used to download the software directly into the MCU [10].

To verify the correct operation of the program, it is necessary to test the device.

To go to the main menu, the user must press the «*» key on the keyboard.
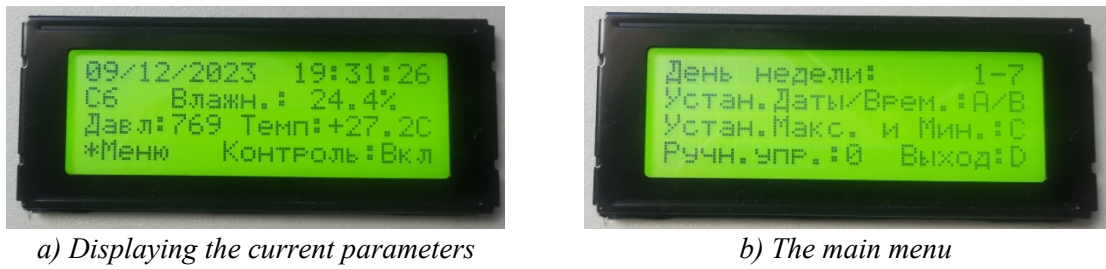
a) Displaying the current parameters

b) The main menu

Fig. 9. Device test results

To proceed to the detailed setting of the control parameters by day of the week, press any key with a number from 1 to 7 for the corresponding day.

To enable or disable climate control for a selected day, the user must press any numeric key. To switch to changing the interval for monitoring the climatic parameters of the current day, the user should press the «*» key. The menu for setting the monitoring time interval will appear. In this menu, the user can set the required time interval for monitoring climatic parameters for the current day of the week. To move through the hours and minutes, use the keys: «*» – forward and «#» – backward. To select a number, use the corresponding numeric key. To go to the main menu, use the "D" key.

The user can also set the current date and time from the main menu. To do this, he will need to press the «A» or «B» keys, respectively. Navigation in this menu is similar to navigation from the previous menu.
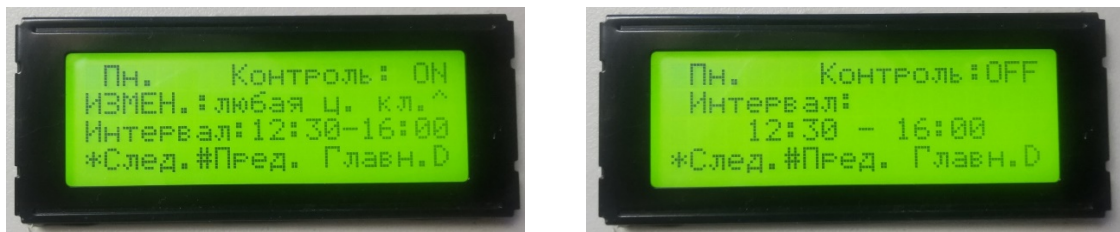
a) Settings menu

b) Control time interval setting menu
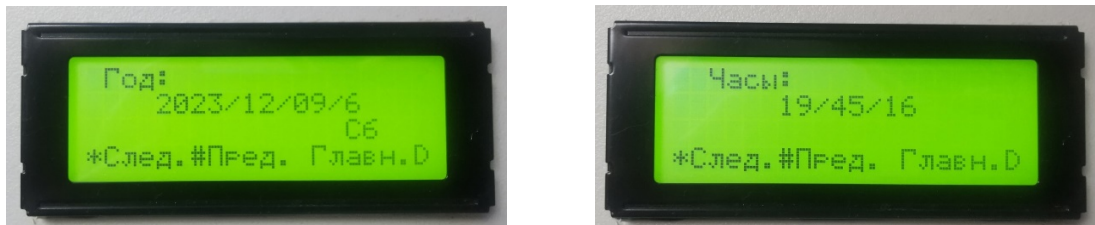
c) Current date setting menu

d) Current time setting menu
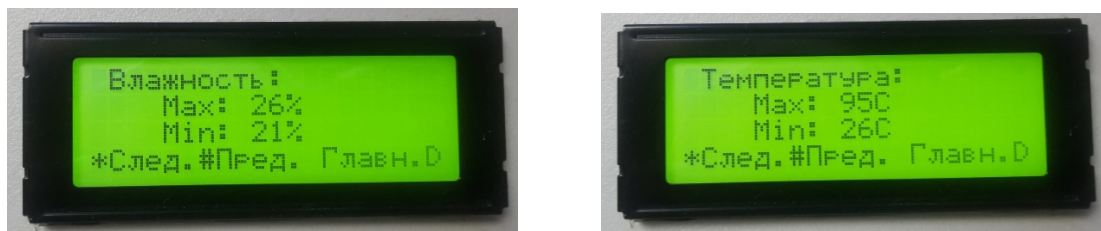
e) The menu for setting the maximum and minimum humidity values

f) The menu for setting the maximum and minimum temperature values

Fig. 10. Device test results

To set the maximum and minimum values of temperature and humidity parameters, the user should press the «C» key from the main menu. Navigation in this menu is similar to navigation from the previous menu.

From the main menu, you can also go to the relay health check menu. To do this, the user must press the «0» key in the main menu.

Fig. 11. Relay health check menu

Conclusion

In this work, a method for constructing an MS for controlling climatic parameters in a room with time functions was considered. Based on the analysis of this method, the block diagram was developed, as well as the circuit diagram and the element base has been selected.

A software model of the MS was developed, an algorithm for executing the program was developed, and its program code was written in C language in the CodeVisionAVR environment. Testing of the device showed that the device operates properly and in accordance with the established requirements.

In the design and technological part, the material and type of printed circuit board were selected, a PCB drawing was drawn up indicating all the seats and with the routing of the conductors. Based on this drawing, an assembly drawing was drawn up with the placement of all elements on the board. The requirements for soldering and assembly of the PCB are specified.

References

1. Government standard 30494–2011 Residential and public buildings. Microclimate parameters for indoor enclosures. URL: https://docs.cntd.ru/document/1200095053?ysclid=lplnkdgroh766714173 (date of request: 30.11.2023).

2. Government standard 2.701–2008 Unified system for design documentation. Diagrams. Kinds and types. General requirements for fulfillment. URL: https://internet-law.ru/gosts/gost/47901/?ysclid=lrns4w56uy633582651 (date of request: 01.12.2023).

3. ATMega128 datasheet. URL: https://static.chipdip.ru/lib/121/DOC004121816.pdf (date of request: 15.09.2023).

4. FM30C256 datasheet. URL: http://datasheet.elcodis.com/pdf2/121/95/1219547/fm30c256s.pdf (date of request: 20.09.2023).

5. Government standard 26246.5-89 Epoxy-impregnated glass cloth foil-clad electrical insulating material of stipulated combustibility for printed circuit boards. URL: https://docs.cntd.ru/document/1200011703?ysclid=lrnsl5nc8e833190834 (date of request: 05.01.2024).

6. Industry standard 4.010.030-81 Installation of mounted elements on printed circuit boards. Construction. URL: https://gostost.com/wp-content/uploads/2017/05/%D0%9E%D0%A1%D0%A2-4.010.030-81-%D1%87%D0%B0%D1%81%D1%82%D1%8C-2.pdf?ysclid=lrnsroi32h878986832 (date of request: 07.01.2024).

7. Government standard 21931-76 Tin-lead solders in the form of products. Specifications. URL: https://stv39.ru/gosts/GOST-21931-76.pdf?ysclid=lrnsvppnoj285030701 (date of request: 07.01.2024).

8. Deitel Harvey. How to program C. URL: https://vk.com/doc7608079_448866001?hash=wXGAxIlNo8Eq1ErMEhVINVavyMaFpoohYREav9BSfzD&dl=Zw8yjOyoriF8oT25M9z2yomSMLXvMjZCx5mkogPrAgL (date of request: 10.08.2023).

9. Lebedev M.B. CodeVisionAVR. A beginner's guide. URL: https://djvu.online/file/YPZ5dusslfZCf?ysclid=lplo7va75h568019595 (date of request: 20.09.2023).

10. AVR-ISP-MK2 datasheet. URL: https://www.olimex.com/Products/AVR/Programmers/AVR-ISP-MK2/resources/AVR-ISP-MK2.pdf (date of request: 03.01.2024).

STUDY OF THE OUTPUT RADIATION CHARACTERISTICS OF A SINGLE-FREQUENCY RING FIBER LASER NEAR THE GENERATION THRESHOLD

G. S. Grishkin[1,2]

[1]*State University of Aerospace Instrumentation, Bolshaya Morskaya str., 67, lit. A, 190000 Saint Petersburg, Russia*
[2]*LLC NordLase, Tallinskoye sh., 206, 198205, Saint Petersburg, Russia*

Abstract: *The study investigated the dependence of the output power of a single-frequency ring fiber laser, operating in continuous mode, on the length of erbium-doped active fiber, which acts as a saturable absorber. The boundaries of stable laser operation mode without signal distortion by beatings were determined. Operating modes of the scheme were explored using different algorithms for initiating the generation process. This allowed optimizing the generation initiation mode to increase the laser output power without sacrificing the quality of the signal's spectral characteristics.*
Keywords: *single-frequency ring fiber laser, erbium-doped active fiber, saturable absorber.*

Composition of the single-frequency ring fiber laser

The narrowband ring fiber laser, which includes a pump diode, fiber resonator, and Peltier element (Fig. 1). The fiber resonator consists of a ring section, which includes a wavelength-division multiplexer, active fiber, radiation coupler, polarization circulator, and a linear section composed of a saturable absorber and fiber Bragg grating [1].
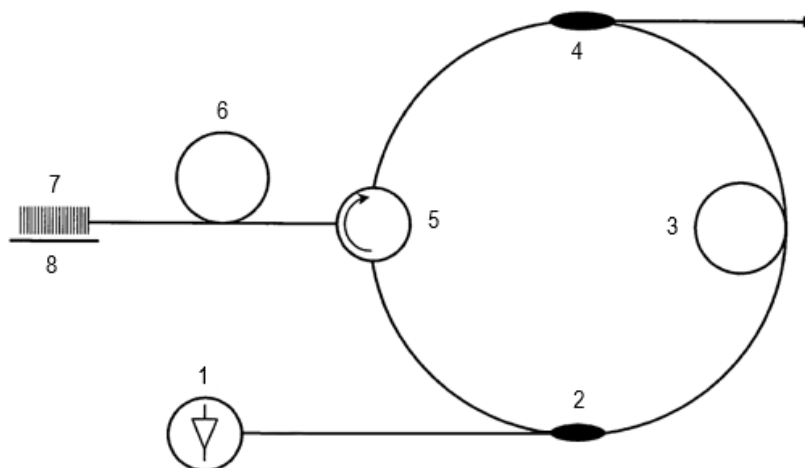
Fig. 1. Scheme of a ring fiber laser
1 – pump diode; 2 – wavelength-division multiplexer; 3 – active fiber; 4 – radiation coupler;
5 – polarization circulator; 6 – saturable absorber (non-pumped active fiber);
7 – fiber Bragg grating; 8 – Peltier element.

This configuration enables high stability operation of the laser at a wavelength of 1550 nm, operating in continuous mode with a spectral width of less than 1 kHz. This achievement is made possible by employing active fiber made of material with a high concentration of doping impurities, allowing the total length of the ring laser resonator to be reduced to 7 meters. Furthermore, this setup remove pump radiation at the laser output by using bidirectional propagation of pump and laser radiation.

The resonator's active medium consists of optical fiber with a high concentration of erbium doping, enabling a short effective length of the resonator's active part (0.5 m). Reducing the total length of the resonator decreases the laser's sensitivity to temperature and vibration. Additionally, it increases the mode spacing, thereby reducing the requirements for the laser's mode filter frequency characteristics and making single-mode generation more stable.

Pumping of the active fiber is achieved through a wavelength-division multiplexer with a single-mode stabilized diode emitting at a wavelength of 976 nm and an output power of up to 600 mW. Unidirectionality of the setup and reduction of polarization mode interaction are achieved by employing a po-

larization circulator. The linear section of the resonator consists of non-pumped active fiber and a blind Bragg grating. Radiation output is facilitated by a 1x2 radiation coupler with a 50:50 splitting ratio. The Bragg grating is installed on a Peltier element, allowing for its thermal stabilization and wavelength tuning if necessary.

The non-pumped active fiber in the linear section of the laser resonator applies as a saturable absorber, where two counter-propagating waves form a narrowband dynamic absorption coefficient grating, effectively selecting and filtering longitudinal modes [2,3]. The bandwidth of such a filter typically ranges from several tens of megahertz and depends on the fiber length, which is determined by the doping level, while the resonant frequency matches to the laser generation wavelength. Moreover, such an absorber provides short-term stabilization of the laser generation wavelength.

The influence of the saturable absorber length on laser operation

In the described laser setup, the initial length of the non-pumped active fiber, L_{sa}, was 6 cm, ensuring an acceptable balance between output power level and stability. This work describes the research results into changes in laser output parameters with varying L_{sa} length.

Increasing the length of L_{sa} to 10 cm led to an increase in the laser generation threshold. By gradually increasing the pump current in 10 mA steps, the minimum output power of the laser radiation, P_{out}, was fixed at 85 mW of pump power, $P_{th\ up}$, instead of 54 mW at L_{sa} length equal to 6 cm. However, upon increasing the pump power further and gradually decreasing the pump current, the minimum pump power at which laser radiation could still be detected, $P_{th\ down}$, was 61 mW (Fig. 2). Afterwards, with an increase in the length of the non-pumped active fiber, the pump power $P_{th\ down}$ increased, reaching 85 mW at $L_{sa} = 20$ cm, while $P_{th\ up}$ was 491 mW.
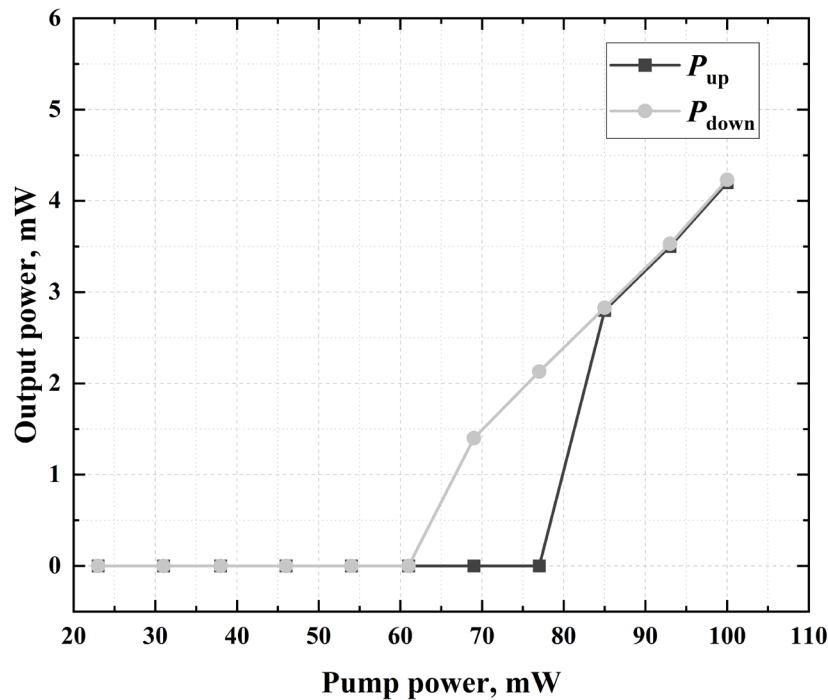
Fig. 2. Dependence of laser generation power on pump power ($L_{sa} = 10$ cm)

With gradual increase in pump power, the quality of the output signal remained unchanged up to a certain level (93 mW for $L_{sa} = 10$ cm). Afterward, periodic beatings, lasting 1-2 seconds and significantly exceeding the signal level in power, were observed on the oscillogram (once a minute). Further increase in pump current led to an increase in the frequency of occurrence of these beatings (Fig. 3).

It is also worth noting that at pump power $P_{th\ down}$, periodic pulses were observed on the oscillogram of the output signal. By adjusting the pump power within a small range near the generation threshold, the frequency of pulse occurrence can be changed (Fig. 4). It is noted that this tuning range increases with the length of the active fiber, which is especially noticeable at $L_{sa} = 20$ cm. The nature and shape of the pulse signals match to relaxation oscillations [4].
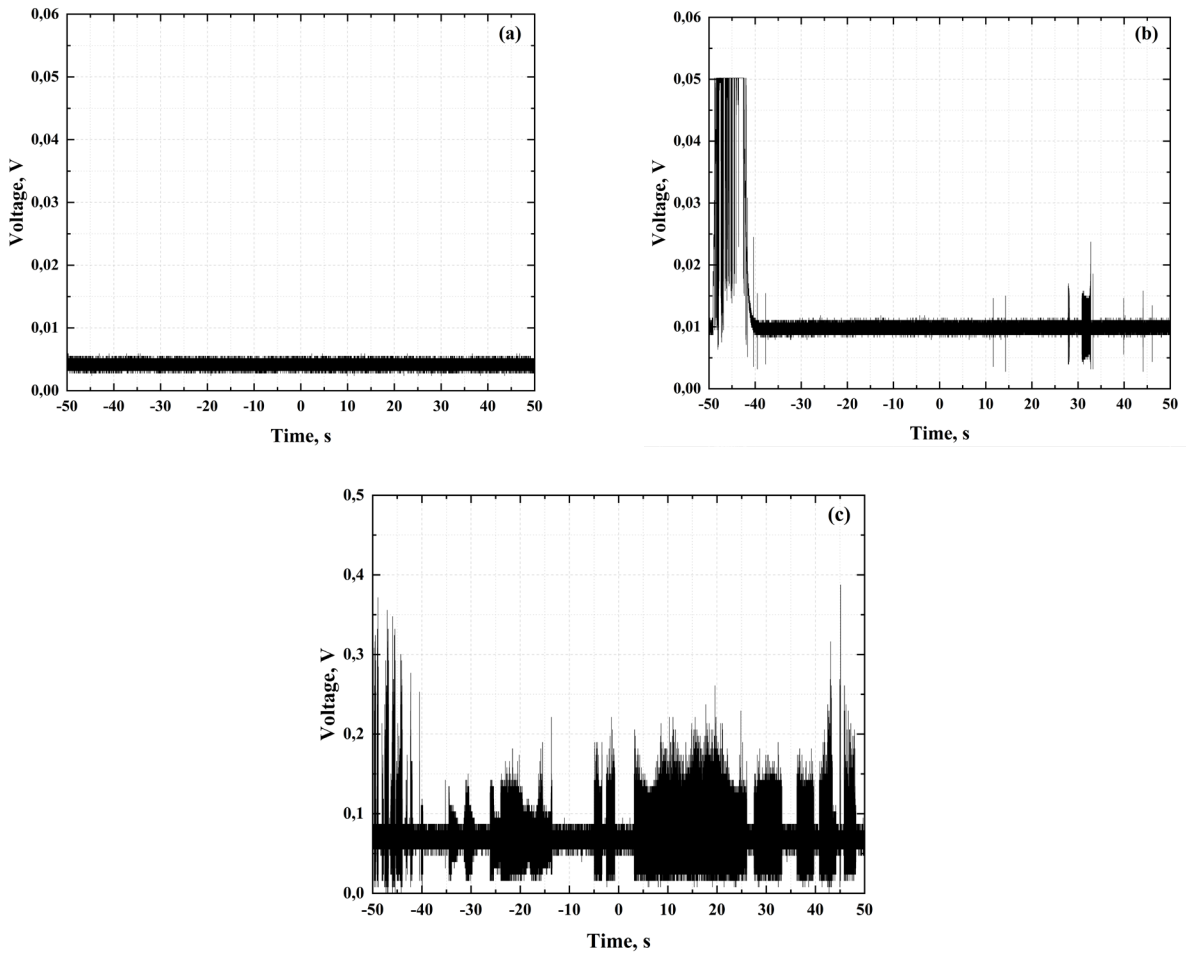
Fig. 3. Oscillograms of the signal at different pump powers ($L_{sa} = 20$ cm)
(a) – 100 mW; (b) – 153 mW; (c) – 491 mW
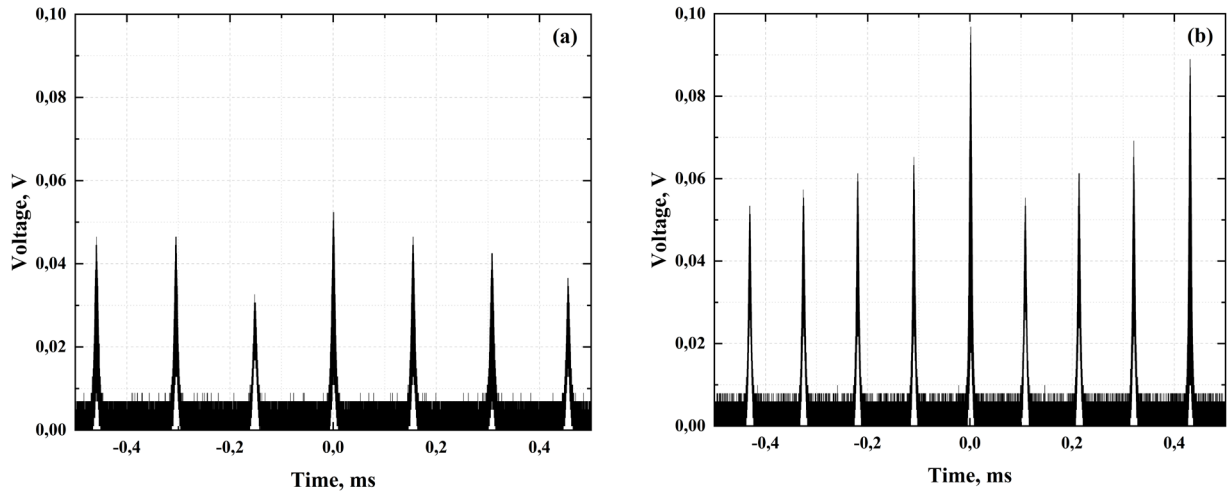
Fig. 4. Oscillograms of pulsed-periodic operation mode ($L_{sa}=20$ cm): (a) – 85 mW; (b) – 93 mW.

Increasing the length of the non-pumped fiber to 50 cm resulted in a significant loss of output signal power. Due to the high concentration of doping impurities, attenuation in the linear part of the laser resonator was sufficiently large, resulting in almost no radiation at the output (measured in tens of micro-watts).

During the research, an inverse dependence between the level of laser output power P_{out} and the length of the non-pumped fiber L_{sa} at constant pump power was established (Fig. 5).
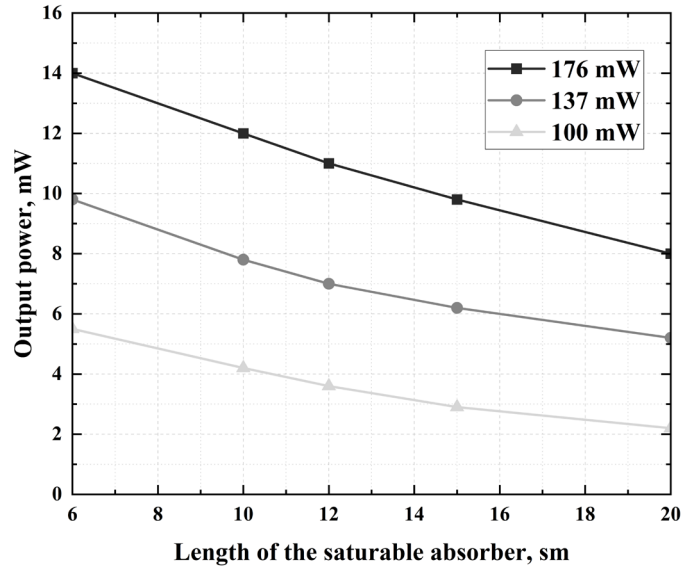
P_{pump}, mW	P_{out}, mW				
	6 sm	10 sm	12 sm	15 sm	20 sm
46					
54	+				
61					
69					
77					
85		+			
93					
100			+		
107					
115					
122					
129					
137					
145					
153					
161					
168				+	
176					

- pulse mode
- without beats
- rare beats
- noisy signal
+ - $P_{th\,up}$
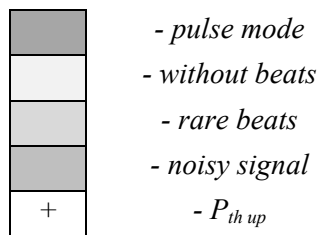
Fig. 5. Dependence of laser output power
on the length Lsa at different pump powers

Upon analyzing the obtained values of output power as the length of the non-pumped fiber increased, it was concluded that the output power of the signal without beatings in the scheme with an L_{sa} length of 20 cm almost doubled compared to the output power in the original scheme with a 6 cm saturable absorber (Fig. 6). Meanwhile, the spectrum of the output signal of the fiber ring laser did not degrade (Fig. 7).
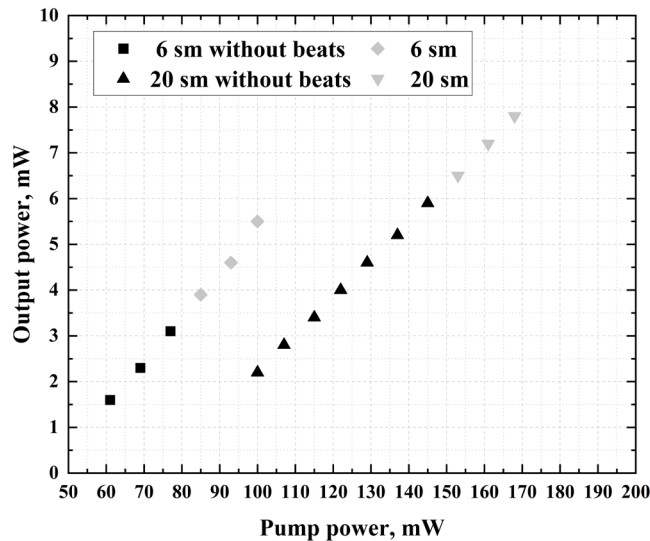
Fig. 6. Dependencies of laser output power on pump power
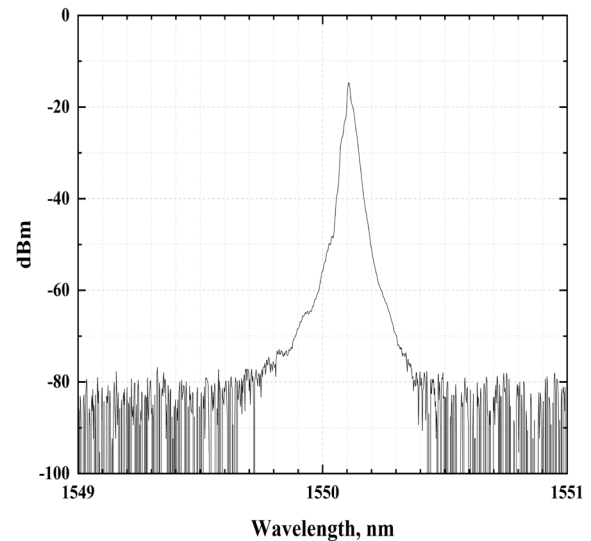
Fig. 7. Spectrum of the signal at the laser output (L_{sa} = 20 cm)

Using a saturable absorber shorter than 6 cm is uselessly, as it leads to a decrease in output power and degradation of the laser's output signal spectrum.

Conclusion

The obtained research results allowed us to determine the boundaries of stable operation mode of the laser at different lengths of the non-pumped active fiber and to select the optimal activation algorithm to increase the output power level without sacrificing signal quality.

References

1. Patent 2554337 Russian Federation, Int. Cl. H01S 3/067, H01S 3/083. Narrowband annular fibre laser / Bochkov A.V., Kolegov A.A., Sofienko G.S., Leshkov A.O.; proprietors: RF, represented by State Corporation "Rosatom", "RFNC-VNIITF named after academician E.I. Zababakhin". – No. 2014100329/28; filed 01/09/2014; published 06/27/2015, Bulletin No. 18 – 6 p.

2. H. Inaba, A. Onae, Y. Akimoto, T. Komukai,"Observation of Acetylene Molecular Absorption Line With Tunable, Single-Frequency, and Mode-Hop-Free Erbium-Doped Fiber Ring Laser"/ QUANTUM ELECTRONICS, VOL.38, NO.10, OCTOBER 2002.

3. Y. Cheng, J.T. Kringlebotn, W.H. Loh, R.I. Laming, and D.N. Payne. Stable single-frequency traveling-wave fiber loop laser with integral saturable-absorber-based tracking narrow-band filter / OPTICS LETTERS. -Vol. 20, No. 8. -1995. –p. 875-877.

4. Lina Ma, Zhengliang Hu, Xun Liang, Zhou Meng, and Yongming Hu. Relaxation oscillation in Er3+-doped and Yb3+/Er3+ co-doped fiber grating lasers / APPLIED OPTICS. -Vol. 49. -No. 10. -2010. –p. 1979-1985.

REVIEW OF CODE SEQUENCES BASED ON HADAMARD DIFFERENCE SETS

Yana Gromysh

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: gromysh03@icloud.com

Abstract. *This paper provides a literature review of methods for constructing code sequences with a single-level periodic autocorrelation function based on Hadamard difference sets. Well-known sequences are considered – Legendre, Jacobi, etc., and new code sequences based on rows of persymmetric quasi-orthogonal matrices are also considered.*

Keywords: *matrices, Hadamard difference sets, code sequences, autocorrelation function.*

Introduction

Currently, discrete-coded signals have gained wide usage in radio technical systems [1], where the encoded parameters change at fixed time intervals. The law governing the variation of the encoded parameter in such signals is defined by a code sequence (CS), which fully determines the signal properties. The application area of such signals with specified correlation and spectral characteristics has significantly expanded from individual tasks of radiolocation, hydrolocation, and navigation [2-3] to issues related to the application of such signals in addressing, information transmission, synchronization, encoding, etc.; their use in neural networks during training stages; in the development of high-resolution and high-precision medical equipment; in computing systems, automation systems, and telematics systems for protection against unauthorized access, ensuring high noise immunity, cryptostability, electromagnetic compatibility, etc [5-6].

Two-level code sequences with single-level periodic autocorrelation function (PACF), constructed based on cyclic difference sets of the Hadamard type, are of particular interest [7].

The aim of this work is to conduct a review of known methods for constructing sequences of this class.

HADMARD MATRICES AND THE THEORY OF DIFFERENCE SETS

Hadamard matrix is a matrix of order n $\mathbf{H}_n$ with two levels of elements $\{+1,-1\}$ and possessing the property

$$\mathbf{H}_n \mathbf{H}_n^{\mathrm{T}} = n\mathbf{I}.$$

Adamar matrices exist for orders of n=4t, where t is a natural number. However, there are problematic orders where Adamar matrices have not yet been found, for example, 668, 716, 892, 1004, 1132, 1244, 1388, 1436, 1676, 1772, 1916, 1948, 1964, etc.

Ryser's hypothesis [8] restricts the existence of cyclic Adamar matrices. Only one cyclic matrix of order 4 is known.

$$\mathbf{H}_4 = \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

However, a large number of Hadamard matrices are known, consisting of a cyclic submatrix of order $(n-1) \times (n-1)$, with a border of elements "1". As an example, let's consider the portrait of an 8th-order Hadamard matrix, where a white square corresponds to an element with a value of "1", and a black square corresponds to an element with a value of "-1".

Such matrices in foreign scientific literature began to be called cyclic Hadamard matrices [9], although formally they are not such. More accurately, they are Hadamard matrices of the core-with border construction [10].
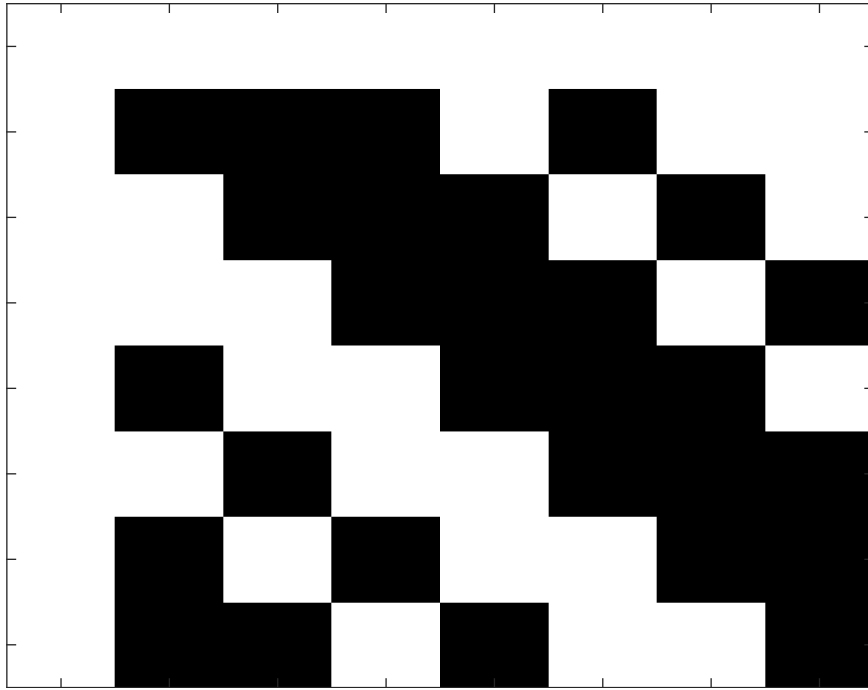
Fig. 1. The portrait of an 8th-order Hadamard matrix.

The theory of their construction is related to the theory of difference sets

$$D(v,k,\lambda) = \{d_1, d_2, ..., d_k\},$$

where d_i – the elements of the difference set, v,k,λ – the parameters of the set [7].

The column indices in the cyclic submatrix can form difference sets – Adamar difference sets, with parameters.

$$(v,k,\lambda) = \{4t - 1_1, 2t - 1, t - 1\}.$$

Golomb, S., in the work [11], notes the correlation between cyclic difference sets of the Hadamard type and binary sequences with one-level periodic autocorrelation function (PAKF). It is observed by them that quadratic residue sequences [12] exhibit a two-level PAKF. Meanwhile, the work [13] sets out the task of determining all constructions that form Hadamard difference sets.

REVIEW OF ADAMAR DIFFERENCE SETS

In the work [11], Raymond Paley discovered a new method for constructing Hadamard matrices using a core with border. The rows of this core can be used as sequences with minimal sidelobe levels in the periodic autocorrelation function. Subsequently, such sequences were named quadratic residue sequences.

In work [14], Zinger difference sets were obtained, in the particular case of the finite field characteristic $p = 2$, these sets lead to M-sequences of length $N = 2^m - 1$.

In work [15], difference sets of Adamar were constructed with parameter $v = 4b^2 + 27 = p$, k – natural number (Hall sequences), and in work [16] with parameter $v = p_1 \cdot p_2$, where $p_2 = p_1 + 2$, p_1 and p_2 – are twin primes (Jacobi sequences).

From the work [17], GMW-difference sets are known, with parameter $v = 2^n - 1$ (Gordon-Mills-Welch (GMW) sequences of four types:
- GMW sequences;
- cascade GMW sequences;
- generalized GMW sequences of type I;
- generalized GMW sequences type II.
Until 1971, only the sequences listed above were known.

In the works [18-21], sequences were obtained that could not be attributed to any of the known constructions for lengths N = 255, 511, 1023.

The intensive search for constructions to build the discovered but unexplained sequences with single-level autocorrelation function began in the early 1990s. During the period from 1997 to 2004, it was possible to systematize and propose efficient constructions for all previously unexplained sequences of length $N = 2^m - 1$, obtained for cases where $m = 7, 8, 9, 10$. The following results were obtained in the development of constructions for sequences of length $N = 2^m - 1$ with a single-level autocorrelation function [22]:

- periodic 3-term sequences of length $N = 2^m - 1$ for odd $m = 2t + 1$, where $m \geq 5$;

- periodic 5-term sequences of length $N = 2^m - 1$ for $m \pmod 3 \neq 0$, where $m \geq 7$;

- WG-sequences of length $N = 2^m - 1$, where $m \geq 7$, obtained by WG-transformations of periodic 5-term sequences.

In the works [23-25], it is shown that Serge [24] and Glynn [25] hyperovals of the first and second types generate cyclic difference sets of Adamar parameters $\left(v = 2^n - 1, k = 2^{n-1} - 1, \lambda = 2^{n-2} - 1 \right)$. In parallel with the study of hyperoval sequences in the work [26], a representation of WG sequences by Kasami power functions was found, and it was demonstrated by modeling that such a construction has a one-level periodic ACF for $m \leq 23$.

In 2002, new sequences with a one-level ACF [27] were obtained using the decimation transformation of Hadamard.

In 2005, the monograph [28] reviewed all known rules at that time for constructing binary sequences based on cyclic difference sets of Adamar.

In 2011 [29], the theory of quasi-orthogonal matrices generalizing Hadamard matrices [30] began to develop. Based on the rows of these matrices, in the work of 2018 [31] it was possible to obtain new code sequences of length 3,7 and 11 with a single-level periodic autocorrelation function and improved properties of the aperiodic autocorrelation function, and in the works of 2020 [32,33] strategies for obtaining such codes whose length is in the order of 4t-1:

- 4t-1=p;
- 4t-1=$p_1 p_2$;
- 4t-1=2^n-1.

Conclusion

The main focus of research in this area is currently on refining the properties of known code sequences, expanding the capacity of established coding rules, and identifying new applications for code sequences.

The conducted review leads to the conclusion that the quest for new code sequences with single-level autocorrelation functions is intertwined with the search for Hadamard matrices and their generalizations. It is advisable to concentrate efforts on methods for obtaining these matrices themselves. These findings are expected to have a lasting impact, particularly in practical applications such as code sequence synthesis.

Acknowledgment

References

1. Gantmakher, V. E. Noise-Like Signals: Analysis, Synthesis, Processing: [monograph] / V. E. Gantmakher, N. E. Bystrov, D. V. Chebotarev; V. E. Gantmakher, N. E. Bystrov, D. V. Chebotarev. – Saint Petersburg: Science and Technology, 2005. – 396 p.

2. Sergeev, M. B. Nested code constructions of Barker-Mersenne-Raghavarao / M. B. Sergeev, V. A. Nenashev, A. M. Sergeev // Information and Control Systems. – 2019. – No. 3(100). – P. 71-81. – DOI 10.31799/1684-8853-2019-3-71-81.

3. Zhuk, A. P. Method of obtaining code sequences for prospective secure satellite radio navigation systems / A. P. Zhuk, E. P. Zhuk // Science-intensive technologies in space research of the Earth. – 2012. – Vol. 4, No. 2. – P. 26-28.

4. Grigoriev, E. K. Analysis of correlation characteristics of new code sequences based on persymmetric quasi-orthogonal circulants / E. K. Grigoryev // Proceedings of communication educational institutions. – 2022. – Vol. 8, No. 2. – P. 83-90. – DOI 10.31854/1813-324X-2022-8-2-83-90.

5. Grigoriev, E. K. Method of protective data encoding obtained by optical sensors of unmanned aviation systems / E. K. Grigoryev, A. M. Sergeyev // Proceedings of MAI. – 2023. – No. 133.

6. Sergeev M. B., Tatarnikova T. M., Sergeev A. M., Bozhenko V. V. Method of ensuring data confidentiality using orthogonal matrices. Engineering Herald of the Don. 2024. No. 2. P. 1-9.

7. Leukhin, A. N. Construction of cyclic difference sets of Adamar / A. N. Leukhin // Mathematical methods of pattern recognition. – 2009. – Vol. 14, No. 1. – P. 395-398.

8. Ryser H. J. Combinatorial mathematics. The carus mathematical monographs // The mathematical association of America. New York. John Wiley and Sons. 1963. No. 14. 162 p.

9. Golomb S.W., Gong G. Signal design for good correlation for wireless communication, cryptography, and radar. // Cambridge Univ. Press, 2006.

10. Balonin, N. A. Special matrices: pseudo-inverse, orthogonal, Hadamard, and Kronecker / N. A. Balonin, M. B. Sergeev. – Saint Petersburg: Polytechnic, 2019. – 196 pages.

11. Golomb S.W. Remarks on orthogonal sequences // The Glenn L. Martin Company, Baltimore, MD, July 28, 1954.

12. Paley R. E. A. C. On orthogonal matrices // J. of mathematics and physics. 1933. Vol. 12. P. 311-320.

13. Golomb S.W. Sequences with randomness properties. Baltimore, Glenn L. Martin Company, 1955.

14. Singer J. A theorem in finite projective geometry and some applications to number theory // Trans. Amer. Math. Soc. – 1938. – Vol. 43. – PP. 377-385.

15. Hall M. Survey of difference sets // Proc. Am. Math. Soc. – 1956. – Vol. 7. – PP. 975-986.

16. Stanton R.G., Sprott D.A. // A family of difference sets. Canadian J. Math. – 1958. – Vol. 10. – PP. 73-77.

17. Scholtz R.A., Welch L.R. GMW sequences // IEEE Trans. Inform. Theory. – 1984. – Vol. IT-30, No. 3. – PP. 548-553.

18. Baumert L.D. Cyclic difference sets, lecture notes in mathematics. – Berlin: Springer-Verlag, 1971. – 182 p.

19. Cheng U.J. Exhaustive construction of (255, 127, 63) – cyclic difference sets // Journal of Combinatorial Theory. – 1983. – Vol. A-35. – PP. 115-125.

20. Dreier R.B., Smith K.W. Exhaustive determination of (511, 255, 127) cyclic difference sets // Manuscript. – 1991.

21. Gaal P., Golomb S.W. Exhaustive determination of (1023, 511, 255) – cyclic difference sets // Mathematics of Computation. – 2001. – Vol. 70. – PP. 357-366.

22. No J.S. et al. A new family of binary pseudo-random sequences having optimal periodic correlation properties and larger linear span / J.S. No, S.W. Golomb, G. Gong, H.K. Lee, P. Gaal // IEEE Trans. Inform. Theory. – 1998. – Vol. 35. – PP. 371-379.

23. Maschietti A. Difference sets and hyperovals // Designs, Codes and Cryptography. – 1998. – Vol. 14. – PP. 89-98.

24. Segre B. Ovals in finite projective plane // Canadian J. Math. – 1955. – Vol. 7. – PP. 414-416.

25. Glynn D.G. Two new sequences of ovals in finite desarguesian planes of even order // Lecture Notes in Mathematics. – 1983. – Vol. 1036. – PP. 217-229.

26. No J.S., Chung H.B., Yun M.S. Binary pseudorandom sequences of period   with ideal autocorrelation generated by the polynomial   // IEEE Trans. Inform. Theory. – 1998. – Vol. 44. – PP. 1278-1282.

27. Gong G., Golomb S.W. The decimation-Hadamard transform of two-level autocorrelation sequences // IEEE Trans. Inform. Theory. – 2002. – Vol. 48. – PP. 853-865.

28. Golomb S.W., Gong G. Signal design for good correlation for wireless communication, cryptography, and radar. – New York: Cambridge University Press, 2005-438 p.

29. Balonin, N. A. M-matrices / N. A. Balonin, M. B. Sergeyev // Information and Control Systems. – 2011. – No. 1(50). – P. 14-21.

30. Balonin, N. A. Matrices of local maximum of determinant / N. A. Balonin, M. B. Sergeyev // Information and Control Systems. – 2014. – No. 1(68). – P. 2-15.

31. Nenashev V.A., Sergeev A.M., Kapranova E.A. Research and Analysis of Autocorrelation Functions of Code Sequences Formed on the Basis of Monocyclic Quasi-Orhtogonal Matrices // Information and Control Systems. – 2018. – No. 4. – P. 9-14.

32. Search and modification of code sequences based on persymmetric quasiorthogonal circulants / E. K. Grigoriev, V. A. Nenashev, A. M. Sergeyev, E. V. Samokhina // Telecommunications. – 2020. – No. 10. – P. 27-33.

ALGORITHM OF MATHEMATICAL MODELING OF STOCHASTIC SPATIAL FIELD WITH GIVEN CORRELATION CHARACTERISTICS

Ilya Ivanov

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia,
E-mail: *b-boy431260@yandex.ru*

Abstract. *This work presents an algorithm for modeling an anisotropic non-Gaussian Markov field, which allows taking into account the laws of the modeled distribution and the correlation and spectral characteristics of the simulated sequences. Such algorithms are widely used in image processing, as well as in the processing of information signals in multichannel complex automatic control systems.*

Keywords: *modeling algorithm, anisotropic field, Markov process, non-Gaussian field, correlation and spectral characteristics, generating field, underlying surface.*

When designing complex systems, in particular automation systems, the method of mathematical modeling is widely used, which allows at all stages of the life cycle of the system to assess its operability and efficiency in various operating conditions[1]. One of the most important tasks in the implementation of mathematical modeling is the task of synthesizing algorithms for simulating input influences on the system, both informational influences and interfering ones[2]. As mathematical models of such influences most often use random processes, which in the case of multichannel systems are vector, i.e. represent random vectors or matrices.

In the case of a vector random process, one of the coordinates represents time. If we replace the time coordinate by a spatial coordinate, we obtain a random matrix, which can be interpreted as a random field. It is this case that is considered in this paper. So, let us synthesize an algorithm for modeling a random field whose spatial coordinates we denote as X and Y, and the values of the field at a particular point (x,y) we denote by Z. In these notations $z=\varphi(x,y)$, that is, z is some function of (x,y). In particular, it can be the brightness in photo-image processing or the power of a locational signal reflected by some surface, for example, a sea surface or a land surface[3,4].

Let the modeled field be defined by boundaries on the coordinate x $\left(x_{\min} \le x \le x_{\max}\right)$ and y coordinate $\left(y_{\min} \le y \le y_{\max}\right)$. Let us divide, respectively, the sides of the rectangle defining the modeled field into n and m segments. Without limiting the generality of the consideration, we will consider the segments for each coordinate to be equal – for the x coordinate, the size of the segment is equal to $\Delta x = \left(x_{\max} - x_{\min}\right)/n$, and on the y-coordinate – $\Delta y = \left(y_{\max} - y_{\min}\right)/m$. In fact, we have discretized the simulated field (Fig. 1) by dividing it into n and m rectangular regions. The vector of the centers of the regions along the X axis is equal to

$$\mathbf{X}_n = \left(x_1, x_2, \ldots x_i, \ldots x_n\right)^{\mathrm{T}}, \, x_i = x_{\min} + \left(i - \frac{1}{2}\right)\Delta x, \, i = 1, 2, \ldots n, \tag{1}$$

where T is the transpose sign. Similarly, the vector of area centers along the Y-axis is equal to

$$\mathbf{Y}_m = \left(y_1, y_2, \ldots y_j \ldots y_m\right)^{\mathrm{T}}, \, y_j = y_{\min} + \left(j - \frac{1}{2}\right)\Delta y, \, j = 1, 2, \ldots m. \tag{2}$$

Taking into account expressions (1) and (2), the modeled field is represented by *nm* samples $z_{ij}=\varphi(x_i,y_j)$, $i=1,2,\ldots n$, $j=1,2,\ldots m$, which, in turn, can be represented as a matrix $\mathbf{Z}_{n,m}$

$$\mathbf{Z}_{n,m} = \begin{Vmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,m} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n,1} & z_{n,2} & \cdots & z_{n,m} \end{Vmatrix}. \tag{3}$$

The columns of the matrix $\mathbf{Z}_{n,m}$ physically represent the field changes along the X coordinate, and the rows represent the field changes along the Y coordinate. The field anisotropy in this case is represent-

ed by various correlation relations between the column elements determined by the field properties in the X coordinate and correlation relations between the row elements determined by the field properties in the Y coordinate.

Fig. 1. Matrix formation of the modeled field

In this way, the representation of the field in the form of the matrix $\mathbf{Z}_{n,m}$ allows us to reduce the problem of modeling the field to the problem of modeling the elements of the matrix $\mathbf{Z}_{n,m}$. In this case, the representation of the field as a matrix allows us to reproduce not only the law of probability distribution of field samples, but also to take into account various correlation and spectral characteristics of the field when observing it from different angles.

In this work, we limit ourselves to the consideration of a Markov field, i.e., a field whose sequence of samples at each coordinate represents a Markov process. With this restriction, to define the algorithm for modeling the field samples, it is sufficient to specify only the distribution law of the field samples and the correlation coefficients for each of the coordinates – for the X coordinate and for the Y coordinate.

When modeling Markov processes, one can use a modeling algorithm based on the conditional distribution of samples, which in this case considered in this paper is defined as

$$f\left(x_{i+1}/x_i,x_{i-1},\ldots x_1\right)=f\left(x_{i+1}/x_i\right)=\frac{f\left(x_i,x_{i+1}\right)}{\int\limits_{-\infty}^{+\infty}f\left(x_i,x_{i+1}\right)dx_{i+1}}=\frac{f\left(x_i,x_{i+1}\right)}{f\left(x_i\right)},\ i=1,2,\ldots n-1, \qquad (4)$$

there $f\left(x_i\right)$, $f\left(x_i,x_{i+1}\right)$, $f\left(x_{i+1}/x_i\right)$ – are one-dimensional, two-dimensional and conditional distribution densities of samples of the Markov process. A similar expression is obtained for the coordinate axis Y

$$f\left(y_{j+1}/y_j,y_{j-1},\ldots y_1\right)=f\left(y_{j+1}/y_j\right)=\frac{f\left(y_j,y_{j+1}\right)}{\int\limits_{-\infty}^{+\infty}f\left(y_j,y_{j+1}\right)dy_{j+1}}=\frac{f\left(y_j,y_{j+1}\right)}{f\left(y_j\right)},\ j=1,2,\ldots m-1. \quad (5)$$

For a Markov field, expressions (4) and (5) actually solve the problem of synthesizing an algorithm for modeling the field, that is, an algorithm for modeling the elements of the matrix $\mathbf{Z}_{n,m}$, which is the mathematical model of the field.

In particular, for a Gaussian field, this algorithm can be represented as follows.

The matrix $\mathbf{Z}_{n,m}$ can be rewritten as

$$\mathbf{Z}_{n,m}=\begin{Vmatrix}z_{1,1} & z_{1,2} & \cdots & z_{1,m}\\ z_{2,1} & z_{2,2} & \cdots & z_{2,m}\\ \vdots & \vdots & \ddots & \vdots\\ z_{n,1} & z_{n,2} & \cdots & z_{n,m}\end{Vmatrix}=\mathbf{X}_{1,m}=\left\|\mathbf{X}_1,\mathbf{X}_2,\ldots\mathbf{X}_m\right\|=\mathbf{Y}_{n,1}=\begin{Vmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\\ \vdots\\ \mathbf{Y}_n\end{Vmatrix}, \qquad (6)$$

where $\mathbf{X}_{1,m}$ is the matrix whose elements are the columns of the matrix $\mathbf{Z}_{n,m}$, that is, the vectors $\mathbf{X}_j=(z_{1,j}, z_{2,j},\dots z_{n,j},)$, $j=1,2,\dots n$, and $\mathbf{Y}_{n,1}$ is the matrix whose elements are the rows of the matrix $\mathbf{Z}_{n,m}$, that is, the vectors $\mathbf{Y}_i=(z_{i,1}, z_{i,2},\dots z_{i,m},)$, $i=1,2,\dots m$.

Vectors $\mathbf{X}_j$ determine the correlation properties of the field along the X axis. Let the correlation coefficient between two neighboring elements of vector $\mathbf{X}_j$ be r_x, then the matrix $\mathbf{R}_x$ of vector correlation coefficients is equal to

$$\mathbf{R}_x = \begin{Vmatrix} 1 & r_x^1 & r_x^2 & \dots & r_x^n \\ r_x^1 & 1 & r_x^1 & \dots & r_x^{n-1} \\ r_x^2 & r_x^1 & 1 & \dots & r_x^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_x^n & r_x^{n-1} & r_x^{n-2} & \dots & 1 \end{Vmatrix}. \tag{7}$$

For vectors $\mathbf{Y}_i$, defining the field properties along the Y axis, similarly, we have the matrix of correlation coefficients

$$\mathbf{R}_y = \begin{Vmatrix} 1 & r_y^1 & r_y^2 & \dots & r_y^m \\ r_y^1 & 1 & r_y^1 & \dots & r_y^{m-1} \\ r_y^2 & r_y^1 & 1 & \dots & r_y^{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_y^m & r_y^{m-1} & r_y^{m-2} & \dots & 1 \end{Vmatrix}. \tag{8}$$

where r_y is the correlation coefficient between neighboring field elements along the y-axis.

Let us fill the matrix $\mathbf{Z}_{n,m}^{(0)}$ with random variables $\eta_{i,j}$, distributed according to the normal law with zero mean and unit variance

$$f_{\eta_{i,j}}(x) = f_\eta(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right), \; i=1,2,\dots n, \; j=1,2,\dots m. \tag{9}$$

$$\mathbf{Z}_{n,m}^{(0)} = \begin{Vmatrix} \eta_{1,1} & \eta_{1,2} & \dots & \eta_{1,m} \\ \eta_{2,1} & \eta_{2,2} & \dots & \eta_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{n,1} & \eta_{n,2} & \dots & \eta_{n,m} \end{Vmatrix}, \tag{10}$$

The random variables $\eta_{i,j}$ are a set of jointly independent, identically distributed random variables.

Then we transform the matrix $\mathbf{Z}_{n,m}^{(0)}$ into the matrix $\mathbf{Z}_{n,m}^{(x)}$, whose $u_{i,j}^{(x)}$ elements are defined by the expressions

$$\begin{cases} u_{1,j} = \eta_{1,j}, \\ u_{i,j} = r_x \cdot u_{i-1,j} + \sqrt{1-r_x^2}\,\eta_{i,j}, \; i=2,3,\dots n, \end{cases} \tag{11}$$

$j=1,2,\dots m$. This transformation provides correlation of column elements of the matrix $\mathbf{Z}_{n,m}^{(x)}$

$$\mathbf{Z}_{n,m}^{(x)} = \begin{Vmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,m} \\ u_{2,1} & u_{2,2} & \dots & u_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n,1} & u_{n,2} & \dots & u_{n,m} \end{Vmatrix}, \tag{12}$$

and the rows of the matrix are independent.

Let's transform the matrix $\mathbf{Z}_{n,m}^{(x)}$ into the matrix $\mathbf{Z}_{n,m}$, whose $z_{i,j}$ elements are defined by the expressions

$$\begin{cases} z_{1,j} = u_{1,j}, \\ z_{i,j} = r_y \cdot z_{i,j-1} + \sqrt{1-r_y^2}\, u_{i,j}, \quad j = 2,3,\dots m, \end{cases} \qquad (13)$$

$i=1,2,\dots n$. This transformation provides correlation of matrix row elements $\mathbf{Z}_{n,m}$

$$\mathbf{Z}_{n,m}^{(x)} = \begin{Vmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,m} \\ u_{2,1} & u_{2,2} & \dots & u_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n,1} & u_{n,2} & \dots & u_{n,m} \end{Vmatrix}. \qquad (14)$$

The generated matrix $\mathbf{Z}_{n,m}$ is a matrix that represents a mathematical model of an anisotropic Markov field [5]. This matrix can be used both directly and as a generating matrix in modeling non-Gaussian Markov fields in the case when a nonlinear functional transformation is used to change the law of distribution of field fluctuations, since at such a transformation the "Markovian" property of the field is preserved [6,7].

Below, Fig. 2 shows a 3D diagram of a homogeneous isotropic normal field formed by elements of the matrix $\mathbf{Z}_{n,m}^{(0)}$, used as a "generating" field to form samples of an anisotropic Markov field with given correlation characteristics, the 3D diagram of which is shown in Fig. 3. The modeling was performed for a normal Markov field with mean equal to zero, mathematical expectations equal to one, and correlation coefficients r_x and r_y on X and Y axes, respectively, equal to 0.9.
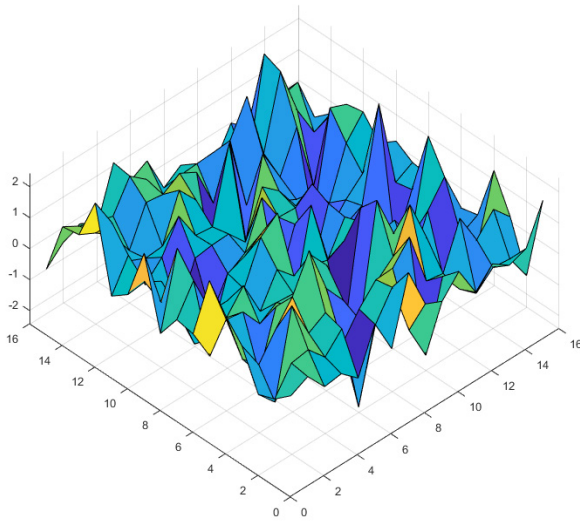
Fig. 2. Homogeneous normal field –
matrix $\mathbf{Z}_{n,m}^{(0)}$
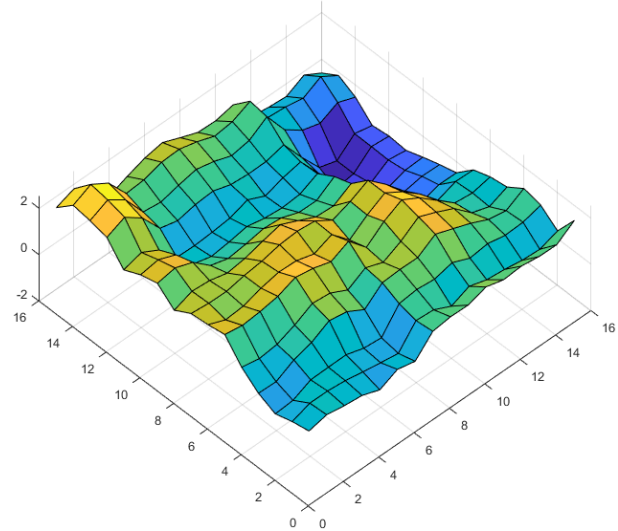
Fig. 3. Anisotropic normal Markov field –
matrix $\mathbf{Z}_{n,m}^{(x)}$

Fig. 4 shows a block diagram of the algorithm for modeling an anisotropic Markov field described by the matrix $\mathbf{Z}_{n,m}^{(x)}$.
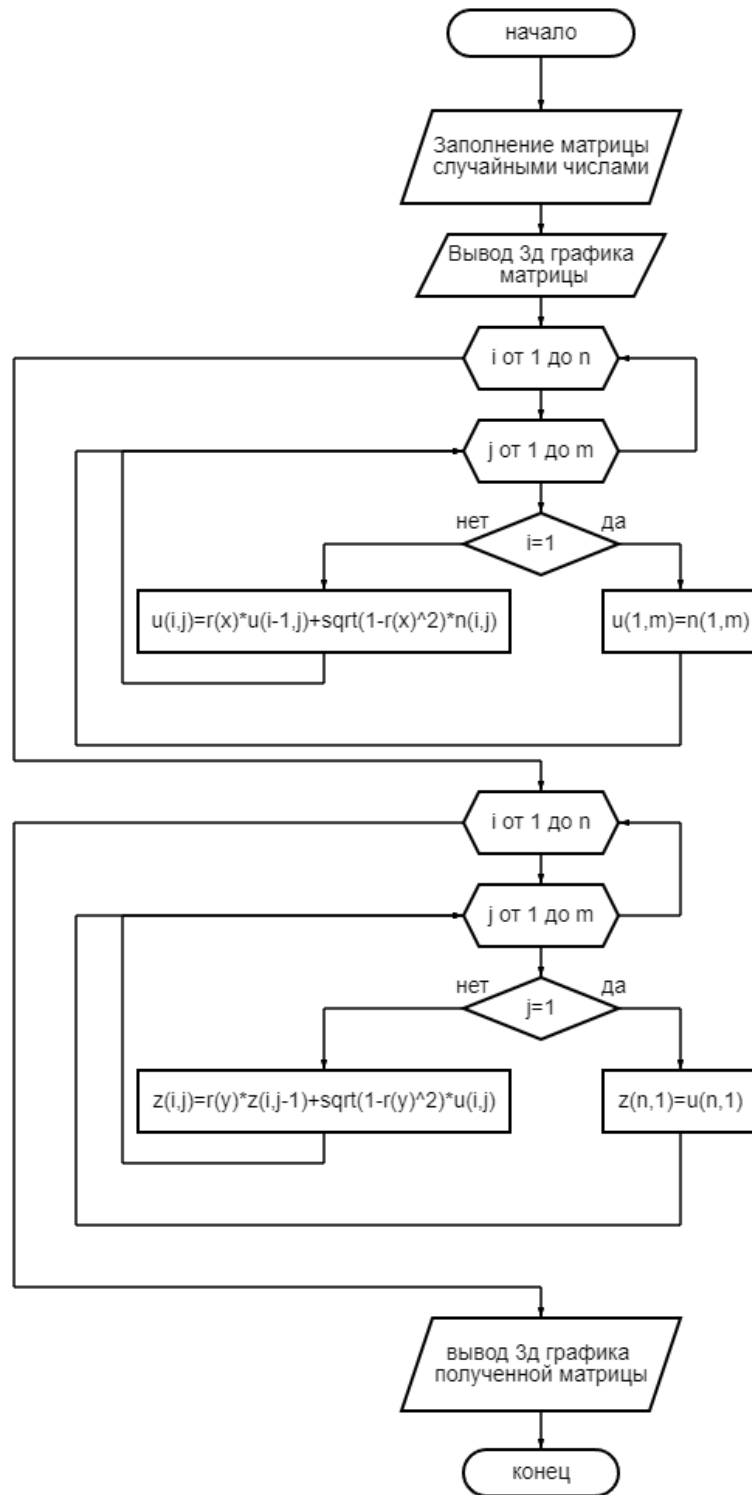
Fig.4. Block diagram of the algorithm for modeling anisotropic Markov field

References

1. Izrantsev V.V., Shepeta D.A. Input signal simulation for the airborne instrument systems of the 5-th generation aircrafts. *Nauchnoe priborostroenie* [Scientific instrumentation], 2000, vol.10, no. 2, pp. 14-19. (in Russian)

2. Shepeta D.A. Development of mathematical models and synthesis algorithms for modeling input signals of onboard information processing and control systems.. *Candidate's thesis*. Saint Petersburg, 2000, 170p. (in Russian)

3. Blaunshtejn N.Sh., Sergeev M.B., Shepeta A.P. *Prikladny`e aspekty` e`lektrodinamiki* [Applied aspects of electrodynamics]. Saint Petersburg, Agraf+, 2016, 272 p. (in Russian)

4. Isakov V.I., Shepeta D.A. Simulation of Radar Signals Reflected from the Edge of Land and Sea. *Informacionno-upravlyayushhie sistemy`* [Information and control systems]. 2017, no. 5 (90). pp. 89-94. (in Russian)

5. Shepeta A.P.. Synthesis of nonlinear shaping filters for modeling input signals of location systems. *Trudy` mezhdunarodnoj nauchno-texnicheskoj konferencii / AN Ukrainy`; NPO Kvant* [Proceedings of the International Scientific and Technical Conference , Academy of Sciences of Ukraine; NPO Kvant]. Kiev, 1994, Vol. 1, pp.81-85. (in Russian)

6. Shepeta D.A., Bozhenko V.V., Dolgov E.N. Algorithm for modeling correlated numerical sequences distributed according to the Weibull law. Volnovaya e`lektronika i infokommunikacionny`e sistemy`. Sbornik statej XXV Mezhdunarodnoj nauchnoj konferencii. V 3-x chastyax [Wave electronics and info-communication systems. Collection of articles of the XXV International Scientific Conference. In 3 parts]. Saint Petersburg, 2022, pp. 130-134. (in Russian)

7. Dolgov E.N., Shepeta D.A., Yakovleva E.A. Modeling of Logarithmic-Normal Processes by the Method of Forming Filters. Molodoj ucheny`j [Young Scientist], 2022, no. S47-1(442-1), pp. 60-62. (in Russian)

TECHNOLOGICAL FEATURES OF CREATING A DIY IMPLEMENTATION OF A FLOCKER

Mikhail Kalinichenko

Saint Petersburg State University of Aerospace Instrumentation,
Ivangorod Branch, Saint Petersburg, Russia,
E-mail: jayli04@mail.ru

Abstract. *The article discusses the technological features of a DIY flocking device. The physical foundations of flocking technology and the principle of operation of the device are described. The concept of the device and the list of components for self-assembly of the flocker are discussed. The features of the development of a DIY flocker automated using the Ardiuno platform are being discussed.*
Keywords: *flocator, flocker machine, Ardiuno, DIY, electronic components, automation.*

The object of automation described in this article is a manual device for applying a flock to a surface.

Flock, or "static grass" that's what the creators of dioramas call it — it consists of finely chopped textile fibers (lint, hairs). It is produced from various types of raw materials such as wool, cotton, polyamide, viscose, nylon, etc.

Immediately after dyeing in an electrostatic field, the fibers undergo chemical activation. This activation consists in processing the fibers with special electrolytes, the need for which is to orient the flock fibers activated in the electrostatic field of the flocker perpendicular to the applied surface, i.e. they lie on the material "hedgehog", digging into the base with glue Fig. 1 [1].

The technical solutions used for flocking in large-scale industrial production have wide functional implementations and appropriate software control.

The application of the pile is carried out in an installation where an electric field is created between two electrodes, the top of which is a hopper with a mesh bottom in which the pile is located. From this hopper, the flock-fibers, by vibration, passes through the mesh between the electrodes, i.e. to the flocking zone. A material with an adhesive layer moves along the bottom electrode, into which the charged pile is embedded [2].

In the industrial production of various decorated surfaces, flocking, as a technology, is used quite widely, as well as in small-scale industries, or diorama lovers. Implementations for manual use have only one on and off button, and the flow is supplied by mechanically shaking the device from an inverted hopper with a grid at the bottom.

Accordingly, it is not possible to control the intensity of surface seeding and pause the process. When the flock-fibers is in a hopper that is not separated from the metal mesh, and it is its bottom, storing the device is quite problematic and requires constant cleaning of the hopper. As part of the design solution, it is required to integrate a hopper with a flock into the device body with the ability to close it. Accordingly, it is required to provide for the possibility of refueling it.

A closed hopper with a flock-fibers raises the question of the need for a hopper opening mechanism and an appropriate mechanism for feeding the flock to the grid. The body of the device, and in particular the area itself from the hopper to the grid, will be made in the form of a pipe with a screw placed in it. The auger has the ability to adjust its rotation speed, which allows you to adjust the density of planting with a flock.

Fig. 1 shows the operating scheme of the device and the physical process that underlies the operation of the device.

The flokers (flocker machine) on sale are mostly non-automated solutions with a single power button. Complex automation of various units of the device with parameter indication on the display is currently not implemented.

The concept of the assembled device provides for the automation of the device using sensors, a motor, a light indication, a motor speed controller, a voltage sensor and a display for indicating parameters. The Arduino Nano will be selected as the board for controlling the device.
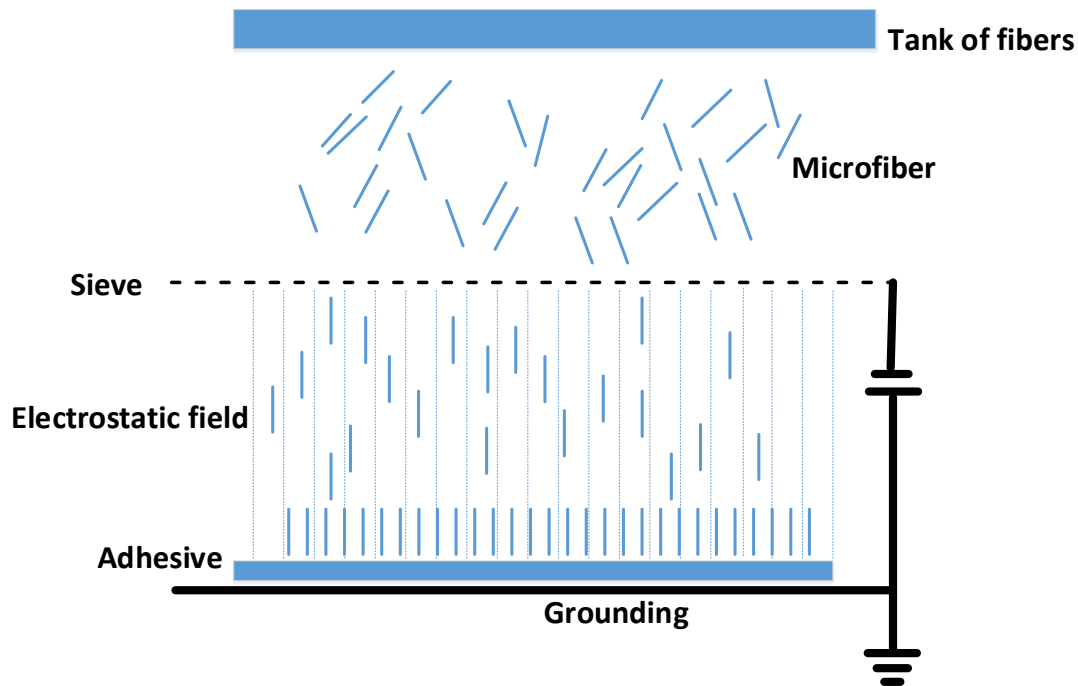
Fig. 1. The flocking process

The basis for automation and assembly, within the framework of the current task, are the following components:
- arduino nano board;
- motor drive expansion board;
- display;
- leds;
- motor;
- servo drive
- button;
- dimmer;
- resistors;
- battery pack;
- screw printed on a 3d printer;
- pulse generator;
- digital current and voltage sensor;
- inductive proximity sensor;
- structural elements for fixing blocks, plugs, lid and sash of the hopper;
- wires.

The functionality achieved in the implementation of this automation project of a manual flocking device is reduced to the following points:
- battery charge monitoring, displaying information about the percentage of battery charge on a digital monochrome display;
- control and display of the flow rate from the hopper, output of data on the selected speed mode on a monochrome display;
- control of the filling of the hopper with a flock, information output to the display (data from sensors);
- adjustment of the supplied voltage by a dimmer and the corresponding indication on the display of the parameters;
- a light indication when the device is turned on and a corresponding indication when the device is ready with a duplicate message on the display.

Regarding the construction of the device, there are several problems with the components. One of the problems that needs to be solved is the selection of a module for generating negative ions. These modules are available ready-made and have different configurations according to their parameters.

This project considers the possibility of using a ready-made or assembling its circuit yourself.

A secondary task regarding the negative ion generation unit is the ability to control it from a regulator.

The appropriate adjustment of the device parameters, within the framework of our automation concept, requires feedback in the form of an indication of the current operating parameters of the device on the display, which implies the presence of a voltage sensor for the generator. We will use a dimmer and a resistor to adjust the voltage.

Prototyping of the project implementation, taking into account the described functions, is performed in the form of an arrangement of blocks on a breadboard: Fig. 2.
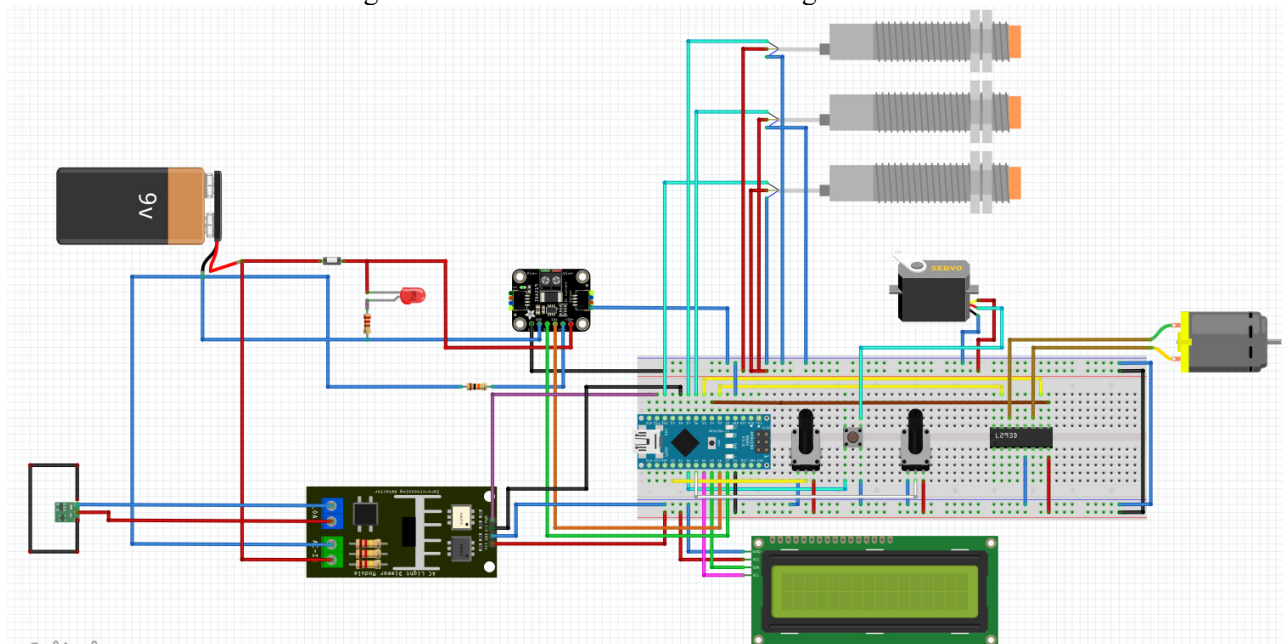
Fig. 2. Automation scheme based on Arduino-compatible components

A 110x2.7 mm sewer pipe is selected as the base for the housing. The remaining components of the housing and internal parts: auger, valves, plugs, mounting and mounting frames for electronic filling are designed for printing on a 3D printer and, accordingly, are made of PLA plastic.

Summing up the consideration of the designed structure and elements for automation of the manual flocker, it is possible to form a conclusion about the possibility of using Arduino to automate the control of the device. The difficulty is caused by the selection of modules of the component base and their coupling. It can also be noted that it is more correct to use mains power for this task.

References

1. Synthetic flock for decoration of light industry products and their properties / I. G. Davletbayev, A. R. Gabdullin, R. R. Khusnutdinova, L. R. Fatkhullina // Bulletin of the Technological University. – 2015. – vol. 18, No. 20. – pp. 177-178. – EDN VBWOMJ.

2. Ivanov, O. M. Analysis of the flocking process based on the boundary conditions of pile charging /Collection of scientific papers on the results of the International Scientific Conference dedicated to the 135th anniversary of the birth of Professor V.E. Zotikov : COLLECTION of SCIENTIFIC PAPERS, Moscow, May 25, 2022. Volume Part 1. – Moscow: Federal State Budgetary Educational Institution of Higher Education "Kosygin Russian State University (Technologies. Design. Art)", 2022. – pp. 109-114. – EDN QHLPUO.

CALCULATION OF PARAMETERS OF SOLAR BATTERIES AS THE MAIN SOURCE OF ENERGY FOR NANOSATELLITES OF THE CUBESAT FORMAT

Mikhail Kalinichev

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: mihail.kalinichev@gmail.com

Abstract: *the issue provides calculations of the main parameters of the energy supply system for student nanosatellites of the CubeSat format, and determines the characteristics of solar panels used as the main source of energy.*
Keywords: *primary energy sources, nanosatellite, CubeSat, mathematical modeling.*

The generation power of nanosatellites depends primarily on the parameters of the primary energy source. As a source of energy on small spacecraft, solar panels are used, consisting of semi-conductor photocells that convert the energy of sunlight into electricity. The use of solar panels is optimal when the required duration of continuous operation of the device is from several weeks to several years, with relatively low energy consumption, since solar panels are reliable and cheap, and their service life is limited by degradation due to meteor erosion, radiation and sudden temperature changes and can reach 10 years.

The requirements of minimal cost and ease of development on the part of educational institutions in relation to the device impose restrictions on the use of active moving parts and mechanisms, in particular, limit the use of actively oriented solar panels. This also has an impact on other systems of the spacecraft that indirectly affect the power supply system, e.g. orientation systems are not used. Summarizing the above, the requirements imposed on the mass and size characteristics of the nanosatellite as a result of the development of the satellite by students of educational institutions limit all possible design options of the device to the CubeSat 1U format with dimensions of 10×10×10 cm and a maximum mass of 2 kg, and rigidly fixed solar panels act as energy sources. Due to the lack of an orientation system to maximize power generation, the panels are located over the entire area of the machine.

Thus, the primary source of electricity is 6 solar panels with sides of 9.8×8.26 cm and therefore an area of approximately 80.9 cm2 each. SXC-SGS-03 solar panels manufactured by SPUTNIX [1] have such dimensions. It should be taken into account that out of 6 panels, no more than 3 can be simultaneously oriented to the sun, and their power directly depends on the orientation of the panels relative to the direction of the sun's rays.

The power output of a P array solar array can be determined by formula (1):

$$P_{\text{array}} = s\eta F_{СБ} S_{СБ} \cos(\theta), \tag{1}$$

where s is the solar radiation flux density, 1000–1500 W/m2; η – efficiency, for gallium arsenide cells can reach 24-32%; $F_{СБ}$ is the degradation coefficient of solar batteries, which takes into account changes in the emission of photovoltaic cells of the battery due to the influence of space factors; $S_{СБ}$ – Solar Array Area; Θ is the angle of orientation of the solar arrays relative to the Sun, defined as the angle between the direction of the Sun and the normal to the surface of the solar array [2].

Since in formula (1) only cos(Θ) depends on the orientation in space, and the rest of the quantities are constant, the calculation of the generation of a randomly oriented apparatus in space is reduced to the calculation of cos(Θ) for six panels depending on the orientation in space. It is convenient to describe the orientation in space of a rigid body using Euler angles, where the rotation of the body and the associated coordinate system (X, Y, Z) relative to a fixed coordinate system (x, y, z) is defined by 3 angles: α, β, γ. The intersection of the xy and XY coordinate planes is called the line of nodes N. The definition of Euler angles is shown in Fig. 1. The definition of the moving coordinate system and the designation of the sides of the spacecraft are shown in Fig. 2.
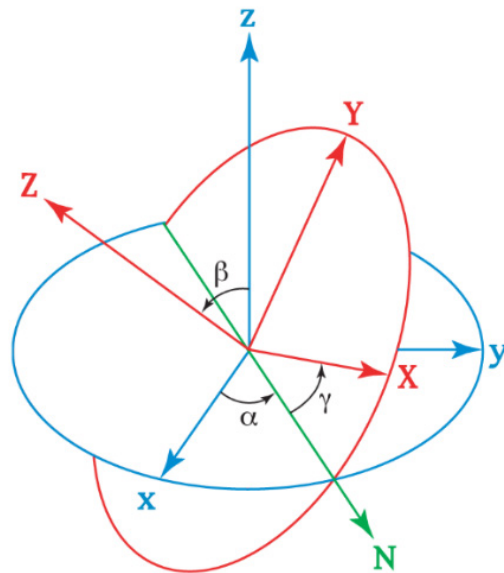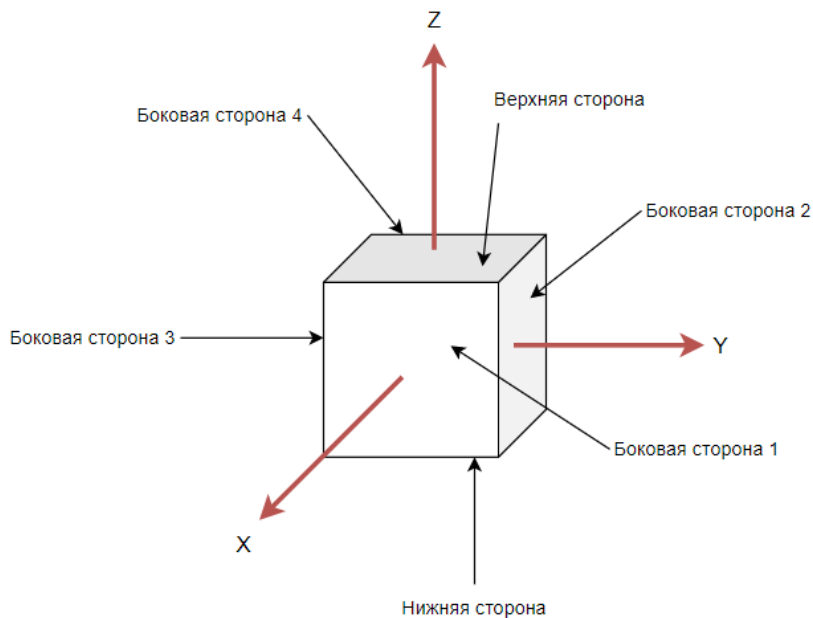
Fig. 1. Euler angles

Fig. 2. Defining a Moving Coordinate System

In the case of free rotation of a solar panel, the determination of its generation via cos(Θ) is inaccurate, since it does not take into account the possibility of rotating the panel by an angle greater than π/2. When rotated by an angle greater than π/2, cos(Θ) becomes negative, which makes sense of negative generation, which in reality cannot be the case, and the lasing is zero. Therefore, it is more convenient to use the formula:

$$\frac{\cos(\theta)+\left|\cos(\theta)\right|}{2} \tag{2}$$

Substituting (2) into (1) we get:

$$P_{array} = s\eta F_{СБ} S_{СБ} \frac{\cos(\theta)+\left|\cos(\theta)\right|}{2} \tag{3}$$

Thus, formula (3) takes into account the possibility of any angle relative to the direction of the Sun. The graph constructed by formula (2) in comparison with the graph cos(Θ) is shown in Fig. 3.
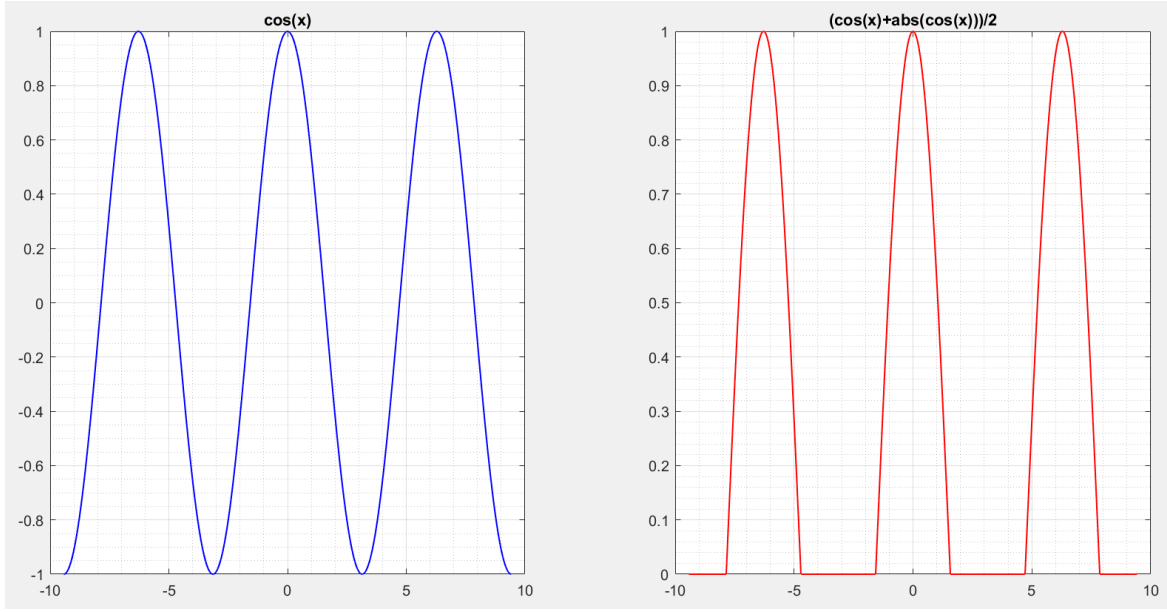
Fig. 3. Graph of the cos(Θ) function – on the left, graph of formula (2) – on the right

It is also necessary to take into account that all six solar panels will have a different angle and depend differently on the rotation angles. Thus, it is necessary to sum up the generation of all the panels of the device:

$$P_{array} = s\eta F_{CБ} S_{CБ} \left(\frac{\cos(\theta_1) + |\cos(\theta_1)|}{2} + \frac{\cos(\theta_2) + |\cos(\theta_2)|}{2} + \frac{\cos(\theta_3) + |\cos(\theta_3)|}{2} + \right.$$

$$\left. + \frac{\cos(\theta_4) + |\cos(\theta_4)|}{2} + \frac{\cos(\theta_5) + |\cos(\theta_5)|}{2} + \frac{\cos(\theta_6) + |\cos(\theta_6)|}{2} \right) = \qquad (4)$$

$$= s\eta F_{CБ} S_{CБ} \sum_{i=1}^{6} \frac{\cos(\theta_i) + |\cos(\theta_i)|}{2}$$

Next, it is necessary to find the dependence of the angle between the direction of the Sun and the normal to the surface Θ_i n the rotation angles α, β, γ, i.e. the functions $\Theta_i(α, β, γ)$ for all six sides of the spacecraft, where the index i is the panel number. The four sides correspond to terms with indices from 1 to 4 in accordance with Fig. 5. The top and bottom sides are indexed 5 and 6 respectively.

For the top side:

$$\theta_5(α, β, γ) = β \qquad (5)$$

For the underside, respectively:

$$\theta_6(α, β, γ) = β + π \qquad (6)$$

Taking into account that β is the angle between the z and z axis, and, consequently, between the xy and xy planes intersecting along the line of nodes, the normals to the side surfaces will lie in the xy plane and move in it depending on the angle of rotation γ counted from the node line, which corresponds to the angle of the side 2. This construction is shown in Fig. 4.

Taking into account the unit radius of the sphere under consideration, the angle Θ of the sides, the normals to which lie in the plane xy, can be considered as the length of the arc AB. The length of this arc can be found by considering the spherical triangle ANB, in which the angle ∠ANB = π/2 is β, and the sides |AN| = π/2, |NB| = γ.

The arc length AB can be found using the spherical cosine theorem:

$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(C), \qquad (7)$$

where a, b are the sides of the triangle, and C is the angle between these sides, and c is the side opposite the angle C.
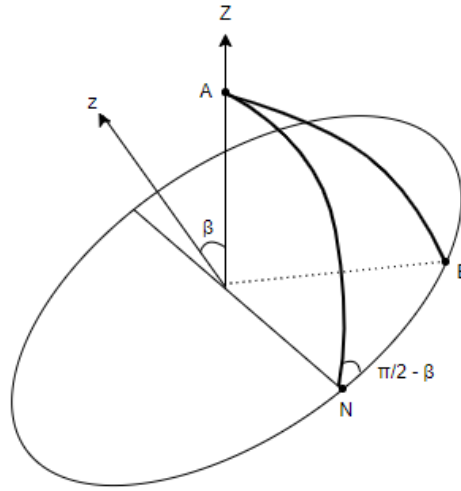
Fig. 4. Constructing a spherical triangle

By transforming (14), we find the side AB and the angle Θ_2:

$$\theta_2 = |AB| = \arccos\left(\cos\left(\frac{\pi}{2}\right)\cos(\gamma) + \sin\left(\frac{\pi}{2}\right)\sin(\gamma)\cos\left(\frac{\pi}{2}-\beta\right)\right) \tag{8}$$

In the same way, the angles Θ_1, Θ_3, Θ_4 can be found, taking into account the relative position of the sides, which is expressed in the shift of the rotation of the γ by $-\pi/2$ for side 1, and by $\pi/2$ and π for sides 3 and 4, respectively. Thus, we get:

$$\theta_1 = \arccos\left(\cos\left(\frac{\pi}{2}\right)\cos\left(\gamma-\frac{\pi}{2}\right) + \sin\left(\frac{\pi}{2}\right)\sin\left(\gamma-\frac{\pi}{2}\right)\cos\left(\frac{\pi}{2}-\beta\right)\right) \tag{9}$$

$$\theta_3 = \arccos\left(\cos\left(\frac{\pi}{2}\right)\cos\left(\gamma+\frac{\pi}{2}\right) + \sin\left(\frac{\pi}{2}\right)\sin\left(\gamma+\frac{\pi}{2}\right)\cos\left(\frac{\pi}{2}-\beta\right)\right) \tag{10}$$

$$\theta_4 = \arccos\left(\cos\left(\frac{\pi}{2}\right)\cos(\gamma+\pi) + \sin\left(\frac{\pi}{2}\right)\sin(\gamma+\pi)\cos\left(\frac{\pi}{2}-\beta\right)\right) \tag{11}$$

Simplifying equations (8), (9), (10), and (11), we get:

$$\theta_1 = \arccos\left(\sin(\beta)(-\cos(\gamma))\right) \tag{12}$$

$$\theta_2 = \arccos\left(\sin(\beta)\sin(\gamma)\right) \tag{13}$$

$$\theta_3 = \arccos\left(\sin(\beta)\cos(\gamma)\right) \tag{14}$$

$$\theta_4 = \arccos\left(-\sin(\beta)\sin(\gamma)\right) \tag{15}$$

Substituting (5), (6), (12), (13), (14) and (15) into formula (4) we get:

$$P_{array} = s\eta F_{СБ} S_{СБ}\left(\frac{\cos\left(\arccos\left(\sin(\beta)(-\cos(\gamma))\right)\right) + \left|\cos\left(\arccos\left(\sin(\beta)(-\cos(\gamma))\right)\right)\right|}{2} + \right.$$

$$+ \frac{\cos\left(\arccos\left(\sin(\beta)\sin(\gamma)\right)\right) + \left|\cos\left(\arccos\left(\sin(\beta)\sin(\gamma)\right)\right)\right|}{2} +$$

$$+ \frac{\cos\left(\arccos\left(\sin(\beta)\cos(\gamma)\right)\right) + \left|\cos\left(\arccos\left(\sin(\beta)\cos(\gamma)\right)\right)\right|}{2} +$$

$$+ \frac{\cos\left(\arccos\left(-\sin(\beta)\sin(\gamma)\right)\right)+\left|\cos\left(\arccos\left(-\sin(\beta)\sin(\gamma)\right)\right)\right|}{2} +$$

$$+ \frac{\cos(\beta)+\left|\cos(\beta)\right|}{2} + \frac{\cos(\beta+\pi)+\left|\cos(\beta+\pi)\right|}{2}\Bigg). \tag{16}$$

By simplifying the expression (16), we get:

$$P_{array} = s\eta F_{СБ} S_{СБ}\left(\left|\sin(\beta)\right|\left(\left|\sin(\gamma)\right|+\left|\cos(\gamma)\right|\right)+\left|\cos(\beta)\right|\right) \tag{17}$$

This formula allows you to find the lasing power of the CubeSat 1U device depending on the orientation in space

The average power generated by all 6 solar panels of such a device when exposed to the sun can be calculated by integrating formula (24) at all three angles of rotation and dividing by the size of the detection area [3]. In the calculation, it is assumed that the device has a random orientation in space, since it does not have an orientation system and all possible positions in space are equally probable. The average power is calculated as follows:

$$P_{avr\ array} = s\eta F_{СБ} S_{СБ}\frac{\displaystyle\int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi}\left(\left|\sin(\beta)\right|\left(\left|\sin(\gamma)\right|+\left|\cos(\gamma)\right|\right)+\left|\cos(\beta)\right|\right)\sin(\beta)d\beta d\alpha d\gamma}{\displaystyle\int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi}\sin(\beta)d\beta d\alpha d\gamma} = \tag{18}$$

$$= s\eta F_{СБ} S_{СБ}\frac{118.435}{8\pi^2} = s\eta F_{СБ} S_{СБ}\cdot 1.5.$$

Taking the average values s = 1300 W/m2, η = 0.28, and SSB = 0.00809 m2 and substituting (25) into the formula, we get:

$$P_{avr\ array} = s\eta F_{СБ} S_{СБ}\cdot 1.5 = 1300\cdot 0.28\cdot 1\cdot 0.00809\cdot 1.5 = 4.416\ \text{Tue} \tag{19}$$

A graph of the function (24) is shown in Fig. 8. The graph is plotted relative to the angles of β and γ, which range from 0 to π for β and from 0 to 2π for γ.

It is easy to see that the graph has 8 maxima, corresponding to such an orientation that the spacecraft is oriented in one of the 8 corners in the direction of the Sun, so that 3 panels are illuminated at the same time. In this case, the lasing is Pmax $_{array}$ = 5.1 W.
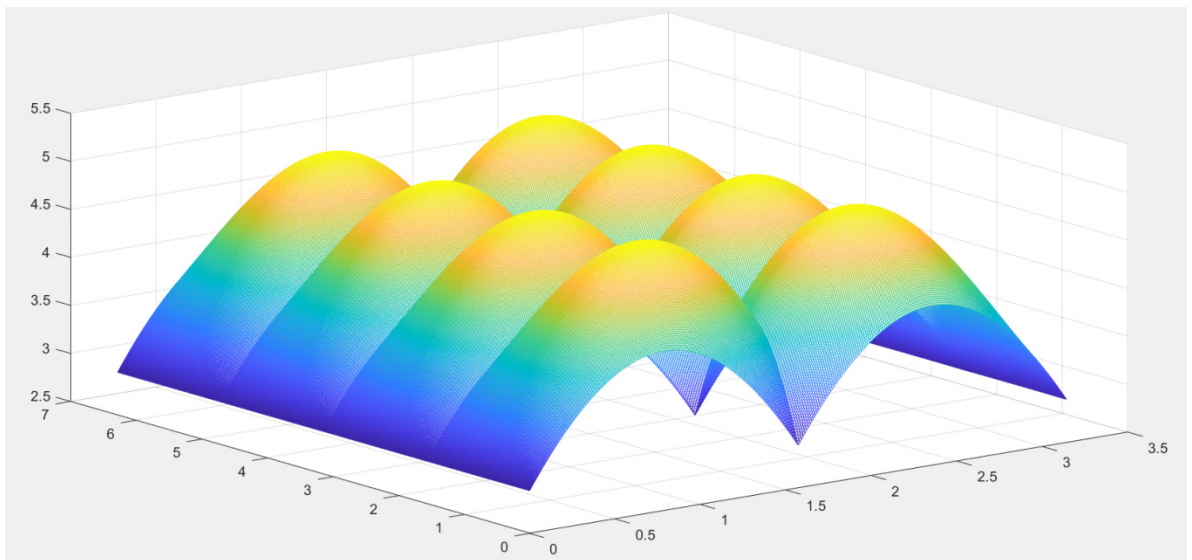
Fig. 5. Function graph (17)

Taking the orientation angles as uniformly distributed random variables, it is possible to construct a histogram of the distribution of the function (17), which will make it possible to approximate the distribution function of the random lasing variable of the apparatus. This histogram is shown in Fig. 6.
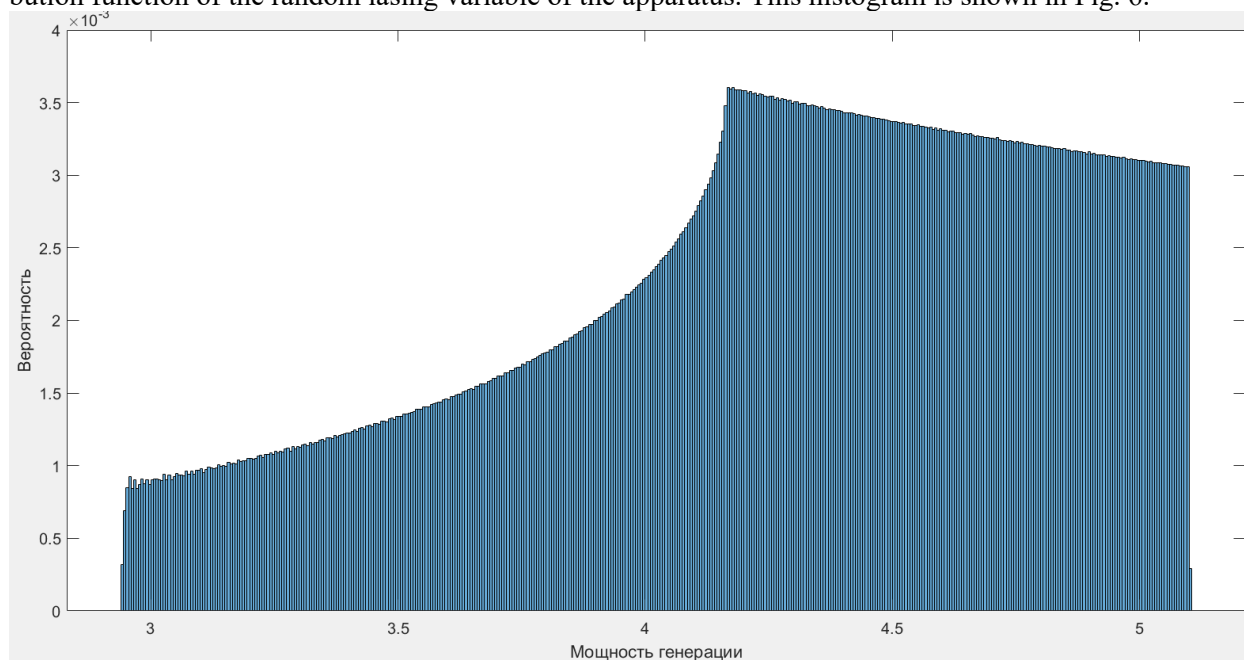
Fig. 6. Histogram of the lasing distribution of the apparatus

The value of the average power of generation $P_{\text{avr array}}$ = 4.416 W obtained in formula (19) is relevant only at the time of the start of the operation of the device and towards the end of its planned service life will be significantly reduced due to the degradation of solar panels. The degradation coefficient of solar cells $F_{\text{СБ}}$ can be estimated by formula (20) [1].

$$F_{\text{СБ}} = \left(1 - YD\right)^{YM} , \qquad (20)$$

where YD is the degradation coefficient of the solar panel per year, which is approximately 0.03; YM is the time of the mission in years from the date of launch.

Thus, at the time of launch, in formula (26), $F_{\text{СБ}}$ = 1, and after 3 years, according to formula (27), it will decrease to fsb = 0.91. The power generated by the device will decrease by about 9% and will be $P_{\text{avr array}}$ = 4.019 W. It is this value that should be used as the estimated average generation capacity, because then the average generation capacity will exceed the calculated one throughout the planned operating time. thus creating an unforeseen surplus of energy. Excess energy in the system can always be simply eliminated. It can even be useful as a contingency reserve. In turn, an unforeseen flaw can lead to premature battery discharge and emergency shutdown of consumers [4].

References

1. Cappelletti C., Battistini S., Malphrus B. K. CubeSat Handbook From mission design to operations. – London: Academic Press, 2021. – 461 p.

2. Ablameyko S.V. Small Spacecraft: A Manual for Students of the Faculties of Radiophysics and Computing. Technologies. Minsk: BGU, 2012. 159 p. (in Russian).

3. Kalinichev M., Power supply system for student nano-satellite// Bulletin of the UNESCO department "Distance education in engineering" of the SUAI: Collection of the papers. Saint Petersburg, Issue 8. – SPb.: SUAI, 2023. – P. 111—119.

4. Kalinichev M.A., Kononov O.A. General Principles of Building Nanosatellites of the CubeSat Format. Devices and Methods of Information Processing: The Fourth All-Russian. Scientific conference. (April 10-17, 2023): Compilation. Reports. – Saint Petersburg. SUAI, 2023. – P. 110 – 117.

ANALYSIS OF THE ABBIGUITY FUNCTIONS OF MODIFIED LEGENDRE AND JACOBI CODE SEQUENCES

Daria Karabaeva

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: karabaevadasha@mail.ru

Abstract. *The article analyzes the ambiguity function for modified code sequences of maximum length obtained from rows of quasi-orthogonal matrices. The procedure for obtaining modified code sequences is considered. A comparison of the estimates of the characteristics of the obtained new code sequences is given. Since these code sequences have potential application in earth remote sensing systems, not only correlation properties are considered, as for communication systems, but also the ambiguity function, which considers not only correlation characteristics, but also frequency characteristics.*

Keywords: *Quasi-orthogonal matrices, earth remote sensing, correlation function, ambiguity function, modified code sequences.*

Introduction

In the process of developing new active remote sensing systems, one of the main tasks is to ensure high noise immunity and interference protection of probing signals used for detecting and extracting information about observed objects. When solving practical tasks of this kind, signals modulated by code sequences (CS) are widely used. These sequences define the basic properties of such signals and simultaneously serve as their main classification feature [1].

An example of a signal modulated by a CS is a phase-manipulated signal. The phase sequence of the signal being 0 or π at different time points can be associated with a CS of +1 and -1, respectively. The alphabet of the sequence provided in this example is binary (± 1), integer, and symmetrical.

It is important to consider that in communication systems, the main focus is on the correlation properties of the sequence, such as autocorrelation and cross-correlation functions. However, when solving remote sensing tasks, it is crucial to assess the properties of the ambiguity function (UF), which takes into account not only signal correlation over time but also over frequency [1,2].

The formula for the UF is given by:

$$R(\tau, f) = \sum_{i=0}^{N-1} y_i y_{i-\tau} e^{-j2\pi f i \Delta_t}$$

where y_i – an element of the codeword sequence; f – the frequency shift; τ – the time shift (a shift of the sequence with duration Δt).

This work aims to conduct a study of the UF for the purpose of comparative analysis and to explore the possibility of using as CS structures based on circulant quasi-orthogonal matrices with an alphabet (1, -*b*) [3] – sequences of maximum length, similar to sequences generated based on quadratic residues and Jacobi symbols [4]. This signifies a departure from the classical approach where the alphabet of the CS is integer and symmetric.

MODIFICATION OF THE CODEWORD SEQUENCE

To convert the original sequence to the alphabet (1, -*b*), it is necessary to carry out its modification. The flowchart of the modification procedure is presented in Fig. 1.

Fig. 1 requires explanations. In this work, sequences of Legendre and Jacobi were chosen as the original sequences. The computation of the unknown value -*b* is carried out based on the theory of quasi-orthogonal matrices [3,5]. As a result, a square matrix is obtained, an example of which is shown in Fig. 2, where a white square corresponds to "1" and a black square corresponds to "-*b*".
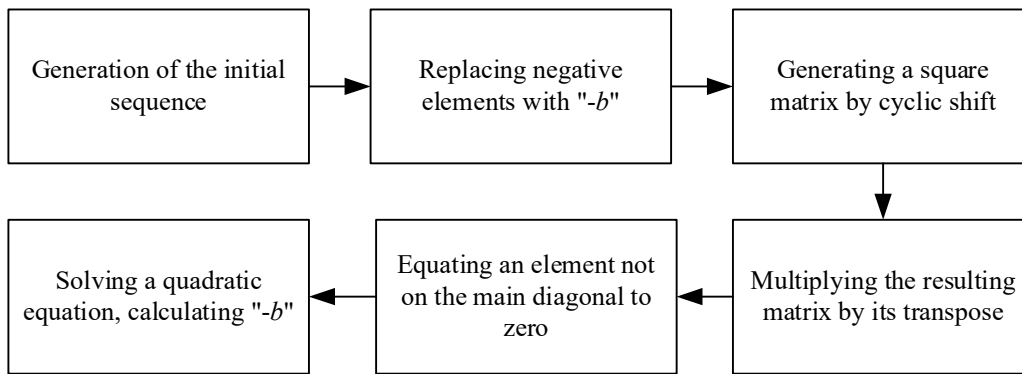
Fig. 1. Procedure for generating sequences with the alphabet (1, -b)
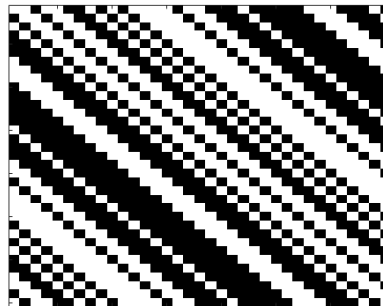
Fig. 2. Portrait of a cyclic quasi-orthogonal matrix based on a Jacobi sequence of length 35.

ANALYSIS OF AMBIGUITY FUNCTIONS DERIVED FROM THE MODIFIED SEQUENCES

The methodology of the experiment was as follows. After generating a cyclic matrix with elements (±1), a line-by-line search for the sequence with the lowest level of sidelobe indecision function at f=0 was conducted. Then, after converting the matrix to the alphabet (1, -b), a search for the sequence with the lowest sidelobe level was repeated, followed by comparison of the values. As an example, Fig. 3a and Fig. 3b are presented, showing the plots of the normalized to one indecision function of the Jacobi code sequence of order 143 with alphabets (1, -1) and (1, -b) respectively.

а) alphabet (1, -1)   б) alphabet (1, -b)
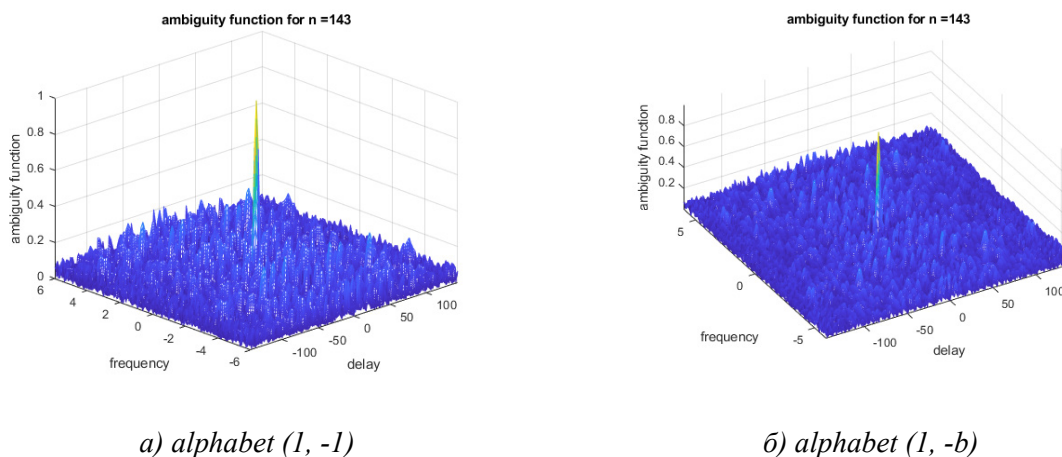
Fig. 3. The normalized-to-one indecision function of the Jacobi code sequence of order 143

When analyzing the indecision function, it is also advisable to consider its cross-section at zero frequency $f = 0$. For example, Fig. 4 shows the cutoff at zero frequency for the indecision function of the Jacobi code sequence of the same order as depicted in Fig. 3. As seen from Fig. 4, the maximum level of the sidelobe of the Jacobi code sequence is lower in the case of alphabet (1, -b) compared to the case of alphabet (1, -1).

a) alphabet (1, -1)   б) alphabet (1, -b)
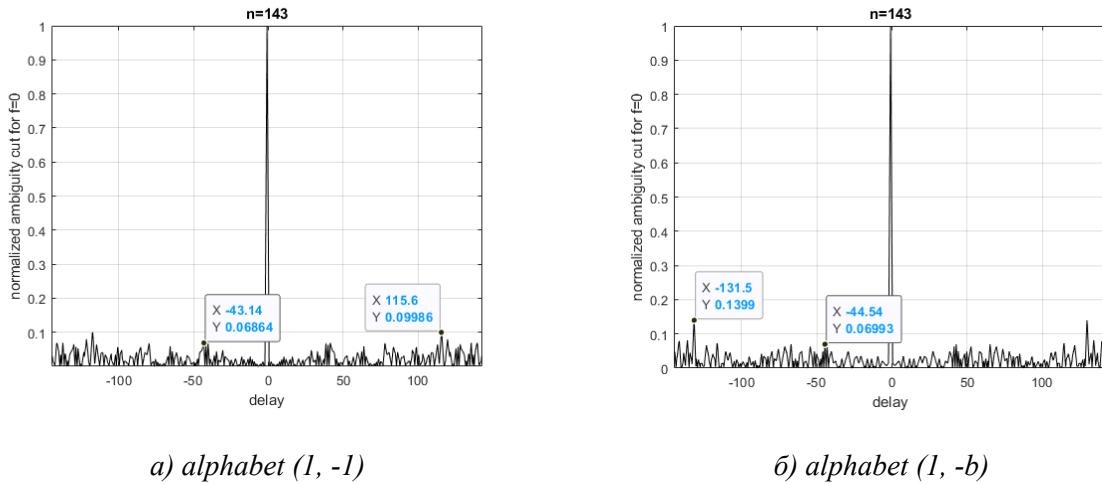
Fig. 4. The normalized-to-one ambiguity functions
of the Jacobi code sequence of order 143 (Cutoff at zero frequency)

Since one of the main criteria for the applicability of signals in remote sensing systems is their resistance to interference, it is advisable to examine the CS mentioned in the paper for the maximum level of the sidelobe of the ambiguity function and the maximum level of the sidelobe at zero frequency, which mean autocorrelation function (AF). The specified values are presented in Tables 1 and 2 for sequences based on Legendre and Jacobi symbols, respectively.

Table 1

Table of side lobe values for Legendre sequence

Order	Alphabet (1,-1)		Alphabet (1,-b)	
	AF	UF	AF	UF
7	0.1714	0.4249	0.1398	0.4981
11	0.1091	0.5815	0.06153	0.5737
19	0.1579	0.4047	0.1666	0.3444
31	0.1548	0.3262	0.1518	0.3588
131	0.0733	0.2059	0.0726	0.1941
143	0.3636	0.3636	0.3996	0.3996
163	0.0761	0.2172	0.0750	0.2246
511	0.4012	0.4012	0.4216	0.4216

Table 2

Table of side lobe values for the Jacobi sequence

Order	(1,-1)		(1,-b)	
	AF	UF	AF	UF
15	0.0667	0.3877	0.0833	0.376
35	0.1829	0.3886	0.1813	0.3633
143	0.1399	0.2958	0.0999	0.2989
323	0.0718	0.3611	0.0701	0.3546

The obtained values prove the feasibility of an approach based on replacing a symmetric alphabet with (1, -1) with an asymmetric one (1, -b).

Conclusion

The results obtained in the study demonstrate the potential of the described approach and allow us to conclude on the feasibility of using modified code constructions when applied in information exchange systems, as well as in distributed radar station remote sensing systems.

Acknowledgment

References

1. Gantmacher V. E., Bystrov N. E., Chebotarev D. V. Noise-Like Signals. Analysis, Synthesis, Processing. Saint Petersburg, Nauka i Tekhnika, 2005. 400 p.

2. Vakman D. E. Complex Signals and the Principle of Uncertainty in Radar. – Moscow, "Sov. Radio", 1973, 312 p.

3. Balonin N. A., Sergeev M. B. Local Maximum Determinant Matrices // Information and Control Systems. 2014. №1. P. 2-15.

4. Grigoriev E. K., Nenashev V. A., Sergeev A. M., Samokhina E. V. Search and Modification of Code Sequences Based on Persymmetric Quasi-Orthogonal Circulants // Telecommunications. 2020. №10. P. 27-33.

5. Balonin N., Sergeev M. Quasi-Orthogonal Local Maximum Determinant Matrices // Applied Mathematical Sciences. 2015. Vol. 9. № 8. P. 285-293. DOI: 10.12988/ams.2015.4111000

AUTOMATION OF MD SIMULATION DATA PROCESSING

K. P. Karasev, D. A. Strizhkin

Peter the Great Saint Petersburg Polytechnic University,
29, Polytechnicheskaya str., Saint Petersburg, 195251, Russia
kir.karasyov2017@yandex.ru

Abstract. *Molecular Dynamics is a simulation method commonly employed in scientific research, specifically in the field of solid state physics. In particular, rapid processes that take place on the surface after cluster ion impact such as sputtering, roughness modification, crater and rim formation can be studied. However, these phenomena implicate the analysis of substantial amounts of data, encompassing the velocity, coordinates, and energies of each atom in the simulation. So that it is necessary to develop scalable and automated methods to treat vast amount of simulated data yielding various parameters of the system including size of impact crater, rim characteristics, surface deformation and the energy and directions of sputtered particles. To illustrate the efficacy of these methods, results obtained during simulation of (100) Si crystal surface bombardment with accelerated C_{60} molecules are presented.*

Keywords: *molecular dynamics simulation, ion irradiation, data processing, cluster ion, silicon.*

Introduction
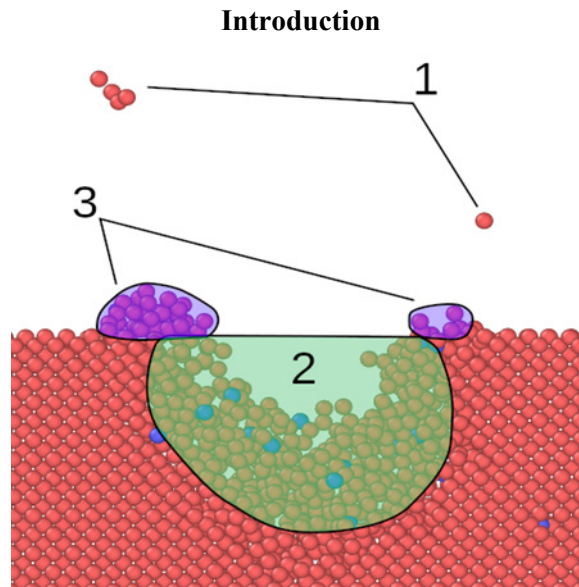
Fig. 1. Cross section of a Si target surface after single 14 keV C_{60} ion impact event at 0K.
1 – sputtered particles; 2 – impact crater; 3 – rim.

  Molecular Dynamics (MD) simulations is a widely used method of investigations in modern science. Previously, it was applied in studies over widespread fields such as biology, surface modification with accelerated ions, etc. [1-2]. The main advantage of MD modeling is the possibility to study position and momentum vector of each atom in the system at every timestep. However, such detailed description yields vast amount of raw data produced in every experiment carried out by MD simulation. These data have to be processed before any type of analysis to be performed. Hence, it is necessary to develop methods of automatic post processing in a scalable manner to avoid doing it manually. Surface modification by accelerated cluster ion irradiation is important tool in modern technology. MD is very versatile tool to get insights to the microscopic mechanisms of the processes taking place. For example, characteristics of crater formation after impact of large argon clusters on silicon targets have been studied using MD [3]. It was demonstrated that incident per-atom energy affects crater formation stronger than total incident cluster energy. On the other hand, only one trajectory run was performed for each simulation case with different size of argon cluster or impact parameters. However, for such type of experiment with single ion impact case, statistics consisting of even 50 trajectories may significantly increase the quality of the results obtained.

  Number of phenomena occur during C_{60} ion impact was studied [4]. Typical snapshot of a crystal surface after a single 14 keV fullerene molecule impact is illustrated in Fig. 1. When fullerene ion impacts

a surface it transfers energy to the nearest atoms who in turn spread energy into the bulk of a target. Some atoms receive enough energy to travel away from the surface either as a single particles or group of atoms called cluster. This effect is known as sputtering. Apart from sputtering, formation of crater and rim around it were observed. All this phenomena underlie modification of the surface relief. Different parameters of abovementioned effects could be determined from MD simulations and compared with experimental data, so microscopic characteristics of the system such as momentum vectors of sputtered clusters are also of interest to be derived from MD data.

Analysis of large amount of data in order to derive various parameters from MD simulation results is hard to be done manually. Therefore, it is necessary to develop algorithms that would automate data processing and show the results in a convenient way. It is also important to make such algorithms independent from the number of initial parameters and amount of statistical data.

Model description

Calculations were carried out using open-source code LAMMPs [5], a special program to run MD simulations. Well established Tersoff [6] potential was used to describe interatomic interactions between all types of atoms. Furthermore, it was splined with ZBL potential to describe interaction between particles with high energy. The initial crystal with size of 40x40x31 unit cells was built from Si atoms according to diamond lattice structure and heated [7] to a specified temperature (0 to 1000 K). Before the start of simulation, a fullerene molecule was located 100 nm above the open surface of target. Periodic boundary conditions were applied in lateral directions of simulation cell. Three bottom layers of atoms were fixed to avoid drift after the impact. Dissipation of excess energy from the system during the simulation was provided by setting Berendsen thermostat [8] to 1 unit cell wide bottom layer of silicon. Integration of equations of motion for all the remaining Si atoms were done as microcanonical ensemble (NVE) and for each time step new positions of atoms were calculated. Electronic energy loss experienced by fast atoms in dense solid target was taken into account for particles of kinetic energy higher than 10 eV and having more than 30 neighbors and implemented as a quasi-friction force. Total time of each simulation was set according to the energy of C_{60} molecule and varied from 5 to 10 ps (for 8 to 14 keV correspondingly). Voronoi-Dirichlet method [9] was used to identify vacancies and interstitial atoms during the simulation. Before the calculations, Wigner–Seitz cells were set by Voronoi tessellation over all crystal volume. Further, each cell was analysed for the number of silicon atoms it contains to derive the point defects that appear after C_{60} ion impact. If there were no atoms, it was identified as a vacancy and extra-atoms in the cells containing more than one particle were identified as interstitials. Atoms, escaped from the surface, were determined as sputtered and analysed apart. Such system conditions were used for both types of experiments, carried out by MD method: single impact events and cumulative irradiation. In the first situation, 50 statistically independent cases were investigated. After every impact, results were saved to special files and the whole system was recovered to the initial state. Then, C_{60} molecule was moved to a small random distance in both lateral directions and the simulation started again. In the case of cumulative bombardment, sputtered atoms were analysed after each event and then deleted from the simulation cell. Further, system was cooled down to the initial temperature during 3 ps and new C_{60} ion was placed above the surface and moved to a small random distance as it was described for single impacts.

Automation of experiments

Important phenomenon caused by accelerated ion impact that have to be analysed is sputtering of particles from the target. Analysis of these particles is common experimental observation that can be crucial for investigations of the composition and structure of the target. For example, in Secondary Ion Mass Spectrometry (SIMS) sputtering yield analysis can be used to study concentration of impurities in doped semiconductors. Therefore, it is essential to develop an algorithm to identify and examine sputtered particles in simulations. Also as a result of fullerene ion impact, surface relief is modified and new structures appear on it during the irradiation. So all atoms in the system could be divided into two categories: sputtered atoms and in-substrate atoms. For this purpose, before the simulation surface of the crystal was divided in XY plane into equal sized square areas of specified dimension that was varied from 0.25 to 1 unit cells. All atoms close to the crystal surface were distributed into the groups, corresponding to each squared area using their X and Y coordinates. The highest particle in each group was pointed out and their mean was set as surface zero level. Such method to find the value of this parameter was used in the previous work to analyse surface oscillations amplitude [7]. After a simulation run is complete, particles situated above this zero level, were marked as sputtered candidates and considered apart. The result of applying

this algorithm is illustrated in Fig. 2. Dark atoms represent the crystal and coloured ones show atoms considered as sputtered candidates. It can be seen, that despite some of the atoms in the coloured group are situated above the surface they are in close contact to the crystal. These atoms were pulled out by the ion impact, but have not enough energy to go away and fell down to the surface near the crater forming structure called rim. So they are not actually sputtered. Other coloured atoms (pointed as #1 in the Fig.2) are clearly sputtered. To treat that, all sputtered candidate particles were divided into clusters, using built-in LAMMPs function. Atoms located within the cut-off distance of 3.0 Å from each other were considered comprising one cluster. All these small groups then were split into two groups: if the group contained atoms, which are located less than 2.0 Å from the zero level, it was recognized as a part of the rim, while the remaining clusters were identified as sputtered. As a result, it is possible to obtain information about type, mass, energy, and direction of emission using IDs of these atoms. Simple math calculations, using Python script and values of mentioned parameters, provide data about sputtering, such as energy and momentum vector of centre of mass of each sputtered cluster.
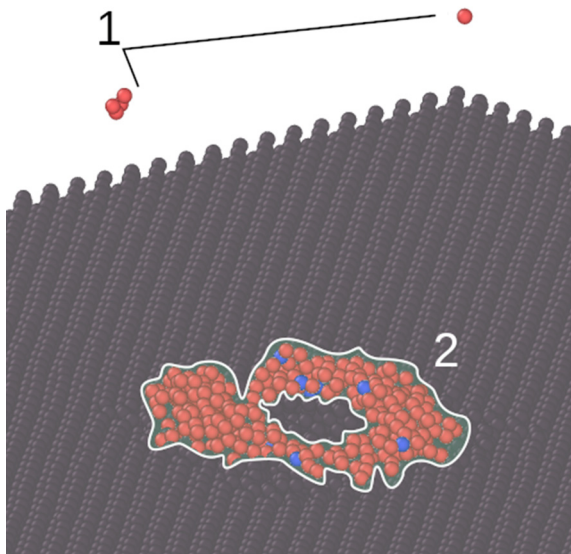
Fig. 2. Snapshot of the surface after14 keV C_{60} ion impact at 0K. 1 – sputtered particles; 2 – rim atoms.
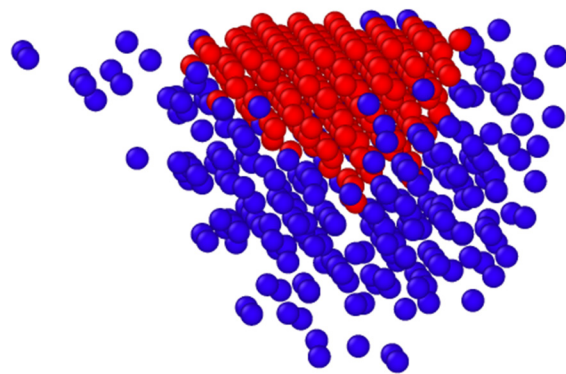
Fig. 3. Snapshot of initial Si crystal atoms whose Voronoi cells will become vacant after 14keV C60 ion impact. Red – cluster with largest number of cells. Blue – other cells that form smaller vacancy clusters around the impact point.

As it was already mentioned, all of atoms in the system can be split in two categories: sputtered and substrate. Surface of the crystal is modified during the simulation and new structures such as rim or, at a higher fluence, nanoripples appear. New relief of the target can significantly affect final properties of material and change its RMS, friction coefficient etc. One of the main characteristics of the surface relief is roughness, so it is necessary to develop an algorithm to calculate it. For this purpose, method that was earlier used to identify zero level was applied for the crystal in its final state. It was divided in XY plane into equal sized squared areas of specified dimension, all the atoms were distributed into the groups corresponding to each squared area using their (X, Y) coordinates and the highest atom in each square was pointed out. Thus, it is possible to create a surface mesh, where max Z coordinates are determined by highest atom position and could be graded by colour in the figure. If selected square had not included any atoms, it means either the square size was too small or data were incorrect. However, surface point for such square was still calculated as an average of its neighbour's Z coordinates. When all of the surface points had values, assigned to them, a 3D or 2D surface depth mapping were plotted. Height distribution histogram of surface points was also charted. Roughness of the surface was calculated as standard deviation of Z values. For cumulative irradiation situation, it makes possible to keep track of surface roughness modification during the simulation automatically and in situ.

One more structure that appear on the surface of a crystal even after single C_{60} ion impact is crater. Therefore, it is important to analyse its parameters, for example depth, and radius, to find out mechanisms of its formation. To do this, silicon atom IDs associated with vacant Wigner-Seitz cells were identified after the simulation. Subsequently, Si crystal was recovered to its initial state to find atoms situated in these cells according to Voronoi decomposition. Further, these atoms were separated using cluster analysis function, similar to sputtering and rim assessment. Cut-off distance of 3 Å was used again and

the largest cluster touching the surface was identified as crater. Output of this algorithm was visualized using Ovito tool and demonstrated in Fig. 3. Largest cluster was highlighted with red colour, so it is possible to see crater shape. Volume of the crater was determined by multiplying the mean Wigner-Seitz cell volume by number of particles corresponding to largest vacancy cluster. Crater opening area was calculated in similar way. The number of atoms in the two top most layers of the crater group was counted, as these atoms represent the surface layer of the crater. Then this number was multiplied by square of the half of Si lattice parameter. As it was done earlier for rim, the average and maximum atom radii from the impact centre in the XOY plane were calculated. Mean and maximum Z coordinates of crates group relative to zero level were also determined for the crater group to derive man and maximum crater depth.

Conclusion

As a result, a set of algorithms for automatic analysis of MD simulation results was developed and tested. With the help of these tools, vast amount of raw data may be processed and presented in a convenient way. Application of these methods provides an opportunity to derive quantitative data for a number of various effects that take place during ion bombardment of a solid surface, e.g. crater volume and mean depth; rim mean radius and height; sputtering parameters, namely, number of sputtered atoms, their possible clustering, and momentum vectors. In addition, the way to calculate surface profile mapping and roughness was also established. All the scripts are based only on atom coordinates and their velocities that were parsed by Python and built-in LAMMPs functions. Effectiveness of proposed methods was demonstrated on C_{60} ion bombardment of silicon crystal with different incident energies and temperatures of the target. Nevertheless, it is just a particular case and procedures described in this work are of a general origin, so they can be applied for any systems that require analysis of sputtered structures and surface deformations of solid crystals.

References

1. K. D. Krantzman and A. Wucher, "Fluence Effects in C_{60} Cluster Bombardment of Silicon", J. Phys. Chem. C 2010, 114, 5480. https://doi.org/10.1021/jp906050f

2. A. Delcorte B. and J. Garrison, "Sputtering Polymers with Buckminsterfullerene Projectiles: A Coarse-Grain Molecular Dynamics Study", J. Phys. Chem. 2007, 111, 15312. https://doi.org/10.1021/jp074536j

3. T. Aoki, T. Seki, J. Matsuo, "Molecular dynamics simulations for gas cluster ion beam processes", Vacuum 2010, 84, 994-998, https://doi.org/10.1016/j.vacuum.2009.11.018

4. K.P. Karasev, D.A. Strizhkin, A.I. Titov, P. Karaseov, "MD simulation of silicon irradiation with 2-8 keV buckminsterfullerene C60 ions", J. Surface Investig. 2023, 17, 66–71. https://doi.org/10.1134/S102745102301010X

5. A.P. Thompson, H.M. Aktulga, R. Berger et al., "LAMMPS – a flexible simulation tool for particle-based materials modelling at the atomic, meso, and continuum scales", Comp. Phys. Comm. 2022, 271, 10817, https://doi.org/10.1016/j.cpc.2021.108171

6. J. Tersoff, "New empirical approach for the structure and energy of covalent systems". Phys. Rev. B 1988, 37, 6991-7000, https://doi.org/10.1103/PhysRevB.37.6991

7. K.Karasev, D.Strizhkin, P.Karaseov, "Convenient Way to Create an MD Model of a Hot Crystal with an Open Surface" Proceedings of the 2022 International Conference on Electrical Engineering and Photonics, EExPolytech 2022, 2022, 242–245 https://doi.org/10.1109/EExPolytech56308.2022.9950888

8. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J.R. Haak, "Molecular dynamics with coupling to an external bath" J. Chem. Phys. 1984, 81, 3684, https://doi.org/10.1063/1.448118

9. F. Aurenhammer, "Voronoi Diagrams – A Survey of a Fundamental Geometric Data Structure", ACM Computing Surveys, 1991, 23(3), 345-405, https://doi.org/10.1145/116873.116880

ABOUT ONE APPROACH TO ASSESSING THE QUALITY OF MASKING VISUAL INFORMATION

Tatiana Kazakevich

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: kazakevitchtatjana@yandex.ru

Abstract. *The paper proposes the approach to assessing the quality of one-side and two-side matrix masking based on the analysis of the two-dimensional correlation coefficient and the protected image pixels correlation. The obtained results allow us to numerically evaluate the masking quality using matrices of various structures. The experiments show that if it is impossible to provide masking matrix size and the original image equality (the best possible masking quality), it is expedient to set thresholds for the two-dimensional correlation coefficient of no higher than 0.5, which ensures acceptable masking quality, and a threshold of no higher than 0.2, which ensures an inability to extract any useful information for an outside observer for each class of images.*
Keywords: *matrix masking, quasi-orthogonal matrices, masking quality, pixel correlation, privacy assurance.*

Introduction

A large number of data transformation and analysis tasks are associated with the processing and transmission of visual information. Due to its ease of perception [1], this type of information is often used, for example, in everyday network communication among Internet users or when consuming various media content [2, 3]. Visual information, like any other, is susceptible to leaks, which requires solving the problem of ensuring the confidentiality of transmitted data.

In this area the majorly used confidentiality ensuring methods are cryptographic methods, however matrix methods are also being actively developed [4], in particular, matrix masking [5], that uses low-level structured quasi-orthogonal matrices [6]. This method, which transforms visual information to a noise-like form and uses a simple mathematical apparatus, is well-developed, since the literature [7-9] covers in detail the issues of compression and resistance to distortion of the transmitted information, as well as the issues of selecting matrices for this procedure. Nevertheless, the question of assessing the quality of masking still remains open.

In work [10], the classical metrics MSE, PSNR, SSIM were used for evaluation. However, due to the specifics of visual information, when masking with small-sized matrices, the MSE and PSNR values can be quite large, but pronounced contours remain in the image, visual analysis of which allows, in some cases, complete restoration of the image/frame of the video [11]. The structural image similarity index (SSIM), in turn, is not suitable for accurately assessing image quality, since it can only assess the similarity of two images/frames of a video stream, and also does not always correctly assess the similarity of the content of visual information [12].

In the research [4], the proposed approach to assessing the masking quality is based on the proximity of the masked data spectrum to white noise, however studies were carried out only for audio data, and further research is required for visual information.

Thus, at the moment, the search for a masking quality metric remains an urgent task.

CORRELATION APPROACH TO ASSESSING THE QUALITY OF MASKING

Any real digital image/video stream frame has a strong spatial correlation of pixels, and, as a result, has redundancy. An integral part of many compression algorithms is the procedure of pixel decorrelation; in the paper [12], it is also proposed to use correlation to assess the similarity of an image to a standard. In this regard, within this work, we will evaluate the influence of the size of the masking matrix on two parameters – the correlation of image pixels and the two-dimensional correlation coefficient.

Matrix masking can be conducted in two versions. The first is one-way masking, in which the original image (or image fragment) $\mathbf{X}_n$ size $n \times n$ is multiplied by a matrix $\mathbf{M}_n$ of the same size in the form:

$$\mathbf{Y}_n = \mathbf{X}_n\mathbf{M}_n, \tag{1}$$

where **Yn** – A protected image transmitted digitally over a communication channel.

Matrix **Mn** henceforth we will call as a key matrix.

The second option is two-way masking, in which the original image (or its fragment) is multiplied by the key matrix $\mathbf{M}_n$ on the left and the transposed key matrix $\mathbf{M}_n^T$ on the right in the form:

$$\mathbf{Y}_n = \mathbf{M}_n \mathbf{X}_n \mathbf{M}_n^T, \tag{2}$$

Inverse transformations to obtain the original image for one-sided and two-sided masking are performed as follows:

$$\mathbf{X}_n = \mathbf{Y}_n \left(\mathbf{M}_n\right)^{-1}, \tag{3}$$

$$\mathbf{X}_n = \left(\mathbf{M}_n\right)^{-1} \mathbf{Y}_n \left(\mathbf{M}_n^T\right)^{-1}. \tag{4}$$

The usage of orthogonal or *quasi*-orthogonal matrices **Mn**, for which $(\mathbf{Mn})^{-1} = \mathbf{M}_n^T$ for masking simplifies the inverse transformations according to (3) and (4).

The classic image Lena [13], in tif format, was chosen as a test image.

For masking, two-level matrices were used:

1) Mersenne structured according to Walsh [14] is a symmetrical structure.

2) Based on the M-sequence [15] – a cyclic structure.

Tables 1-3 show estimates of vertical (r_{vert}), horizontal ($r_{horizon}$), and diagonal (r_{diag}) correlations of pixels. Moreover, Table 1 shows the characteristics of the original image, and Tables 2 and 3 for matrices of symmetric and cyclic structure, respectively.

Table 1

	r $_{vert}$	r $_{horizon}$	r $_{diag}$
Original image	0.9846	0.9668	0.9595

Table 2

Masking matrix size	One-sided			Two-sided		
	r $_{vert}$	r $_{horizon}$	r $_{diag}$	r $_{vert}$	r $_{horizon}$	r $_{diag}$
7x7	0.9849	0.8535	0.8526	0.9076	0.8477	0.7990
15x15	0.9849	0.7711	0.7723	0.8624	0.7812	0.7185
31x31	0.9870	0.5863	0.5865	0.7210	0.6431	0.4818
63x63	0.9873	0.5290	0.5345	0.7739	0.6628	0.5892
127x127	0.9848	0.2365	0.2392	0.5213	0.2380	0.1758
255x255	0.9849	0.1304	0.1350	0.3100	0.1876	0.1983
511x511	0.9848	-0.0890	-0.0902	0.2564	-0.0314	0.0567

Table 3

Masking matrix size	One-sided			Two-sided		
	r $_{vert}$	r $_{horizon}$	r $_{diag}$	r $_{vert}$	r $_{horizon}$	r $_{diag}$
7x7	0.9848	0.9199	0.9100	0.9532	0.9191	0.8875
15x15	0.9849	0.9107	0.9000	0.9427	0.9048	0.8587
31x31	0.9862	0.9221	0.9111	0.9488	0.9292	0.8823
63x63	0.9858	0.9397	0.9276	0.9533	0.9386	0.8949
127x127	0.9849	0.9549	0.9423	0.9700	0.9564	0.9298
255x255	0.9849	0.9642	0.9516	0.9806	0.9667	0.9498
511x511	0.9850	0.9691	0.9565	0.9841	0.9691	0.9556

To calculate the pixel correlation, 5000 two adjacent image elements were randomly selected, and the calculation was made using the following formula:

$$r_{x,y} = \frac{\sigma_{x,y}}{\sqrt{\sigma_x^2 \sigma_y^2}},$$

where x and y are the values of two adjacent pixels in the grayscale image. $\sigma_{x,y}$ – covariance, σ_x^2 and σ_y^2 – variances of random variables x and y, respectively.

Thus, analysis of the results obtained in Tables 1-3 allows us to numerically confirm the statements that two-side masking made in [5] gives better results in comparison with one-side masking. Additionally, we can conclude that, in the general case, masking with matrices of a symmetrical structure provides better decorrelation of image pixels, and, as a consequence, better destruction of the structure of the original image.

The following formula was used to calculate the bivariate correlation coefficient:

$$r = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X_{ij} - \overline{X}\right)\left(Y_{ij} - \overline{Y}\right)}{\sqrt{\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X_{ij} - \overline{X}\right)^2\right)\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(Y_{ij} - \overline{Y}\right)^2\right)}},$$

where X and Y are the original and masked images, respectively. $\overline{X}$ and $\overline{Y}$ are the average values of X and Y respectively. M and N are the vertical and horizontal dimensions of the original and masked images.

The values of the two-dimensional correlation coefficient for images masked with a symmetric and cyclic matrix are given in Table 4 and Figures 1–3.

Table 4

lena_std.tif

Masking matrix size	Symmetrical structure		Cyclic structure	
	One-sided	Two-sided	One-sided	Two-sided
7x7	0.8701	0.8360	0.9425	0.9160
15x15	0.7583	0.6992	0.8594	0.8010
31x31	0.5718	0.4839	0.5136	0.4624
63x63	0.3416	0.2537	0.1998	0.2129
127x127	0.2275	0.1618	-0.1781	0.0678
255x255	0.1498	0.1142	-0.1530	0.0919
511x511	0.0259	-0.0074	-0.5327	0.1175

The results, presented in Figures 1 – 3, are graphs of the dependence of the two-dimensional correlation coefficient on the size of the masking matrix for three classes of images: aerial photographs, textures and pronounced objects. Upon completion of image processing, a visual experiment was carried out, during which an observer, unfamiliar with the source material, was shown the results of masking in descending order of matrix size. On the graphs, green markers indicate those values of the correlation coefficient at which the observer could already accurately determine what was depicted, and yellow markers indicate values at which there were still doubts about the correctness of the assumption or other versions were present.

The results obtained in Table 4 are proposed to be interpreted using the Chaddock scale [16, 17]. It is assumed that moderate (0.3 – 0.5) and weak correlation (0.1 – 0.3) will provide sufficient destruction of the structure of the original image for any masking option (two-way or one-way). The tabular data and graphical masking results should be compared to enable joint analysis. Fig. 4 and 5 show, respectively, the results of one-way and two-way matrix masking using symmetrical Mersenne matrices structured according to Walsh.
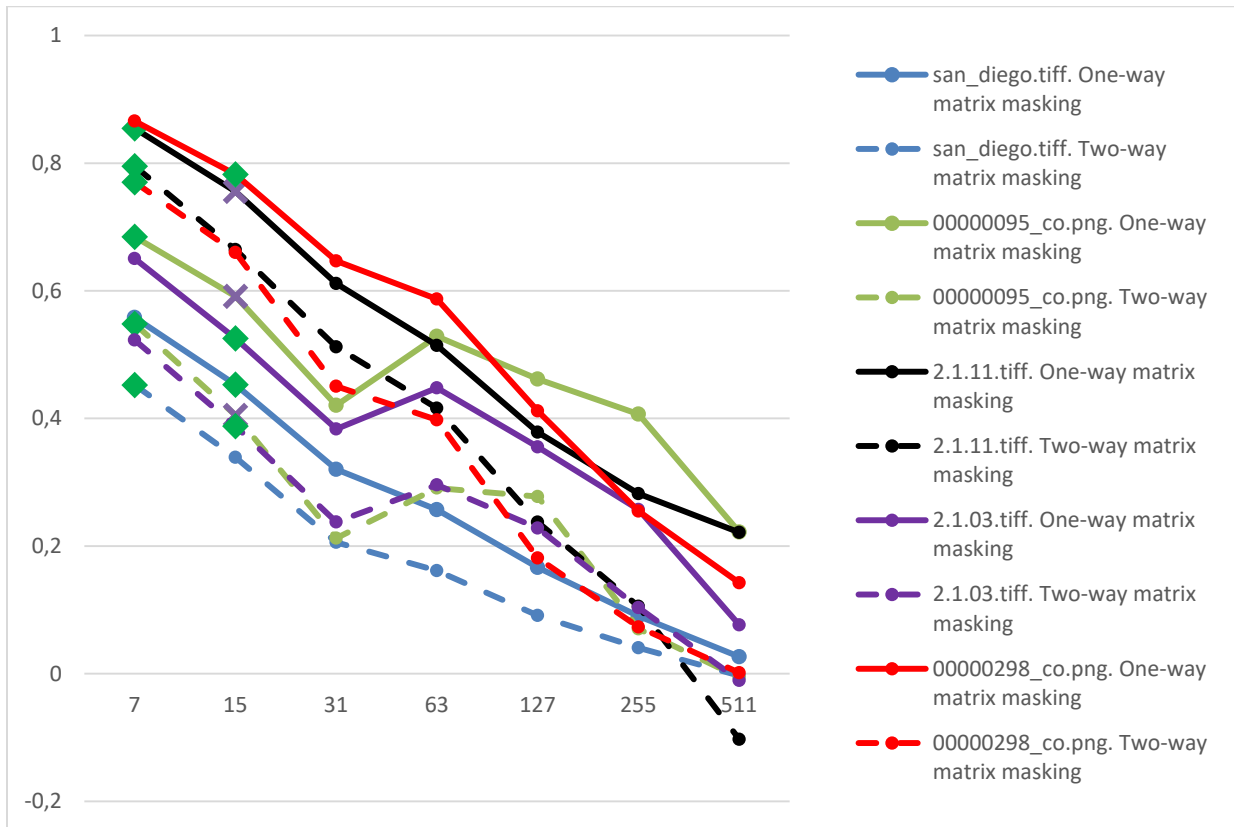
a)
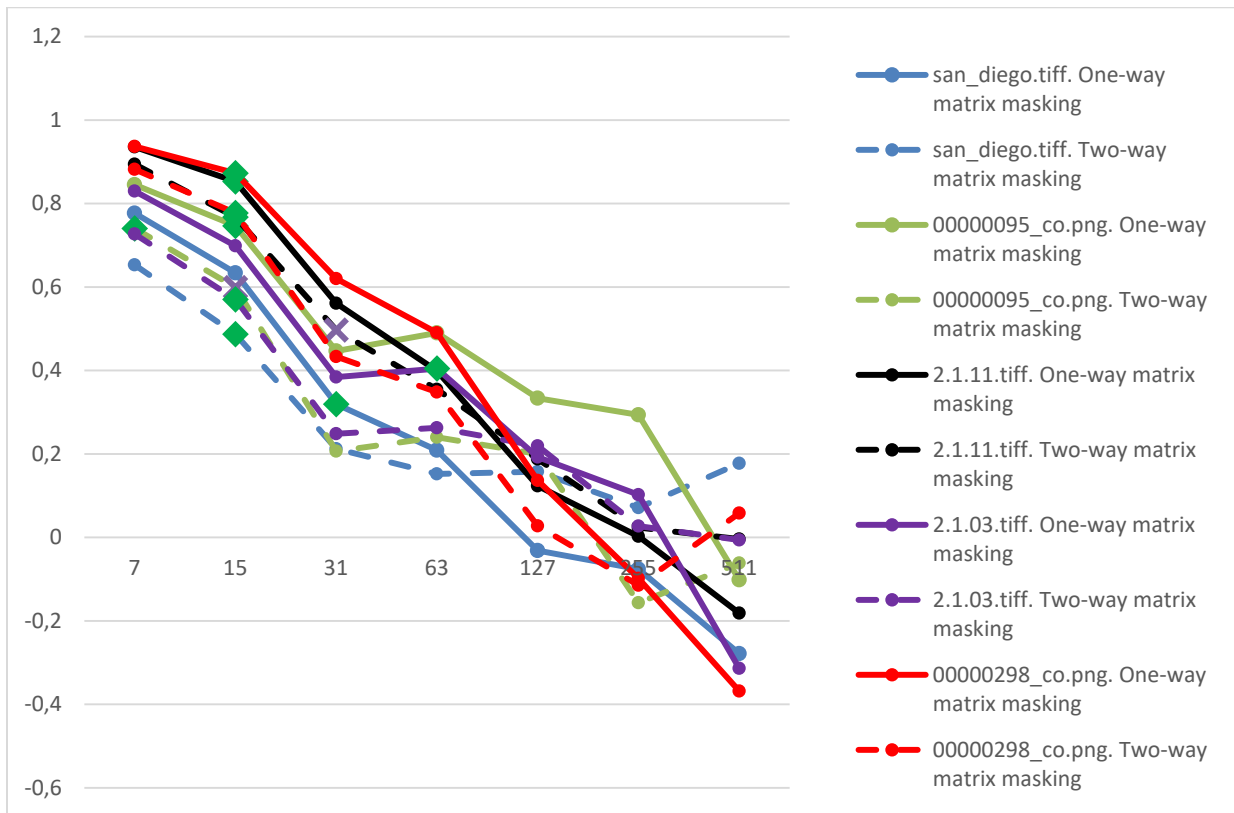
b)

Fig. 1. Graph of the dependence of the two-dimensional correlation coefficient on the size of the matrix of a symmetrical structure (a), cyclic structure (b) for aerial photographs
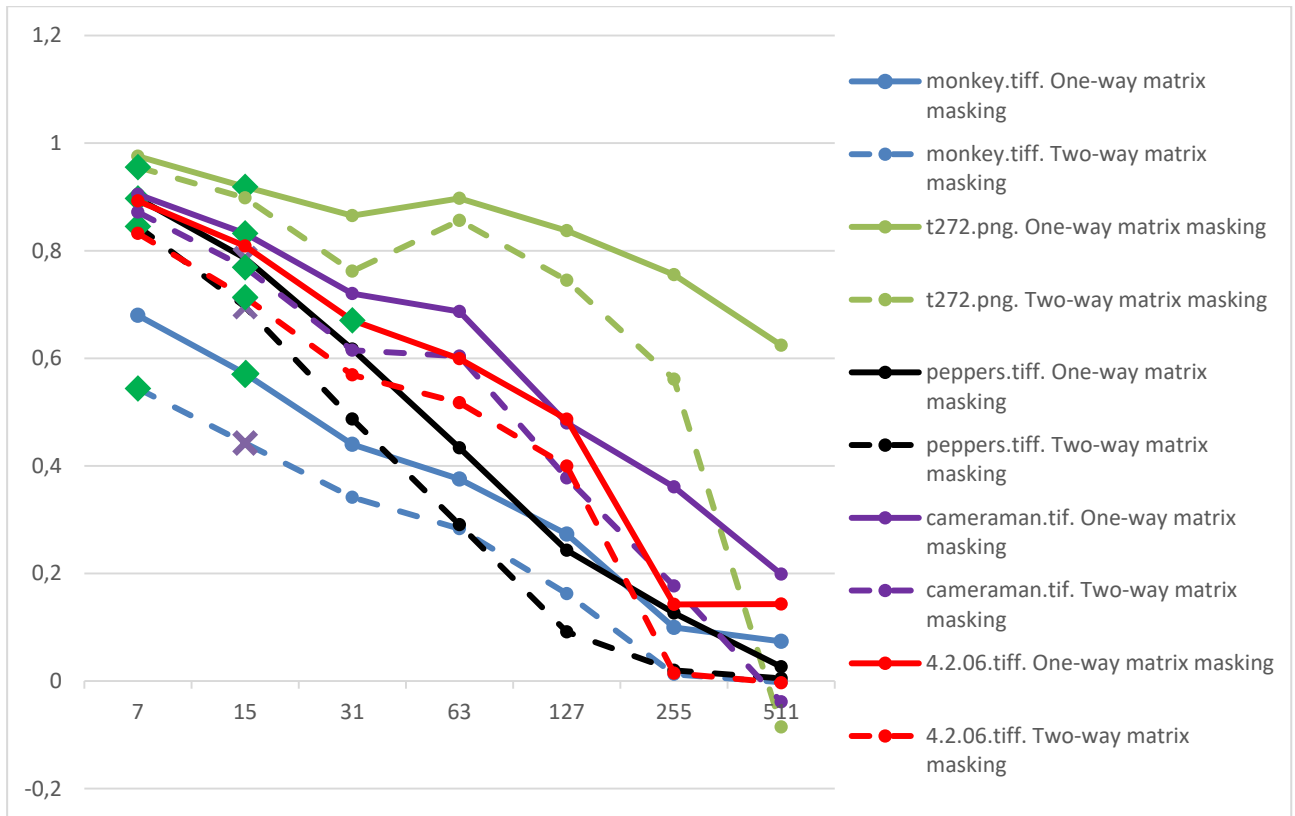
a)
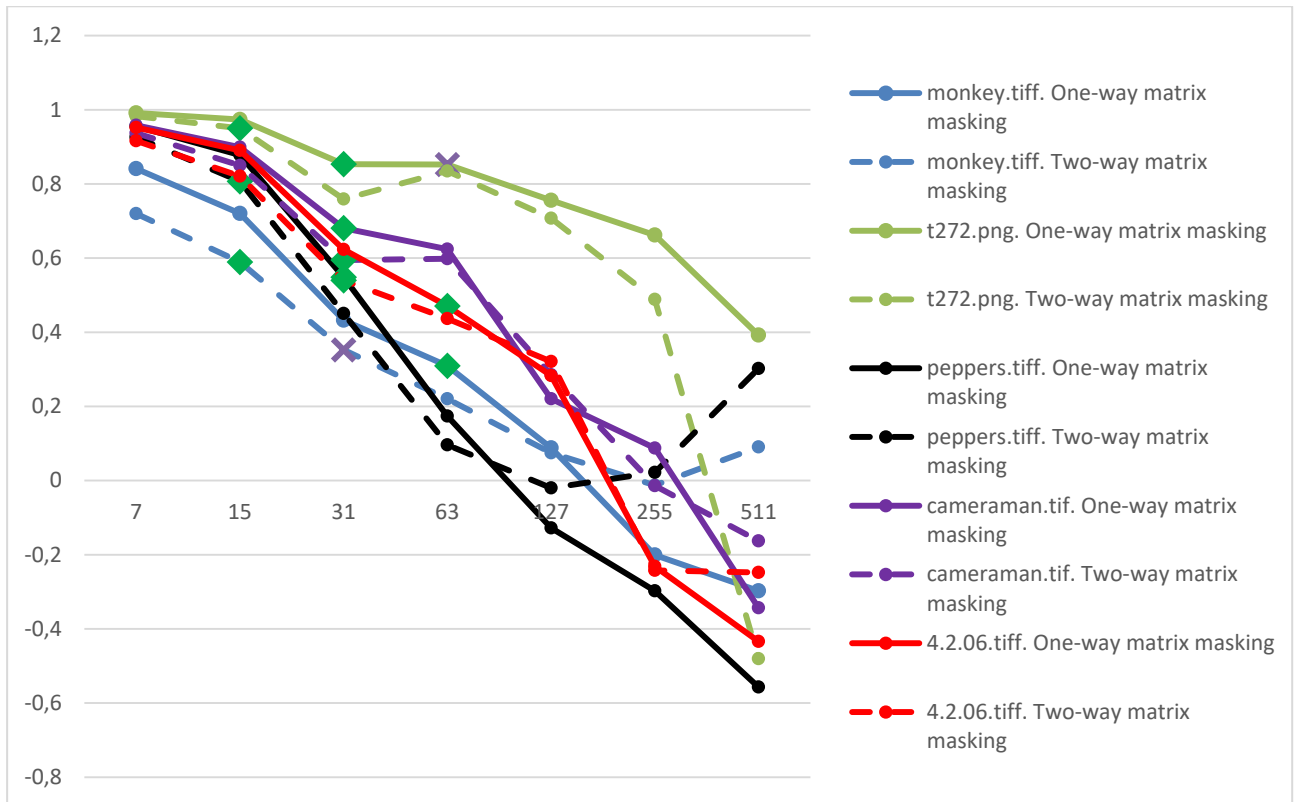
b)

Fig. 2. Graph of the dependence of the two-dimensional correlation coefficient on the size of the matrix of a symmetrical structure (a), cyclic structure (b) for images of pronounced objects
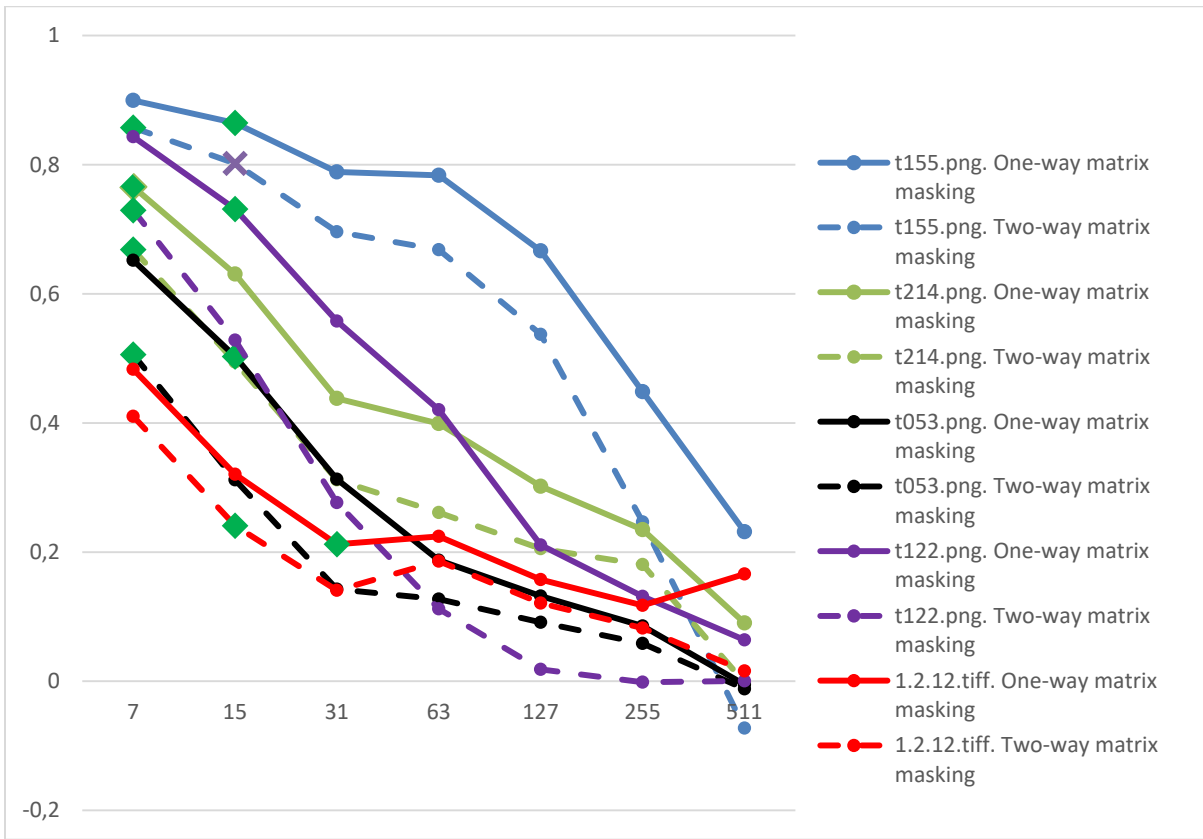
a)
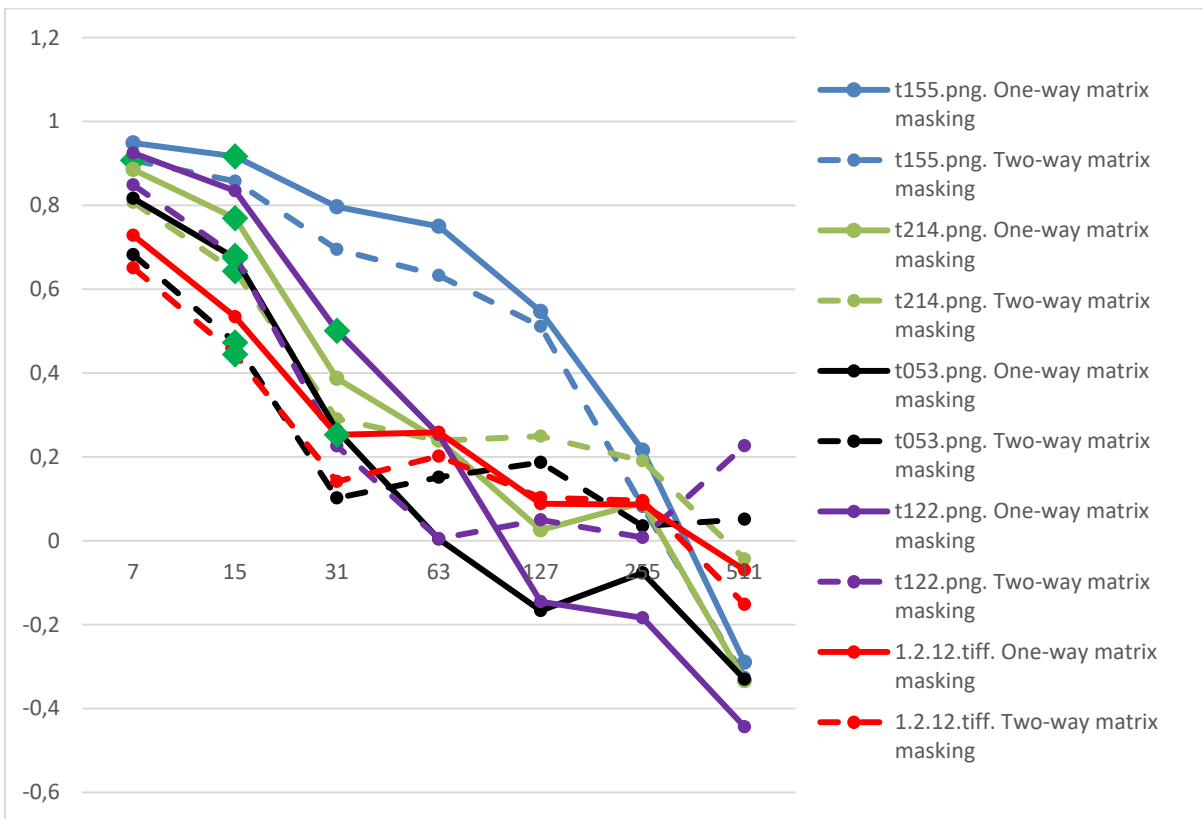
b)

Fig. 3. Graph of the dependence of the two-dimensional correlation coefficient on the size of the matrix of a symmetrical structure (a), cyclic structure (b) for texture images

| *a) 15x15, r=0.7583* | *b) 63x63, r=0.3416* | *c) 31x31, r= 0.5718* |
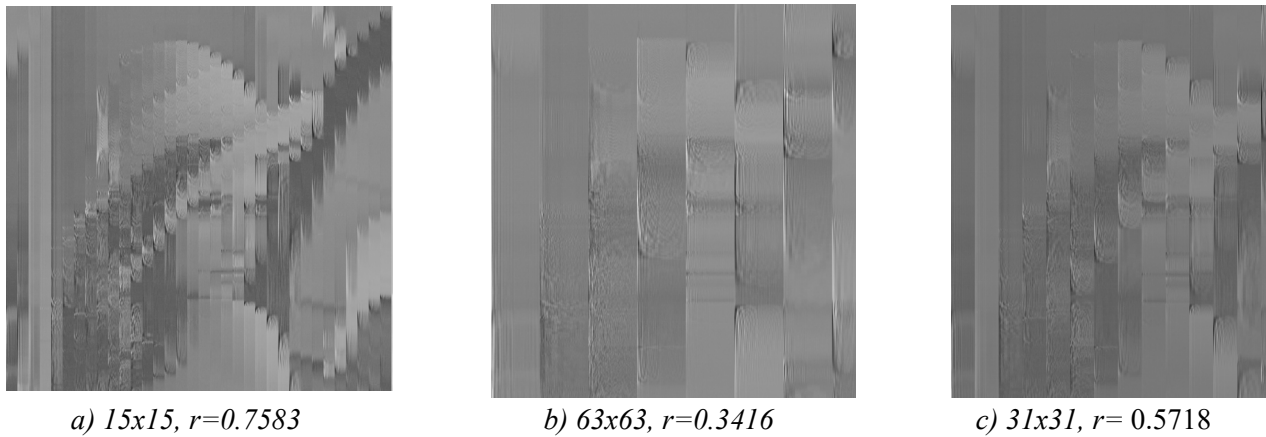
Fig. 4. Result of one-way masking with matrices of different orders

| *a) 15x15, r=0.6992* | *b) 63x63, r= 0.2537* | *c) 31x31, r=0.4839* |
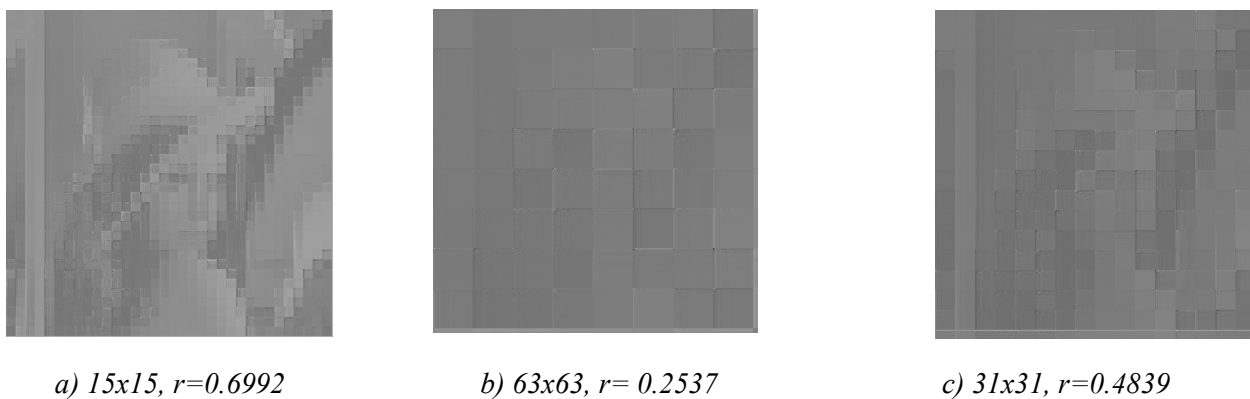
Fig. 5. Result of two-way masking with matrices of different orders

The results of the experiment with an observer showed that a sufficient degree of destruction of the original image structure was achieved when the two-dimensional correlation coefficient was no higher than 0.65 for pronounced objects and textures and from 0.54 to 0.78 for aerial photographs.

The results of visual analysis allow us to make the assumption that if for some reason it is impossible to provide the masking matrix sizes and the original image sizes equality (the best possible masking quality), then setting a threshold for the two-dimensional correlation coefficient of no higher than 0.5 will ensure sufficient masking quality.

Conclusion

The paper proposes a simple approach to assessing the quality of image masking, based on estimates of the vertical, horizontal and diagonal correlation of masked image pixels, as well as an estimate of the two-dimensional correlation coefficient. This approach is acceptable due to the fact that when masking an image/frame of a video stream, there is always access to an unprotected image to calculate the necessary parameters.

It has been numerically confirmed that two-side masking with matrices of symmetrical structure gives better results for this procedure in comparison with the use of one-sided masking or matrices of cyclic structure.

Based on the analysis of the results of masking all three classes of images together, as well as a visual test using the example of the image "Lena" and an experiment with an independent observer, two thresholds for the bivariate correlation coefficient can be preliminarily established. The first threshold is not higher than 0.5, at which acceptable masking quality is ensured, and the threshold is not higher than 0.2, at which for each class of images an outside observer could not extract useful information from the masked image. These thresholds can be set if for some reason it is impossible to ensure equality in the size of the image and the masking matrix.

Acknowledgment

References

1. Ziatdinov R. Visual Perception, Quantity of Information Function and the Concept of the Quantity of Information Continuous Splines // Scientific Visualization. 2016. Vol. 8, Iss . 1. P. 168-178.

2. Gradyushko A.A. Instagram and TikTok platforms in the digital space: a comparative aspect // Proceedings of BSTU. Ser. 4, Print and media technologies. 2021. No. 1(243). pp. 12 -19.

3. Kruglova L.A., Konovaltseva A.O. Russian television channels on the YouTube platform // Bulletin of the Russian Peoples' Friendship University. Ser.: Literary Studies. Journalism. 2020. T. 25, No. 2. P. 351-359.

4. Grigoriev E.K., Sergeev A.M. Assessing the quality of matrix masking of digital audio data // Proceedings of educational institutions of communication. 2023. 9(3). pp. 6-13.

5. Vostrikov A. A., Sergeev M. B., Litvinov M. Yu. Masking of digital visual information: term and basic definitions // Information and control systems. 2015. T. 5. No. 78. P. 116-123.

6. Balonin N. A., Sergeev M. B. Matrices of the local maximum of the determinant // Information and control systems. 2014. No. 1(68). pp. 2-15.

7. Nenashev V.A., Sergee A.M., Kapranova E.A., Ryzhov K.Yu. Experiments on replacing the DCT with a quasi-orthogonal transformation in image compression algorithms // Scientific session of the SUAI: pp . report _ At 3 o'clock Part II . Technical science. SPb.: GUAP. 2018. pp. 369-373.

8. Vostrikov A.A., Chernyshev S.A. On assessing the resistance to distortion of images masked by M-matrices // Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics. 2013. T. 5. No. 87. P. 99-103.

9. Vostrikov A. A., Mishura O. V., Sergeev A. M., Chernyshev S. A. On the selection of matrices for image masking and unmasking procedures // Fundamental Research. 2015. T. 2. No. 24. P. 5335-5339.

10. Chernyshev S. A. Development and research of a method of matrix masking of video information in globally distributed systems: diss . ...cand. tech. Sci. Saint Petersburg, 2018. 120 p.

11. Erosh I. L., Sergeev A. M., Filatov G. P. On the protection of digital images during transmission via communication channels // Information and control systems. 2007. T. 5. No. 30. pp. 20-22.

12. Starovoitov V.V. The ssim index is not a metric and poorly assesses the similarity of images // Informatics. 2019. no. 2. pp. 1–17

13. The Lenna Story. URL: http :// lenna . org / (date appeal 11/15/2023).

14. Sergeev A. M. Walsh-structured two-level and modular two-level quasi-orthogonal matrices for image masking // News of higher educational institutions. Instrumentation. 2023. T. 66, No. 5. P. 399-408

15. Nenashev V.A., Grigoriev E.K., Sergeev A.M., Samokhina E.V. Strategies for calculating persymmetric cyclic quasi-orthogonal matrices as the basis of codes // Electrosvyaz. 2020. No. 10. pp. 58-61.

16. Chaddock RE. Principles and methods of statistics. Boston, New York, [etc.]. 1925. 471 p

17. Koterov A.N., Ushenkova L.N., Zubenkova E.S., Kalinina M.V., Biryukov A.P., Lastochkina E.M., Molodtsova D.V., Vainson A.A. The power of connection. Message 2. Gradations of the correlation value // Medical radiology and radiation safety. 2019. T. 64. No. 6. P. 12–24.

QAM-64 IMPLEMENTATION IN MATLAB ENVIRONMENT

Boris Kleshnin

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
e-mail: boria456@gmail.com

Abstract: *the paper considers the process of creation and the principle of operation of QAM modulation. The theoretical signal modeled in MATLAB is calculated in detail. The advantages and disadvantages of this type of modulation are described.*
Keywords: *signal constellation, additive white Gaussian noise, QPSK, spectrum*

In modern radio electronics, the question of ways to protect the transmission of information from interference caused by natural conditions or created by human hands (EW) is becoming more urgent. Modulation is used to solve this problem. In addition to amplitude, frequency and phase modulation, there is also QAM (QPSK, QAM-16, QAM-64, etc.) modulation. This type of modulation has great noise immunity, allows for more efficient use of the spectrum, and is applicable to various standards[1].

Quadrature Amplitude Modulation (QAM) is a modulation technique that provides multiple amplitude and phase levels for efficient data transmission. The key element when using QAM is the constellation of signals (Fig. 1a). They display the amplitude and phase values for each symbol in the modulated signal. Next, the process will be simulated using QAM-64 in the MATLAB environment. It is worth noting that the qammod command already exists in MATLAB, however, due to the specifics of the task, its use will be minimized in order to show the entire process in detail[2].

Any signal consists of harmonics. Each harmonic has its own amplitude and phase. According to this data, encoding/decoding takes place by means of QAM modulation (Fig. 1b).
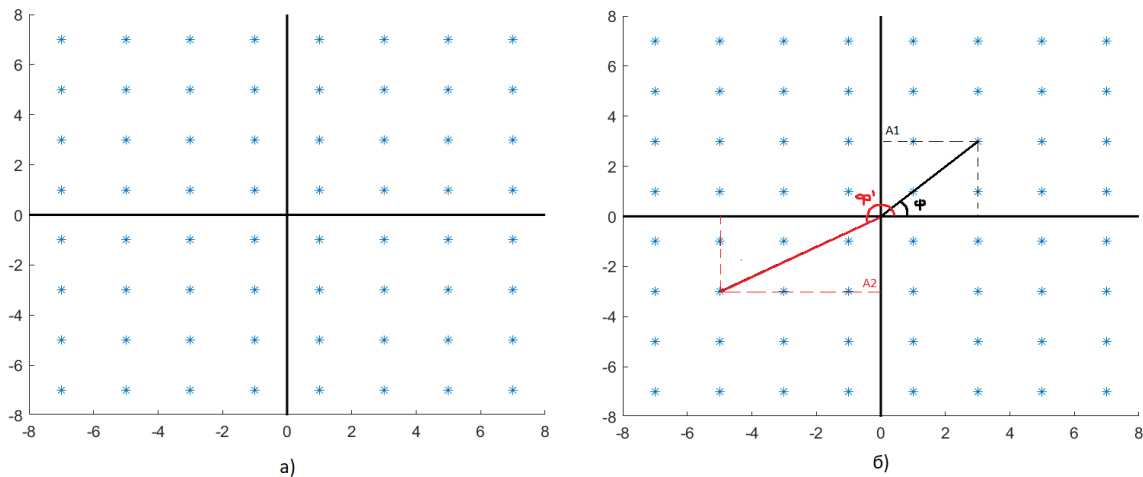
Fig. 1. a) A signal constellation; b) A graphical representation of the method of encoding amplitudes and phases of harmonics using QAM-64.

First, you need to create a vector of complex numbers (signal simulation) with a specified length parameter (vector_size)[3]. To create such a vector, the complex_vector function was created.

Listing of the complex_vector function:

```
function random_complex_vector = complex_vector(vector_size)
    a = -7:2:7;
    real_part64 = [];
    for i = 1:vector_size
        d = datasample(a, 1);
        real_part64 = [real_part64, d];
    end

    imag_part64 = [];
```

```
   for i = 1:vector_size
      d = datasample(a, 1);
      imag_part64 = [imag_part64, d];
   end

   random_complex_vector = complex(real_part64, imag_part64);
end
```

It is noteworthy that the signal constellation uses a signal with a maximum amplitude of 7. Increasing the amplitude of the signal can improve its noise immunity, since higher amplitudes provide wider distances between the points of the constellation, which makes it possible to better distinguish signals in conditions of noise and interference[4]. However, an increase in amplitude also contributes to an increase in signal energy, which can be costly in terms of transmission power and resource requirements. Thus, there is a compromise between noise immunity and power consumption, which is selected depending on the specific requirements of the communication system.

The signal energy is calculated using the formula:

$$E_s = \sum_{n=1}^{N} |x[n]|^2$$

Where N is the number of samples in the signal

Hence, the energy of the received signal is 42.27 cu.

After creating a theoretical signal, it is necessary to apply additive white Gaussian noise. As a result, the previously modeled signal is given interference (Fig. 2). The recovery of the received signal occurs by finding the error vector (Fig. 3). That is, from each value obtained in the vector of complex numbers with noise, it is necessary to plot a vector to each point of the signal constellation. Since with QAM-64 there will be 64 such vectors for each of the 800 preset values of the signal vector, their location will not be graphically clear, so an example with QPSK modulation is used here.
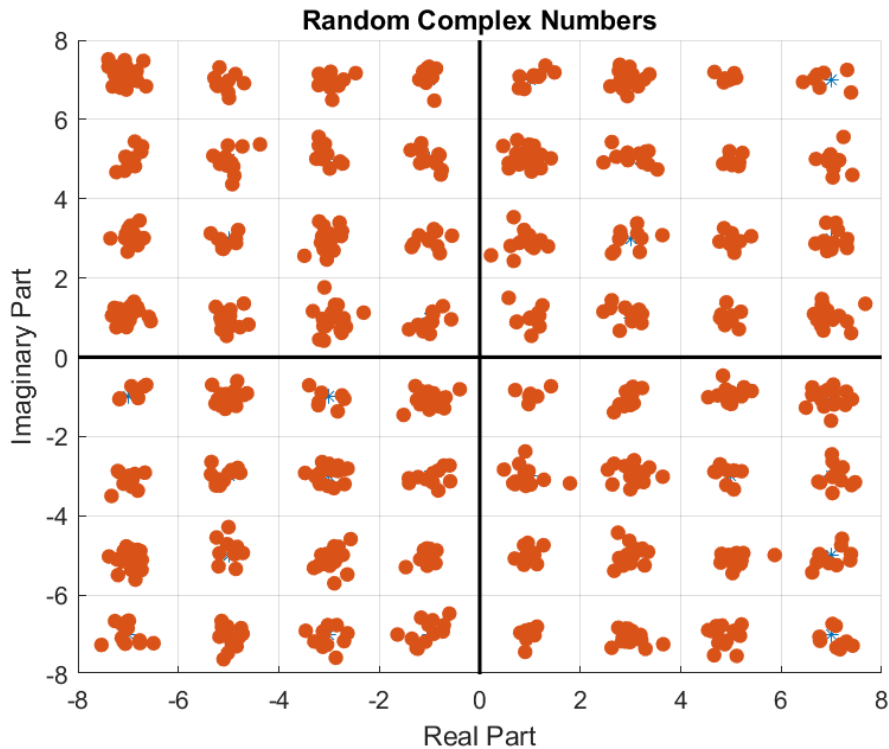
Fig. 2. Simulated theoretical signal with interference on a complex plane

Next, a matrix is created indicating the data of all error vectors for each of the 800 specified points. After that, the smallest ones are determined – they will indicate the initial (before interference) values of the signal[5].

Listing the code in MATLAB:

```
clc
close all
clear all
finaly=[];
vector_size = 800;
m1=complex_vector(vector_size);
y = awgn(m1,10);
qam64_points=[];
qam4_points = [1 + 1i, 1 – 1i, -1 + 1i, -1 – 1i];
for k=0:7
    for n=0:7
qam64_points = [qam64_points,qam4_points(1,1)+2*(k-4)+(n-4)*2i];
    end
end
figure
scatter(real(qam64_points), imag(qam64_points),'*');
hold on
axis;
line([0, 0], ylim, 'Color', 'k', 'LineWidth', 1.5);
line(xlim, [0, 0], 'Color', 'k', 'LineWidth', 1.5);
hold on
scatter(real(y), imag(y), 'filled');
xlabel('Real Part');
ylabel('Imaginary Part');
title('Random Complex Numbers');
grid on;
for u=1:vector_size
random_point = y(u);
distance_to_qam64 = abs(qam64_points – random_point);
g=min(distance_to_qam64);
for i=1:64
if g==distance_to_qam64(1,i)
    g=qam64_points(1,i)
end
end
finaly=[finaly,g];
end
signal_energy = sum(abs(m1).^2) / length(m1);
sum(finaly-m1)
```
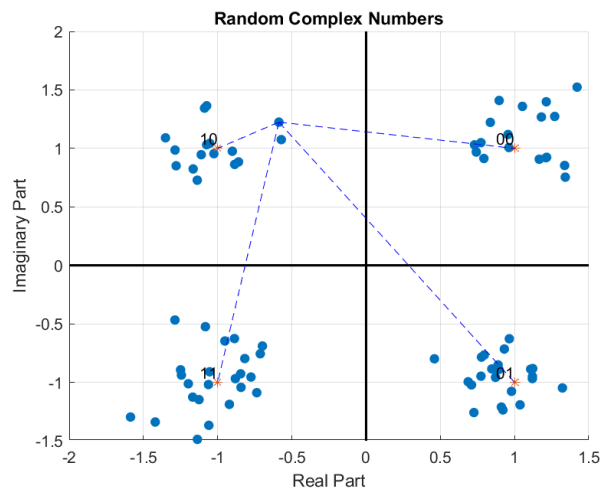

Fig. 3. Error vectors for one of the signal points in QAM-4.

To verify the correctness of the decoding results, it makes sense to output the result of the difference between the two vectors "sum(finaly-m1)". The result is 0. Which means that the signal is completely restored without loss. If you add a level of Gaussian noise, this result will no longer work. We will have to increase the energy (increase the coordinates of the points in the signal constellation).

References

1. Proakis, J. G., & Salehi, M. (2002). Contemporary Communication Systems Using MATLAB. PWS Publishing – pp. 115-178

2. Beklemishev, D. V. (2006). Digital modulation and parameterization of signals. Solon Press. – pp. 11-98

3. Makarenko, A. S. (2016). MATLAB and its tools. BHV-Petersburg

4. Haykin, S., & Moher, M. (2018). Modern Wireless Communications. Pearson – pp. 123-214

5. Zubarev, V. P., Zubareva, L. P., & Zubarev, V. V. (2003). Theory and technology of ultra-wideband radio systems. Radio and communications – pp. 143-203

COMPARATIVE ANALYSIS OF NETWORK LAYER HEADER COMPRESSION

Anastasiya Kolesnikova

Saint Petersburg State University of Aerospace Instrumentation,
B.Morskaya 67, 190000, Saint Petersburg, Russia
E-mail: anast.kolesnikova@mail.ru

Abstract. *In today's world, where data transmission plays a key role in many aspects of life and business, efficient use of network resources has become increasingly important. In this context, the development of methods for compressing IP packet headers is an urgent task, since it allows reducing the amount of transmitted data and, therefore, improving network throughput. The need to minimize transmitted data is especially acute in the face of growing traffic volumes and increasing requirements for information transfer speeds. Thus, the development and application of IP packet header compression algorithms has direct practical significance for optimizing the operation of networks and increasing their performance.*
Keywords: *compression, decompression, header, RoHC*

Introduction

In modern network technologies, efficient management of transmitted information is becoming a key aspect of ensuring high performance and optimal use of resources. One method for optimizing network data transmission is to compress packet headers at the network level. This paper provides a comparative analysis of various packet header compression algorithms, focusing on their efficiency, compression ratio, and impact on network performance. An analysis of the methods used and the research results will help to understand which approaches to header compression best combine the trade-off between saving bandwidth and ensuring the integrity of the transmitted data.

TCP/IP header compression for low-speed serial links (CTCP)

The compressor converts each input packet into a TYPE_IP, UNCOMPRESSED_TCP, or COMPRESSED_TCP packet:

- TYPE IP – the packet is an unmodified copy of the input packet, and its processing does not change the state of the compressor in any way;

- UNCOMPRESSED TCP – this packet is identical to the input one, except that the IP protocol field is changed from "6" (TCP protocol) to the connection number.

- COMPRESSED TCP – IP and TCP headers are completely replaced with a new compressed header;
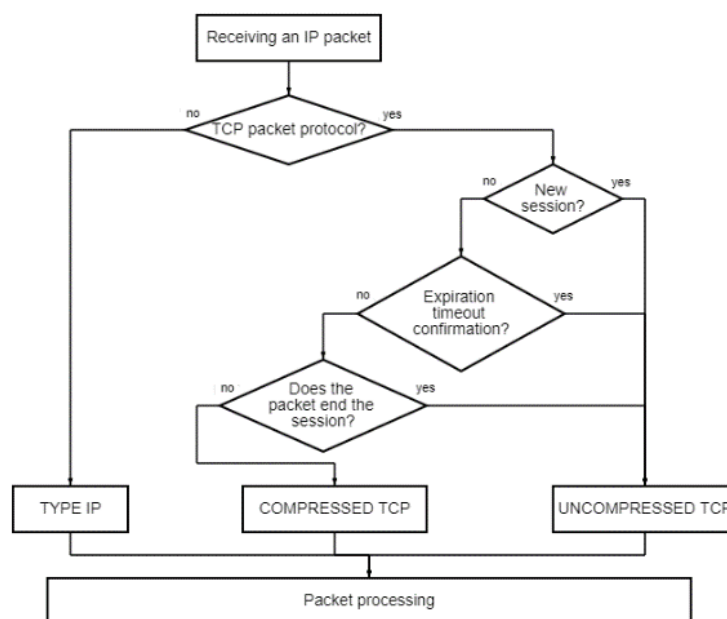
Fig. 1. Algorithm of operation of the CTCP compressor

IP Header Compression (IPHC)
IP Header Compression (IPHC) uses different types of headers during packet transmission
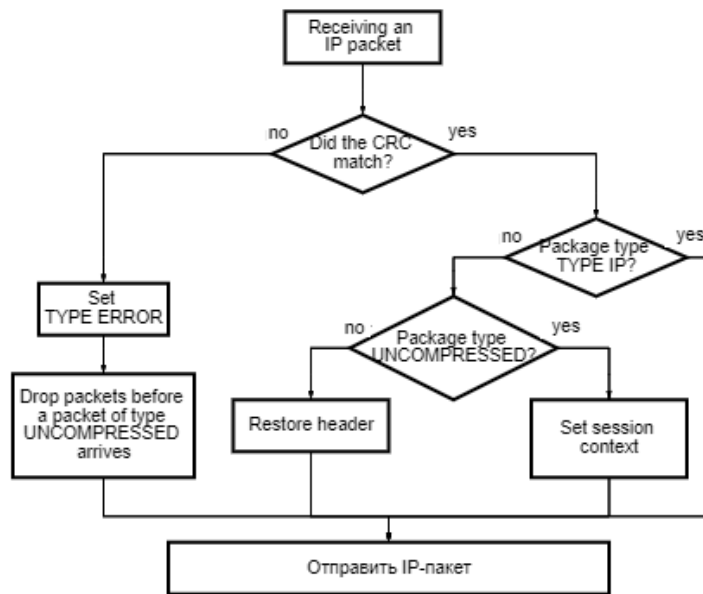
Fig. 2. Algorithm of operation of the CTCP decompressor

between the compressor and decompressor:

FULL_HEADER – indicates a packet with an uncompressed header, including the connection number and, if the packet is not TCP, then the source of generation is additionally included.

COMPRESSED_NON_TCP – Indicates a packet with a compressed header that is not TCP.

COMPRESSED_TCP – Indicates a packet with a compressed TCP header containing the connection number, a flag indicating which fields were modified, and the modified fields encoded differently from the previous value.

COMPRESSED_TCP_NODELTA – indicates a packet with a compressed TCP header in which all fields are unchanged and sent as is.

CONTEXT_STATE – indicates a special packet sent from decompressor to compressor to carry a list (TCP) of CIDs for which synchronization has been lost.
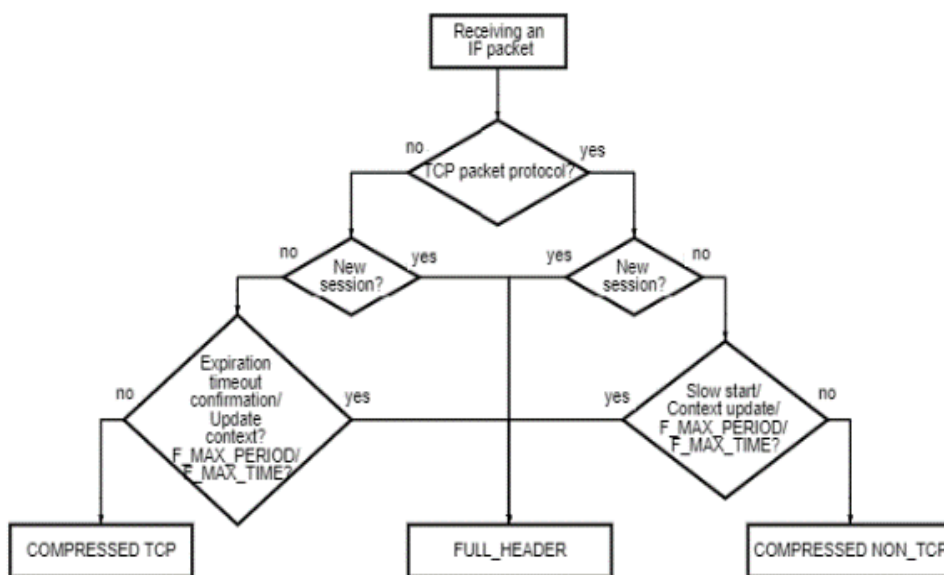
Fig. 3. IPHC compressor operation algorithm
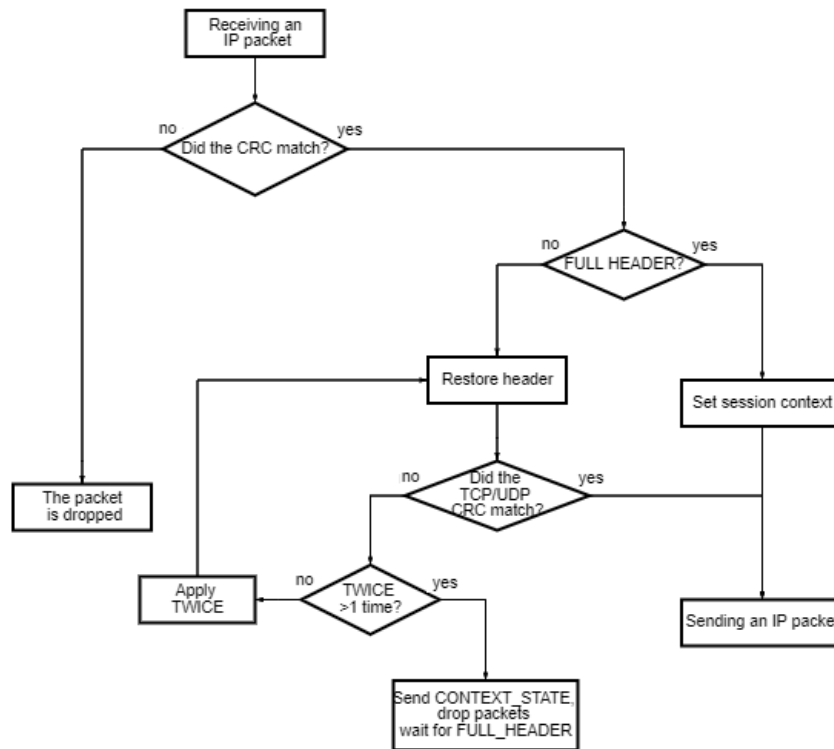
Fig. 4. IPHC decompressor operation algorithm

IP/UDP/RTP header compression for low-speed serial links (CRTP)

Compression RTP (CRTP) requires the link layer to provide an indication of four new packet formats in addition to the usual IPv4 and IPv6 formats for the transmission of uncompressed and compressed packets: COMPRESSED_RTP – the RTP header is compressed along with the IP and UDP headers. The size of this header is two bytes or larger, depending on the need to convey field differences. This type of packet is used when the second order difference in constant fields is zero. In this case, delta encoding is applied to fields that have changed by an amount different from what was expected to determine the difference of the first-order fields after the uncompressed RTP header was sent.
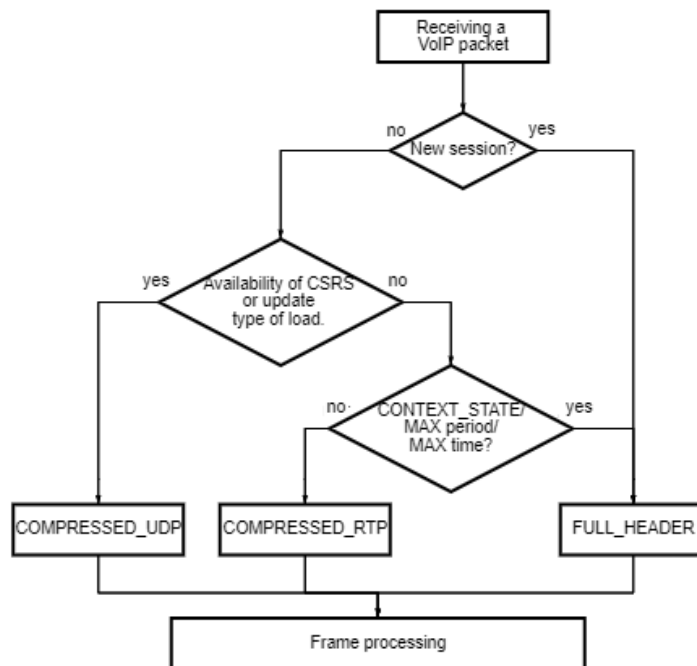
Fig. 5. Operation algorithm of the CRTP compressor
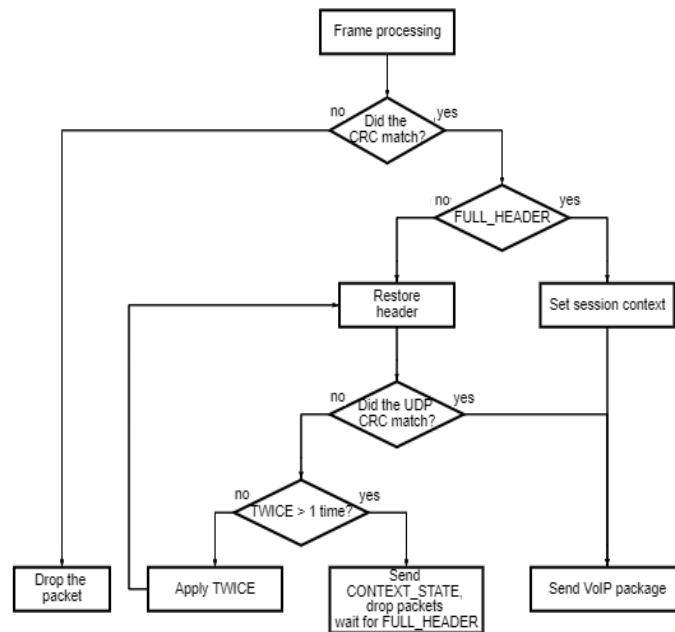
Fig. 6. Operation algorithm of the CRTP decompressor

CONTEXT_STATE – indicates a special packet sent from decompressor to compressor to convey a list of context identifiers where synchronization has been or could be lost. This packet is sent exclusively over a point-to-point link, so it does not require an IP header.

COMPRESSED_NON_TCP – transmits compressed IP and UDP headers without differential encoding. If this type were used for IPv4, storing the identifier field would require one or two more bytes than COMPRESSED_UDP.

Compression of RTP/UDP/SCP headers on channels with high packet loss (RoHC)

In RoHC there are several states that a compressor can be in during a session:

Initialization and Refresh (IR) – in this state, the static components of the context are initialized in the decompressor or to recover from a failure.

First Order (FO) – in this state, information about violations in the packet flow is sent.

Second Order (SO) is a state of optimal compression.

The compressor can operate in three modes:

Unidirectional Mode – In this mode there is no feedback channel between the decompressor and the compressor.

Bi-directional Optimistic Mode – This mode has a feedback channel, but it is only used when requesting a context update if the decompressor fails.

Bidirectional Reliable Mode – In this mode, the compressor waits for confirmation from the decompressor that the headers have been restored correctly.
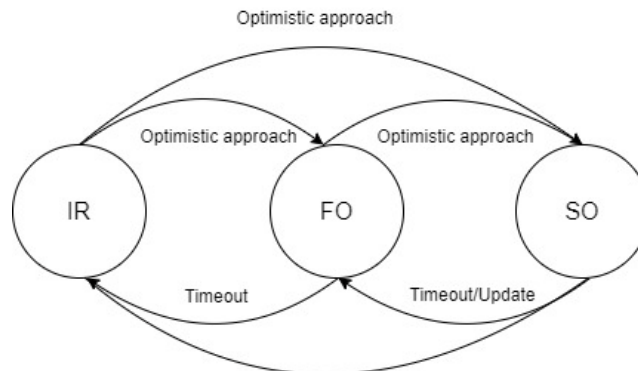
Fig. 7. RoHC compressor state diagram in U-Mode
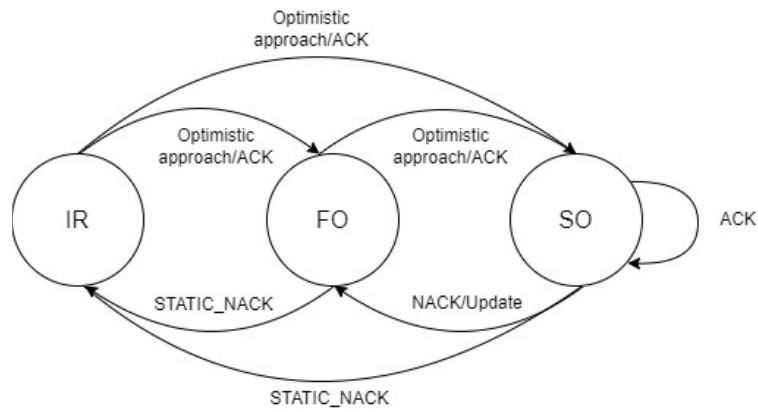
Fig. 8. RoHC compressor state diagram in O-Mode

The decompressor also has three states: NO_CONTEXT – the state of the decompressor when initializing the session, STATIC_CONTEXT – state corresponding to the first compression order for the compressor. FULL_CONTEXT – decompressor state corresponding to the second compression order for the compressor.

Conclusion

CTCP compresses the 40- byte IP+TCP header into 4 bytes.

he CTCP compressor detects retransmissions at the transport layer and sends an updated header, completely updating the context when they occur. This recovery mechanism does not require explicit signals between the compressor and decompressor. General IP header compression scheme improves on CTCP, capable of compressing arbitrary IP, TCP and UDP headers.
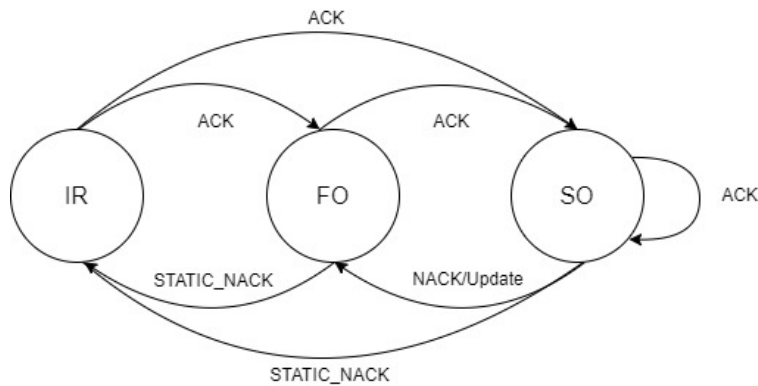
Fig. 9. RoHC compressor state diagram in R-Mode
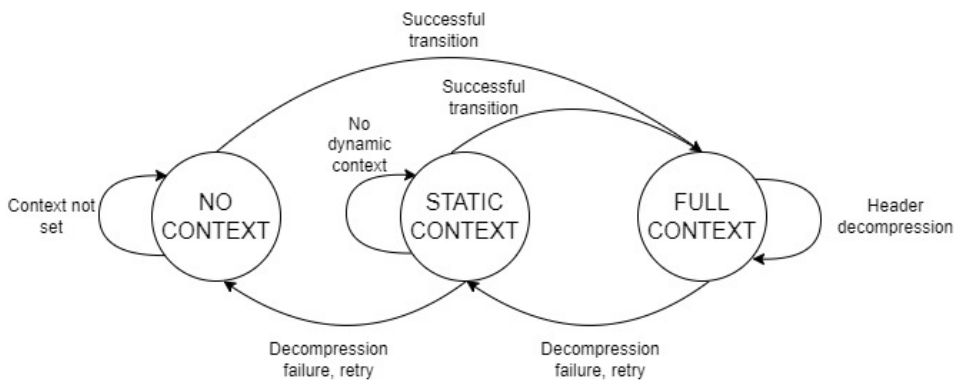
Fig. 10. RoHC decompressor status diagram

When compressing non-TCP headers, IPHC does not use delta encoding and is considered more reliable. In the case of TCP compression, the CTCP recovery mechanism is complemented by a Nack scheme at the link level, which speeds up recovery. However, IPHC does not compress RTP headers.

Technique	Headers	Compressed header size (bytes)	Availability of feedback	Noise-resistant	Error correction method
VJNC	IPv4, TCP	4-7	No	low	No
IPHC	IPv4, IPv6, UDP, TCP	2-5 (not TCP) 4-7 (TCP)	Yes	low	TWICE
cRTP	IPv4, IPv6, UDP, RTP	2 without CRC 4 with CRC	Yes	low	TWICE
RoHC	IPv4, IPv6, UDP, RTP, ESP	1-4	Yes	high	Local correction

Fig. 11. Characteristics of Header Compression Techniques

CRTP is a header compression scheme that reduces 40-byte IPv4/UDP/RTP headers to a minimum of 2 bytes if the UDP checksum is not enabled, and 4 bytes if enabled. CRTP cannot use the same recovery mechanism as CTCP due to the lack of retransmissions in UDP/RTP.

Instead, CRTP uses explicit

decompressor-to-compressor signaling messages called CONTEXT_STATE messages to indicate that the context is out of sync. Thus, CRTP has limited effectiveness on lossy links with long transit times. To avoid losses due to context desynchronization, CRTP decompressors can locally restore context using the TWICE mechanism – each CRTP packet contains a counter that is incremented by one for each packet sent by the CRTP compressor. If the counter increases by more than one, at least one packet is lost on the communication link. It greatly improves the performance of CRTP, but there are some problems when using it.

The way ROHC compression works is to establish functions from sequential numbers (SN) to other fields and then transmit the SN in a reliable form.

Every time a function associated with an SN is changed to a different field, that is, when the current result of the function differs from the corresponding field in the header that is subject to compression, additional information is sent to update the parameters of that function. An assessment of the characteristics of different header compression techniques is given in Fig. 11.

References

1. S. Casner and V. Jacobson Compressing IP/UDP/RTP Headers for Low Speed Serial Links, Request for Comments 2508, Feb. 1999

2. M. Degermark, B. Nordgren, and S. Pink IP Header Compression, Request for Comments 2507. February 1999.

3. V. Jacobson Compressing TCP/IP Headers for Low-Speed Serial Links, Request for Comments 1144, February 1990. [4] C. Bormann RObust Header Compression (ROHC). Request for Comments 3095, July 2001. IETF Network Working Group.

ANALYSIS OF RELIABILITY CALCULATION METHODS

Timofey I. Komarov

student
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
tim1kom@yandex.ru

Abstract. *Ensuring reliability is one of the most important problems in the creation and operation of any technical system. It is especially relevant for complex systems, such as power supply systems, consisting of a large number of elements and having extensive internal and external connections.*
Reliability calculation is necessary in such industries as the automotive industry, the radio-electronic industry, the aviation industry, construction, the energy industry and others. Reliability is an indicator of structural strength. Due to correct reliability calculations, it is possible to assess what loads an object or system can withstand, whether its operation is possible under certain temperature conditions. The paper focuses on the importance of calculating reliability as indicators of durability, safety, and safety of equipment and systems.
Keywords: *reliability, industrial industry, calculation, system, object, method, probability.*

Reliability is understood as an indicator that reflects the degree of quality of an industrial facility throughout the entire time of its operation from the moment of its creation to the moment of unsuitability. Reliability calculation is a method of obtaining an indicator of the reliability of an individual object or an entire system [1, 264].

We focus on the heat and power industry. It is necessary to understand that energy-important facilities must remain operational in different conditions of their installation and operation, respectively, it is impossible to do without calculating reliability in this industry.

There are various classifications of reliability calculation methods. One of them is shown in Fig. 1 [2, 352].
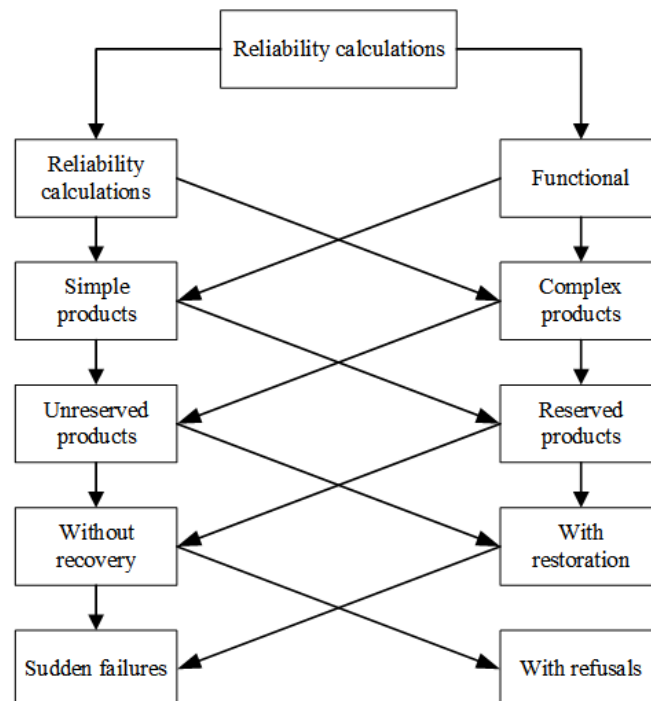
Fig. 1. Classification of reliability calculations in the thermal power industry

The classification of calculation methods is diverse. So, for example, there are methods for calculating reliability indicators. Among them:
− the method of calculating durability;
− the method of calculating reliability;

− the method of calculating forecasting , etc.

According to the principles of calculating reliability properties, forecasting methods, physical methods, and structural methods are distinguished.

Depending on the "period of existence" of the object (system), methods for calculating reliability are distinguished at the design, operation and testing stages. Based on the type, simple calculations and complex calculations are distinguished. Each of these types of calculation has its own variations (calculations of system states, calculations of redundant systems, matrix method, logic circuit method, etc.) [3, 75].

The main indicators of reliability of thermal power facilities are reliability, maintainability [4, 126]. The main methods for calculating these indicators, which in turn mean determining reliability, are structural methods. The essence of such methods is to represent an object (system) in the form of a logical circuit and further calculate reliability indicators.

Let's consider the principle of calculating the sequential connection of elements, which is a subspecies of the structural method for calculating reliability [5, 172]. For example, the system consists of M separate elements.

Let the system consist of N separate elements.

Event Ci is the trouble-free operation of the i–th individual element i = 1,2, ..., M. Event D is the trouble-free operation of the system. The system will work flawlessly when all the individual elements are functioning at the same time, that is (1):

$$D = C_1 * C_2 * ... * C_M = \prod_1^M C_i .\tag{1}$$

For independent events, the formula takes the form (2):

$$P(D) = P(C_1) * P(C_2) * ... * P(C_M) = \prod_1^M P_i(C_i).\tag{2}$$

Accordingly:

$$P(t) = P_1(t) * P_2(t) * ... P_m(t) = \prod_1^M P_i(t) .\tag{3}$$

Where P(t) is the probability of system uptime, $P_1(t) * P_2(t) * ... P_m(t)$ - the probability of operation of individual elements of the object (system).

In the special case, with the same reliability of the elements $(P_1(t) = P_2(t) = ... P_M(t)$ the probability that the system will work flawlessly (that is, it will be reliable) is determined by the following expression:

$$P(t) = \{P_i(t)\}^M .\tag{4}$$

The modern radio-electronic industry is very promising. New more complex devices and equipment are being created, the element base is being updated, and production schemes are being developed. And special requirements are also imposed on the objects of this industry, the key of which is reliability. Reliability tests are necessary to determine how reliable an object (system) is under certain test conditions. As a rule, when calculating reliability, a batch is taken, for example, of products, which are subsequently subjected to calculations and tests.

So, in radio electronics, such reliability calculation methods are used as:

− calculation of the failure rate;

− a consistent method for determining reliability;

− single and double sampling method;

− simulation modeling, etc.

The method of calculating the failure rate was deeply studied by the Russian scientist, Professor N.K. Yurkov. In his work devoted to the analysis of the risk of failure of electronic equipment, a methodology for assessing the reliability of radio electronics is provided. The following formula (5) is used to estimate the failure rate (which is one of the methods for calculating reliability) of such mechanical elements as couplings, bearings, and watch devices:

$$\lambda_{\ni} = \lambda_0 * \prod_{i=1}^{1} a_i, \qquad (5)$$

where λ_0 – is the failure rate of a part, shaft, joint or complex product in nominal mode and under normal conditions (ambient temperature 20±10ºC; relative humidity 30...70%; atmospheric pressure 0,825...1,06*105 Pa; absence of vibration and shock); a_i – are coefficients that take into account the design features of the part, conditions production and operation of the part.

In turn, the last coefficient is calculated using the formula:

$$a_i = \prod_{i=1}^{1} K_{ij}, \qquad (6)$$

where K_{ij} – are coefficients that take into account changes in the failure rate of a part, system, or product under the influence of various factors [6, 225].

Conclusion

Thus, reliability is an important property of technology that characterizes its operability during operation. Reliability is calculated so that the equipment of a particular industry (machine, object, system) retains its operational characteristics in accordance with regulatory and technical documentation.

When calculating reliability, the main indicator is the probability of failure-free operation of the object. In turn, trouble-free operation means, again, operability without failures, without interruptions and without forced stops of the process.

Reliability calculation is important and necessary in every industrial sector. For example, in the aviation industry, reliability is calculated to determine the degree of safety of an airplane or helicopter. Structural strength is assessed in order to predict how the machine will behave under a particular load. Of course, a high degree of reliability will reduce the number of accidents and ensure a high level of safety [7, 266].

In the construction industry, it is also important to check the reliability of structures, bridges, dams, buildings and other structures. Using calculations, they determine the load that building structures can withstand. All this makes it possible to ensure the safety and security of property.

Cars are also subjected to reliability calculations. The reliability of the car's design is calculated in order to reduce the risk of accidents and increase the level of safety and durability.

The energy industry is unthinkable without calculating reliability. Power plants, wind turbines, heat exchangers, reactors are only a small component of this industry. In order to ensure the safety of equipment, personnel safety, reliability calculations are performed and check what loads the facilities can withstand. The examples given show that reliability calculation is extremely important in all industries.

References

1. Невзоров, В. Н. Надежность машин и оборудования. Проектирование, эксплуатация, экспериментальные исследования [Текст] / В. Н. Невзоров, Е. В. Сугак. – Красноярск: Сиб. гос. технологич. ун-т, 1998. – С. 264

2. Дорохов, А. Н. Обеспечение надежности сложных технических систем [Текст] : учебник / А. Н. Дорохов, В. А. Керножицкий, А. Н. Миронов и др. – СПб.: Изд-во «Лань», 2017. – С. 352.

3. Шульц, В.В. Теория надежности машин [Текст] / В. В. Шульц. – Л.: Ленинградский инж.-строит. ин-т, 1983. – С. 75

4. Кравченко, Е. Г. Надежность технических систем в машиностроении [Текст]: учеб. пособие / Е. Г. Кравченко. – Комсомольск-на- Амуре: ФГБОУ ВПО «КнАГТУ», 2014. – С. 126

5. Жаднов, В.В. Прогнозирование надежности электронных средств с механическими элементами: научное издание. / В.В. Жаднов. – Екатеринбург: Изд-во ООО «Форт Диалог-Исеть», 2014. – С. 172.

6. Юрков, Н.К. К проблеме моделирования риска отказа электронной аппаратуры длительного функционирования / Н.К. Юрков, И.И. Кочегаров, Д.Л. Петрянин // Прикаспийский журнал: управление и высокие технологии. 2015. № 4 (32). С. 220-231.

7. Лушпа, И.Л. Исследование надежности механических компонентов антенно-фидерного устройства системы управления беспилотным летальным аппаратом. / И.Л. Лушпа, М.А. Монахов. // Труды Международного симпозиума «Надежность и качество»: в 2т. Т.2. Пенза: ПГУ, 2014. С. 264-266

ANALYSIS OF THE EXPERIENCE OF INTRODUCING DIGITAL TWINS IN DOMESTIC PRODUCTION

Victoria S. Komarova
student
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
vikap1999@mail.ru

Abstract. *The article describes the experience of introducing digital twins in domestic production and their potential to improve the efficiency and safety of production processes. The article examines the successful application of the digital twin in industry, where it helps to reduce the time and cost of developing new installations, as well as train personnel and conduct training on production process management.*
Keywords: *digital twin, Internet of things, digital twin functions*

By the end of 2023, about 150 foreign companies had left Russia, including those working in significant sectors of the domestic industry, so it is especially important to replace the departed companies with Russian analogues, using the most effective methods and tools of design and production, for example, digital twin technology.

GOST 2093-82 prescribes the use of sieve analysis to study the granulometric composition. However, the laser diffraction method has the largest measurement range, which, based on data on the intensity of the projected rays, is able to conclude about the particle size.

Digital doubles are a technology that allows you to create an exact copy of an object in virtual space. This technology has found wide application in the Russian oil and gas industry, where it is used to control production processes, improve safety and work efficiency.

One of the main advantages of digital twins is the ability to conduct experiments and testing virtually, which reduces the risks and costs of real-world testing. For example, when designing new wells or oil pipelines, you can create a digital copy of the object and conduct various tests on it to identify possible problems and shortcomings. [1]

In addition, digital twins allow you to monitor the condition of the equipment in real time. With the help of sensors and monitoring systems, it is possible to obtain data on temperature, pressure, flow rate and other parameters of equipment operation. This data is processed in a digital double, which allows operators to quickly respond to possible problems and prevent accidents.

Digital doubles are also used for staff training. With their help, you can create simulators where new employees can practice managing equipment or solving various tasks. This allows you to reduce the training time and improve the skills of the staff. A digital double can be created in various programs, depending on the specific object or process for which it is being created. For example, modeling and simulation programs such as Siemens PLM Software, Dassault Systemes, Autodesk and others can be used to create a digital twin of industrial equipment. To create a digital twin of a transport system, programs for modeling traffic flows and simulation, such as PTV Group, Aimsun, TransModeler and others, can be used. In addition, there are specialized programs for creating digital twins, such as ThingWorx from PTC, Azure Digital Twins from Microsoft and others.

The digital twin model was first presented at the Society of Technological Engineers (SME) conference in Troy, Michigan, in October 2002 at one of the presentations. The presentation was dedicated to the support/operation stage of the product lifecycle. The model didn't even have a name. The slide was simply titled "The Conceptual Ideal of PLM".

In Russia, the technology of digital twins began to be introduced relatively recently. The first information about such digitalization in large companies began to appear in 2015.

The purpose of this article is to analyze the experience of introducing digital twins into domestic production.

The use of digital twin technology in the domestic industry

In Russia, the digital technology usage index is distributed extremely unevenly (Fig. 1). If "Big Data" has already become quite widespread in domestic industries, then "Digital twins" are only gaining momentum.
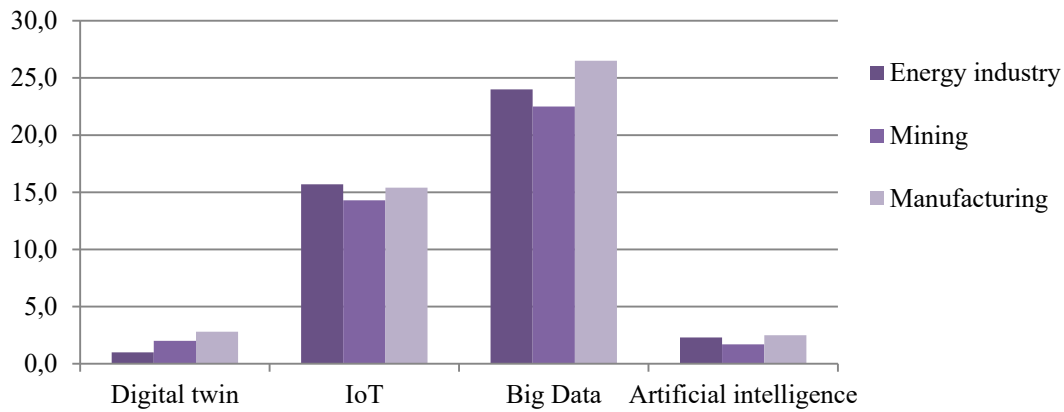
*Fig. 1. Index of the use of digital technologies in traditional industries -
the flagships of Russian industry [2]*

One of the successful examples of a digital twin within the oil and gas industry is the Gazprom-Neft project to create a digital twin of the Alexander Zhagrin field in the Khanty-Mansiysk Autonomous Okrug. This deposit is one of the largest in Russia and has a complex geological structure.

The digital twin of the field was created on the basis of data obtained as a result of geological exploration and drilling of wells. It allows you to simulate various scenarios of field operation, conduct experiments and testing virtually, as well as monitor the condition of equipment in real time.

With the help of a digital twin, the optimization of production processes at the field was carried out. Opportunities have been identified to increase oil and gas production, reduce operating costs and improve operational efficiency. New mining technologies have also been developed, which will increase the output of products.

In addition, the digital twin of the field was used for staff training. Based on it, simulators have been created where new employees can practice managing equipment and solving various tasks.

There is an increase in the volume of the global digital twins market: for example, in the period from 2020 to 2022, the increase was 71%, and by 2028 an additional increase of 61.3% is expected at an annual growth rate (Fig. 2).

The introduction of a digital twin of the Alexander Zhagrin field has allowed Gazprom Neft to reduce the cost of operating the field and improve operational efficiency. This project has become an example of the successful implementation of digital twin technology in the Russian oil and gas industry.
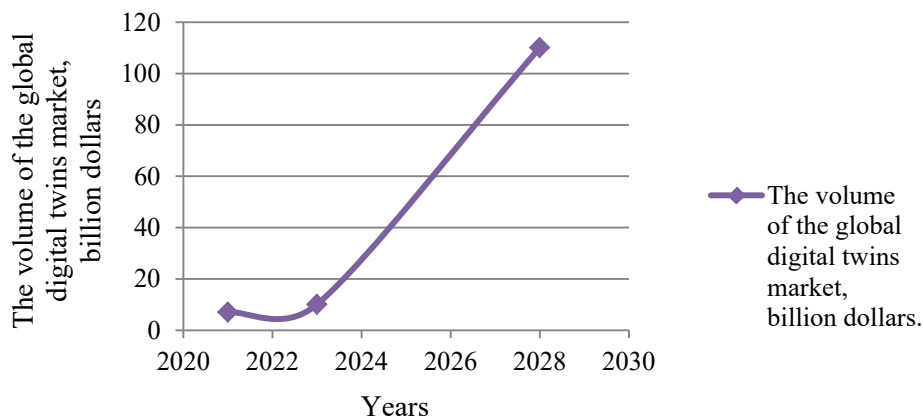
Fig. 2. The volume of the global digital twins market, billion dollars

There are also examples of the introduction of digital twins technology in healthcare in Russia. For example, the company "Computer Technologies in Medicine" has created a digital human double that

allows you to simulate the work of organs and body systems. This project was implemented as part of the program for the development of innovative technologies in medicine.

The digital human double is used to teach medical students, as well as to perform virtual surgeries and treat patients. It allows you to reduce the time and cost of training and examination of patients, as well as improve the accuracy of diagnosis and treatment.

The main advantage of a digital human double is the ability to perform virtual surgeries and treat patients. This reduces the risks for patients and increases the effectiveness of treatment. In addition, the digital double of a person can be used to teach students of medical universities, which allows to improve the quality of medical education.

One of the examples of the use of a digital double in Russia within the framework of construction is the "Digital Double" project of the SCANEX company. It allows you to create a virtual model of a construction site and conduct various tests and tests on it, which reduces the time and cost of real construction. [3]

In Russia, the digital twin is actively used in the automotive industry. For example, AvtoVAZ uses a digital twin to create virtual car models and conduct various tests on them. This reduces the time and cost of developing new car models.

In addition, the digital twin is used to monitor production processes and optimize production. For example, with the help of a digital double, you can monitor the condition of equipment, control the quality of products and predict possible problems in the production process.

Similarly, the digital twin can be used for staff training and training. For example, using a virtual car model, you can train mechanics and engineers in various technical skills and conduct training on car maintenance and repair.

Such technologies are actively used in the automotive industry by various companies such as GAZ, KAMAZ, UAZ and others.

In Russia, the digital twin is also actively used in the framework of nuclear energy. For example, Rosatom uses a digital twin to create virtual models of nuclear reactors and conduct various tests on them.

This reduces the time and cost of developing new nuclear reactors, as well as improving the safety and efficiency of their operation. In addition, the digital twin is used to monitor production processes and optimize nuclear fuel production.

In addition, the digital twin can be used for staff training and training. For example, using a virtual model of a nuclear reactor, operators and engineers can be trained in various technical skills and conduct training on the management of nuclear installations.

Such technologies are actively used in the Russian nuclear power industry and help to improve the safety and efficiency of nuclear installations.

Conclusion

Thus, based on the analysis of the experience of introducing digital twins in domestic production, we can say that this technology has great potential to increase the efficiency and safety of production processes. In Russia, the digital twin has already been successfully used in many areas of industry, which will have a positive impact on the import substitution process. However, it is necessary to continue research and develop this technology in order to expand its application to other domestic production facilities and improve the quality of their work.

References

1. Сосфенов Д.А. Цифровой двойник: история возникновения и перспективы развития//Интеллект. Инновации. Иивестиции. 2023. №4. С.35-43.

2. Loginom: Цифровизация промышленности. – Москва. – URL: https://loginom.ru/blog/digitalization-industry (дата обращения: 19.02.2020). – Текст: электронный.

3. Г.Б. Коровин. Возможности применения цифровых двойников в промышленности. Вестник ЗабГУ. 2021. Т.27. №8.С. 124-133.

MATHEMATICAL MODEL OF AN ACTIVE VOLTAGE RECTIFIER

Mikhail Kozlov

Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
kozlov_31kaf@bk.ru

Abstract. *The paper discusses the construction of a mathematical model of an active voltage rectifier (AVR).*

Keywords: *automatic control system (ACS), active voltage rectifier, voltage and current equilibrium equations, column vectors.*

To create a mathematical model of an active voltage rectifier, it is necessary to set the equilibrium equations for voltage and currents. When writing these equations, it is usually assumed that the rectifier is built on ideal power switches in the case of using IGBT modules [1, 2].

Fig. 1 shows the equivalent circuit of the AVN power section.

In this work, we will consider the case of a symmetric system with the following equalities

$r_a = r_b = r_c = r$ – active resistance of the reactor;

$k = l_b = k = l$ - reactor inductance.
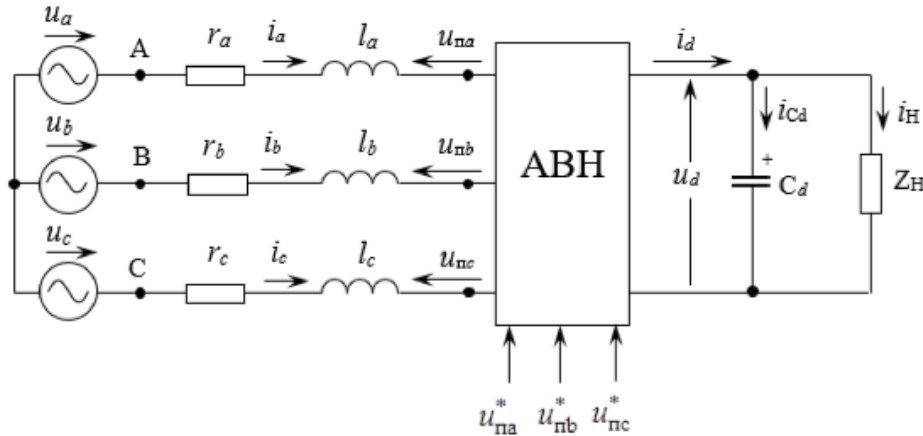
Fig. 1. Equivalent circuit of the AVN power section

Next, we introduce some abstract algebraic column vectors, which consist of relative parameters of voltages and currents in the network determined at some point in time, voltage values at the power input of the converter and information data flows at the input of the rectifier

$$\mathbf{U}_{df}^{F} = \begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix}; \quad \mathbf{I}_{df}^{F} = \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}; \quad \mathbf{U}_{\Pi}^{F} = \begin{bmatrix} u_{\Pi a} \\ u_{\Pi b} \\ u_{\Pi c} \end{bmatrix}; \quad \mathbf{U}_{\Pi}^{*F} = \begin{bmatrix} u_{\Pi a}^{*} \\ u_{\Pi b}^{*} \\ u_{\Pi c}^{*} \end{bmatrix}$$

The variables specified above are phase quantities and are designated by the letter "F" in the superscript. And information data flows are designated "*".

AVN can be described by the following equations

$$\mathbf{U}^{F} = \mathbf{U}_{\Pi}^{F} + \mathbf{R}^{F} \cdot \mathbf{I}^{F} + \mathbf{L}^{F} \frac{d\mathbf{I}^{F}}{d\tau}; \quad \mathbf{U}_{\Pi}^{F} = u_d \cdot \mathbf{\Phi}_{\Pi}^{F}; \quad \mathbf{\Phi}_{\Pi}^{F} = \frac{1}{2} \cdot \mathbf{M} \cdot \mathbf{F}_{\Pi}^{F} \cdot (\mathbf{U}_{\Pi}^{*F}, \tau); \quad i_d = \mathbf{\Phi}_{\Pi}^{F} \cdot \mathbf{I}^{F}$$

$$i_d - i_{\text{н}} = C_d \cdot \frac{du_d}{d\tau}$$

where

$$\mathbf{L}^F = \begin{bmatrix} l_a & 0 & 0 \\ 0 & l_b & 0 \\ 0 & 0 & l_c \end{bmatrix} \text{ and } \mathbf{R}^F = \begin{bmatrix} r_a & 0 & 0 \\ 0 & r_b & 0 \\ 0 & 0 & r_c \end{bmatrix}$$ - matrices of reactor inductances and active resistances

$$\mathbf{M} = \frac{1}{3} \cdot \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$ - some previously specified matrix constant;

$$\mathbf{\Phi}_\Pi^F = \begin{bmatrix} f_{\Pi a} \\ f_{\Pi b} \\ f_{\Pi c} \end{bmatrix}$$ - discrete switching vector function;

$\mathbf{F}_\Pi^F(\mathbf{U}_\Pi^{*F}, \tau)$ - discrete vector function that denotes information signals and the state of the keys. For pulse-width systems with a main signal $U_0(\tau)$ it can be expressed as follows

$$\mathbf{F}_\Pi^F(\mathbf{U}_\Pi^{*F}, \tau) = \begin{bmatrix} f_{\Pi a}^* \\ f_{\Pi b}^* \\ f_{\Pi c}^* \end{bmatrix} = \begin{bmatrix} sign(u_{\Pi a}^* + u_0(\tau)) \\ sign(u_{\Pi b}^* + u_0(\tau)) \\ sign(u_{\Pi c}^* + u_0(\tau)) \end{bmatrix}$$

where τ is time.

Further, due to the fact that the analyzed quantities in the alternating current circuit change in time, the AVN equation can be written using an orthogonal coordinate system with axes *x, y* and *z*, which rotates at a certain speed ω_k.

Thus, the previously indicated column vectors, taking into account the transformed values, can be written as

$$\mathbf{U} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}; \quad \mathbf{U}_\Pi = \begin{bmatrix} i_{\Pi x} \\ i_{\Pi y} \\ i_{\Pi z} \end{bmatrix}; \quad \mathbf{I} = \begin{bmatrix} i_x \\ i_y \\ i_z \end{bmatrix}; \quad \mathbf{U}_\Pi^* = \begin{bmatrix} u_{\Pi x}^* \\ u_{\Pi y}^* \\ u_{\Pi z}^* \end{bmatrix}$$

Next, using a special transformation matrix

$$\mathbf{A} = \mathbf{A}(\Theta_k) = \frac{2}{3} \cdot \begin{bmatrix} \cos\theta_k & \cos(\theta_k - \Delta) & \cos(\theta_k + \Delta) \\ -\sin\theta_k & -\sin(\theta_k - \Delta) & -\sin(\theta_k + \Delta) \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix}$$

we obtain equations for the transformed variables

$$\mathbf{U} = \mathbf{A}(\theta_k) \cdot \mathbf{U}^F$$

$$\mathbf{U}_\Pi = \mathbf{A}(\theta_k) \cdot \mathbf{U}_\Pi^{F}$$

$$\mathbf{I} = \mathbf{A}(\theta_k) \cdot \mathbf{I}^F$$

$$\mathbf{U}_\Pi^* = \mathbf{A}(\theta_k) \cdot \mathbf{U}_\Pi^{*F}$$

where θ_k – is the total transformation number associated with the orthogonal coordinate system. Moreover, it is comparable to the phase coordinates (*a, b, c*).

Taking into account the discreteness of processes in ABN and the description of its orthogonal system, we can write a mathematical model described by the following equations

$$\mathbf{U} = \mathbf{U}_\Pi + \mathbf{R} \cdot \mathbf{I} + \mathbf{L} \cdot \frac{d\mathbf{I}}{d\tau} + \omega_k \cdot \mathbf{B} \cdot \mathbf{L} \cdot \mathbf{I}; \ \mathbf{U}_\Pi = u_d \cdot \mathbf{\Phi}_\Pi;$$

$$\mathbf{\Phi}_\Pi = \frac{1}{2} \cdot C \cdot F_\Pi \cdot (U_\Pi^*, \tau); \ i_d = \frac{3}{2} \cdot \mathbf{\Phi}_\Pi \cdot \mathbf{I}; \ i_d - i_n = C_d \cdot \frac{du_d}{d\tau}; \ \omega_k = \frac{d\theta_k}{d\tau}.$$

(1)-(6)

where

$$\mathbf{B} = \omega_k^{-1} \cdot \mathbf{A} \cdot \frac{d\mathbf{A}^{-1}}{d\tau} - \text{ ome specific matrices;}$$

R,L – some new matrices of active resistances and inductances of a symmetrical inductor;

$$\mathbf{\Phi}_\Pi = \begin{bmatrix} f_{\Pi x} \\ f_{\Pi y} \\ f_{\Pi z} \end{bmatrix} - \text{ the resulting final commutator vector function of the rectifier;}$$

$$\mathbf{F}_\Pi(\mathbf{U}_\Pi^*,\tau) = \begin{bmatrix} f^*_{\Pi x} \\ f^*_{\Pi y} \\ f_{*\Pi z} \end{bmatrix} = \mathbf{A} \cdot \mathbf{F}_\Pi^F(\mathbf{A}^{-1}\mathbf{U}_\Pi^*,\tau) - \text{ the resulting finite vector function representing the in-}$$

formation signals and the state of the keys.

Taking into account some specific operating modes of the AVN, we determine the speed value as

$$\omega_k = \omega_u, \tag{7}$$

If we take into account that the switching frequency of switches in PWM mode is quite high, then to study the transient and steady-state modes of an active voltage rectifier it would be logical and correct to use its equivalent and lightweight mathematical model.

Under the above conditions, the equivalent mathematical model will be some version of the model written according to equalities (1) – (6), which were obtained from the results of a continuous approximation of the discrete commutator function of the rectifier.

Taking the above into account, the mathematical model can be represented in the following form

$$u_x = u_{nx} + r \cdot i_x + l \cdot \frac{di_x}{d\tau} - \omega_k \cdot l \cdot i_y, \quad u_y = u_{ny} + r i_y + l\frac{di_y}{d\tau} - \omega_k l i_x,$$

$$f_{nx} = \frac{u^*_{nx}}{2 \cdot U}, \quad f_{ny} = \frac{u^*_{ny}}{2 \cdot U} \quad u_{nx} = f_{nx} \cdot u_d, \quad u_{ny} = f_{ny} \cdot u_d, \quad i_d = \frac{3}{2} \cdot (f_{nx} \cdot i_x + f_{ny} \cdot i_y); \quad i_d - i_{\text{н}} = C_d \cdot \frac{du_d}{d\tau};$$

$$\omega_k = \frac{d\Theta_k}{d\tau}, \tag{8) – (14}$$

where U_0 is the difference in the subcarrier PWM signal.
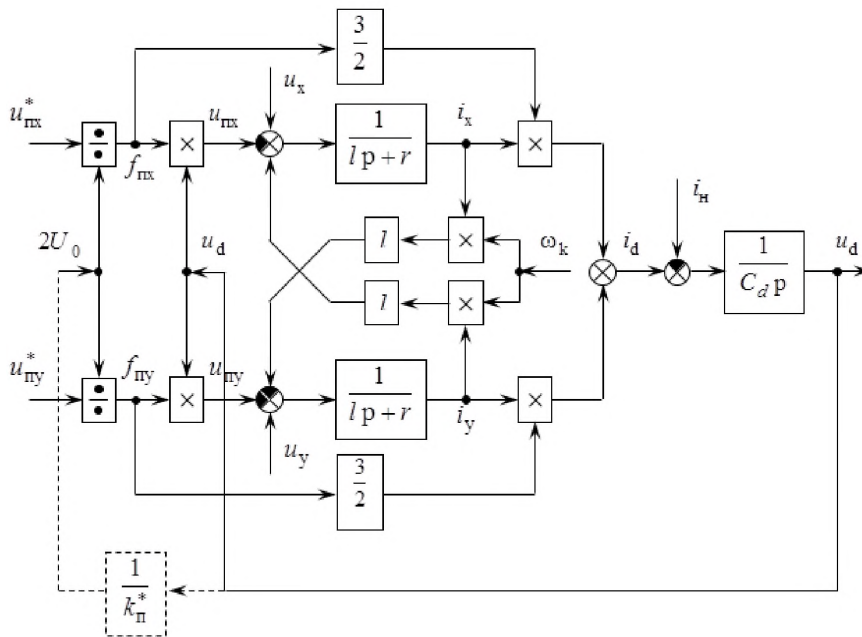
Fig. 2. Block diagram of an equivalent continuous AVN model

The continuous approximation equations are correct if

$$\left| u_{nj}^{*} \right| \leq U_0, j = a, b, c$$

The power factor AVN, provided it is in the power supply system, can be determined by the formula

$$\cos \varphi = \frac{u_x \cdot i_x + u_y \cdot i_y}{\sqrt{(u_x^2 + u_x^2) \cdot (i_x^2 + i_x^2)}}$$

Fig. 2 shows the block diagram of an equivalent continuous AVN model corresponding to equations (8) – (14).

References

1. Shreiner R.T. Mathematical modeling of AC electric drives with semiconductor frequency converters. Ekaterinburg: Publishing house URO RAS, 2000. – 654 p.

2. Shreiner R.T., Efimov A.A., Mukhamatshin I.A. Predictive relay vector control of an active current rectifier // Automation and advanced technologies: Proceedings of the III interindustry scientific and technical conference. Novouralsk: Publishing House NGTI, 2002. – P. 178 – 182.

DEVELOPMENT OF ALGORITHMS FOR LIGHT MONITORING AND INTELLIGENT LIGHTING CONTROL

Yuri Kuzmenko

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: spider22boy@mail.ru

Abstract. *This article discusses the design of a lighting system that is automatically controlled using an algorithm based on the assessment of one sensor of the ambient light outside the room to adjust the internal illumination.*
The purpose of the presented research is to develop an algorithm for intelligent control of artificial lighting based on a system with one light sensor installed outside the room.

Introduction

In the energy consumption of office, residential, public buildings, schools, hospitals, etc.п. a significant share is occupied by electric lighting 40-60% of the total consumption, so существуеттересть is a needto reduce the consumption of lighting installations.

The most commonly used solution for improving the efficiency of a lighting system is to simply replace old lighting devices with new and modern ones with greater energy efficiency. At the moment, such light sources are LED lamps. The use of LEDs can significantly reduce the energy consumption of the lighting system. Additional economic benefits возсan also be obtained by using an intelligent lighting control system. This system allows you to combine lighting fixtures and control circuits into a single network that maintains the required level of illumination in all rooms involved in the system. Monitoring inside the system is carried out using various sensors: light, motion, presence, acoustic, etc. The use of such a system allows you to reduce electricity consumption by using lighting devices in the right place, at the right time, with the right illumination. Such a system allows you to create usage scenarios, with the help of such scenarios, it is possible to control both one and many lamps.

The problem of energy efficiency of lighting installations is particularly acute in enterprises, in office and public buildings, as it is necessary to illuminate a large number of premises with different areas and tasks of their use. As it was mentioned above, electricity consumption for lighting needs occupies a significant share of the total consumption at such facilities. Recently, this area has been actively developing, as equipment has become more affordable both in the physical sense and in the economic sense, therefore, the payback period has become much shorter than it was 10-15 years ago. The high cost of equipment was associated with several factors: novelty, complexity, high cost of manufacturing microprocessor controllers and chips, the lack of a single or unified protocol, and the complexity of creating new software. As areas such as computer science, IT technology and industrial electronics are actively developing, the question of the high cost and complexity of equipment is gradually decreasing. A large number of different manufacturing companies, both Russian and foreign, appear, which favorably affects the price of equipment due to market competition.

Analysis of the literature on the issue under study

The issues of lighting control and related problems of energy saving have been considered at various times by a large number of researchers. In the information space of works in Russian, a significant contribution was made by A. K. Solovyov, in the textbook "Environmental Physics" [1], the reference book "Energy Saving in lighting" and" Reference Book on Lighting " edited by Professor Yu. B. Aizenberg [2, 3]. In his works, A. K. Solovyov addresses the issues of calculating the coefficient natural light intensity (KEO), the permissible limits of this coefficient, and the resulting dependencies when considering the issue of illumination of a horizontal surface. In the reference book of Yu. B. Aizenberg, the issues of energy saving for lighting purposes are considered. Currently, the most common issues of energy saving in lighting systems are related more to the solution of architectural problems by increasing KEO [1, 4, 5]. To a much lesser extent, the issues of solving technical problems and algorithmization of lighting control are addressed [6, 7, 8, 9, 10].

When designing and implementing lighting control systems, first of all, they rely on the characteristics declared by the manufacturer, which claim up to 60% savings, compared to the period before the

introduction of the lighting control system, or use the results of prototype systems and single implemented systems [11, 12]. At the same time, it should be noted that due to insufficiently formalized and defined quality criteria of lighting devices themselves and methods for determining the quality of lighting devices, there is an overestimation of characteristics, which leads to a decrease in the economic effect of the introduction of an intelligent lighting system [13].

Lighting system control parameters

Energy saving in lighting installations with controlled LED lamps is achieved due to the fact that with a decrease in the luminous flux of lamps, the power of the system with LED lighting also decreases, as shown in Fig. [14].
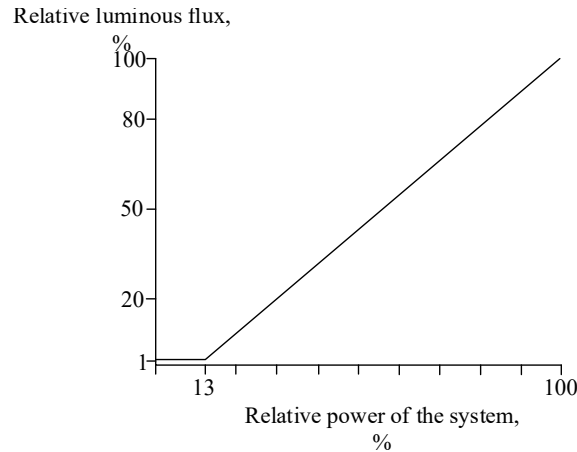
Fig. 1.Diagram of the ratio of the power of the LED lighting system and the luminous flux of lamps.

Fig. 1 shows the relationship between the power consumption of an LED lighting device and the luminous flux. So, for example, when the luminous flux value is reduced to 1%, the lamp consumes about 13% of the power relative to the rated power. This dependence is true for most types of LED lighting devices. The lower and upper power limits can be higher or lower depending on the specific lamp [14]. To control the light flow of LED lamps, special drivers are available with the ability to control. The most common control signals are analog 1 ... 10 V and digital DALI [14].

Algorithm for smooth lighting control using a single light sensor

For the end user, after the introduction of a smart lighting system, important parameters are to reduce energy consumption and find the level of illumination within the normalized limits. Therefore, the implemented algorithms should not only reduce the consumption of electrical energy, but also maintain an acceptable level of lighting in the room. To implement such operation of the algorithm, it is necessary to determine the dependences of illumination in workplaces for a specific number of specifically located lamps at any level of the control signal of the control value. At the same time, the system must be able to calculate the level of the control signal based on the required light level.

The uneven distribution of the light flux from individual lighting devices is a major challenge when creating control algorithms. As a consequence, when changing the level of the control signal for a particular workplace to achieve the required illumination, it does not guarantee that the illumination value will be within the normalized limits at another workplace. To implement this behavior of the lighting system, it is necessary to introduce some optimization function $opt(P_{kj})$. This function for a set of values P_{kj}, where k is the number of the workplace, j is the number of the lamp, can optimize the parameters and, accordingly, select the level of the control signal to maintain the illumination at a particular workplace (k-th), so that the illumination is within the normalized limits, while achieving the greatest possible reduction in illumination and as a result, the consumption of electrical energy. In order to exclude the operation of the lighting system during times of emergency when people are in the room, it is necessary to introduce an additional switch or sensor for forced shutdown.

Control algorithm for controlling outdoor lighting

For this algorithm, in addition to information about the light distribution from the op amp at different levels P_j of, you need to know e_k for each workplace. Finding these values is possible by direct

measurement, or through calculation methods, and it is also necessary to determine the margin coefficient $K_{s,\,e}$, since this coefficient takes into account the decrease in KEO and illumination, as a result of operation and factors associated with it [15]. The value $of\ e_k$ can be found using the formula:

$$e_k = \frac{r_0^2}{r_k^2} \cdot \frac{KEO}{K_{3,e}}.$$

Where: r_0 – normalized distance; $r_{\text{к},}$ – distance to a specific workplace.

When the control system is not activated? it is in command standby mode, and system activity begins when the operation signal is received. I.e., simultaneously with the lamp's power supply. To begin with, the controller uses a sensor to read information about outdoor lighting e. In this case, the current value of the control signal pj from each power source of the luminaire is also processed. P_j Then, based on the received data, the system simultaneously calculates for each workplace:

- the level of natural light at each workplace;

$$E_{ek} = e_k \cdot E_{\text{н}},$$

- the total level of illumination from artificial lighting at *the k*-th workplace $E_{\text{and}k}$ based управляющей функцииon the illumination levels of each specific lighting installation (j-th) obtained by the control function $E_{kj} = f\left(P_j\right)$.

$$E_{\text{и}k} = \sum_{j=1}^{M} E_{kj}.$$

After рассчитанныеthat, the illumination values calculated by the controller are added up, and then the system determines the value ΔE_k, which must be compensated by increasing the artificial illumination, in the case $\Delta of\ E_k > 0$, in the case $\Delta of\ E_k < 0$, reducing the natural illumination of the workplace

$$E_k = E_{ek} + E_{\text{и}k},$$
$$\Delta E_k = E_{\text{ном}} - E_k.$$

Together the system calculates the required level of the control signal P_{kj} from the lamp to the workplace for full compensation ΔE_k

$$\begin{cases} \Delta E_{kj} = \dfrac{E_{kj}\left(P_j\right)}{E_{\text{и}k}} \cdot \Delta E_k \\[2mm] \Delta E_{kj}^{'} = E_{kj} + \Delta E_{kj} \\[2mm] P_{kj} = f\left(\Delta E_{kj}^{'}\right) \end{cases} .$$

The system analyzes the received p_{kj} and uses the optimization function to determine the required signal level to send to the driver. Further, the algorithm is repeated, in standby mode, the system works according to the same principle, but the signal to the driver is sent only at the time of closing the circuit going to the LED lamp

The advantages of this algorithm are:

• There is no influence of external illumination, because the sensor only tracks changes in internal illumination.

• This eliminates the possibility that the sensor area will not be visible or obscured from the outside.

However, the implementation of this algorithm has the following disadvantages:

• When using the KEO value in the normalized parameters, the control error may increase.

• Higher cost of the sensor due to the need for its greater protection from the external environment;

• When using any curtains or screens, they must always be in the same position, or it is necessary to make additional clarification of the KEO and the position of curtains or screens.

- This eliminates the possibility that the sensor area will not be visible or obscured from the outside.

Conclusion

Improving the efficiency of LED lighting systems should be carriedьout in a comprehensive manner, i.e., it is necessary to increase the efficiency of both individual luminaires and the control systems of these luminaires.

The use of the developed algorithm for controlling LED lighting allows you to increase the efficiency of the lighting system and allows you to more flexibly adjust the illumination in accordance with the biological indicators of a person, as well as for better compliance with the requirements of regulatory documentation in terms of workplace illumination.

The applied algorithm has a significant advantage in the face of using outdoor illumination indicators to change indoor illumination indicators, which allows you to more accurately adjust the illumination required at a given time.

The described algorithm can effectively solve lighting control problems, while it does not require large financial and labor costs for implementation.

References

1. Solov'ev A. K.-Environmental physics. Textbook: Moscow: Publishing House DIA, 2008. – 344s.

2. Spravochnaya kniga po svetotekhnika [Reference book on lighting engineering], edited by Aizenberg, Moscow: Energoatomizdat, 1995, 972 p.

3. Energy saving in lighting / Edited by Prof. Yu. B. Aizenberg, Moscow: Znak Publishing House, 1999, 264 p.

4. Solov'ev A. K. Distribution of brightness across the sky and its consideration in the design of natural lighting of buildings / / Svetotekhnika-2008-No. 6-pp. 18-27.

5. Korzhneva T. G. – Study of the efficiency of combined lighting taking into account the energy balance of the room: dissertation ... Candidate of Technical Sciences: 05.23.03 / Korzhneva Tatyana Gennadyevna; [Place of defense: Tyumen State University of Architecture and Construction. un-t]. – Tomsk, 2015. – 110 p.: ill. Heat supply, ventilation, air conditioning, gas supply and lighting.

6. Werner V. Intelligent control systems for internal lighting/ / Lighting Engineering. -1993. – No. 4. – from 15-19.

7. Vystavskaya E. V., Kazarinov L. S. Metod adaptivnogo upravleniya osveshcheniem razredlozhennykh obektov [Adaptive lighting control method for distributed objects]. Vestnik SUSU, 2006, no. 23, pp. 70-74.

8. Likhotkin V. S., Rodin V. V., Gubanov D. V. Avtomatizatsiya upravleniya i kontrolya osveshcheniya obshchestvennykh zdaniy [Automation of management and control of lighting of public buildings]. Sovremennye problemy nauki i obrazovaniya [Modern problems of Science and education]. – 2014. – No. 4. Available at: www.science-education.ru/118-14107 (accessed: 21.01.2016).

9. Slobodnik E. B., Allash E. Kh., Kazakov V. A., Kazantsev S. B. Sistema upravleniya osveshcheniyem na ledovakh [Electronic resource] / / AVOK – Noncommercial Partnership "Engineers for heating, ventilation, air conditioning, heat supply and construction thermophysics" Access mode: http://www.abok.ru/for_spec/articles.php?nid=4181.

10. Fomin A. G.-Research and optimization of energy consumption in installations of combined lighting with automatic control: dissertation ... Candidate of Technical Sciences : 05.09.07. – Moscow, 2000. – 174 p.

11. Catalog Northcliff 2013 [Electronic resource] / / Russian-language website of the company "Nordcliff". – 2013. Access mode: http://www.northcliffe.ru/files/367/Book_forsite1.pdf.

12. Benoit ROISIN, Arnaud DENEYER, Peter D'HERDT, Christian EUGENE. Optimization of lighting power consumption in offices // SINAIA 2006 International Lighting Symposium « Modern Quality Solutions for an Efficient Lighting» 12-14 October 2006, Sinaia, Romania.

13. Kuzmenko V. P. Modeli i metodiki obespecheniya kachestva lednykh osvetitel'nykh priborov: dis. kand. teh. nauk [Models and methods of ensuring the quality of LED lighting наук.- devices], Moscow, 2021. https://fs.guap.ru/dissov://fs.guap.ru/dissov

14. Gvozdev-Karelin S., Novozhilov S. Osram lighting control systems for solving energy saving problems [Electronic resource] // News and analytical portal "Time of Electronics". Access mode: http://www.russianelectronics.ru/leader-r/review/47098/47101/doc/47154/

15. SP 52.13330.2011 Natural and artificial lighting. Updated version of SNiP 23-05-95*. – M.-2011.

HARDWARE IMPLEMENTATION OF A PLANT GROWTH MONITORING SYSTEM

Sofya Lisovenko
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: lisovenko123123123@gmail.com

Abstract*. The purpose of this article is to design and assemble a device for monitoring plant growth. The device must monitor the state of soil moisture, air temperature and adjust the lighting depending on the set mode. If there are problems with the temperature or soil, an audible and visual alarm appears, letting the observer know that something is wrong. This article focuses specifically on the hardware of the device, selecting all the necessary components and connecting them to Arduino.*
Keywords: *automation, Arduino, sensors, agricultural technologies, growing plants.*

Introduction

Nowadays, when technological progress is rapidly increasing and interest in a healthy lifestyle is constantly increasing, home growing of plants is becoming more important. It ceases to be just a pleasant hobby and becomes an integral part of caring for the environment. For those looking to create a welcoming green space in their home and monitor the health and development of their plants, a home grow control system is an invaluable tool.

This system automates and optimizes plant care processes, providing them with ideal conditions for growth and development. From controlling temperature and soil moisture to providing optimal lighting, this system gives the user complete control over the environment to grow plants efficiently and gently.

Justification of the technical solution

To implement a plant growing control system, we will use sensors, a button, a microcontroller, an LCD display, LEDs, and a lighting system. The structure of this device is shown in Fig. 1.
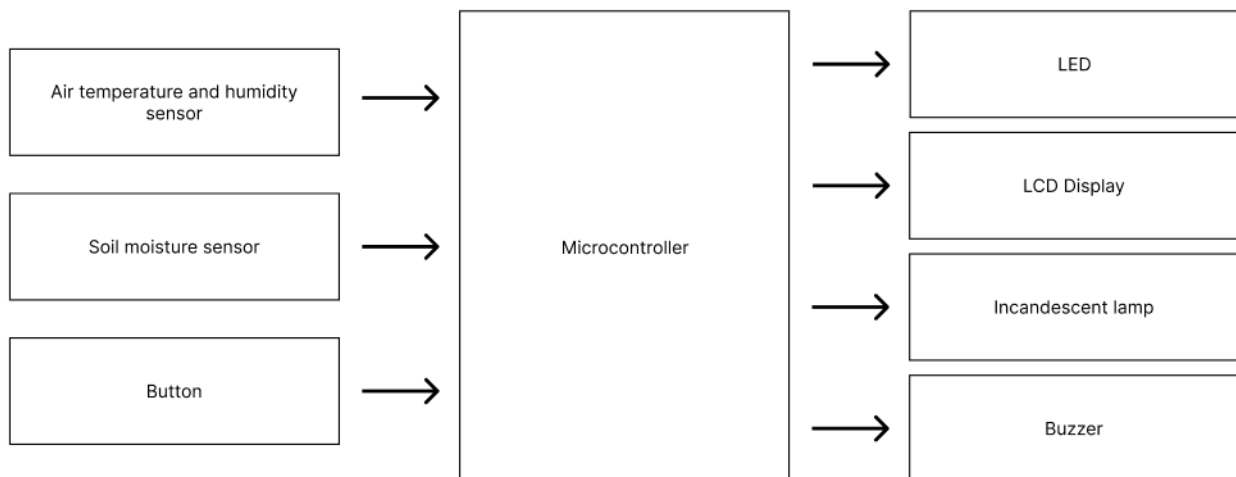
Fig. 1. Block diagram of the device

The circuit includes three main information receivers. The first is a complex sensor that combines an air humidity sensor and a temperature sensor. The second is a soil moisture sensor. The third is a button with which you will need to switch the station's operating modes. All three components are connected to a microcontroller, which processes the information, displays the data on the screen and, thanks to LEDs and a sound element, reports problems in the microclimate, and also regulates the lighting system.

Selection of element base

The main component of the system is the Arduino UNO microcontroller – the brain of the device. Arduino Uno was chosen because of its ease of connection and compatibility with various devices via

connectors. Arduino Uno combines ease of use and the necessary functionality for the successful implementation of our ideas. Arduino coordinates the operation of all connected sensors and actuators, processes the received data and makes decisions about the necessary adjustments in growing conditions [1].

Sensors are used to monitor the condition of plants.

• soil moisture sensor (YL-69) – determines the soil moisture level, allowing the system to maintain the optimal moisture level for each type of plant [2];

• air temperature and humidity sensor (DHT11) – measures temperature, which is critical for creating a suitable microclimate; the humidity measurement function is not used in this system [3].

Also important for plant growth is the lighting system – it is adjusted depending on the needs of the plants and the current level of natural light. Under ideal conditions this should be a UV lamp, but when modeling, an incandescent lamp is used, and when prototyping, an LED with a resistor is used, due to limitations in software capabilities and budget.

To warn the observer about critical changes in soil temperature and moisture, sound and light signals are used. If the temperature is increased, the red LED lights up, if it is low, the blue LED lights up. If problems arise with soil moisture, turn on the orange LED. The sound signal is activated in all of the above cases.

An LCD display is also used for monitoring, displaying current temperature and humidity readings, allowing the user to monitor the status of the system in real time.

A button is used to switch different operating modes, depending on the stage of plant growth, and the RGB module indicates which mode is currently enabled [4].

System implementation

The TinkerCad program was used to construct the diagram. Power to the circuit will be supplied via the Micro USB connector. The power and ground inputs of all components must be connected to the GND ground and VCC power of the Micro USB connector. The operating voltage of the circuit is 5 V, so the power supply must be no more than 5 V from the Micro USB connector.

The functional diagram of the system is presented in Fig. 2.
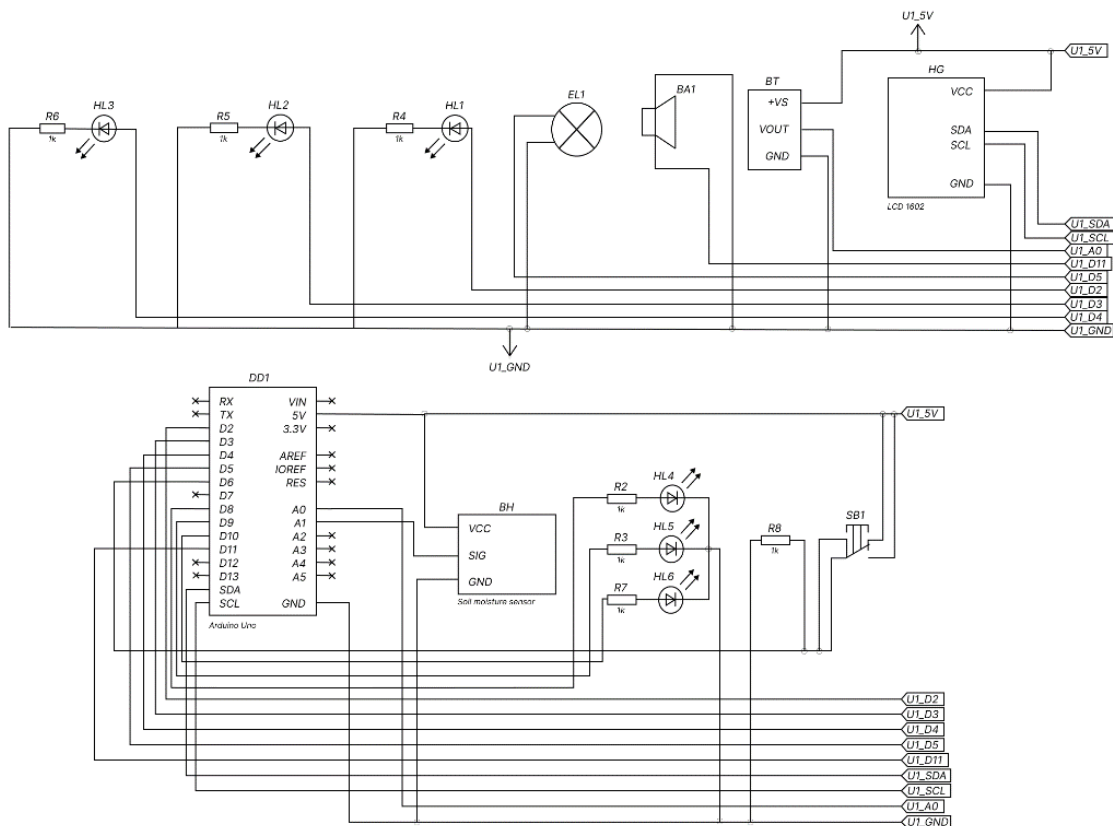
Fig. 2. Functional diagram of the control system

For better understanding, it is necessary to divide the functional diagram into 2 blocks, shown in Figures 3 and 4.
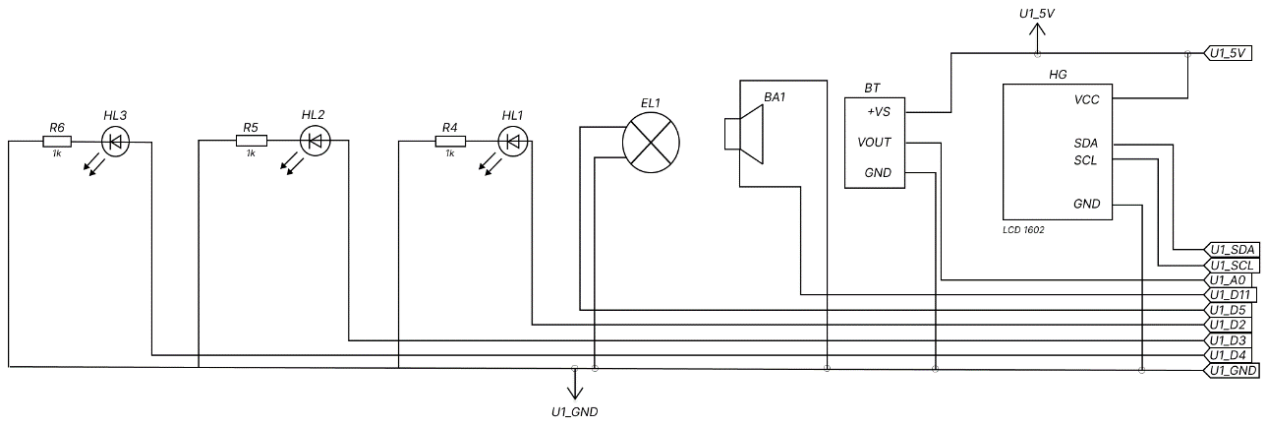
Fig. 3. LEDs, lamp, sound element, temperature sensor and LCD display
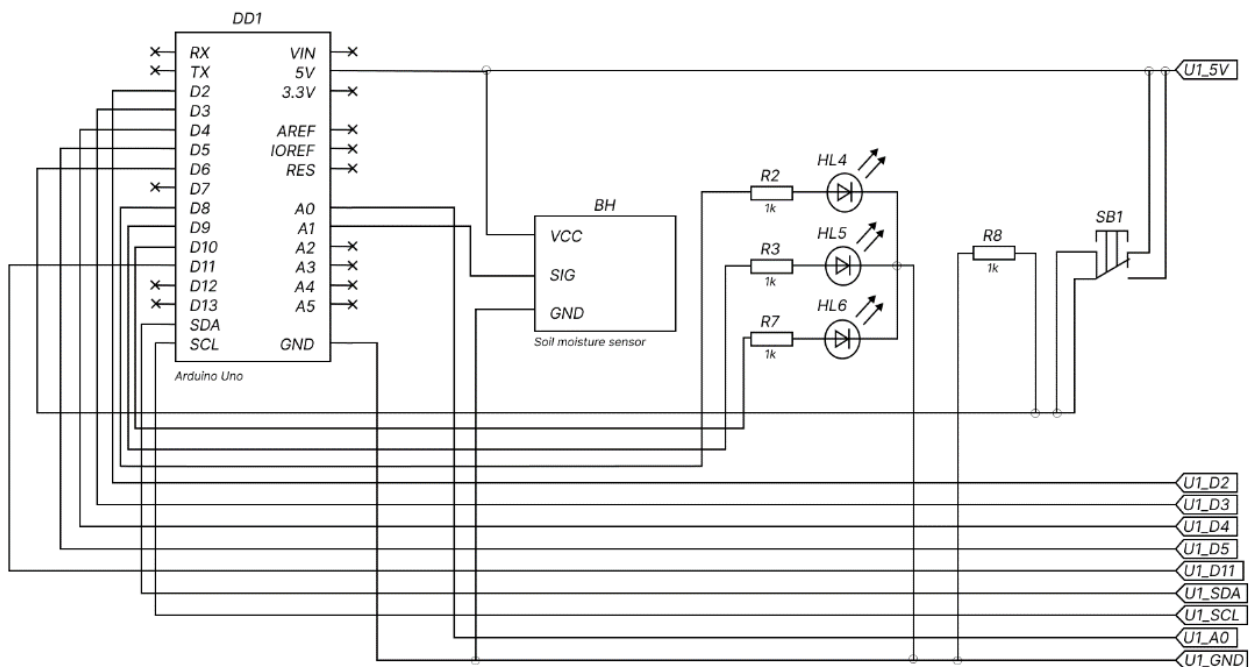
Fig. 4. Arduino, soil moisture sensor, RGB module, button

Next, modeling was carried out in the TinkerCad programming environment. Fig. 5 shows operation at normal temperature but high humidity.
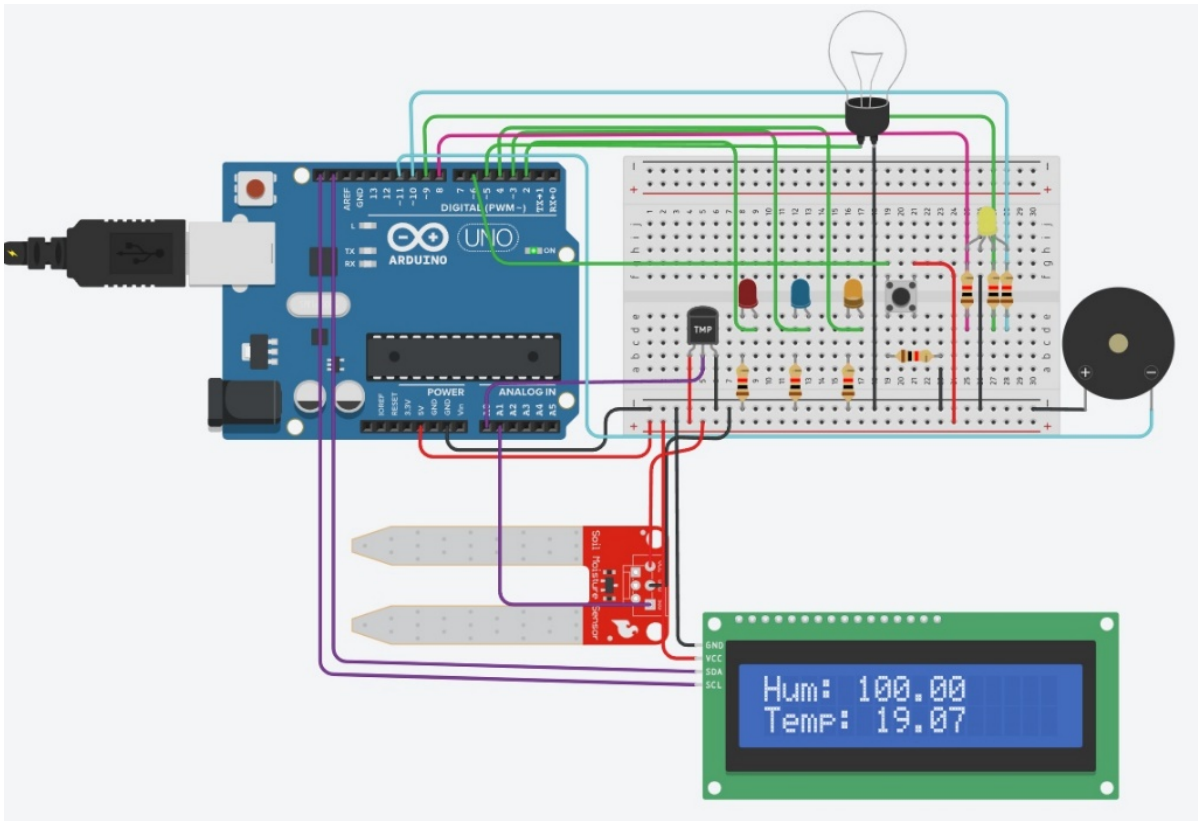
Fig. 5. Circuit operation at normal temperature but high humidity

Circuit operation at low temperature and low humidity presented in Fig. 6.
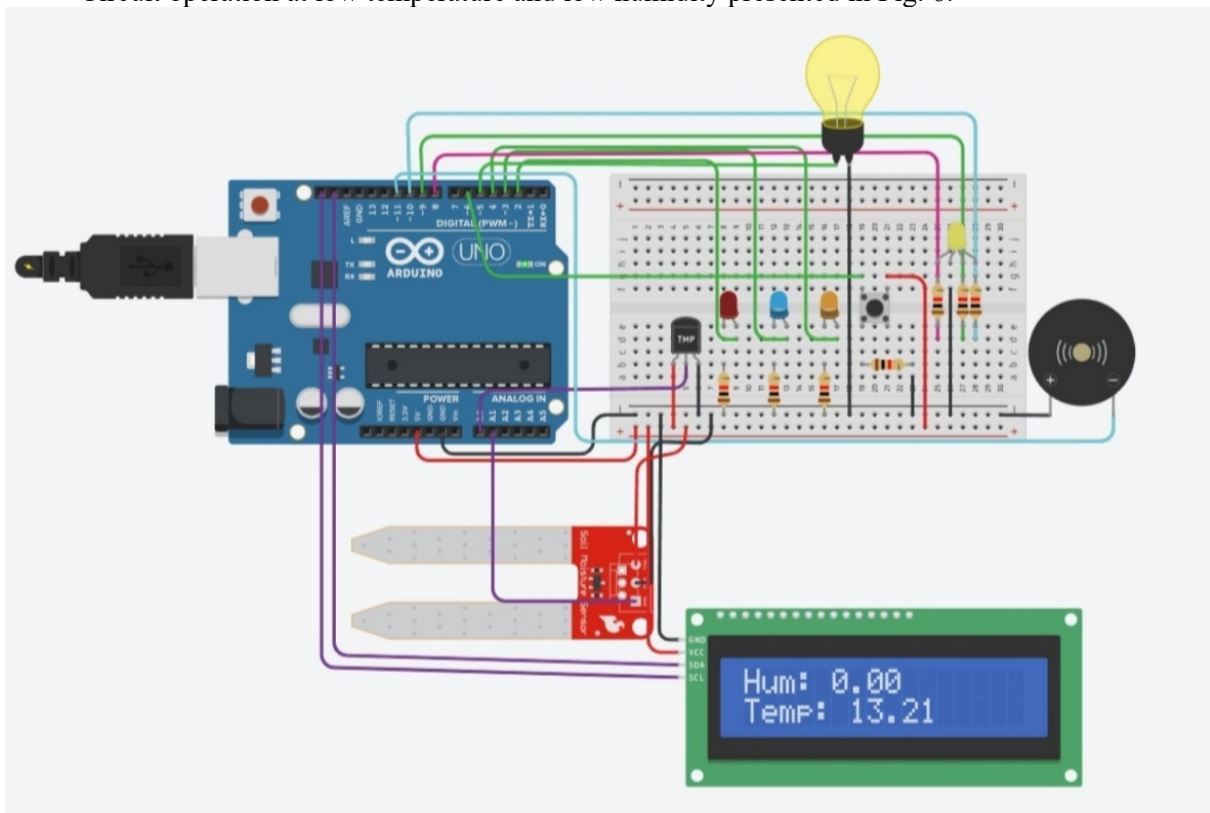
Fig. 6. Circuit operation at low temperature and low humidity

And the last step is to prototype the system in real conditions. A layout of the system is shown in Fig. 7.
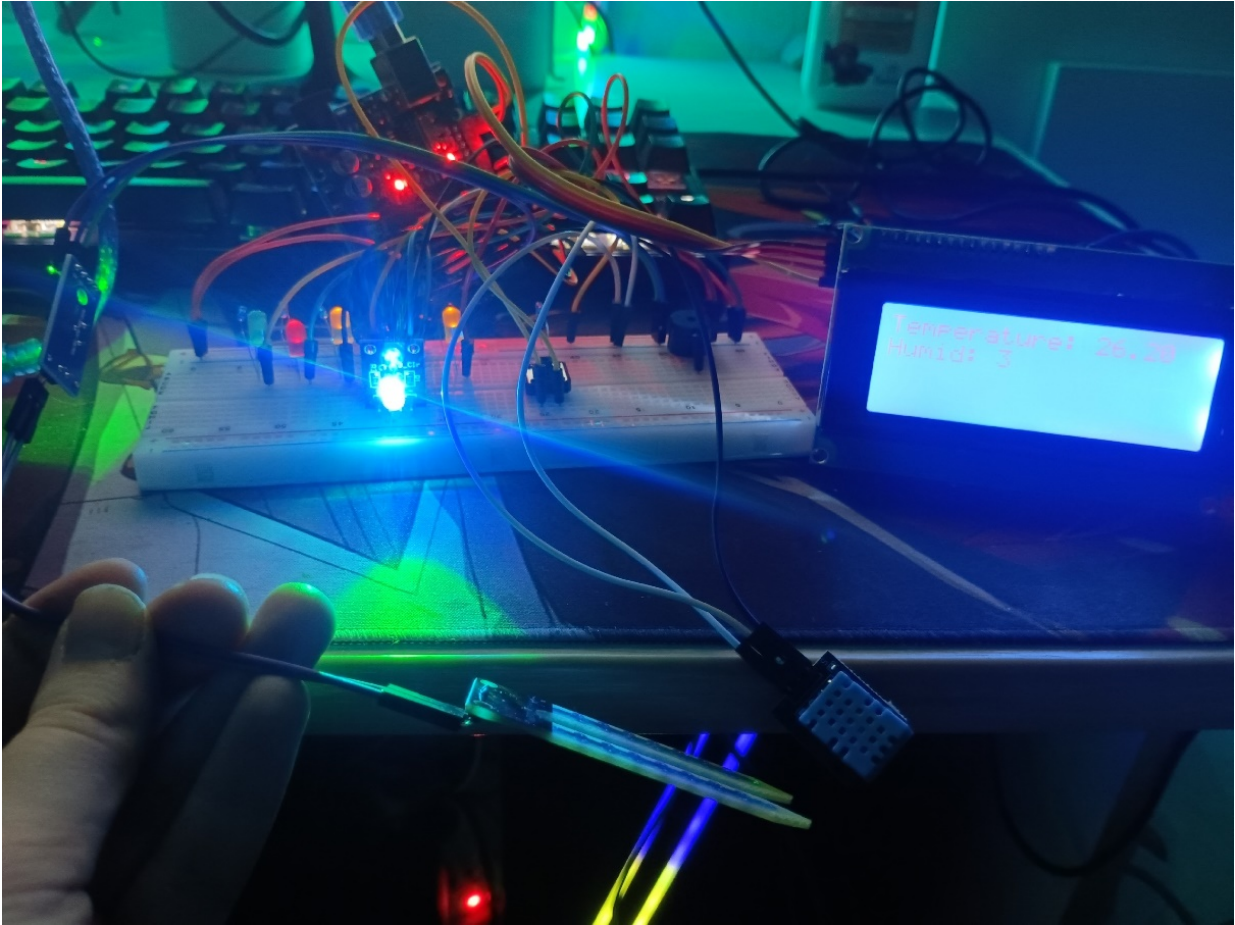
Fig. 7. System layout

Conclusion

The process of modeling and creating a model of a plant growing control system was successfully completed. Further improvements to this design provide the opportunity for automation. Automatic watering, a temperature control system, and a real-time clock can be implemented to provide complete and effective control over the plant growing process.

References

1. Arduino Uno. URL: https://3d-diy.ru/wiki/arduino-platy/arduino-uno.
2. Soil moisture sensor. URL: https://3d-diy.ru/wiki/arduino-datchiki/datchik-vlazhnosti-pochvy-arduino.
3. Humidity and temperature sensor DHT11: URL: https://3d-diy.ru/wiki/arduino-datchiki/datchik-vlazhnosti-i-temperatury-dht11.
4. Plant growth stages. URL: https://gidrostore.ru/Stadii-rosta-rasteniy.

HOME AUTOMATION SYSTEM

Marc'Antonio Lopez, Filippo Simone Iannello, Vito Cammarata

Computer Engineering – "Kore" University of Enna

marcantoniolopez0@gmail.com , filippo.iannello1@gmail.com , vitocammarata00@gmail.com

Abstract. *The evolution of modern technol- ogy has brought forth new perspectives in im- proving quality of life, thanks in part to the au- tomation of systems. With the advent of the In- ternet of Things, the domestic environment be- comes technologically advanced through sensors and controllers, enabling the gathering of in- formation and monitoring of our homes with- out requiring direct human action. In the con- text of this progress, our goal is to develop an advanced home automation system using Mat- lab/Simulink/Truetime, simulating various sce- narios for monitoring our homes.*

Introduction

A domotic environment does not represent a standard solution, as the scalability margin is really high, which leaves a wide margin of choice on the components to be used. Our aim is to in- troduce as many people as possible to the fantas- tic world of domotic, aware of the fact that this can be a source of great help in managing our homes, as well as alerting us to the occurrence of any anomalous situations. In the scenario hy- pothesized by us, the house has various sensors, from temperature to humidity, from smoke to movement...Each of them collects data on the internal (or, in one case, external, as we will see later) state of the house and forwards it to the relevant controller through the network, with the aim of making adjustments on the parameter it- self or triggering an alarm in the most serious case. In section II we illustrate abstractly the structure of the project, the various connections

between devices, how information is transferred from one node to another, as well as nature it- self of measurements carried out. In section III we will carry out a technical analysis on the de- sign on the design of the simulated scenario on the Matlab/Simulink environment, illustrating the various blocks, scripts, subsystems and out- puts implemented. In section IV, network per- formance will be evaluated by referring to three different performance indices, which will provide us with feedback on the quality of the network service and speed of adjustments by our environ- ment.

Related Works

The latest advances in artificial intelligence and machine learning are taking home automation to the next level. Home automation systems can now learn from users' habits and preferences, anticipating their needs and offering personal- ized solutions. A smart home automation sys- tem could automatically adjust the home's tem- perature based on the inhabitants' preferences or initiate certain actions when it detects ab- normal behavior. For example, [1] shows how a home automation gateway can be used and con- figured to meet the needs of homeowners. An- other example is [2] which shows how to use a fuzzy model to make different systems automat- ically adapt to each situation.

Proposed Approach

Our network scenario involves the use of two dif- ferent networks:

1. A wireless network that uses the Zigbee

802.15.4 protocol, since, in a domotic en- vironment, it brings several advantages:

• Zigbee brings extremely reduced en- ergy consumption compared to the use of the WiFi protocol, allowing the sen- sors to be located in different rooms without the need to always be con- nected to the electricity grid, but sim- ply using a battery integrated into the sensor itself which can last several years.

• Zigbee is known for having low latency and this is crucial when performing de- tections on data such as smoke or mo- tion.

• Zigbee offers solutions such as low-cost devices and sensors, in order to facil- itate the scalability of a system and make this world more desirable even to the most sceptical.

The various sensors will send the data from the measurements to one of the two pos- sible net- work gateways. The decision to implement two gateways was made to dis- tribute the computational load for process- ing the packets across two separate devices. One gateway will handle tasks of low im-

portance, such as humidity and external brightness, while the second gateway will exclusively process high-priority tasks such as smoke detection, temperature monitor- ing, and motion detection. This setup aims to minimize the risk of data loss resulting from network congestion.

2. A wired CSMA/CD Ethernet network is employed to achieve maximum data trans- fer speed from the network gateways to five potential controllers, each dedicated to pro- cessing a specific type of data. We imple- mented six sensors in the wireless network:

- Temperature sensor: it detects values between 15C° and 26C° with a sample time of 1 second.
- Humidity sensor: it detects values be- tween 20% and 80% with a sample time of 60 seconds.
- Outdoor brightness sensor: it detects values between 0 lux and 100.000 lux with a sample time of 60 seconds.
- Smoke sensor: it detects values (mea- sured in parts per million "ppm") be- tween 0 ppm and 50 ppm with a sam- ple time of 1 second
- Motion sensor: it detects presence val- ues (1 for motion detection otherwise 0) with a sample time of 0.5 seconds.
- Security camera: it produces a 100 bit/s bitstream (for simulation neces- sities).

Scenario

Using Matlab/Simulink as simulation environ- ment and Truetime library to establish network connections, we provide detailed technical as- pects of the project.
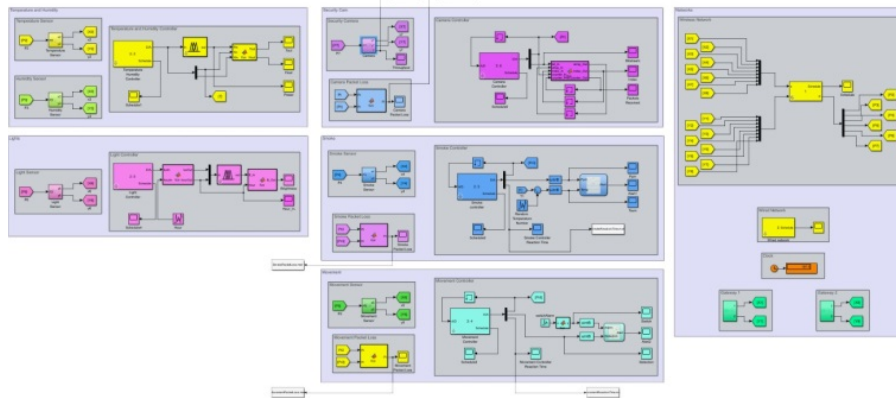
Fig. 1. Project structure.

All the sensors used have the same Simulink structure, and have been implemented in the fol- lowing way:

- ttKernel, which receives a random value as input (in order to simulate every type of possible situation).
- ttBattery, which contains the energy capac- ity of the sensor, measured in mAh.
- Two Const blocks, which we will use to communicate the sensor coordinates to the wireless network.
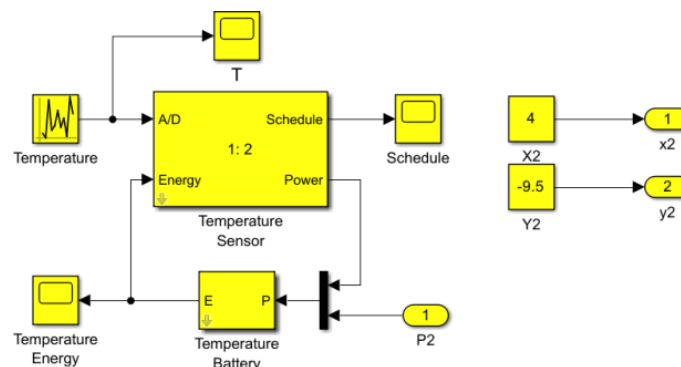
Fig. 2. A sensor structure.

Speaking of controllers, each device has a dif- ferent method of adjusting values and varies depending on the received data. Now, let's de- scribe the five different controllers:

● Temperature and humidity controller. It's made using several blocks:

a) ttKernel, for receiving temperature and humidity data from the wired net- work and sent from the gateway 1.

b) Fuzzy controller, which receives tem- perature and humidity data from ker- nel and returns a power value as out- put.

c) Matlab function, which receives tem- perature, humidity and fuzzy's output as inputs, for making adjustments de- pending on inputs.
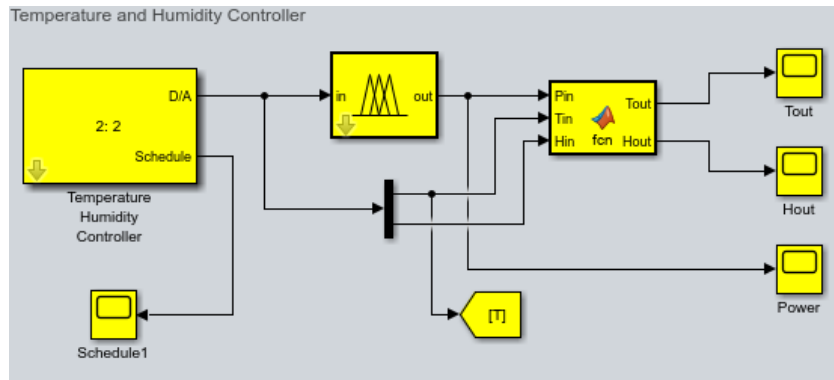
Fig. 3. Humidity and temperature sensor.

● Brightness controller. The blocks used in this case are:

a) ttKernel, to receive external brightness data from gateway 1.

b) Matlab function 1, which receives as input the external brightness and the time of day of detection, and modifies the brightness value according to the time of day (for simulation needs and to obtain the most truthful data).

c) Fuzzy controller, which receives the ex- ternal brightness and the time as input (output of the previous Matlab func- tion) and returns a percentage value of the power of the house lights (ex. lights at 75%).

d) Matlab function 2 receives the output power of the fuzzy controller and the time as input, so as to set the power of the house lights to 0 (0%) between 00:00 and 05:00.
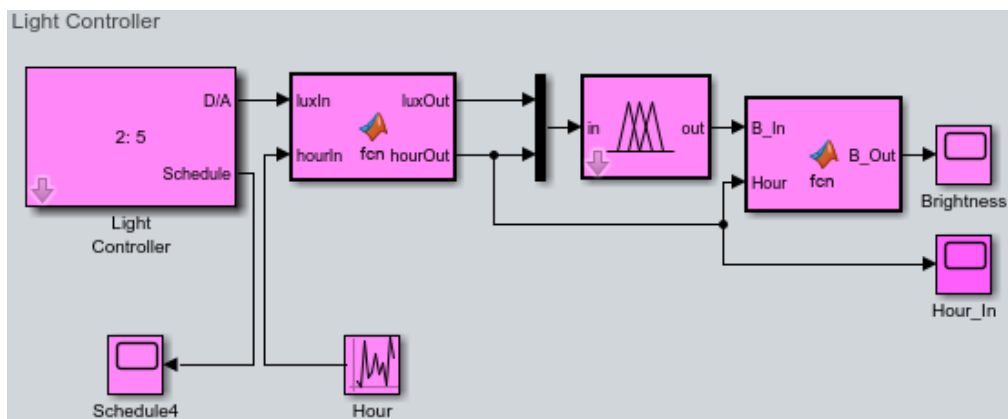
Fig. 4. Light sensor.

● Security camera. In this case we have:

a) ttKernel, to receive the bitstream ar- riving from gateway 1. In addition to the bitstream, it returns the number of packets received.

b) Matlab function, which receives as in- put: bitstream, an array for storing the bitstream, the scroll index of the array, two counters for synchronous ac- tivation of the script (it will only work when a new packet is received); at the output we have the index, the ar- ray and a counter, all feedback. The Matlab function script also provides a function for removing the oldest data: once the capacity of 95% of the array is reached, the oldest data are elimi- nated and the newer data are shifted to the left.
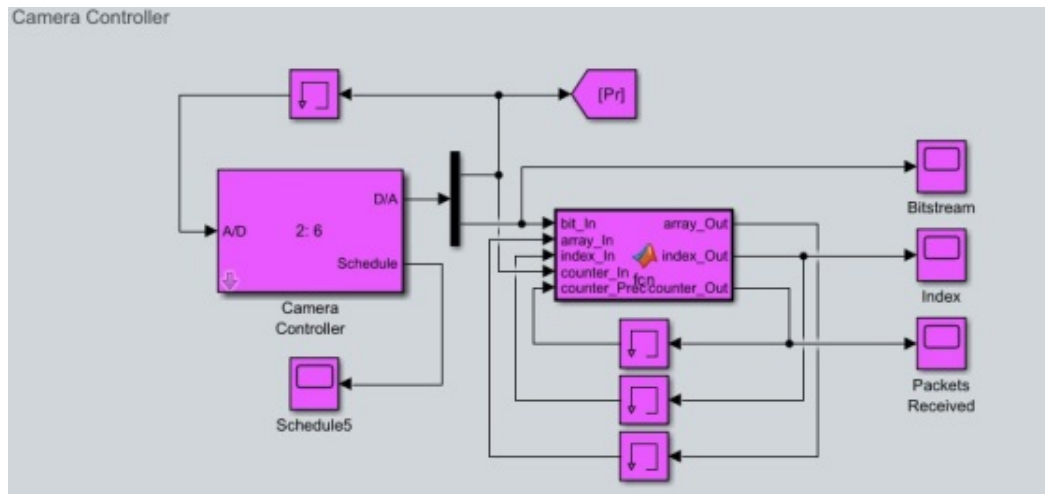
Fig. 5. Security camera.

```
function [array_Out, index_Out, counter_Out] = fcn(bit_In, array_In, index_In, counter_In, counter_Prec)

    % Inizializzazione predefinita
    array_Out = zeros(1, 100);
    index_Out = index_In;
    counter_Out = counter_In;

    % Condizioni
    if counter_In == 0
        % Nessuna modifica necessaria, array_Out già inizializzato
    elseif counter_In ~= counter_Prec
        if index_In == 1
            array_Out(1) = bit_In;
            index_Out = index_In + 1;
            counter_Out = counter_In;
        else
            array_Out = array_In;
            if (length(array_Out) - index_In) == 5
                array_Out = circshift(array_Out, [0, -60]);  % Shift circolare di 60 posizioni
                index_Out = index_In - 59;
            else
                index_Out = index_In + 1;
            end
            array_Out(index_In) = bit_In;
        end
    else
        % Nessuna modifica necessaria, array_Out già inizializzato
    end
end
```

Fig. 6. Data saving script.

● Smoke controller. The blocks are:

a) ttKernel, to receive smoke detections from gateway 2.

b) MBSD, which carries out checks on the amount of smoke in the environment and the temperature, and will issue an alarm at four possible levels, depend- ing on the input values. A further check is carried out using debounce to obtain the most reliable output.
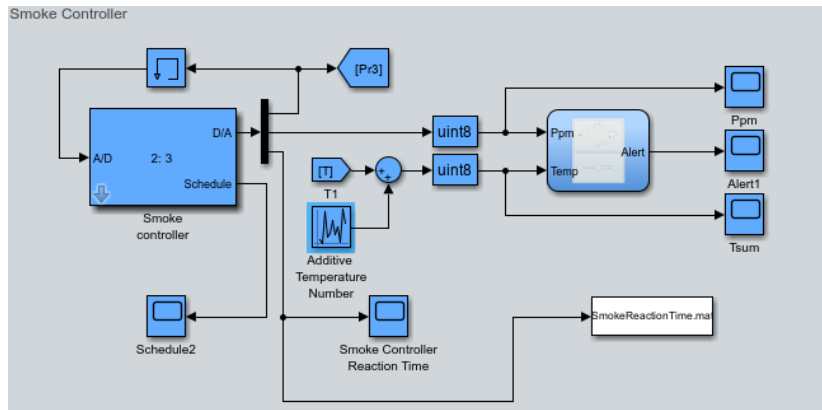
Fig. 7. Smoke controller
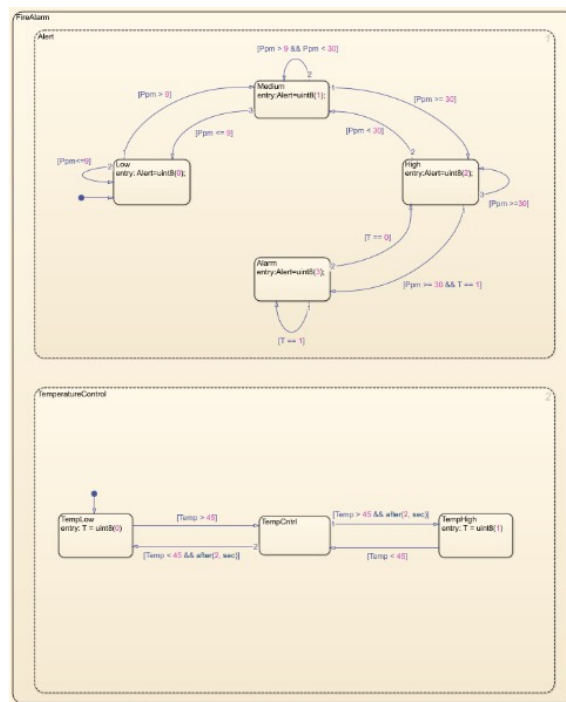
Fig. 8. Smoke controller MBSD.

• Motion controller. Here we have:

a) ttKernel, to receive motion detection values from gateway 2.

b) MBSD, which receives the detection values as input together with a switch that represents the switch-on state of the anti-theft alarm. The output value will be a three-level alarm, depending on the values assumed by the inputs. Also in this case an additional check is carried out using debounce.
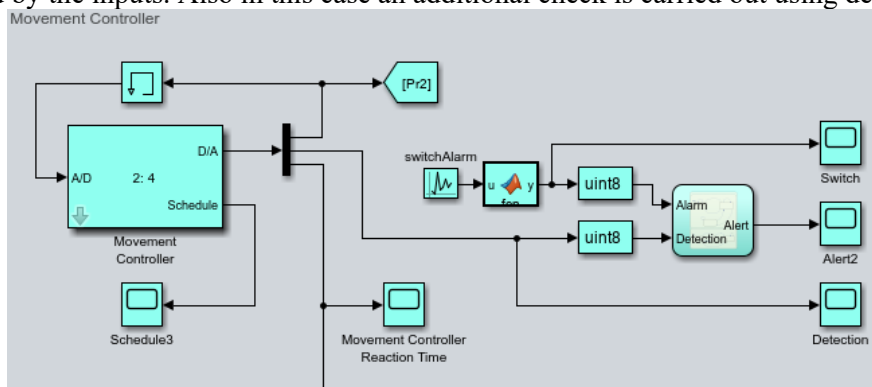
Fig. 9. Motion controller.
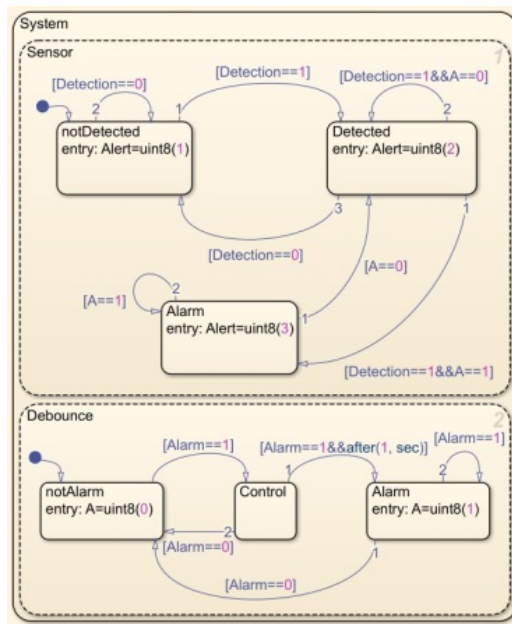
Fig. 10. Motion controller MBSD.

Concerning the networks, these were created us- ing ttWirelessNetwork blocks for the one which uses Zigbee protocol (data rate 80000 bit/s and loss probability 0.001), while ttNetwork for the Ethernet one (data rate 10,000,000 bit/s and loss probability 0). The gateways are made with ttKernel blocks, each of which has its own han- dler, allowing the sensors to receive the relevant packets and forwarding them to the various con- trollers, each one with its own handler.
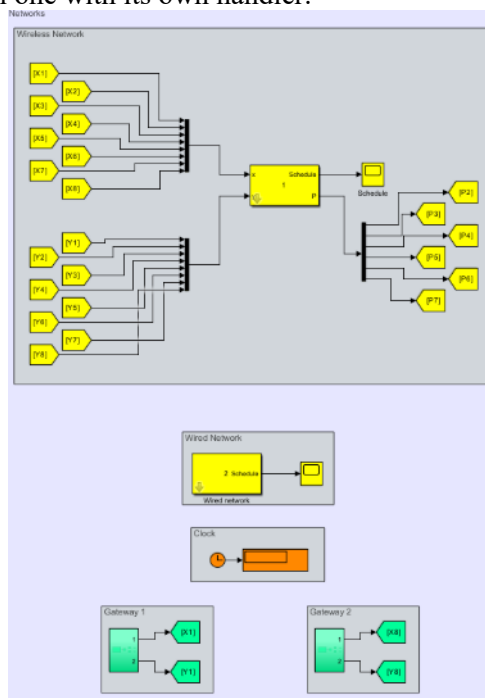
Fig. 11. Networks structure.

Performance Metrics

In the context of a home automation context, it seemed appropriate to carry out the performance evaluation by selecting in particular three differ- ent indices measured on certain sections of the project:

- Throughput. Represents the quantity of data transmitted for unit of time. The choice to measure the throughput fell on se- curity camera, in this we can ensure that the network can support the constant flow of data generated by the camera, without loss of quality or delays. From the graph below

we notice how, in a one hour simula- tion, the throughput tends asymptotically towards a value equal to 2000 bit/s.
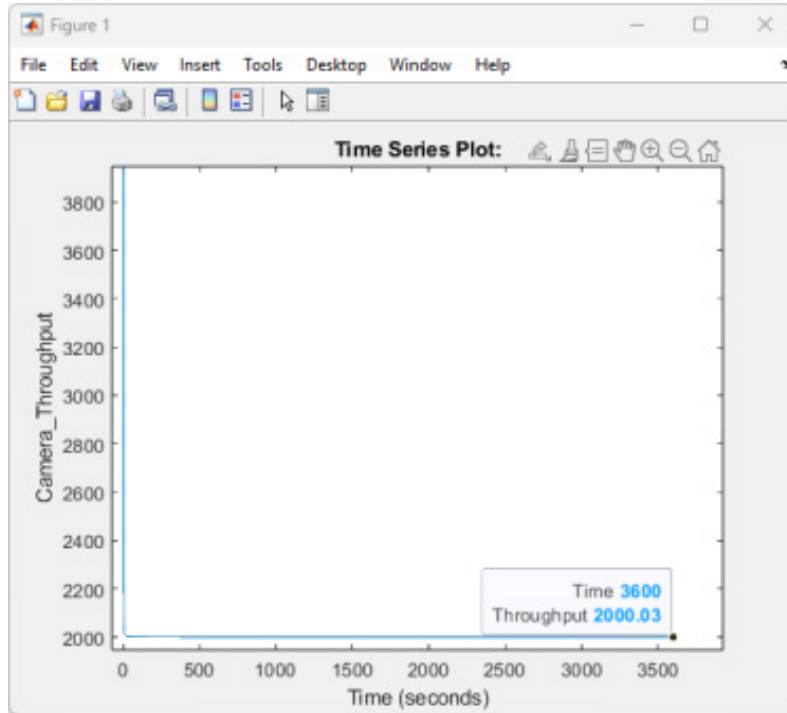
Fig. 12. Security camera throughput.

- Packet loss. This is the percentage of loss of transmitted packets. We decided to mea- sure the packet loss on three different sec- tions, in particular on the security cam- era, smoke controller and motion controller: this is because, in the case of the cam- era, we would like the lowest packet loss to have less buffering and latency. ; regard- ing smoke and movement, we know that these are extremely important tasks where we want the measurement to be as pre- cise as possible. The three values measured in the various devices appear to be differ- ent: in smoke, we have a value of approxi- mately 2.6%, in movement the value is ap- proximately 0.85%, finally in the camera the value settles at around 0.13%. The values obtained show how little data is lost com- pared to that received.
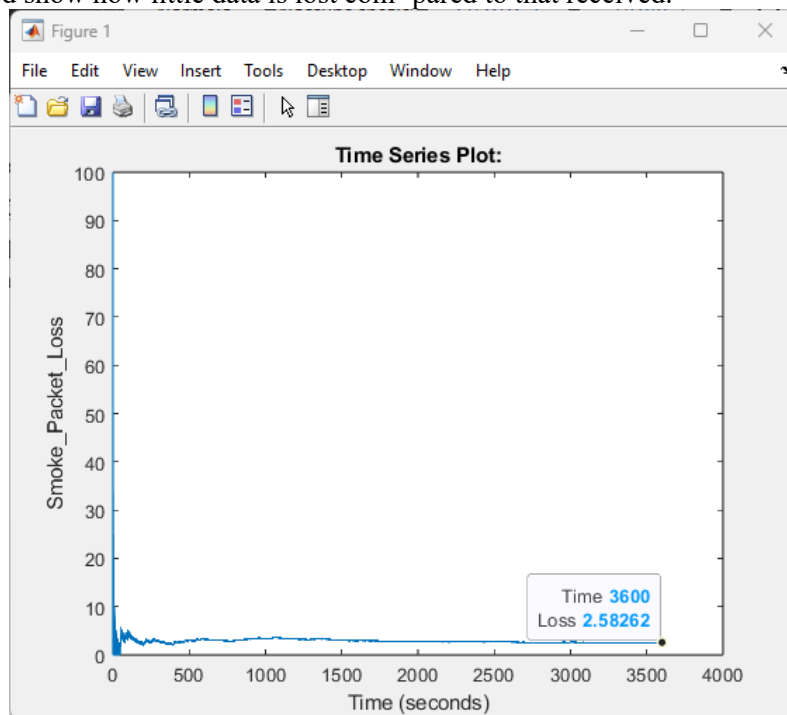
Fig. 13. Smoke packet loss.
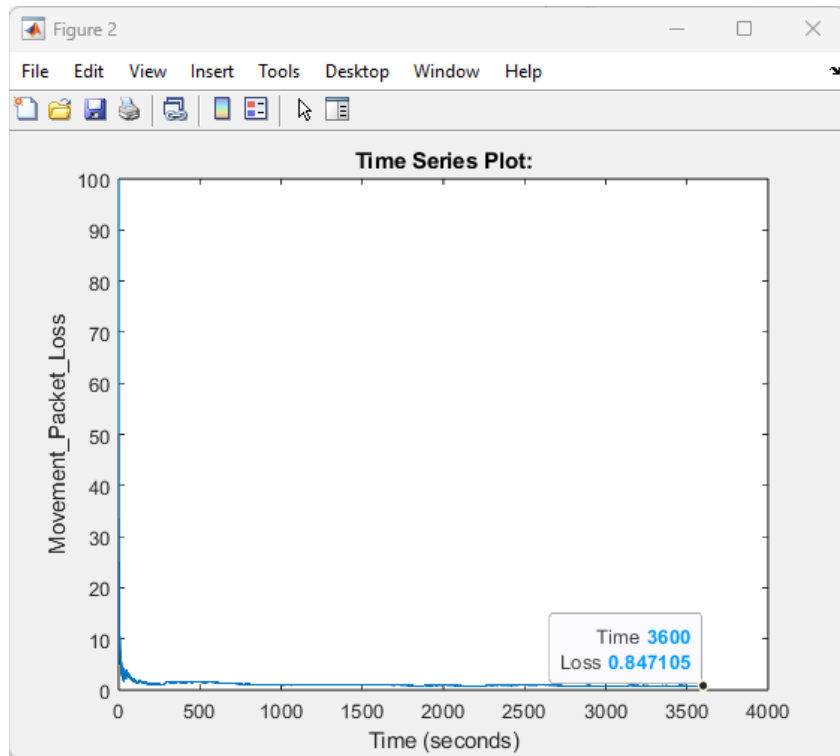
Fig. 14. Motion packet loss.
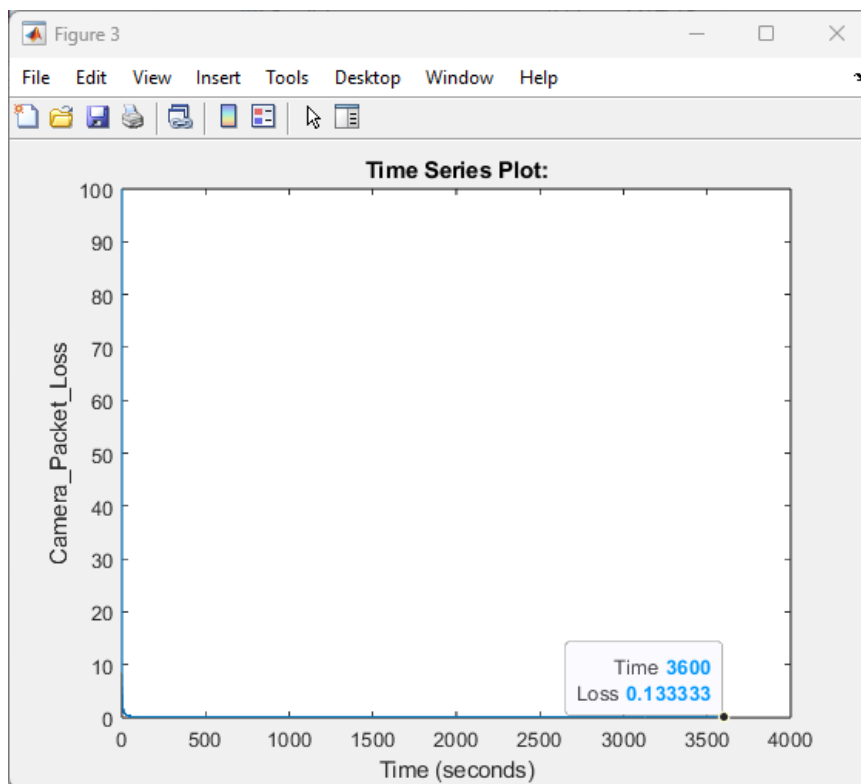
Fig. 15. Camera packet loss.

- Response time represents the difference be- tween reception and transmission time, and is a crucial parameter for evaluating the ef- fectiveness of smoke and motion controllers. These controllers must be able to respond quickly to emergency situations. In our scenario, we measured the reaction time of both controllers and obtained significant re- sults, showing a timely response from both, with average times between 0 and 0.05 sec- onds, well within acceptable limits to ensure the safety of the home environment.
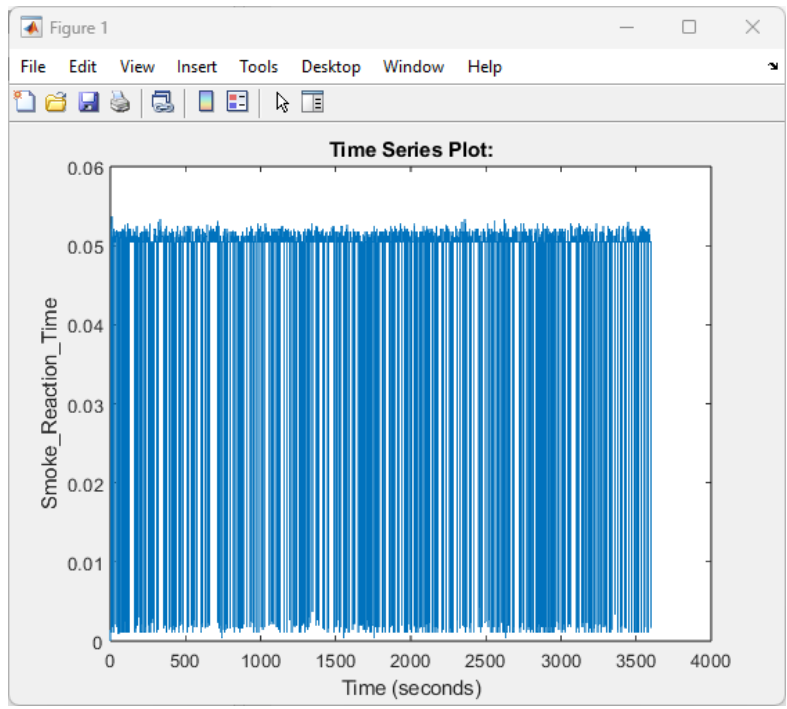
Fig. 16. Smoke reaction time.
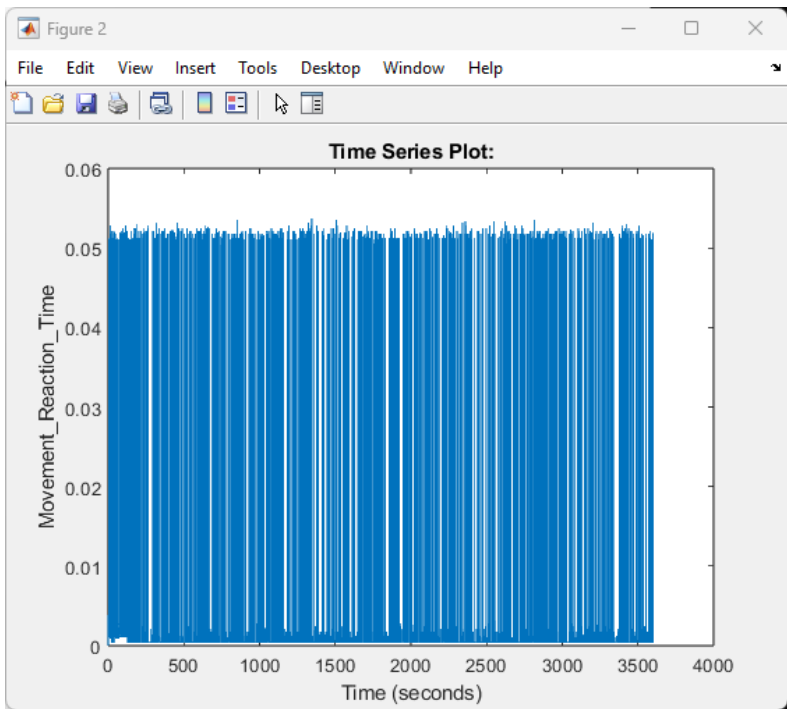
Fig. 17. Motion reaction time.

Conclusions

In conclusion, this home automation project has demonstrated itself as a robust and scalable system. The utilization of both a wireless Zigbee and a wired Ethernet network has provided nu- merous advantages in terms of energy consump- tion, latency, and costs. The strategically placed sensors throughout the house, as shown in the simulation, reliably transmit data to their re- spective controllers. Each controller is designed to handle specific data from the sensors, show- casing good responsiveness and adaptability to various situations. The performance analysis justifies the system's efficiency, achieving suf- ficient throughput to support security camera, low packet loss, and adequate reaction times for critical controllers. We hope that our project has effectively highlighted how implementing a home automation system

can enhance home en- vironment management, providing valuable data for monitoring and automatic regulation.

References

1. Paolo Pellegrino, Dario Bonino, and Ful- vio Corno. "Domotic house gateway". In: *Proceedings of the 2006 ACM Symposium on Applied Computing*. SAC '06. Dijon, France: Association for Computing Machin- ery, 2006, pp. 1915–1920. isbn: 1595931082. doi: 10 . 1145 / 1141277 . 1141730. URL: https : / / doi . org / 10 . 1145 / 1141277 . 1141730.

2. R. S´aenz L´opez, E. Jim´enez Mac´ıas, and M. P´erez de la Parte. "Fuzzy Model for Decision Taking of Technologies in Home and Building Electronic Systems". In: *2008 Second UKSIM European Symposium on Computer Modeling and Simulation*. 2008, pp. 15–20. doi: 10.1109/EMS.2008.18.

OVERVIEW OF 5G CORE NETWORK MONITORING METHODS

Ekaterina Marienko

Saint Petersburg State University of Aerospace Instrumentation,
B.Morskaya 67, 190000, Saint Petersburg, Russia
E-mail: marienkokate77@gmail.com

Abstract. *The 5th Generation (5G) mobile network is the new global standard for wireless communications after 1G, 2G, 3G and 4G networks. 5G provides the new type of network needed to connect virtually everyone and everything, including machines, objects and devices. 5G wireless technology is designed to deliver higher peak data rates of several Gbps, ultra-low latency, higher reliability, greater network capacity, increased availability and a more uniform user experience for more users. This paper presents 5G core monitoring methods and services.*
Keywords: *5G core, observability, monitoring methods, monitoring services.*

1. Overview of 5G core monitoring methods

Currently, 5G networks are being actively developed, which is accompanied by increasing complexity and the number of components in the network. This creates difficulties in evaluating the effectiveness of new technological solutions and increases the time to find the root causes of failures. In addition, fifth-generation mobile networks play a key role not only in connecting people, but also in interconnecting devices, and serve as the basis for the development of areas such as industrial automation, the Internet of Things and smart cities. Due to the maintenance of a large number of connections in these areas, the application of complexes that allow the system to possess the property of observability seems logical and appropriate.

2. Observability

In the past, the term "observability" has been used in control theory to refer to how the state of a system can be determined from its external outputs. As applied to IT, we will use the following definition:

Definition: *observability — is a property that allows estimating the current state of an infocommunication system or its part on the basis of the data generated by it [1].*

Applications and the IT components they use provide output in the form of metrics, events, logs, and traces. Metrics, logs, and traces are analyzed to assess the state of the target application. A system is considered "observable" if its current state can be assessed only from its external outputs – without regard to its composition or architecture.

Observability tools are designed to collect and aggregate as much information as possible from every component of the system, including infrastructure, applications, serverless services, middleware, and databases, to provide a comprehensive view of the internal state of the system at the most critical time: when data is sent to another system for processing and use.

3. Отличия наблюдаемости от мониторинга

In this subject area, monitoring is defined as follows:

Definition: *monitoring is the process of collecting, analyzing, and using information to track a program's progress toward its goals and make management decisions. Monitoring focuses on observing specific metrics.*

Monitoring lets you know something is wrong, and observability will let you know why. Monitoring is a subset and a key action for observability [2].

Monitoring tracks the overall health of the application. It collects data on how the system is performing in terms of access speed, connectivity, downtime, and bottlenecks. Observability, on the other hand, allows you to drill down into what has been happening in the application by providing detailed and contextual information about its specific failure modes.

4. The three pillars of observability

4.1 Basic information of the section

Understand how observability works by examining the telemetry data on which it is built – the so-called three principles of observability [3]. These principles act as the connecting element that allows each of them to provide an almost infinite number of levels of real-time insight into a wide range of system performance across the enterprise.

4.2 Metrics

The term "metrics" in this subject area are defined as follows:

Definition: *Metrics are typically numerical time series data for calculation, aggregation or averaging purposes.*

By analyzing the metrics at the beginning of the time series, we can predict the metrics at the subsequent time interval.

While most monitoring tools can gather metrics from popular platforms and systems to report trends or anomalies over time, they often provide limited insight into when something is broken.

With an observability solution, metrics can now provide critical data to build a response by measuring accurate system performance values. Observability offers hard facts about elements such as service level metrics, latency and downtime. Metrics built from this system data provide actionable visualizations of overall or specific system performance, helping to predict potential system trouble spots.

4.3 Logs

This subject area is defined as follows:

Definition: *Logs are detailed records of each software's events, user actions, and network activity.*

All logs provide a temporary, immutable, step-by-step record of every event a component sees. Along with this detailed information, the logs contain valuable metadata. To get an observable system, each of these logs must be collected and correlated with an event. However, logs alone cannot provide a complete picture of system performance.

The observability solution is designed to centralize event and log data along with other performance data, giving teams enterprise-wide visibility.

Unlike metrics, logs record the historical events of the system, while metrics are numerical indicators of its various parts.

4.4 Tracing

Traces record the end-to-end path of every call made in the distributed system architecture during the execution of a unit of work or transaction. A trace clearly shows every touchpoint that a transaction interacted with during the execution of an action. In other words, a trace records every call made to fulfill a request, the chain of calls from one touchpoint to another, the time of the calls, and the delay between each transition.

Tracking down problems to determine the root cause can be a tedious manual task in distributed networks. And as networks have expanded to cloud, edge, and the Internet of Things, this task has become even worse. There are many more routes into, out of, and through each infrastructure than there were a few years ago.

However, observability makes it possible to centralize these tasks to perform tracing quickly. This method, called distributed tracing (or distributed query tracing), can span the entire enterprise and provide domain- and system-independent visibility of system functionсемы.

The observability solution provides IT organizations with the foundation needed to quickly detect application, network and system failures. It provides a single console to continuously monitor affected systems until a solution is reached, which is critical to enabling IT operations to deliver service delivery capabilities.

5. The role of observability in the 5G backbone network

Observability is the ability of a system to provide detailed, complete information about its internal state based on external observations. In the context of 5G, where networks are becoming more complex and dynamic, providing a high degree of observability becomes critical to ensure efficient operation and problem detection.

Metrics monitoring: In the context of 5G, observability involves monitoring various metrics such as throughput, latency, signal strength, errors in the interaction layer of one core network device with another and other parameters to ensure stable and high performance connectivity.

Event tracing: Tracing in the 5G core network allows tracing the processing of a request from the UE (user terminal) in different nodes of the fifth-generation core network, which is important to ensure timely data analysis and error localization, in case a problem occurs during the processing of a request.

Transaction Logging: Tracking and logging of various events and operations in the 5G network provides the ability to analyze and quickly respond to abnormal situations.

6 Review of existing tools for providing a system with the observability property

There are several services in the observability domain that provide tools for monitoring, tracing and logging in distributed systems. Here is a comparative characterization of some of them:

Prometheus [4]:
Type of service: Monitoring.
Programming language: Go.
Scalability: Suitable for scaling in highly loaded environments.
Advantages: Easy to install, has a wide community, good integration with Kubernetes.
Disadvantages: Lack of tracing and logging support.
Elastic Stack (Elasticsearch, Logstash, Kibana) [5]:
Type of service: Logging and searching.
Programming language: Java.
Scalability: Scales well for data logging and retrieval.
Advantages: Powerful search and visualization capabilities, support for multiple data sources.
Disadvantages: Not as effective for monitoring metrics.
Grafana [6]:
Type of service: Data visualization.
Programming language: Go.
Scalability: Scalable to visualize data from a variety of sources.
Advantages: Support for various data sources, wide dashboard customization options.
Disadvantages: Not a complete monitoring and tracing tool.
Dynatrace [7]:
Type of service: Full observability (monitoring, tracing, logging).
Programming language: Unknown (proprietary solution).
Scalability: High scalability for complex and distributed systems.
Advantages: Automatic dependency detection, extensive analytics capabilities.
Disadvantages: High cost, proprietary solution.

In conclusion, observability plays a key role in ensuring the efficiency and reliability of the 5G core network. Monitoring, traceroute and logging systems work together to provide visibility and control over network operations, which is essential for successful deployment and management of high-performance 5G networks.

References

1. Yoram Mireles. "https://newrelic.com/blog/best-practices/what-is-observability" (circulation date: 16.12.2023)

2. Andrew Magnusson, "https://www.strongdm.com/blog/observability-vs-monitoring" (circulation date: 17.12.2023)

3. https://www.eginnovations.com/blog/the-three-pillars-of-observability-metrics-logs-and-traces (circulation date: 18.12.2023)

4. https://prometheus.io/ (circulation date: 17.12.2023)

5. https://www.elastic.co/elastic-stack (circulation date: 17.12.2023)

6. https://grafana.com/ (circulation date: 17.12.2023)

7. https://www.dynatrace.com/ (circulation date: 17.12.2023)

ANALYSIS OF NETWORK PERFORMANCE ON AN AUTOMATED SWIMMING POOL USING VARIOUS SOFT COMPUTING TECHNIQUES

Miccichè Giovanni

Computer Engineering and Networks Laboratory – Kore University of Enna – Italy
Email: giovanni.micciche@unikorestudent.it

Abstract. *Maintenance of a swimming pool requires monitoring of certain factors such as: temperature, humidity and water level. This is an activity that can be time consuming and labor intensive. To automate this process and make it more efficient, we can use a soft computing technique.*
In fact, the following article will be concerned with presenting a soft computing approach for controlling temperature, humidity, and water level in swimming pools. The approach uses wired and wireless networks to interconnect various devices, including sensors, controllers, and actuators,so as to optimize the performance of the pool control system.
This approach has been tested on a simulated pool, and the results have shown remarkable performance without the need for a support operator.

Introduction

The first step of this project includes the establishment of 3 networks, 2 of which are wired and one wireless. As for the monitoring system, it includes 4 elements: temperature, humidity, water level inside the pool, and the number of people entering and leaving the facility. Then a control system for the first 2 variables was created using a Fuzzy controller and an actuator. Next, I entered the MBSD to generate the 4 states in which the pool can be found. Finally, I analyzed the performance of the 3 networks.

RELATED WORKS

Pool monitoring is a topic that is covered in many projects.

The author of [1], for example, analyzed not only water, temperature and humidity levels but also chlorine and ph. These elements are connected via a network: the various sensor and actuator nodes are connected to the gateway, which in turn connects to a server where all the values it is receiving will be saved. To improve security, an application is integrated that receives the data from the server, showing it via graphs to the admin who, in case of anomalies, will be promptly notified by notification. The advantages of this approach can be related to security as it will be greatly improved.

The author of [2] created a low-cost system based on a wireless network of sensors and actuators. The main goal of this system is to save economic and natural resources for its users, contributing to a more sustainable environment.

He has developed an Android application that allows remote monitoring and control of data collected from swimming pools in real time. For this project, a Deep Sleep function was used on the sensor nodes, which allowed him to reduce energy spent as it put them in a sleep state when no data collection was planned while preserving battery power.

THE PROPOSED APPROACH

The scenario of this project involves the implementation of an architecture comprising the *control center*, where the controllers reside, the *pool* on which our sensors will be installed, the *turnstiles* that record the flow of people in and out by sending the information to the control center, and finally the *maintenance room* that receives the data from the control center and activates the actuators.

In the following architecture, 3 networks will be used: the first wired of CSMA/CD (Ethernet) type, as well as for the second, while for the third a Wireless network of 802.11b (WLAN) type was used.

It should be pointed out that different control methodologies will be used according to the variables to be analyzed: the first control methodology used is *Fuzzy*, which from the levels of temperature and humidity generates a power level, taken into account by an algorithm to quantify whether the value of these variables should increase or decrease. Then another methodology used takes into account an algorithm that calculates the water to be introduced from the ideal and actual levels. Finally, an additional control technique was used within the *MBSD* to prevent unintended switching on or off of the general system or the turnstile system alone.
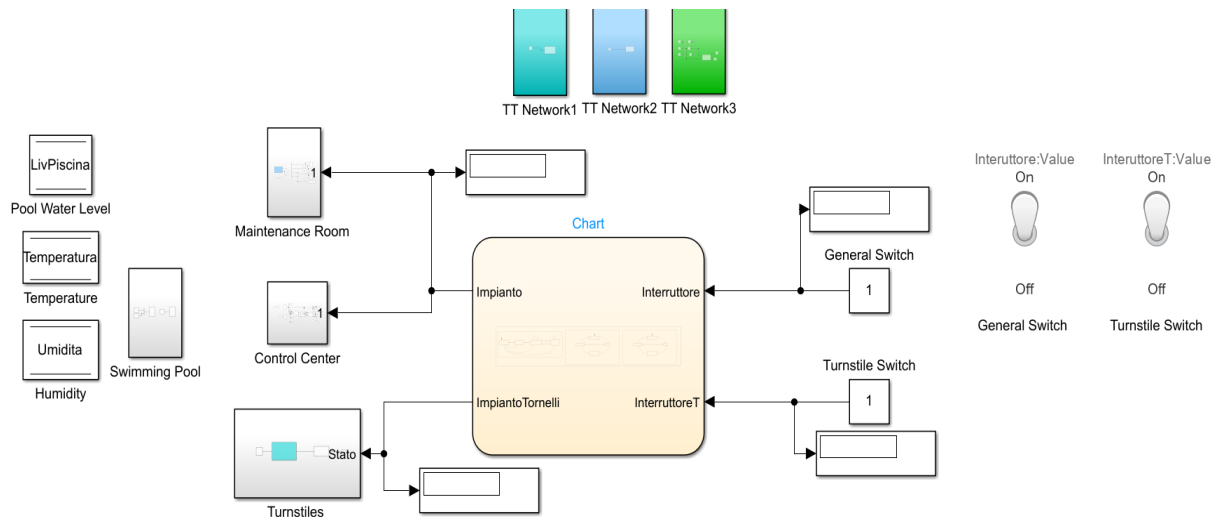
To simulate the three variables-Water Level, Temperature, and Humidity: this block was exploited to store data, read data, and enter new data. In fact, the Data Store Read was exploited in the sensors to simulate reading, and the Data Store Write in the actuators to enter the new level of the 3 variables.

Let us now analyze the 7 subsystems of which the project is composed.

3 of these subsystems were used to create 3 networks, exploiting TrueTime blocks:

the first was used to create the first wired network that interconnects the TT Kernel Turnstiles and the TT Kernel Controller Turnstiles; while in the second subsystem it will be used to connect TT Kernel SM Controller and the TT Kernel Maintenance Room; In the third subsystem instead, the TT Network Wireless was used to create the Third WLAN having 3 nodes, the first one being TT Kernel SensorT/U, the second TT Kernel Pool Controller and the third TT Kernel SensorLivA.

The fourth subsystem was used to contain 2 other subsystems representing the two sensors SensorT/U and SensorLivA. The first one receives as input the temperature and humidity values, which are to be directly read by the kernel, which will create the msg packet. this packet will contain in the .payload the two variables and sends msg to node 2 of network 3.

The second subsystem takes as input the water level, which is directly read by the kernel, which will create the msg packet by inserting in the .payload the variable LivWater. finally it sends msg to node 2 of network 3 that is TT Kernel Pool Controller.

Let us now go on to analyze the remaining 3 subsystems, which include that of the turnstiles, control center, and maintenance room. The first consists of another subsystem that generates the number of people entering and leaving, which will be read by the kernel. The Turnstiles kernel in particular will create the msg packet by inserting the two variables (InPerson and OutPerson) in the .payload and send msg to the control center.

The second is formed by 3 subsystems, representing the 3 Controllers, and also by a control system,formed by the fuzzy and the MatlabFunction. as for the controllers we have the Pool Controller, formed by the Pool Controller kernel, which receives the packets from the two sensors and forwards the output values,and then generates the response time that will be sent to the workspace; the Turnstiles Controller is formed by the Turnstiles Controller kernel, which receives the packets from the Turnstiles, sums the read values to the total number of people and outputs the value of the people inside,the total number of people entered and exited and finally the response time, which will be sent to the workspace; the SalaM Controller is formed by the SM Controller kernel, which receives the levels of the 3 variables as input.These determine the increase or decrease of the 3 variables. Finally, the controller creates the msg packet that it will send to the TrueTime Kernel Maintenance Room node.The fuzzy controller, which as was mentioned before is part of the control system along with the Matlab Function, based on the temperature and humidity generates an appropriate power level, trying to get the humidity to a value of about 60% and the temperature on the 26 degrees. As for the Matlab Function, several thresholds have been set: low, medium, high, very high. Each threshold corresponds to an action to be performed.

the last subsystem is formed by the kernel Maintenance Room that receives the packets sent by the ControllerSM; adds the increments to the variables water, temperature and humidity; finally updates the variables (it happens only if the system is on since the threshold value intervenes, which if less than 1 does not allow the update).

This chart consists of 3 additional charts executed in parallel.
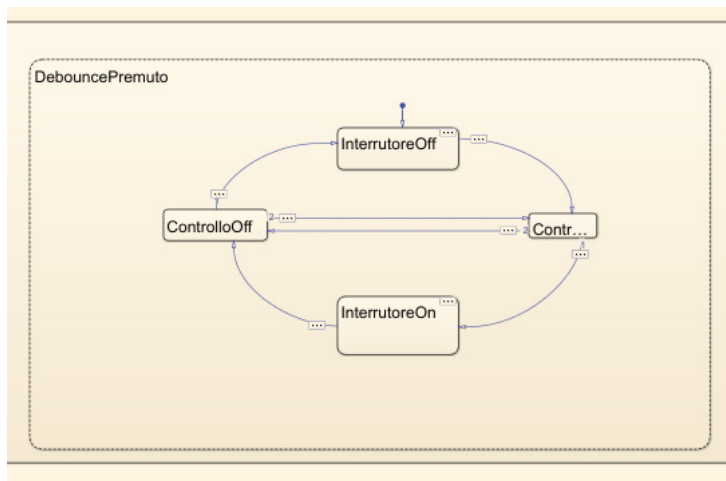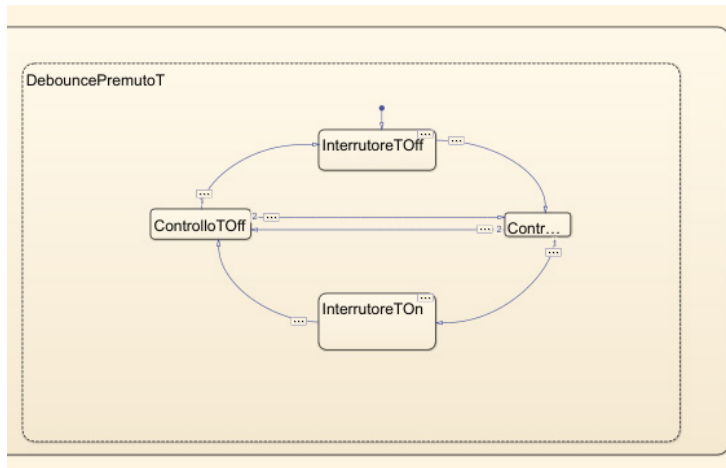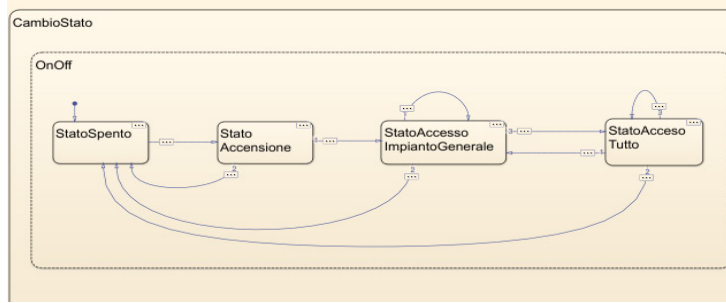
The First char deals with the various states in which the plant can be found. we can distinguish different states.

In state 1 The system is in the "StateOff". in state 2 when the system switch goes to 1 the variable pressed goes to 1 and the system goes into the "StateOn", which stays there for 600 seconds, at the end of which it will go to the "StateOn". If during that time interval the general system is turned off, it will go to "StateOff". in state 3 it is at "StateOnGeneralPlant" and 3 cases can occur: If the turnstiles are turned on, it will go to "StateOnGeneral"; If the general plant is turned off, it will go to "StateOff"; otherwise it will remain in the current state. in state 4 When it is at "StateOnGeneral" 3 cases can occur: if the turnstiles are turned off, it will go to "StateOnGeneral Plant" (see state 3); If the general plant is turned off, it will go to "StateOff"; otherwise it will remain in the current state; Second is Third chart are very similar: they take care of the control on the power on/off key by making the key pressed and pressedT change value only if the switch changing value remains for more than 5 seconds,on or off, to avoid unintended power on.

PERFORMANCE EVALUATION

In this project, the Response Time and Throughput of various elements were evaluated
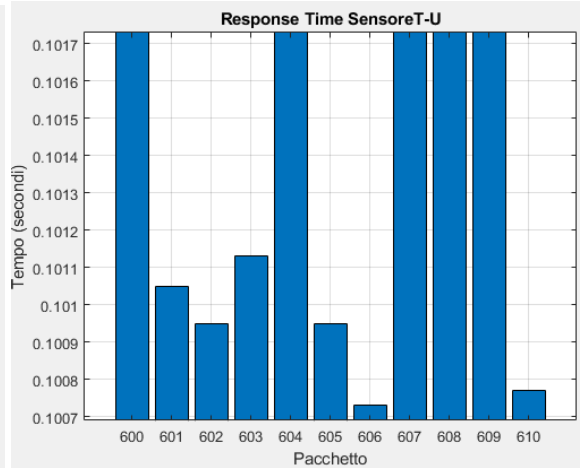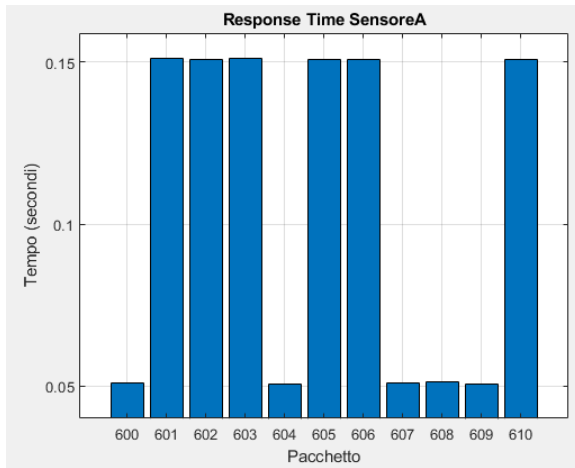
In these two bar graphs we see the time between the acquisition of the value of the 3 variables in the sensors and their reception in the controller pool. As we can see these values are not always equal but if we look closely at the second graph we could almost consider the values equal to each other since they differ at most by 0.0009 of a second
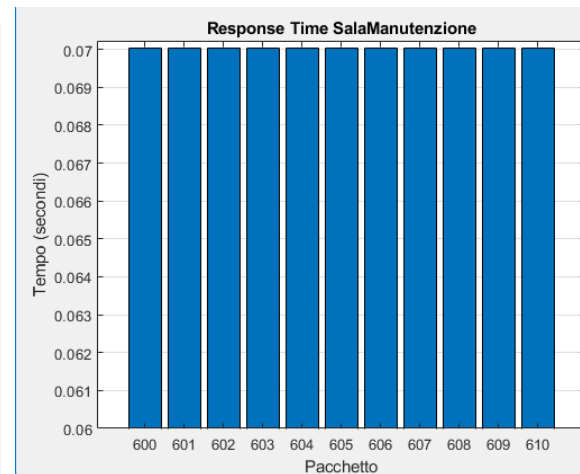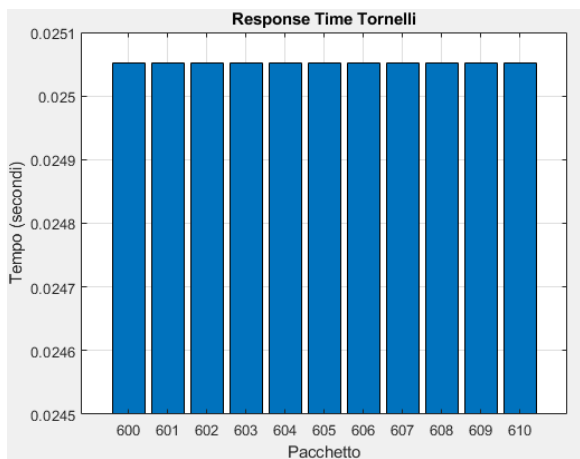
From these bar graphs we can see that it contains constant values, in fact the Response Time Turnstiles, which is the time between capturing people in and out, and receiving and processing these values from the Turnstile Controller is always the same. A similar case is the Response Time Room-Maintenance since the time between sending packets from the Controller and receiving and processing these values from the Actuator is always the same.
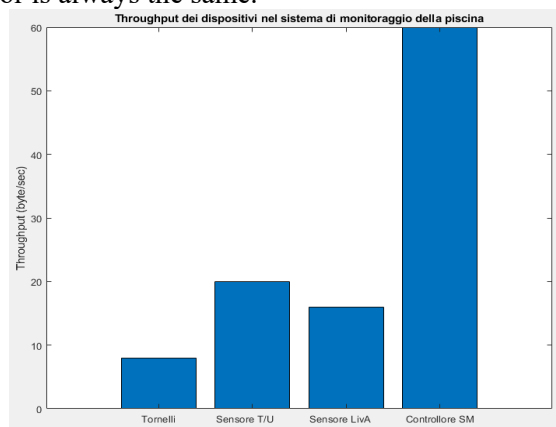
The Throughput, which would be the number of bytes sent in an interval of time, given the nature of the design, is always constant since the various kernels that send the fixed size packets being periodic send them at regular intervals. the throughput for Turnstiles is 8 bytes per second since it will send an 8-byte packet every second, the throughput for T/U Sensor is 20 bytes per second since it sends a 10-byte packet every 0.5 seconds; the throughput for LivA Sensor is 16 bytes per second since it sends an 8-byte

packet every 0.5 seconds; the throughput for SM Controller is 60 bytes per second since it sends a 30-byte packet every 0.5 seconds

Conclusion

In conclusion, this project shows us how these systems can help us design and implement complex systems by first starting to study and analyze them on the computer without them being physically implemented.

In the following project we have analyzed some variables such as water level,temperature, humidity and flow of people in and out but in future developments we could expand this project by analyzing other variables such as ticket cost as a function of people flow through fuzzy controller, or how the workload at the control center would change as the number of pools increases

References

1. G. Simões, C. Dionísio, A. Glória, P. Sebastião and N. Souto, "Smart System for Monitoring and Control of Swimming Pools," 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, 2019, pp. 829-832, doi: 10.1109/WF-IoT.2019.8767240.

2. Gonçalo ARS, Smart System for Monitoring and Control of Swimming Pools, iscte , 2019.

PYROMETRIC SENSOR FOR A SYSTEM FOR REMOTE TEMPERATURE CONTROL OF A FERROELECTRIC SAMPLE

Rostislav Passet

undergraduate student
*Saint Petersburg State University of Aerospace Instrumentation (SUAI),
Saint Petersburg, Russia
rostislav.passet@mail.ru*

Abstract. *An experimental setup has been created that based on a non-contact method for monitoring the temperature response of a sample to external thermal and/or electrical influence using a high-speed IR temperature sensor. The accuracy and reliability of a quantitative assessment of the magnitude of the electrocaloric effect is ensured by corresponding characteristics of the temperature sensor and developed measurement techniques based on the principles of the theory of information-measuring systems. The device provides the detection of thermal radiation from the surface of an object with a diameter of up to 1 mm in the temperature range of 20-200 ℃ and registration of its changes with a speed of up to 1 ms with temperature fluctuations from 0.1 ℃.*
Keywords: *Electrocaloric effect, pyroelectric effect, radiation IR-temperature sensor, ferroelectric materials.*

Introduction

In recent years, there has been a significant increase in interest in ferroelectric materials that have an essential electrocaloric effect (ECE), which manifests itself in a temperature change of a sample under the influence of an electric field and is inverse to the pyroelectric effect change in the polarization of a material P with a change of its temperature. The surge of interest in ECE materials is due to the prospects for their use as cooling systems for microelectronic elements. The progress of microelectronics technology in recent years has led to the emergence of crystals with extremely high current densities. However, the losses generated by semiconductor crystals during operation lead to an increase in their temperature and, as a result, decrease in performance and reliability. There is an empirical relationship according to which, when the average operating temperature of a power crystal increases by 20ºC, its service life is halved. Therefore, the problem of the heat removal is one of the most important for the further development of microelectronics. The expected efficiency of ECE coolers should significantly exceed the currently widely used Peltier thermoelements, which is due to the fact that the ECE is based on a change of the applied field, rather than on the of current flow. Typical ECE values for ferroelectric materials are fractions of a degree [1] and significantly depend on temperature and external heat transfer conditions, which requires the use of precision methods and means of temperature control. At the same time, methods for ECE measuring are not standardized; the vast majority of researchers quantify the ECE indirectly, neglecting a number of influencing physical factors, which leads to a significant scatter of data in available scientific literature [1-3]. First of all, this is due to the limited measuring capabilities of contact temperature sensors (inertia, additional mass, self-heating, etc.) and the significant error of existing IR temperature sensors.

1. MEASURED QUANTITY

A quantitative characteristic of the ECE is the change of the sample temperature, δ T, upon application/removal of an electric field. However, the difficulties noted above in measuring the sample temperature lead to the fact that most researchers determine this value indirectly through measurements of the dependence of the pyroelectric coefficient on the applied field, and Maxwell's relations, that presuppose accurate knowledge of the various characteristics of the ferroelectric material and their field and temperature dependencies:

$$\delta T = -\int_{E_2}^{E_1} \frac{T}{\rho \bullet \bullet C_\rho(E,T)} \cdot \left(\frac{\partial P(E)}{\partial T}\right)_E dE \cong \frac{T}{\rho} \int_{E_2}^{E_1} \frac{1}{C_\rho(E,T)} \rho(E,T) dE$$

where $C_p(E, T)$ is the specific heat capacity; ρ is the density of the sample material; T is the temperature; E is the electric field strength; $P(E)$ is the polarization; $p(E, T) = dP(E)/dT$ – pyroelectric coefficient.

Obviously, the most objective data can be obtained using non-contact (radiation) sensors that do not introduce methodological errors associated with the presence of additional heat removal from the sample, inertia, and the dependence of contact sensor readings on heat transfer conditions. This is especially important when studying film materials. However, the use of non-contact radiation sensors also has its limitations associated with the interpretation of measurement results of the recorded thermal power, its functional connection with the true temperature of the sample, with and taking into account external illumination.

2. BLOCK DIAGRAM OF A PYROMETRIC SENSOR

This paper presents a pyrometric sensor specifically designed for monitoring the temperature of small crystals in the SuperNova X-ray diffractometry system. The sensor is made on the basis of an uncooled IR photodiode (developed by the A.F. Ioffe Physicotechnical Institute , www.ioffeled) sensitive in the region of $4.25\pm0.25\mu m$. To deliver the thermal radiation from the object, a microstructured cladding optical fiber (OF) based on silver halide crystals with a light-conducting core diameter of 750 μm optimized for the IR spectral range of 2-10 μm was used [4]. The experimentally measured value of the optical transmittance in the region of 4 μm at the length of 1 meter was $\tau = 0.5$. Connectors of the SMA-905 type are installed, on both sides of the OB cable, protected by a PEEK polymer tube with a diameter of 2 mm, which allow the matching with the input of the pyrometric sensor and measurement object. The diameter of the light-conducting core is 750 μm with a numerical aperture of NA~0.22, which makes it possible to collect the thermal radiation from the surface of objects with linear dimensions from ~1 mm depending on the distance ΔL [mm] between the object and the end of the optical fiber: d [mm]$\approx2\cdot NA\cdot\Delta L+$ 0.75.

The PD switching circuit converts the current generated in it under the influence of radiation into a measured voltage signal for the subsequent digitization, transmission to a computer and calculation of the absolute temperature of the object in accordance with the Planck's radiation law.

To calculate the true temperature of an object, the sensor is accompanied by a calibration characteristic for a reference object (blackbody model), algorithms for the calibration and calibration for a real object, as well as methods for calculating the temperature of an object in various measurement modes.

The block diagram of the pyrometric sensor structurally contains analog and digital parts and is shown in Fig. 1.
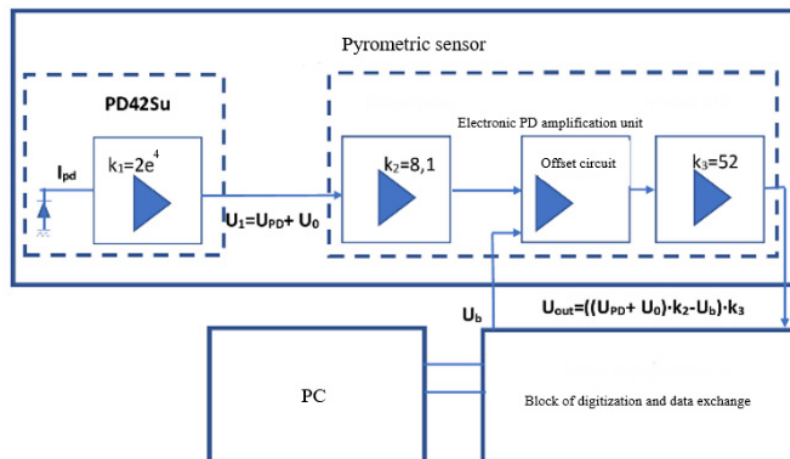

Fig. 1. Block diagram of the pyrometric sensor

The digital part of the pyrometric sensor is made on a microcontroller (MC) ADuC7026. The analog part of the circuit for detecting and amplifying the signal of the pyrometric sensor contains 3 amplification stages with gain factors $k_1 = 2 e^4$ (made on a low-noise op-amp chip of the ADA 4897-1 type, located in close vicinity to the PD chip) $k_2 = 8.1$ and $k_3 = 52$. Circuit is made between the 2nd and 3rd amplifier stages, that allows the introduction of additional bias necessary for further amplification. The total amplification of the PD current is 8.425e6, which theoretically makes it possible to meet the requirement for the sensor sensitivity ($\pm0.1°C$) when measuring temperatures from 20°C for samples with linear dimensions up to 2 mm, losses in the optical circuit of the sensor up to 25 dB, and response speed up to a few milliseconds.

To ensure the required gain, a bias is provided in the electrical circuit of the pyrometer, that is generated by a 12-bit DAC, being a part of the microcontroller (MC) ADuC 7026. The bias allows one to adjust the transfer characteristic of the sensor in an *on – line* mode to match the measured PD signal and the range of the ADC input signal, which is illustrated in Fig.2.
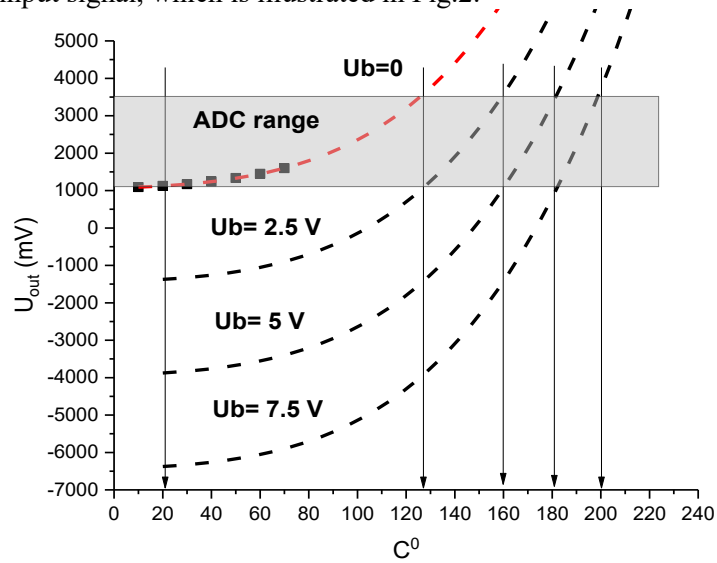

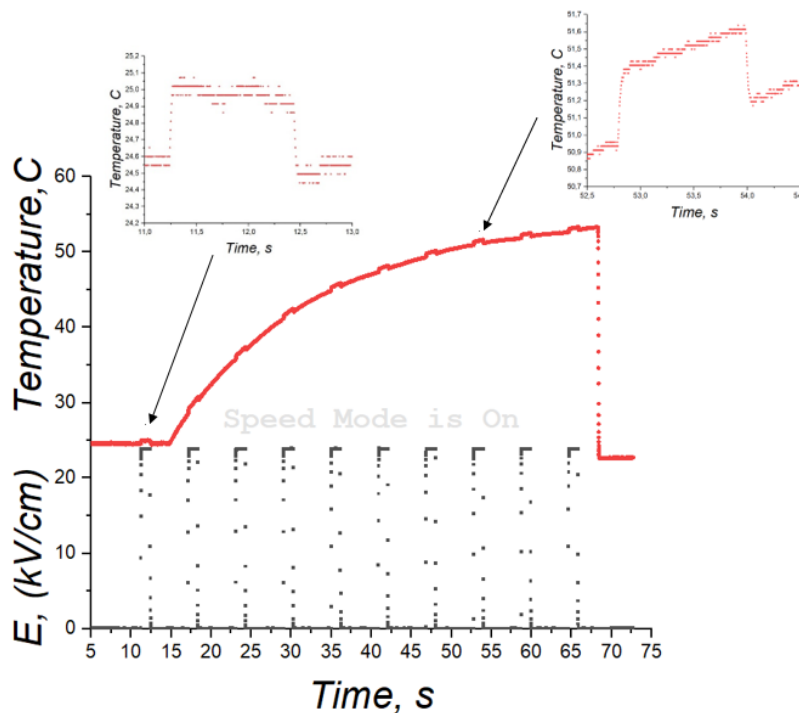*Fig. 2. Transfer (calibration) characteristic of a sensor with an optical circuit (red dotted line)*

Fig. 3. Results of an experiment on recording the ECE in a PMN-PT crystal measuring 3·3 mm^2 at a temperature of 20-60 °C when heated by laser radiation with λ = 0.98 μm.

The resulting output signal, Uout, is fed to a 12-bit ADC included in the microcontroller. The response speed of the pyrometric sensor is ~ 3 ms per count. During this time, the PD signal is measured and the data on the value of the output signal and displacement (U_{out} and U_b) is transferred to the computer for further processing and calculation of the corresponding surface temperature values of the measurement object.

3. Experiment

The goal of the research was to study the ECE in PMN-PT crystals near a phase transition induced by an external influence using the thermal radiography method. The peculiarities of the diffractometric experiment related to the protection of the operator from X-ray radiation, the mobility of the goniometer with the sample and the lack of direct visibility of the object make it possible to implement the pyrometric methods of temperature measurements only with using fiber optics (OF). High requirements for the sensitivity and speed of response for small sizes of the emitting surface in the temperature range up to 200 °C cannot be met by pyrometers available on the domestic and world markets [5].

Th high sensitivity in the temperature range of 20-200 °C is obtained due to original algorithms for detecting and processing the pyrometric sensor signal, including the ability to adjust the offset in 'online' mode during the experiment, which allows the use of standard multi-channel interface boards for digitization and signal transmission using standard 12–14 bit ADCs with a range of analog input signals up to 2.5 V [6,7].

As an example, Fig. 3 shows the results of an experiment on recording the ECE in a PMN-PT crystal with a size of $3 \cdot 3$ mm^2 in the temperature range of 20-60 °C when heated by laser radiation with $\lambda = 0.98$ μm. The experimentally confirmed accuracy of temperature measurement in the considered measuring system was 100 mK in the object temperature range of the order of 20°C, 50 mK at 100°C and 10 mK at temperatures above 200°C at the sensor response speed 10 ms.

Conclusion

The presented experimental results and created experimental setup clearly demonstrate the effectiveness of IR radiometry as a fast, reliable and sensitive method of electrocaloric measurements in a wide range of changes in the temperature of the sample and the external electric field applied to it [6,7].

Th high reliability, stability and reproducibility of measurements are ensured by an algorithm for calibrating the temperature sensor to a real object before starting measurements.

The unique measuring capabilities of the developed sensor are due to the use of modern domestic element base in the field of IR photonics: quasi-monochromatic IR photodiode and microstructured cladding fiber based on silver halide crystals optimized for the mid-IR spectral range.

References

1. Z. Kutnjak; B. Rožič; R. Pirc: Electrocaloric Effect: Theory, Measurements, and Applications // Wiley Encyclopedia of Electrical and Electronics Engineering. 2015. doi.org/10.1002/047134608X.W8244

2. Kar-Narayan S.; Mathur N.D.: Direct and indirect electrocaloric measurements using multilayer capacitors, J.Phys D., Appl. Phys.. 2010. Vol. 43, 32002. doi.org/10.1088/0022-3727/43/3/032002

3. S. Kar-Narayan; S. Crossley; X. Moya; V. Kovacova; J. Abergel; A. Bontempi; N. Baier,E. Defay; and N.D. Mathur: Direct electrocaloric measurements of a multilayer capacitor using scanning thermal microscopy and infra-red imaging, Appl. Phys.Lett.. 2013. Vol.102, 032903. doi.org/10.1063/1.4788924

4. L. N. Butvina, O. V. Sereda, A. L. Butvina [et al.], Large-mode-area single-mode microstructured optical fibre for the mid-IR region // Quantum Electronics. – 2009. – Vol. 39, No. 3. – P. 283-286. – doi: 10.1070/QE2009v039n03ABEH013982.

5. Sotnikova G.Yu., S.A. Alexandrov, Gavrilov G.A., Medium-wave IR pyrometry using photodiodes based on InAs and InAsSb (review) // Advances in Applied Physics, v. 10, no. 4, pp. 389-391, 2022. doi: 10.51368/2307- 4469-2022-10-4-389-403 (in Russian).

6. G. Y. Sotnikova, G. A. Gavrilov, A. A. Kapralov [et al.], Electrocaloric studies of bulk materials and multilayer structures by dynamic infrared radiometry // Ferroelectrics. 2022. – Vol. 591, No. 1. – P. 157-165. doi: 10.1080/00150193.2022.2041934.

7. Sotnikova G.Yu.; Gavrilov G.A.; Muratikov K.L., Passet R.S., Smirnova E.P. Temperature response features of ferroelectric ceramics in electrocaloric effect study, 2022, Tech. Phys. Lett., v.48, 15, doi: 10.1134/S1063783421060226

ON THE CALCULATION OF THE CHARACTERISTICS OF A WAVEGUIDE IN A METAL SHIELD WITH A HOMOGENEOUS DIELECTRIC COATING

Gleb Revunov
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
revunpank@gmail.com

Abstract. *The article presents a method for calculating the main characteristics of an antenna, which is the open end of a waveguide of arbitrary cross-section with a perfectly conductive metal shield covered with a homogeneous dielectric layer of α thickness*
Keywords: *antenna, waveguide, dielectric, open resonator.*

The article presents a method for calculating the main characteristics of an antenna, which is the open end of a waveguide of arbitrary cross-section with a perfectly conductive metal shield covered with a homogeneous dielectric layer of α thickness. For a circular waveguide, the radiation pattern coincides with the one obtained in the article [1]. In the calculation, we assume that the main mode is excited in the waveguide.

A monochromatic electromagnetic field with a frequency ω is described by Maxwell's system of differential equations

$$\triangle \vec{E} + \varepsilon_r \vec{E} = 0, \tag{13}$$

$$\operatorname{div} \vec{E} = 0, \tag{14}$$

$$\vec{H} = i\sqrt{\frac{\varepsilon_0}{\mu_0}} \operatorname{rot} \vec{E}, \tag{15}$$

where $\varepsilon_r = \begin{cases} \varepsilon, & 0 < z \le d \\ 1, & z > d \end{cases}$ is the relative permittivity; ε_0 and μ_0 are the absolute permittivities of free space.

The coordinates x, y, z, and the thickness of the layer d are considered dimensionless, obtained by multiplying the corresponding values by the – wave number of free space [2]. The time factor in the expressions components $k_0 = \frac{2\pi}{\lambda} = \omega\sqrt{\varepsilon_0\mu_0}$ of the electromagnetic field is taken as $\exp(j\omega t)$.

In equations (1-3) we perform the Fourier transform on x and y

$$\vec{E}(t_x, t_y, z) = \int\int_{-\infty}^{\infty} \vec{E}(x, y, z)e^{i(xt_x + yt_y)}dxdy \,; \, \vec{H}(t_x, t_y, z) = \int\int_{-\infty}^{\infty} \vec{H}e^{i(xt_x + yt_y)}dxdy \,;$$

$$\frac{d^2\vec{E}}{dz^2} + (\varepsilon_r - t_x^2 - t_y^2)\vec{E} = 0, \tag{16}$$

$$-it_x\,\mathrm{E}_x - it_y\,\mathrm{E}_y + \frac{d\,\mathrm{E}_z}{dz} = 0, \tag{17}$$

$$\vec{H} = i\sqrt{\frac{\varepsilon_0}{\mu_0}} \begin{vmatrix} \vec{x}_0 & \vec{y}_0 & \vec{z}_0 \\ -it_x & -it_y & d/d_z \\ \mathrm{E}_x & \mathrm{E}_y & \mathrm{E}_z \end{vmatrix} \tag{18}$$

The notations used here are the $\vec{x}_0$, $\vec{y}_0$, $\vec{z}_0$ orts of the coordinate axes E_x, E_y, E_z; H_x, H_y, H_z, the components of the vectors $\vec{E}$ and $\vec{H}$. Let's introduce more notations $\beta = \sqrt{\varepsilon - t_x^2 - t_y^2}$ and $\beta_0 = \sqrt{1 - t_x^2 - t_y^2}$. If the root expression is negative, then $\beta = i|\beta|, \beta_0 = i|\beta_0|$. It follows from (4) that

$$\frac{d^2\,\vec{\mathrm{E}}}{dz^2} + \beta^2\,\vec{\mathrm{E}} = 0, \text{ at } 0 < z \le d,$$

$$\frac{d^2\,\vec{\mathrm{E}}}{dz^2} + \beta_0^2\,\vec{\mathrm{E}} = 0, \text{ at } z > d .$$

We will look for a solution to the system (4-6) taking into account the emission condition in the $z \to +\infty$ form of

$$\vec{\mathrm{E}}(t_x,t_y,z) = \begin{pmatrix} A_x(t_x,t_y) \\ A_y(t_x,t_y) \\ A_z(t_x,t_y) \end{pmatrix} \cos\beta z + \begin{pmatrix} B_x(t_x,t_y) \\ B_y(t_x,t_y) \\ B_z(t_x,t_y) \end{pmatrix} \sin\beta z \text{ at } 0 < z \le d,$$

$$\vec{\mathrm{E}}(t_x,t_y,z) = \begin{pmatrix} C_x(t_x,t_y) \\ C_y(t_x,t_y) \\ C_z(t_x,t_y) \end{pmatrix} \exp\left(-i\beta_0(z-d)\right) \text{ at } z > d .$$

In this case, equation (4) is satisfied. From (5) weget

$$A_z = -i\frac{t_x B_x + t_y B_y}{\beta}, B_z = -i\frac{t_x A_x + t_y A_y}{\beta},$$

$$C_z = -i\frac{t_x C_x + t_y C_y}{\beta}.$$

(19)

From (3, 7) it follows that when $0 < z \le d$

$$H_x(t_x,t_y,z) = i\frac{\sqrt{\varepsilon_0/\mu_0}}{\beta}\begin{pmatrix} \left(A_x t_x t_y + A_y(\varepsilon - t_x^2)\right)\sin\beta z - \\ -\left(B_x t_x t_y + B_y(\varepsilon - t_x^2)\right)\cos\beta z \end{pmatrix},$$

$$H_y(t_x,t_y,z) = i\frac{\sqrt{\varepsilon_0/\mu_0}}{\beta}\begin{bmatrix} -\left(A_x(\varepsilon - t_y^2) + A_y t_x t_y\right)\sin\beta z + \\ +\left((\varepsilon - t_y^2)B_x + B_y t_x t_y\right)\cos\beta z \end{bmatrix},$$

(20)

And at the $z > d$

$$H_x(t_x,t_y,z) = -\frac{\sqrt{\varepsilon_0/\mu_0}}{\beta_0}\left(C_x t_x t_y + C_y(1 - t_x^2)\right)e^{-i\beta_0(z-d)},$$

$$H_y(t_x,t_y,z) = -\frac{\sqrt{\varepsilon_0/\mu_0}}{\beta_0}\left(C_x(1 - t_y^2) + C_y t_x t_y\right)e^{-i\beta_0(z-d)}.$$

(9)

From the Continuity Condition [3] of the Tangent Components of the Electric and Magnetic Vectors at $z = d$

$$A_x \cos\beta d + B_x \sin\beta d = C_x,$$
$$A_y \cos\beta d + B_y \sin\beta d = C_y,$$

$$\left(A_x t_x t_y + A_y(\varepsilon - t_x^2)\right)\sin\beta d - \left(B_x t_x t_y + B_y(\varepsilon - t_x^2)\right)\cos\beta z = i\frac{\beta}{\beta_0}\left(C_x t_x t_y + C_y(1 - t_x^2)\right),$$

$$-\left(A_x(\varepsilon - t_y^2) + A_y t_x t_y\right)\sin\beta d + \left(B_x(\varepsilon - t_y^2) + B_y t_x t_y\right)\cos\beta z =$$

(10)

$$= i\frac{\beta}{\beta_0}\left(C_x(1 - t_y^2) + C_y t_x t_y\right). =$$

From the initial assumption we can see that in the opening of the waveguide

$$E_x(x,y,0) = (1+R)F_x(x,y),$$
$$E_y(x,y,0) = (1+R)F_y(x,y),$$

where F_x and is the components of the F_y electric field for the fundamental mode of the waveguide, and R is the complex reflectance coefficient.

$$E_x(x,y,0) = A_x(t_x,t_y) = (1+R)\mathrm{F}_x(t_x,t_y),$$
$$E_y(x,y,0) = A_y(t_x,t_y) = (1+R)\mathrm{F}_y(t_x,t_y),$$

(11)

where F_x and F_y are the Fourier transformations of functions F_x and F_y respectively.

Let us substitute the expressions (10) into the equations (9). The solution of the resulting system, solved with respect to B_x and B_y, is conveniently written in matrix form. Let's denote

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix},$$

Where is

$$m_{11} = -\beta\beta_0(\varepsilon+1)\cos\psi\sin\psi + i((\varepsilon-h^2)\cos 2\psi - h^2(\varepsilon-1)\sin(\psi+\alpha)\times\sin(\psi-\alpha)),$$
$$m_{12} = m_{21} = i(\varepsilon-1)h^2\sin\alpha\cos\alpha,$$
$$m_{22} = -\beta\beta_0(\varepsilon+1)\cos\psi\sin\psi + i((\varepsilon-h^2)\cos 2\psi - h^2(\varepsilon-1)\cos(\psi+\alpha)\times\cos(\psi-\alpha)).$$
$$D(h) = (\beta\sin\psi - i\varepsilon\beta_0\cos\psi)(\beta_0\sin\psi - i\beta\cos\psi),$$

(12)

then

$$\begin{pmatrix} B_x \\ B_y \end{pmatrix} = \frac{(1+R)}{D(h)} M \begin{pmatrix} \mathrm{F}_x \\ \mathrm{F}_y \end{pmatrix}.$$

(13)

The following notations are used: $\psi = \beta d$, $t_x = h\cos\alpha$, $t_y = h\sin\alpha$, $h \geq 0$, $-\pi \leq \alpha \leq \pi$. From the first two equations (10) it follows that

$$\begin{pmatrix} C_x \\ C_y \end{pmatrix} = \left(\frac{1}{D(h)} M\sin\psi + E\cos\psi \right)(1+R) \begin{pmatrix} \mathrm{F}_x \\ \mathrm{F}_y \end{pmatrix},$$

(14)

where is $E = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$.

Note that in formulas (12-14) only the F_x and F_y reflectance coefficient R, which is still uncertain, depends on the geometry of the waveguide.

To determine the reflectance R we use [4] power continuity at z = 0

$$\frac{1}{2}\iint_S (E_x^* H_y - E_y^* H_x)\,dx\,dy\big|_{z=-0},$$

$$\frac{1}{2}\int\int_{-\infty}^{\infty} (E_x^* H_y - E_y^* H_x)\,dx\,dy\big|_{z=+0},$$

where S is by opening the waveguide by calculating the integral on the right using the Parseval formula through Fourier transforms.

The integral on the left can be computed directly through the known characteristics of the field inside the waveguide. If, for the main mode, the tangent components of the magnetic field are equal $G_x(x,y)$ to and $G_y(x,y)$, then according to Parseval's formula:

$$\frac{1}{2}\iint\limits_{S}(E_x^* H_y - E_y^* H_x)dx\,dy\big|_{z=-0} = \frac{1}{2}(1+R^*)(1-R)\iint\limits_{R}(F_x^* G_y - F_y^* G_x)dxdy =$$

$$= \frac{1}{2}(1+R^*)(1-R)P$$

$$\frac{1}{2}\int\limits_{-\infty}^{\infty}\int(E_x^* H_y - E_y^* H_x)dxdy, z = +0$$

$$\frac{1}{8\pi^2}\int\limits_{-\infty}^{\infty}\int(E_x^* H_y - E_y^* H_x)dt_x dt_y, z = +0$$

(15)

Using (11), (13-15), we get

$$\frac{1}{2}(1+R^*)(1-R)P = \frac{i\sqrt{\varepsilon_0/\mu_0}}{8\pi^2}(1+R^*)(1+R)\int\limits_{0}^{\infty}\frac{hdh}{\beta(h)D(h)}\times$$

$$\times\left[\int\limits_{-\pi}^{\pi}\Big((m_{21}F_x + m_{22}F_y)F_x^* - (m_{11}F_x + m_{12}F_y)F_y^*\Big)d\alpha\right] = \frac{1}{2}(1+R^*)(1+R)Q.$$

(16)

From this equation we get

$$Y = \frac{1-R}{1+R} = \frac{Q}{P}; R = \frac{1-Y}{1+Y},$$

(17)

where Y is the internal conductivity of thesystem.

The integral function in (16) has poles at points $h_1, h_2...h_n \in (1,\varepsilon)$ such that $3(h) = 0$. These points are defined only by the coating and are independent of geometry [5]. In the calculation, the integral should be understood as a contour integral, in which the contour bypasses the poles along the lower half-plane.

If φ and θ are the spherical coordinates of the observation point, then the radiation pattern can be calculated using the pass method, using formulas (7) and (14)

$$|E_\perp|R \sim \Phi(\varphi,\theta) = \frac{\cos\theta}{2\pi}\sqrt{\begin{array}{l}|C_x\sin\varphi - C_y\cos\varphi|^2 + \\ |C_z\sin\theta - C_x\cos\varphi\cos\theta - C_y\sin\varphi\cos\theta|^2\end{array}}.$$

(18)

In formula (18), the arguments of the functions $C_x(h,\alpha)$, $C_y(h,\alpha)$, $C_z(h,\alpha)$ will be respectively $h = \sin\theta$, $\alpha = \varphi$.

Formulas (17) and (18) were used to perform numerical calculations for rectangular and circular waveguides.

References

1. Kryachko A.F., Likhachev V.M., Ovodenko A.A. – Saint Petersburg: POLYTECH-PRESS, 2020. 268 p. (in Russian).

2. Kryachko A.F. Diagnostics of hypersonic plasma by the method of natural electromagnetic oscillations of an open resonator with inhomogeneous inclusions/ Kryachko A.F., Kovalev M.A., Revunov G.M., Tyurina A.I. // Electromagnetic waves and electronic systems. – Moscow: OOO "Publishing Enterprise of the Editorial Office of the Journal "Radiotekhnika", 2021. – №26. P.13-20.

3. Danilin A.A. Izmereniya v tekhnike SVCH: Uchebnoeposobiedlyavuzov [Measurements in microwave technology: Textbook for higher education institutions].- Moscow, Radiotekhnika Publ., 2008.-184 p. (in Russian).

4. Vorontsov A.A., Mirovitskaya S.D. O diffraction ploskoyvolnynamnogoslotelnomdielektrichekomkarline [On the diffraction of a flat wave on a multilayer dielectric cylinder]. – 1986. – 31, No. 12. P. 2330-2334.

5. Mikhailov V.F., Pobedonostsev K.A., Bragin I.V. Prognozirovaniyekharakteristiki antenna s teplozashchita [Forecasting of antenna characteristics with heat protection]. Saint Petersburg, Sudostroenie Publ., 1994. – 304 p. (in Russian).

INVESTIGATION OF CYCLIC LOADS OF GEARS PERFORMED USING ADDITIVE TECHNOLOGIES

Mariya Rassykhaeva
Saint Petersburg State University of Aerospace Instrumentation
E-mail: mitschiru@gmail.com

Cyclic Loads in Additive Manufacturing

Prototyping is an integral part of almost any job. New technologies and materials make it possible to create new products and open up new markets. New production and information technologies are also significantly transforming our sustainable industrial world of work. Production management tasks have become more diverse and complex, which cannot be completed without a full-scale simulation of all production systems.

Different types of equipment are used in mechanical engineering to transmit movements. One of the most well-known types is the gear drive, in which the teeth of a pair of gears are locked in shape and thus transmit the rotational motion of the drive gear with a certain gear ratio to the second gear. Gears are among the most important elements of a machine and are used in a variety of devices that are used by most people on a daily basis. They connect drive lines to output shafts and adjust rotational speeds and torques, they are available in a wide variety of designs for a wide range of applications, for uniform or uneven ratios. The properties according to which suitable gear pairs are selected are extensive: noise and temperature rise, efficiency, self-locking, load-bearing capacity, center distance tolerance, wear and much more.

Selecting and designing a suitable gear pair or an entire gearbox is an elementary task in many R&D projects. In a professional environment with specialized engineers, this task is part of the daily routine and is carried out with the help of sophisticated software systems that are adapted to users with in-depth knowledge, high professional competence and sufficient experience. However, if a small team or a person with little or no experience wanted to build a simple gear drive concept in the form of a functional model, for example, this task can be complex and time-consuming.

Designs usually only provide space for standard gears to be mounted on shafts. The necessary gears are purchased and installed. If you have access to AM technology to produce a functional model or prototype, instead of using purchased gears, it is also possible to manufacture them directly with other necessary components. In order to take full advantage of AM's potential, gears can also be manufactured mounted directly on the shaft. In addition to the structural design, however, three-dimensional gear models must also be calculated [1,2].

Gears are the functional elements of gear drives. They are used to form rotational motions and torques and transmit them in shape and without slipping. The individual teeth of the gear are locked and transmit the rotational motion of the corresponding gear (pinion) to the second gear. Due to the shape conformity (usually uniform), the ratio is independent of the torque transmitted. Central to the subcontracting of the mechanism is the shape of the tooth, on which a large number of gear pair properties depend. Depending on the shape, the teeth are synchronized at a point or on a line and thus have a low or high load-bearing capacity, roll or slide with each other and thus have different levels of wear and different noise and temperature changes. In practice, the involute lock is most often used, also because of the ease of manufacture. Therefore, a distinction is made between involute and non-involutional gears. Non-involutional gears are used in mechanical engineering, where their individual properties are advantageous and justify the high cost of gear plastering Fig. 1.

Gears are functional elements of gear drives. They are used to form rotational motions and torques, and transmit them in shape and without slipping. The individual teeth of the gear are locked and transmit the rotational motion of the corresponding gear(s) to the second gear. Due to the shape conformity (usually uniform), the ratio is independent of the torque transmitted. The shape of the tooth is of central importance in the subcontracting of the mechanism, on which a large number of properties of the gear pair depend. Depending on the shape, the teeth are synchronized at a point or on a line and thus have a low or high load-bearing capacity, roll or slide with each other and thus have different levels of wear and different noise and temperature changes. In practice, the involute lock is most often used, also because of the ease of manufacture. Therefore, a distinction is made between involute and non-involutional gears. Non-involutional gears are used in mechanical engineering, where their individual properties are advantageous and justify the high cost of gear plastering [4,5].
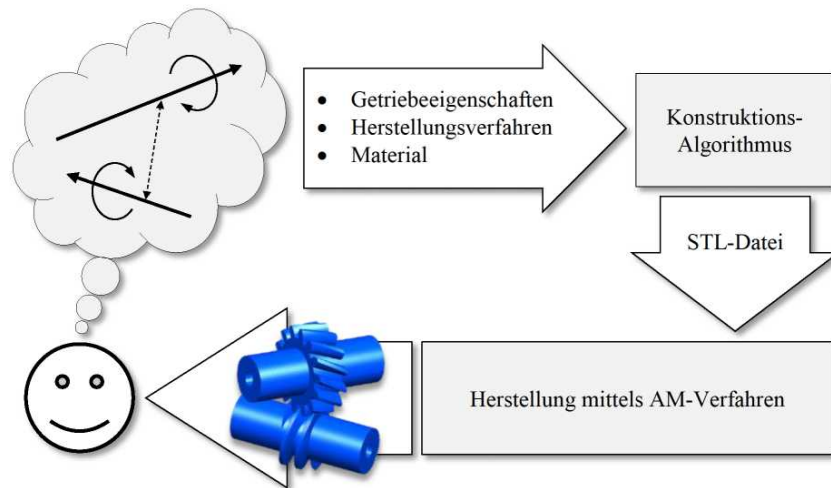
Fig. 1. Computer-aided design concept for gear drives

The quasi-standard for the transmission of 3D models in additive manufacturing is the STL format, which was introduced in 1988 by 3D-Systems (Rock Hill, SC, USA) for their stereolithography systems, hence the abbreviation STL for STereoLithography Fig. 2.
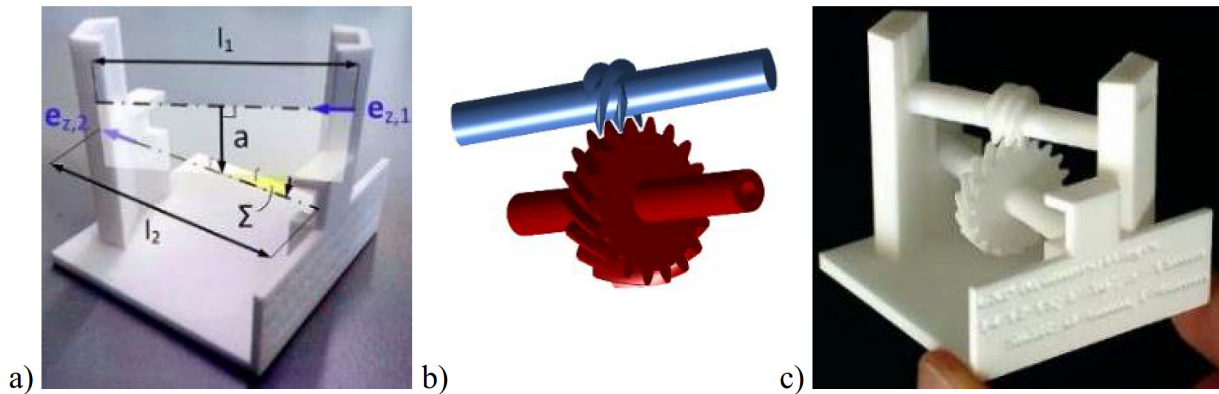

Fig. 2. Example of CAD application:
a) demonstration status, b) computational surface models, c) sintered functional model

Computer-aided design of gear stages for additive manufacturing investigates what gear geometry can be obtained using the AM process. Due to the limited resolution (diameter of the laser focus, layer height), it may happen that dimensions that are too small cannot be manufactured as desired. For straight and cloth gears, the workmanship of the SLS system type Formiga P100 (EOS GmbH, Krailling, Germany) is checked below with the PA2200.

The STL file describes three-dimensional solids with triangular planes. Each triangular surface (facet) is described by three vertices and a unit normal vector. The order of the vertices, as shown on determines the orientation of the normal vector using the right-hand rule. The order of the vertices is chosen so that the normal vector points outward from the surface of the body. This information is defined individually in the STL file for each surface of the triangle. STL files can be saved in ASCII or binary format.

In ASCII format, an STL file always starts with the solid keyword followed by the name of the object. Subsequently, for each triangular surface, the coordinates of the standard vector (faceted normal) and vertices are specified in the appropriate order (outer loop). The end of the STL file is denoted by the endsolid keyword and the name of the object.

In binary format, this information is stored in a more compact way. The 80-byte header (which should not start with the solid keyword) is followed by the total number of triangular planes. This is followed by the coordinates of the normal vector and the three vertices for each triangle. At the end, there is a positive number of 2 bytes (the byte count attribute), which currently has no particular value and is therefore set to zero by default.

Formula for calculation of modulus m, mm, for given input data at pitch circle diameter,

$$m = \frac{2a_w}{z_1 \cdot (1+i)}$$

The values of the modules are regulated by GOST 9563-60. By changing the number of teeth

The gears Z1 (17, 18...) and the center distance aw are determined by the modulus m of the standard series, since the gear ratio i is known. At the same time, it should be taken into account that the more teeth the gear has, the smoother the gear stroke, but at the same time, the resistance of the teeth when bending will be weakened.

The initial (initial) width of the toothed rings (in millimeters) is specified as

$$B_B = \frac{1}{2}(\text{m}\sqrt{d_a + 2}) - \text{h} ,$$

where h = 2.25m is the height of the tooth.

Conclusion

For additive manufacturing of gears with SLS, two rough sizes have been identified with the thickness of the tooth head in the forehead section and the width of the buckle on the base circle in the normal mean value, with which it can be verified depending on the manufacturing tolerances or boundary conditions of the relevant process, whether the desired gear geometry lying in the installation space can be obtained, So that the rotational motion can be reliably transmitted with the tooth. The definable limit values for the SLS system of the Formiga P100 are s ta,min ~ 0.4 mm for the thickness of the tooth head and e nb,min " 0.35 mm.

Due to the process, the achievable quality of the SLS gear is affected by various factors, such as the .dem gear ratio of the used and new powder, the thermal preload of the powder used, or the position and orientation of the gear in the installation space. In this context, the influence of grain size on the surface roughness and melting of the surface layer of the surrounding powder should be emphasized, resulting in dimensional variations. Without special post-processing – apart from shot blasting of gears – gears with gear quality in the range of 12 or lower can be manufactured on the Formiga P100 with PA2200, which according to DIN 3961 can only be used in gearboxes with low requirements. It can be assumed that the deviations, which depend in particular on the size of the powder grains, are more important for smaller gears than for coarse gears. It seems, therefore, expedient to make a special study of gears of various coarseness with regard to the attainable qualities of gears in the future. If necessary, the impact of different post-processing methods on gear quality should also be investigated. [7].

In addition to surface quality, another process-related property of SLS and other AM processes is that sharp edges or notches cannot be obtained. In order to address the resulting stresses at the base of the tooth in the simulation without tooth rounding presented in this paper, an estimate was presented that establishes a correlation between the production curvature at the notch as a function of the grain size of the powdered starting material. In order to make more accurate predictions of the resulting notch stresses, the radii on the notches at different angles should be measured using suitable methods in the future.

References

1. Chabanenko A V, Kurlov A V 2021 Control the quality of polymers based on the model of Dzeno Journal of Physics: Conference Series

2. Chabanenko A V, Kurlov A V and Tour A C 2020 Model to improve the quality of additive production by forming competencies in training for high-tech industries *J. Phys.: Conf. Ser.* 1515 052065.

3. Chabanenko A V and Yastrebov A P 2018 Quality Assurance of Hull Elements of Radio-Electronic Equipment by Means of Control System *J. Phys.: Conf. Ser.* 1515 052065.

4. Chabanenko A V, Kurlov A V 2019 Construction of mathematical model of training and professional development of personnel support of additive production of REA IOP Conference Series: Materials Science and Engineering

5. Quality Management of Additive Products / A. V. Chabanenko // RIA: Journal: "Standards and Quality". 2018. №2. P. 90-94.

6. Implementation of the information PLM system in the organization Rassykhaeva M.D., Chabanenko A.V. In the collection: Selected scientific works of the twentieth International Scientific and Practical Conference "Quality Management". Moscow, 2021. P. 285-289.

7. Plasticity Estimation with Combined Hardening for the Study of Deformation Processes of Structural Materials under Different Modes of Low-Cycle Loading Chabanenko Alexander Valerievich, Rassykhaeva Maria Dmitrievna Certificate of registration of the computer program 2021619545, 11.06.2021. Application No. 2021618914 dated 11.06.2021.

COMPUTER PROCESSING OF LASER BEAM PROFILE IN THE TASK OF EARLY FIRE DETECTION

Yana Ryvkina

undergraduate student
Saint Petersburg State University of Aerospace Instrumentation (SUAI),
Saint Petersburg, Russia
yana.ryvkina@bk.ru

Abstract. *One of the earliest signs of fire is thermal convective flow, which appears already at the smoldering stage, before the flames appear. Registration of heat flow is a way to quickly and effectively fight a fire. This paper presents a laser fire detection system using computer algorithms of laser beam profile processing.*
Keywords: *fire detector, thermal convective flow, laser beam, fire, correlation function, integral-difference function.*

Introduction

The majority of the fire detector market is occupied by heat and smoke detectors. Smoke detectors work on the principle of light scattering [1]. Such systems are triggered when a certain threshold concentration of particles is reached, in addition, dust and dirt can accumulate inside the device, which leads to false alarms. Heat detectors are triggered when the temperature rises rapidly or when a set threshold is reached [2]. For both types of detectors, the main disadvantages are a high threshold when there is an active combustion process and false alarms.

The use of laser technologies for fire detection has great prospects [3-5]. Thermal convective flow (TCF) is one of the early signs of the onset of fire. Its appearance leads to a random change in the refractive index, and the spatial characteristics of the laser beam will also change in such a space.

1. STRUCTURAL DIAGRAM AND TECHNICAL IMPLEMENTATION OF THERMAL CONVECTIVE FLOW DETECTION SYSTEM

It is possible to detect TCF by means of laser radiation, for this purpose a structural scheme of the system is proposed, the principle of operation of which consists in the formation of a laser beam of large size (a few centimeters) and the subsequent registration of the spatial characteristics of the laser beam (intensity profile distribution). The structural scheme is shown in Fig. 1.

The principle of operation of the circuit is as follows. The laser (1), beam expander (2) and linear–CCD (3) are located on the same optical axis. The optical beam from the laser (1) passes through the beam expander (2) and enters the linear–CCD (3). The beam diameter at the output of the beam expander is matched to the size of the linear–CCD. The distance between the beam expander (2) and the linear-CCD (3) can be varied, during the experiment it was 50 cm. From the output of the linear-CCD (3) the counting values of the beam intensity profile distribution are formed, which are supplied to the PC (5). It should be noted that in order to reduce power consumption of the system, the laser (1) can operate in pulse mode with matching of time samples of registration by the linear-CCD (3). The control unit (4*) will be responsible for harmonization of the operation mode.
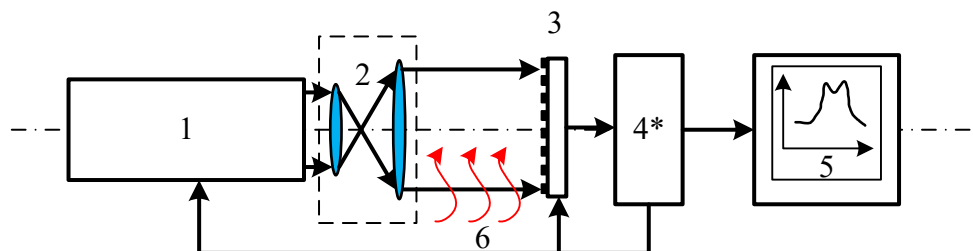
Fig. 1. Block diagram of the TCF detection system:
1 – laser; 2 – beam expander; 3 – linear–CCD; 4 – registration block;*
5 – PC with pre-installed software for beam profile processing; 6 – TCF

To study the dynamics of TCF interaction with the laser beam, a laboratory layout was developed based on the structural scheme (Fig. 2). The laboratory layout used a laser module KLM-D650-16-5 (1) with a wavelength of 650 nm and a power of 16 mW, a tunable beam expander – GBE10-B from "THORLABS" (2) and a linear-CCD SL-TCD-VI (3) from "Avesta".
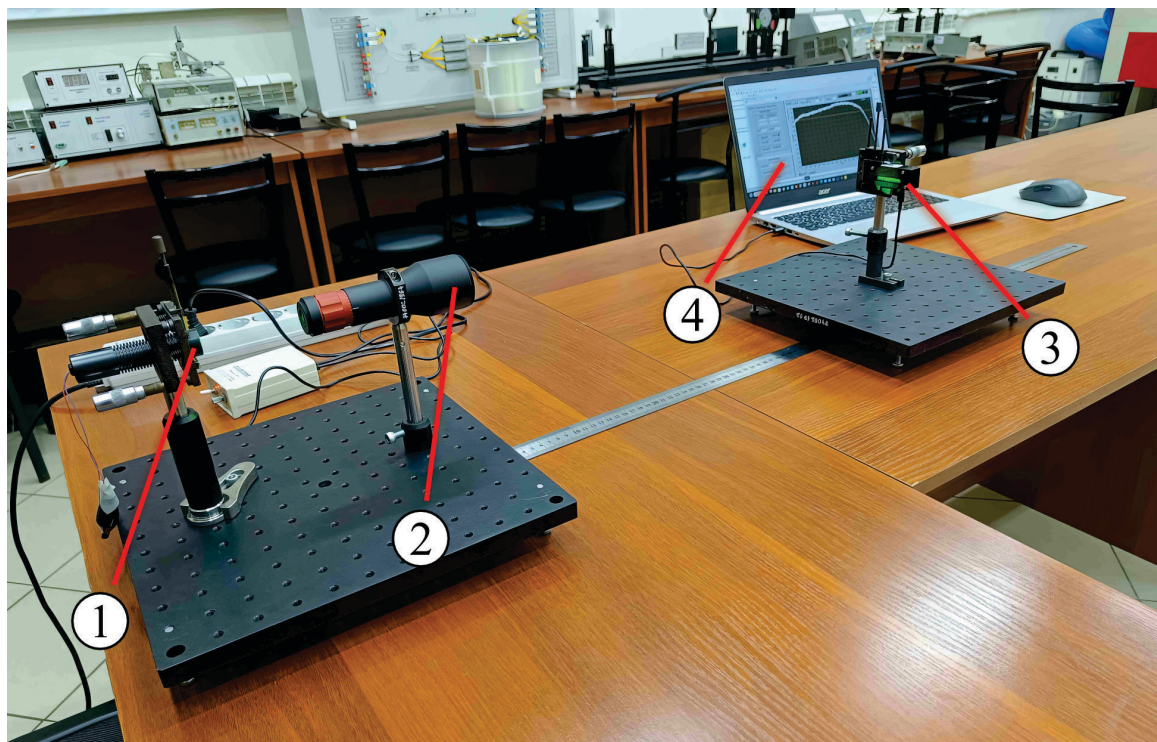

Fig. 2. A system for recording the spatial characteristics of optical beams:
1 – diode laser; 2 – beam expander; 3 – linear–CCD; 4 – PC

The PC was equipped with software that allows mathematical processing of the recorded data. The in-house developed software allows changing the parameters of linear-CCD operation. The software records the beam intensity profile distribution as an unformatted text RAW-file for further computer processing.

To confirm the effectiveness and operability of the proposed scheme for registering the TCF appearance, experimental studies were carried out as follows:

1. Simulation of the normal operation mode of the system. Samples of the beam intensity profile distribution were recorded for 20 seconds at 1 second intervals. The accumulation time of each sample was 10 µs. The sample registration interval was chosen heuristically based on the convenience of the experiment and peculiarities of the software operation.

2. TCF formation. In this series, fluctuations of the beam intensity profile distribution were registered. A candle flame was used as a TCF source. As a result, samples were recorded for 20 seconds.

As a result of the experiment, a RAW file containing 40 columns (20 for the normal mode, 20 for the TCF simulation case) with data on the beam intensity distribution profile was generated.

2. COMPUTER PROCESSING AND INTERPRETATION OF THE OBTAINED DATA

Two approaches were used for computer processing of the obtained data: correlation and integral-difference.

In normal operation mode, the value of the correlation function will remain maximum. Appearance of TCF leads to fluctuations, so there is a decrease in the value of the maximum of the correlation function.

Mathematical processing of the obtained data consisted in the formation of the dynamic maximum of the mutual correlation function of the discrete signal $W(nT)$, which can be represented in the form:

$$W(nT) = \max_l \left[\frac{1}{L} \sum_{i=0}^{L-1} X_i(nT) X_{i-l}([n-1]T) \right], \qquad (1)$$

where T is the sampling registration interval (1 second), $X(nT)$, $X([n-1]T)$ – samples of beam intensity profile distributions registered in neighboring intervals T.

Also, the integral-difference function reflecting the peak-to-peak difference of the recorded distributions of beam profile intensities can be used to process the obtained results. In mathematical form, the proposed integral-difference function for discrete samples of beam profile intensity distributions can be represented as:

$$R(nT) = \sum_{i=1}^{L} |\Delta X_i(nT)| = \sum_{i=1}^{L} |X_i(nT) - X_i([n-1]T)|. \qquad (2)$$

A programme has been implemented in MatLab environment, which processes the received data from RAW-file in two considered ways. A fragment of the programme code is shown in Fig. 3.

```
3       ts=0:511;
4       %Entering a data array in normal mode
5       X=vertcat(x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16); %matrix formation
6       nfig = 1;
7       %Entering a data array when a TCF appears
8       Y=vertcat(y1, y2, y3, y4, y5, y6, y7, y8,y9, y10, y11, y12, y13, y14, y15, y16);
9
10      figure(nfig); nfig = nfig + 1;
11      clf;
12      hold on;
13      plot(y14,'g'), grid on; % Intensity distribution
14      plot(y15,'r'), grid on; % Intensity distribution
15      plot(y16,'b'), grid on; % Intensity distribution
16      title ('Beam intensity distribution in the presence of TCF')
17      xlabel ('x, pixel number');
18      ylabel ('Intensity, a. u.');
19      legend ('Sampling 1', 'Sampling 2 (+1 c)','Sampling 3 (+2 c)' );
20      hold off;
21
22      for i=1:15
23        CorrX=xcorr(X(i+1,:), X(i,:));
24        CorrY=xcorr(Y(i+1,:), Y(i,:));
25      MAX_X(i)=max(CorrX);
26      MAX_Y(i)=max(CorrY);
27      RaznX(i)=sum(abs(X(i+1,:)-X(i,:)));
28      RaznY(i)=sum(abs(Y(i+1,:)-Y(i,:)));
29      end;
30
31      Plot_Max_X=MAX_X/max(MAX_X);
32      Plot_Max_Y=MAX_Y/max(MAX_Y);
33
34      Rez1=horzcat(Plot_Max_X, Plot_Max_Y);
35      Rez2=horzcat(RaznX, RaznY);
```

Fig. 3. Program code fragment for processing data on beam intensity distribution

The results of processing by correlation and integral-difference method are shown in Fig. 4 *a* and *b*, respectively.

The presented graphs clearly show that both considered methods of data processing allow registering the appearance of TCF. When a TCF is created between the beam expander and the linear-CCD (starting at 20 seconds), the value of the target function changes dramatically (decreases in the case of the correlation method, increases in the case of the integral-difference method). In the case of the integral-difference method of processing, the target function increases at least 3.5 times, so this method is more preferable.
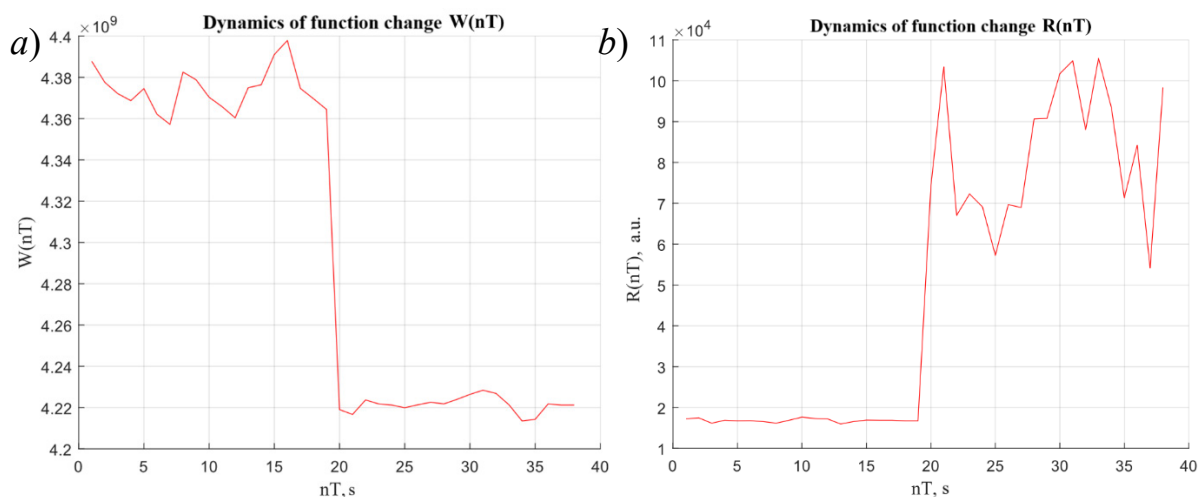
Fig. 4. Results of processing the laser beam intensity profile distribution by correlation (a) and integral-difference (b) methods

Conclusion

The results obtained from experiments and computer modelling show that the method of detecting fires by recording the appearance of thermal convective flow by changes in the distribution of the laser beam intensity profile is promising and effective. The proposed processing methods, correlation and integral-difference, have demonstrated their effectiveness for recording fluctuations in the beam intensity profile distribution.

In the future it is planned to develop more sophisticated processing algorithms to minimize false alarms due to the presence of natural convective flows in the monitored object.

References

1. Menkeev A.I., Darbakov D.V., Chlenov A.N. Current state of the market of smoke point fire detectors // Fire safety: problems and prospects. 2018. №9. С. 592-595.

2. Yurtaev S. V. Principle of detectors operation. Selection of the type of fire detector // Vestnik nauki. 2022. Т. 1. №. 2 (47). С. 227-232.

3. Kulakov S.V., Moskaletz O.D., Preslenev L.N., Shabardin A.N. Fiber–optic linear smoke fire detector // Materials and Devices for Photonic Circuits II, Proc. SPIE 4453. 2001. https://doi.org/10.1117/12.447645

4. Udd E., Spillman Jr W. B., et al. Fiber optic sensors: an introduction for engineers and scientists. John Wiley & Sons, 2011. 512 p.

5. Kazakov V.I., Moskalets O.D., Preslenev L.N. Explosion-proof fibre-optic fire detector. Mathematical model of the sensitive element // Sensors and Systems. 2015. № 2. С. 19-22.

DETERMINING A PERSON'S MEMBERSHIP IN THE MUSIC COMMUNITY USING THE VGG-FACE AND FACENET MODELS

Yana Senichenkova
SUAI
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
senichenkova02@inbox.ru

Abstract. *The use of machine learning in pathognomy is being investigated to test the hypothesis about the influence of classical music education on a person's face. A sample of images of the faces of symphony orchestra musicians and people without musical education was collected, which is used to solve the binary classification problem. The quality of the sample is analyzed using dimensionality reduction algorithms. The pre-trained VGGFace ame FaceNet model is used to obtain embeddings. The values of the quality metrics in the test sample allow us to state with some degree of confidence that it is possible to determine whether a person is a musician of a symphony orchestra based on the image of his or her face.*
Keywords: *classification; image recognition; VGGFace; FaceNet, support vector machine; principal component method; tSNE*

Introduction

Modern facial recognition technologies can serve as an effective identity management tool for law enforcement agencies and businesses and this can facilitate the personalization of services such as advertising, customer service or restaurant menus, or even personalize products such as games[1]. Using machine learning algorithms, you can extract such information from a photo of a face, such as the presence of mental disorders [2] (including the presence of schizophrenia [3] or bipolar affective disorder [4]), political preferences [5]. There are questions about the ethics of extracting personal information from photographs of people's faces.

The determination of personal characteristics, such as sexual orientation and political preferences, using physiognomy is questionable, as this is an unconfirmed science. However, the definition of emotions based on pathognomy can be considered as a more scientific task [6]. Pathognomy is based on the reflection of emotions on the face, traces of experiences, lifestyle and other biological processes. Unlike classical physiognomy, which involves determining character traits and behavior based on the shape of a face without scientific justification and biological interpretation, pathognomy allows you to draw conclusions based on biological processes.

In this paper is being tested the hypothesis that musical education and the profession of a symphony orchestra musician leave a biological mark on a human face. From the point of view of the authors of this article, the task of determining a person's profession by photographing his or her face belongs to the task of pathognomy and, therefore, should not lead to the question of "pseudoscience".

The profession of a symphony orchestra musician was chosen specifically for this study, not only because it is very difficult to present any accusations of racism based on lack or availability of musical education, but also because photographs of musicians are in the public domain and they are quite easy to collect.

Model

The 224x224 pixel image is sent to the input of the pre-trained VGGFace network [7]. This neural network was trained by the authors of the publication on a closed dataset and demonstrates an accuracy of 97.27% in the task for the task of identifying a person from a photo of a face on the dataset "Labeled Faces In The Wild" [8]. The weights of the pre-trained model are available on the Internet. Next, the image is transformed into a vector of dimension 2622, called embedding, containing informative features. After that, the dimension of the vector is reduced to 128 using the principal component method [9]. 128 components describe 86.5% of the variance. It follows from this that with a decrease in the embedding dimension, a significant part of the information is preserved, which allows reducing the dimension of the vector. Next, embedding is transferred to the classifier implemented by the support vector machine [10]. The final pipeline is shown in Fig. 1.
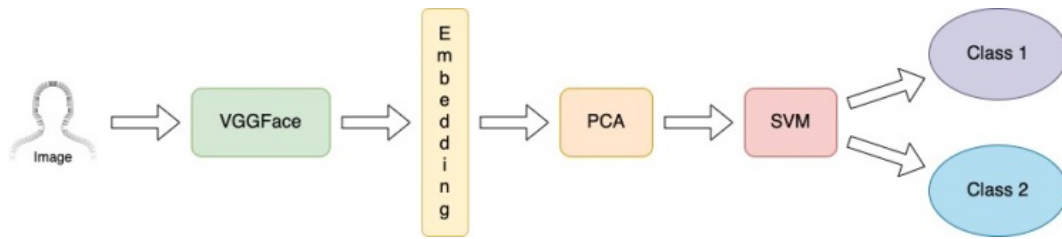
Fig. 1. A model for binary image classification using a VGGFace embedding support vector machine

Dataset

For training and testing of the classifier, photographs corresponding to the classes "Musician" and "Non-musician" were collected. For the "Musician" class, images of employees from the symphony orchestras website were taken from the Internet; for the "Non-musician" class, images from the datasets "Labeled Faces In The Wild" [8], CelebA[11], flckr-faces[12] were taken. Photos of teachers of higher educational institutions were also taken. There are 1318 images in the final version of the dataset. Examples of images from the dataset are shown in Figures 2.a, 2.b.

a)   b)
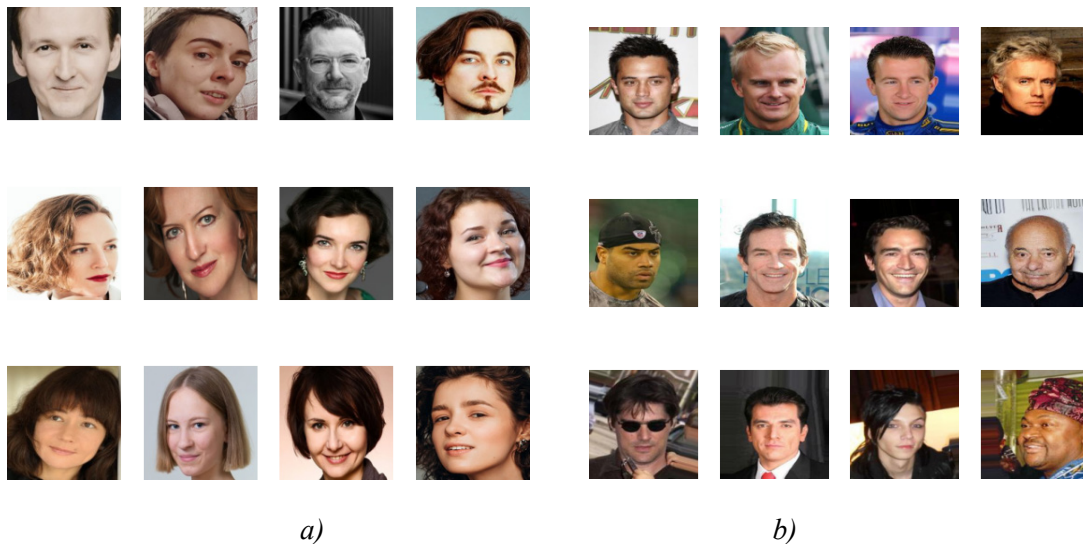
Fig. 2. Examples of images in the collected dataset: a) musicians; b) non-musicians.

Classifier training

Grid search was used to select the optimal value of the classifier hyperparameter based on the support vector machine. The optimal value of the regularization hyperparameter found by this method is 5. The radial basis function (RBF) was used as the core function. Since the data set is balanced, the accuracy metric was used to evaluate the quality of classifier training. Based on the results of training the classifier, the following values of quality metrics were obtained on a test sample of 200 images: accuracy 0.70, F1-score 0.734 ROC AUC score 0.699. The error matrix for the test sample is shown in table I.

Table 2

Error Matrix For A Test Sample Of 200 Images

	Predicted Class	
	Not musician	*Musician*
Not musician	53	47
Musician	17	83

Experiment with FaceNet

This experiment was also conducted on embeddings obtained using the FaceNet neural network [13]. The length of FaceNet embeddings is 512

The optimal value of the regularization hyperparameter found by this method is 9.

Based on the results of training the classifier, the following values of quality metrics were obtained on a test sample of 200 images: accuracy 0.70, F1-score 0.734 ROC AUC score 0.699. The error matrix for the test sample is shown in table II.

Table 2

Error Matrix For A Test Sample Of 200 Images

	Predicted class	
	Not musician	*Musician*
Not musician	55	45
Musician	17	83

Analysis of the collected dataset

To verify that the built classifier has really learned to distinguish the faces of symphony orchestra musicians from people who are not related to the performance of symphonic music, an in-depth analysis of the collected data set (training sample) was carried out.

An attempt was made to visualize the dataset by converting embeddings into two-dimensional vectors. To reduce the number of informative features, dimensionality reduction algorithms were used: the principal component method (PCA) [9] and the t-SNE algorithm [14]. Both approaches were implemented using the standard tools of the sklearn library of the Python programming language.

From Fig. 3.and it can be concluded that the two classes are not linearly separable in the feature space of the two main components, however, there is a pronounced cluster of images corresponding to the "Musician" class. The first two components describe 15.7% and 7.1% of the variance of the data, respectively, which in total gives less than 22.8% of the explained variance.

Visualization of the dataset using t-SNE in Fig. 3.b does not show any explicit clusters. A potentially possible interpretation may be that either the embeddings generated using VGGFace are a set of random numbers and do not make any sense, or the data set is quite wide, diverse and both classes are well represented. From the point of view of the authors of this work, the latter is more likely, since images from various sources were taken when collecting data.

a)   *b)*
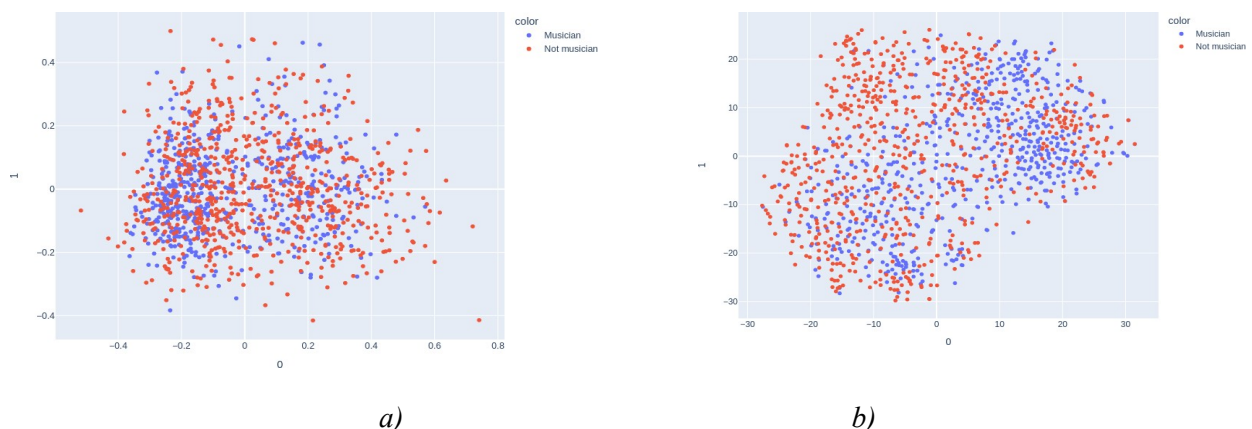
Fig. 3. Visualization of image embeddings: a) The principal component method; b) the tSNE method

Discussion

The metrics obtained by training the classifier on embeddings obtained through FaceNet and VGGFace can be considered identical. It follows from this that in order to improve the quality of the pipe-

line, it is necessary to conduct experiments with other models that will be used to extract the feature vector from photographs.

Conclusion

Using machine learning algorithms, a classifier was built that allows you to determine with 70% accuracy from one photo of a face whether a person's profession is performing as part of a symphony orchestra. Despite the presence of small biases in the training sample, the classifier based on the support vector machine was able to find differences between photographs of symphony orchestra musicians and photographs of "non-musicians". This study shows that there are features in the images of faces that are invisible to the human eye, which can easily be found by machine learning algorithms.

The results obtained indicate that facial recognition technologies pose a real threat to privacy, since in some cases a person's profession can be determined by a photo of his or her face without prior consent to that. This research can have practical application in the field of marketing for the formation of personal recommendations (advertising) of goods and services, for example, in shops, shopping malls, etc., by processing images of visitors' faces obtained from existing video surveillance cameras.

Gratitude

References

1. P. Blom, S. Bakkes, C. Tien Tan, S. Whiteson, D. Roijers, D. Valenti, and T. Gevers, "Towards personalised gaming via facial expression recognition," Proceedings of the tenth annual AAAI conference on artificial intelligence and interactive digital entertainment (AIIDE 2014), pp. 30–36, 2014.

2. H. An et al "Mental health detection from speech signal: A convolution neural networks approach," 2019 international joint conference on information, media and engineering (IJCIME), pp. 436-439, 2019.

3. M.Bishay, et al, "Schinet: Automatic estimation of symptoms of schizophrenia from facial behaviour analysis" IEEE Transactions on Affective Computing, vol. 12, 2019, pp. 949-961.

4. N. Abaei, H. Al Osman, "A Hybrid Model for Bipolar Disorder Classification from Visual Information," ICASSP, vol 2020. pp. 4107-4111, 2020

5. M.Kosinski, "Facial recognition technology can expose political orientation from naturalistic facial images" Scientific reports, vol. 11, 2021, pp. 100.

6. O. Bendel, "The uncanny return of physiognomy," 2018 AAAI Spring Symposium Series, 2018.

7. O. Parkhi, A. Vedaldi, A. Zisserman, "Deep face recognition," BMVC 2015-Proceedings of the British Machine Vision Conference 2015, 2015.

8. LFW Face Database : Main," vis-www.cs.umass.edu. http://vis-www.cs.umass.edu/lfw/ (date of the application: 12.02.2024)

9. K. Pearson, "Principal components analysis," Lond. Edinb. Dublin philos, vol. 2, 1999, pp. 559-572.

10. O. Chapelle, P. Haffner, V. Vapnik, "Support vector machines for histogram-based image classification," IEEE transactions on Neural Networks, vol. 10, 1999, pp. 1055-1064.

11. "CelebA Dataset," mmlab.ie.cuhk.edu.hk. https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html (date of the application: 12.02.2024)

12. "NVlabs/ffhq-dataset," GitHub, Apr. 09, 2021. https://github.com/NVlabs/ffhq-dataset (date of the application: 12.02.2024)

13. F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in IEEE CVPR, 2015, pp. 815-823.

[14] L. Van der Maaten, G. Hinton, "Visualizing data using t-SNE," J. Mach. Learn. Res., vol. 9, 2008.

SMART GREENHOUSE SIMULATION WITH SIMULINK AND TRUETIME

Giovanni Giuseppe Iacuzzo, Francesco Paolo Severino

Computer Engineering and Networks Laboratory – Kore University of Enna – Italy
Email: {giovanni.iacuzzo, francescopaolo.severino }@unikorestudent.it

Abstract. The project consists in the creation of a MATLAB/Simulink model, using TrueTime as well, that simulates devices in a smart greenhouse, interconnected by cabled and wireless connections. Simulated devices include: soil humidity, air humidity and temperature sensors, a controller which analyzes sensors data using fuzzy logic, an actuator which intervenes on the values with soft computing techniques and a gateway which handles the communication between devices. The performance metrics evaluated are response time, reaction time and throughput, plus other metrics derived from them.

Introduction

The project implements a smart greenhouse that works with a net of sensors, controllers and actuator to optimize the management of agricultural resources. It was developed using MATLAB/Simulink and the TrueTime library, used to simulate control systems in real time. The paper is divided in sections that deal the approach and the architecture of the project, the smart greenhouse simulated scenario, results of system performance evaluation and conclusions.

RELATED WORKS

The paper highlights the fundamental role that advanced technologies are playing in the evolution of contemporary agriculture, pushing it towards innovative paradigms like smart and sustainable agriculture. Emerges the magnitude and the urgency to embrace innovative solutions that integrate in a synergistic way artificial intelligence, advanced connectivity and real time simulation to develop efficiency and promote environmental sustainability of agricultural practices. This approach not only tries to optimize production processes, but also to mitigate negative impacts on the environment, paving the way to a new era of agriculture, more aware and responsible.

THE PROPOSED APPROACH

The project implements two types of connection: wireless, with the IEEE 802.15.4 *ZigBee* standard, and cabled, with the CSMA/CD protocol.

In networks 1 and 2, composed respectively by sensors for reading data and by the controller of power management, communication is mainly wireless, while in network 3, where the actuator that modifies and updates input data is located, communication is exclusively cabled.

The three networks have as a common centre the gateway, which in each of them is node number 4. The networks are composed in the following way:

1. Sensors network:
1.1 **Temperature sensor**: Measures temperature in a range from 10 to 45 °C, consumes 10 mW/h of power, and carries a battery of 10.000 mA.
1.2 **Air humidity sensor**: Measures air humidity in a range from 30% to 90%, consumes 10 mW/h of power, and carries a battery of 10.000 mA.
1.3 **Soil humidity sensor**: Measures soil humidity in a range from 30% to 90%, consumes 10 mW/h of power, and carries a battery of 10.000 mA.
1.4 **Power generator**: wired connection. It inputs a current value (a sine wave with an amplitude of 0.05A) and a voltage value (a sine wave with an amplitude of 18 V and a phase displacement of $\pi/2$), it outputs a power calculated as the product of voltage and current, which will be transmitted to every sensor of the network.

2. Controller network:
2.1 **Controller**: Its purpose is to output a power, depending on the value of temperature and on a value computed by a *fuzzy* controller from the two humidities. It carries a 50.000 mA battery and consumes 0,04 W/h of power.

3. Actuator network:

3.1 **Actuator**: Its purpose is to modify, depending on the value of power received, input values to bring them more or less to nominal values. Depending on the value of power it sets an irrigation flag, and with *soft computing* techniques it achieves automatic irrigation. There's another value, which could be set in case of extreme values, that activates an alarm, bringing irrigation to level 5.

4. Network centre:

4.1 **Gateway**: We used a common centre to transmit data to achieve a simpler communication, making the system more fluid and efficient.

SCENARIO

The implemented scenario is the following:
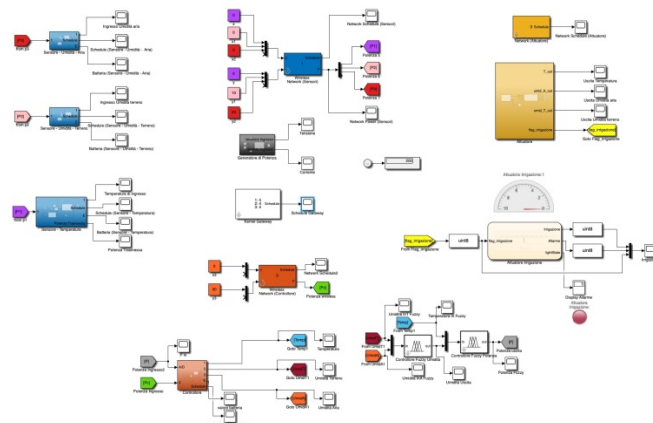
As it can be seen, the sensors network starts with a power generator attached to the sensors, which send data read to the gateway, which will forward it to the controller, that starting from the data received will choose a value of power and will send it to the gateway. Lastly, the gateway forwards the entire data packet to the actuator which will apply the desired changes.

In detail, the functioning of the system can be divided in 4 levels:

Level 1 – Sensors

Wireless sensors take as an input random values generated by an Uniform Random Number object. Being wireless, each one of them has a battery, with some starting energy that will be consumed over time. Each sensor has a TrueTime Kernel that chooses the power consumption and creates a Mail-Box named "power_signal" to receive a power value from the generator. The power generator is connected to sensors via cable; it takes a voltage and a current as an input and multiplying them it calculates the amount of power to send to each sensor in the network.

Every sensor implements via a subsystem power management, subtracting it from the power taken as an input from the sensor battery. This power management is implemented using a *fuzzy* controller, which takes as an input the power sent by the kernel and the one sent by the power generator, so as to make an estimate. This estimated power will go through a MATLAB function, that, depending on the value, will output a new power subjected to amplification given by impulses of extremely short duration, in order to never overcome a value of 0.09, which is the maximum value that can be subtracted. This is done to achieve a longer lasting battery.

Each sensor executes a periodic task, during which it receives a message containing the data read. After that the sensor takes the received message, sets the type as his own specific type, takes note of the time (to calculate reaction time) and sends the entire packet of data to the gateway.

Level 2 – Gateway

The gateway is connected to every network, and has the only job of forwarding data in order to facilitate communication between the various components of the system.

It creates 4 MailBoxes to enter data, one for every sensor and one relative to power management of the controller.

It executes two separate tasks: the first function of the gateway is to read data sent by the sensors, to verify how many packets have been sent from each sensor in certain time intervals and to forward to the controller the entire data packet. The second function of the gateway is to receive the data packet from

the controller along with the chooses power and to forward everything to the actuator. It too will verify how many packets it received.

Level 3 – Controller

The controller's purpose is to take data sent by the gateway and use them to select a suitable power in real time.

It retrieves data sent by the gateway and with two *fuzzy* controllers in series it chooses the power to forward to the gateway.

Level 4 – Actuator

The actuator is an Ethernet component utilizing the CSMA/CD protocol. Receives data forwarded by the gateway and with a MATLAB function, based on the power received as an input, modifies the values to bring them to nominal levels. The editing happens subtracting from the input value the power received multiplied by some constants. Another job of the actuator is to handle automatic irrigation. The irrigation flag is set up by the actuator using *soft computing* techniques inside a *state chart*, that uses it to update the irrigation level. Inside the *state chart* there are two systems working in parallel: the first one handles the irrigation increasing, decreasing or maintaining still the irrigation level, the second one informs the user, in case of error, when the system tries to rise above the maximum value or to go below the minimum value, lighting up a lamp bringing the irrigation value to 5.

PERFORMANCE EVALUATION

Three separate simulations of the model were performed, with a duration of 600 seconds each, modifying appropriately the random value generators seed in every simulation. The results of each performance metric measured coincide in each of the three simulations.

Below are the results of each performance metric:

- **Nominal capacity**

In the context of this project and this ideal simulation, the nominal capacity, or bandwidth, coincides with the throughput.

- **Usable capacity**

What was said for the nominal capacity also applies, in this ideal simulation, for the usable capacity.

- **Efficiency**

The efficiency, or the ratio between nominal capacity and usable capacity, in this ideal simulation, is 1 to 1.

- **Response time**

The response time was measured using the TrueTime function *ttCreateLog*. Below are the results for each task:

- ❖ *Air humidity sensor*: constant at 0,05 seconds
- ❖ *Soil humidity sensor*: constant at 0,05 seconds
- ❖ *Temperature sensor*: constant at 0,05 seconds
- ❖ *Gateway*
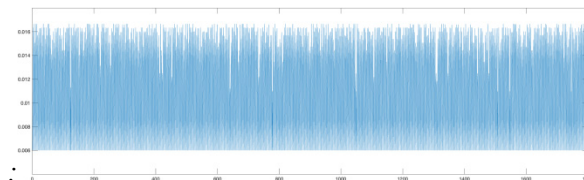
- ❖ *Gateway2*: constant at 0,03 seconds
- ❖ *Controller*:
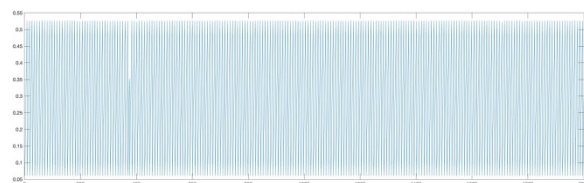
❖ *Actuator*:
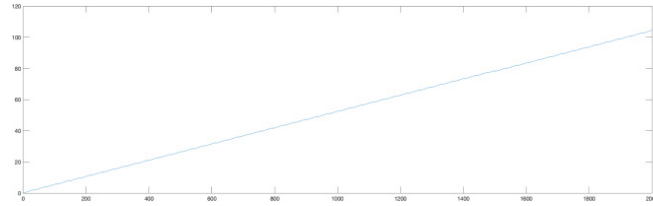
● **Sensors reaction time**

The reaction time of the three sensors, measured as the difference between the moment in which the controller starts executing a request and the moment in which the sensor sends it, appears to be periodic for the entire duration of the simulation. The amplitudes in seconds are different for each sensor, but each is constant over time.
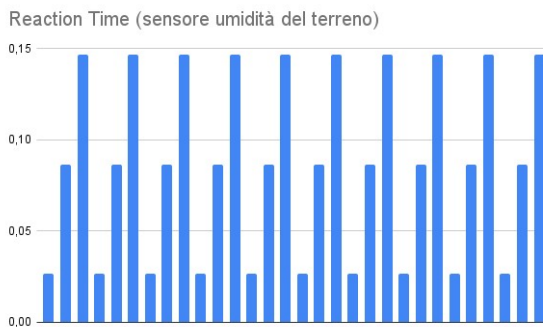
● **Stretch factor**
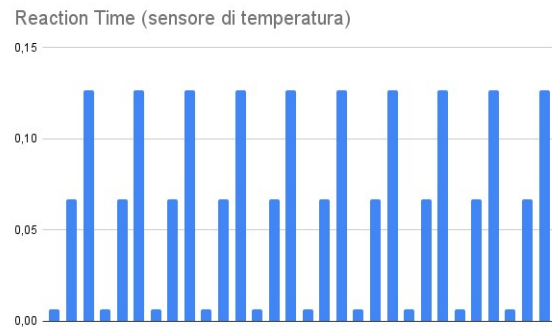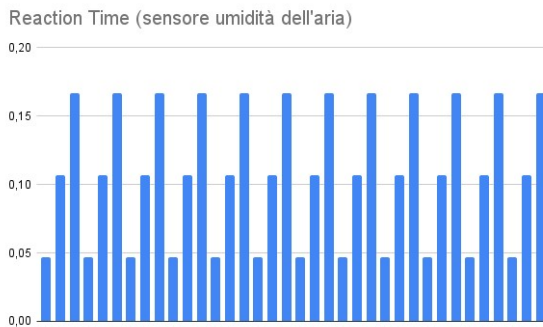
In the simulations of this project there is no deterioration in performance, as the reaction time is periodic and is not subject to loads of different magnitudes.

● **Gateway throughput**
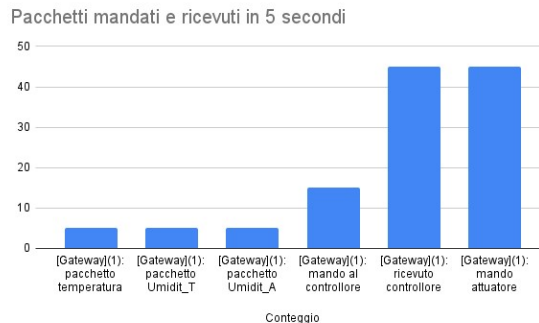
The number of packets received and sent by the gateway was measured, printing a message on the console for each of them, during three 5 seconds intervals at the beginning, middle, and end of the simulation. The results in the three intervals are identical, so the throughput is constant. All packets are 10 bytes in size.

In the interval of 5 seconds the gateway receives 5 packets from each sensor and 45 packets from the controller, the input throughput is therefore 120 bytes/sec. In the same interval it sends 15 packets to the controller and 45 to the actuator, so the output throughput is still 120 bytes/sec.

VI. ConclusionS

The project is a simulation of a smart greenhouse implemented using MATLAB/Simulink and TrueTime, integrating cabled and wireless networks with control algorithms based on *fuzzy* logic e *soft computing*. The system proves its stability and consistency in the simulations, with consistent results in diverse conditions. Performance metrics show an efficient and reactive behavior. The model could be improved incorporating more ambientalo parameters, integrating data from sensors in a real greenhouse to achieve an higher precision. This work provides a solid base for further studies in the field of computer networks, offering a flexible model for optimized management of smart greenhouses, with a focus on performance evaluation.

References

1. Anton Cervin, Dan Henriksson, Martin Ohlin, "*TrueTime 2.0 – Reference Manual*", Department of Automatic Control, Lund University, February 2016 http://archive.control.lth.se/media/Research/Tools/TrueTime/report_2016-02-10.pdf

FORMATION OF CONSUMER VALUES BASED ON TRENDS

Daria Shchukina

Saint Petersburg State University of Aerospace Instrumentation
E-mail: shchukinadaria00@gmail.com

Abstract. *The work defines the concepts of "value innovation" and "customer value". Based on the analysis of user search queries, news publications, the number of patent applications, research publications, the number of new vacancies in the labor market and the volume of investments in innovation, the main technological trends that exist today and value innovations related to them were identified. Methods for creating consumer value are considered and the problem of contradictions in creating value for the company and the consumer is identified.*

The concept of consumer value of innovation is formed on the basis of key trends formed through the development of modern technologies, depending on the need to improve the quality of life not only of the population, but also the production needs of leading enterprises in a particular region.

Consumer value is understood as a set of main advantages of an innovation that satisfy the demands and needs of consumers in various consumption segments. Basic consumer values will be considered within the framework of the organization of large-scale production and the requirements for the formation of a favorable and efficient production environment, as well as for the sale of manufactured products.

New technologies and products are required to improve the quality of organizational production processes and solve specific problems, and obtain new opportunities and advantages.

A trend is understood as the direction of development of the most popular areas of life, supported by objective research and statistical data on the demand for solving problems in a certain area.

The formation of technological trends occurs on the basis of an analysis of the country's economic processes. The formation of a trend is usually associated with certain fundamental and technical factors that affect supply and demand in the market. Trends are documented in the form of government documents, such as forecasts, development strategies, and not only government regulation of development directions, but also reports of large companies and international financial institutions have a significant influence. These research-based documents are highly reliable and comparable. And today, this list can also include influential bloggers who set directions for the mass use of innovations within their audience [1;2].

The main factors influencing the identification of the leading and most significant trends are:

1. Research and development;
2. World market needs;
3. Investments and financing;
4. Regulatory policies;
5. Globalization and international competition;
6. Social, environmental and geopolitical problems.

Instrumentally, the presence of any trend is proven by numerical confirmation of the main characteristics:

- Tilt angle;

The trend slope angle refers to the difference between the trend line and the horizontal line from which it is drawn. The larger the angle, the stronger the trend.

- Time scale;

The longer a trend remains relevant, the stronger and more significant it is. It is believed that if a trend line is 5 years long, then it is more stable than an annual one.

- Duration;

The duration confirms the reliability of the trend.

- Number of touches.

The more times the point touches the trend line, the more stable this trend is considered.

Today, about 10 of the most relevant and popular trends in technology development related to widespread digitalization and automation of production processes are being considered [3]:

1. Use of applied artificial intelligence;
- Economic benefit from 17 to 26 trillion dollars;

- The number of organizations using artificial intelligence has grown from 20% in 2017 to 50% in 2022;

- Reducing investment from $146 billion to $106 billion.

2. Introduction of machine learning;

- The volume of investment in 2021 is 4.7 billion dollars, in 2022 – 3.4 billion dollars;

- 85% of organizations are budgeting for expanding machine learning;

- High percentage of interaction between technology platforms and organizations;

- The number of vacancies in the labor market increased by 23% from 2021 to 2022.

3. Generative artificial intelligence;

- Growth of venture investments by 425% since 2018;

- Widespread use of GPT-4;

- About 10 billion investments in OpenAI;

- 80% of research in the field of artificial intelligence relates to generative models.

4. Development of low-code platforms;

- By 2026, 80% of personnel without programming skills will not work in IT organizations;

- Creation of GitHub CoPilot.

5. Security design based on the inherent vulnerability of the company's systems;

- The volume of investments in 2022 is $47 billion;

- Increase in the number of security-related vacancies by 16% from 2021 to 2022;

6. Web3 concept;

- The number of downloads of Ethereum tools is actively increasing;

- The trend is confirmed by the desire to regulate the cryptocurrency market;

- Zero-knowledge systems.

7. Development of fiber optic networks;

- The volume of investments in 2022 is $118 billion;

- The number of vacancies in the labor market has increased by 7% over the past year;

- Increasing the number of organizations working in the field of fiber optic connections;

8. Cloud and edge computing;

- More than 400 cases of edge computing in various industries, and over the next five years this trend is expected to grow double-digit worldwide;

- Volume of capital investments in 2022: $84 billion;

- The number of vacancies increased by 12% from 2021 to 2022 as part of the trend;

- Since 2020, the need for development engineers has increased by 2.5 times, and other positions have also demonstrated multiple growth.

9. Quantum technologies;

- Creation of devices based on quantum technologies (for example, Google IBM)

- The volume of investments in 2022 is about $2 billion;

- Increase in the number of vacancies on the market by 12%.

10. Bioengineering of the future.

- The volume of investments for 2022 reaches 43 billion dollars;

- Development of vaccines against COVID-19;

- The most common topic for scientific research over the past 2 years among those presented.

The most significant and influential trends and tendencies that bring socially significant improvements to the quality of life of a wide audience are capable of forming new views and values among a significant percentage of the population. Existing approaches to the formation of consumer value imply some contradictions in views on the factors of value of a product/service/technology for the consumer and for the company. For a company, value can be quantified based on various metrics such as profitability, market growth, market share, etc. The company can strive to improve these metrics and create a value proposition that is competitive in the market. For the most part, the process of creating value for the consumer is complex and is associated with product quality, satisfaction of their needs, ease of use, brand reputation, etc. To assess customer value, many criteria must be taken into account, which can often be difficult to quantify [4].

The concept of "provided customer value", which was described by F. Kotler, is that the consumer acts rationally, choosing a product in accordance with the maximization of value due to the difference between customer value and the total costs of the consumer. However, if we consider the B2B market, and the company as a consumer of innovation, then the decisive decision is not based on standard marketing laws and cannot take into account all the standard rules for choosing a new product/service or tech-

nology. The fundamental factor will be the degree of demand for innovation to solve a specific production or social problem in the organization.

Returning to the issue of creating consumer value, we can identify the main value innovations in the form of products for each of the previously proposed technological trends.

An example of value innovation that creates consumer value for the "use of applied artificial intelligence" trend can be considered Natural Language Processing (NLP) systems, a computer vision system, and an automatic decision-making system. For the trend associated with the implementation of machine learning, a value innovation that changes the views of consumers on the formation of the value of a product will be machine learning programs such as TensorFlow, scikit-learn, PyTorch, Keras, Microsoft Azure Machine Learning. A valuable innovation for generative artificial intelligence is modern neural networks and chat bots capable of generating and communicating with the user, answering queries, learning, conducting a dialogue and generating appropriate responses, for example, Chat GPT, Writesonic, YouChat and their analogues.

Conclusion

Based on the analysis, the main technological trends that exist today were formulated and value innovations that reflect them were identified. Further research involves identifying and assessing consumer characteristics and values for innovation and its widespread integration. Trends were identified based on an analysis of user search queries, news publications, the number of patent applications, research publications, the number of new vacancies on the labor market and the volume of investment in innovation.

References:

1. Korobkin D., Fomenkov S., Golovanchikov A. Method of identification of patent trendsbased on descriptions of technical functions. // J. Phys. Conf. Ser.1015, 7, 2018, pp.1-6.

2. Mysior M., Hnat W., Koziołek S. Method of Identification of Useful Functions in the Scope of Technical System Development. // New Opportunities for Innovation Breakthroughs for Developing Countries and Emerging Economies. 2018. pp. 204-215.

3. Технологические тренды 2023 — версия McKinsey Digital // Цифровая экономика URL: https://d-economy.ru/news/tehnologicheskie-trendy-2023-versija-mckinsey-digital/ (дата обращения: 09.12.2023).

4. Сологуб Елена Владимировна Теоретические подходы к формированию потребительской ценности товара/услуги // Russian Journal of Economics and Law. 2019. №2.

A SURVEY OF EXISTING METHODS OF SCHEDULING IN LTE SYSTEMS

Sofya Smorodskaya

Saint Petersburg State University of Aerospace Instrumentation,
B.Morskaya 67, 190000, Saint Petersburg, Russia
E-mail: sofi.smorodskaya@gmail.com

Abstract. *The work describes and compares various planning methods in LTE wireless network, focusing on the efficient utilization of available frequency and time resources. An evaluation is conducted on the network's average throughput under load and the fairness of resource distribution among users.*
Keywords: *scheduling, LTE, network, throughput, fairness.*

Introduction

With the advancement of wireless technologies and the constant increase in demand for high-speed mobile internet, the evolution of LTE (Long-Term Evolution) networks has become an integral part of study and development. One key aspect of ensuring optimal performance and efficient resource utilization is channel planning.

The work provides an overview of existing methods for channel resource planning in LTE networks. This review not only elucidates the fundamental principles behind various approaches to radio resource management but also analyzes their advantages and limitations. Adding to the relevance is the fact that LTE standards lack a single prescribed method for resource allocation, leading to a large number of mobile network operators conducting independent research to find the most optimal resource distribution in an attempt to define and solve the informal task of ensuring a specified quality of service.

System Model

The LTE standard defines two directions of data transmission: uplink and downlink, which correspond to data transmission from mobile devices to the base station and in the reverse direction. This division is based on several parameters: frequency, time, and multiple access methods. Frequency division is implemented using Frequency Division Duplex (FDD) technology, where the frequency of the uplink channel is chosen higher than that of the downlink. Time division is achieved through Time Division Duplex (TDD) technology, where time intervals for uplink and downlink channels are allocated based on a schedule managed by the base station. For the uplink channel, the multiple access method employs Single-Carrier Frequency Division Multiple Access (SC-FDMA) with orthogonal frequency division, while for the downlink channel, Orthogonal Frequency Division Multiple Access (OFDMA) with orthogonal frequency division is used [1].

Resource planning in the uplink and downlink channels differs due to unique requirements and characteristics. However, common methods are applicable to both channels.
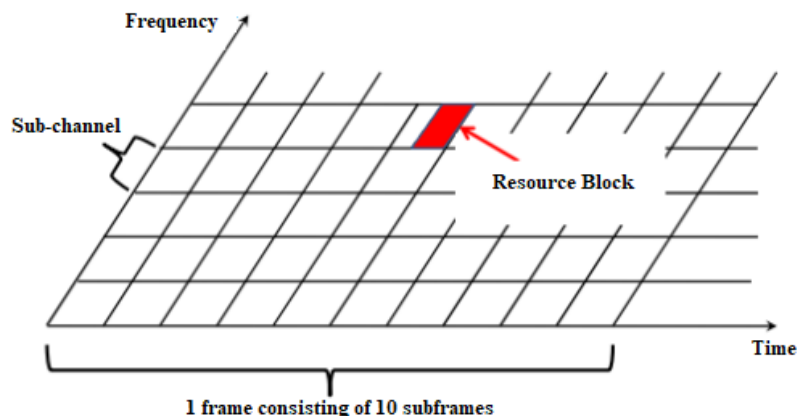
Fig. 1. Radio resource grid

As mentioned earlier, radio resources are allocated among subscribers in the time and frequency domains. To illustrate, let's refer to Fig. 1. In the time domain, resources are allocated through each time transmission interval (TTI). One TTI lasts for 1 ms. TTIs consist of two slots of 0.5 ms each. Ten consecutive TTIs form a frame with a duration of 10 ms. In the frequency domain, the total bandwidth is divided into 12 sub-channels, each with a bandwidth of 180 kHz. The minimum unit that can be allocated to a subscriber in the frequency-time resource for data transmission is called a resource block (RB) [2]. The duration of one resource block is 0.5 ms. In the frequency domain, a resource block occupies one sub-channel, which is 180 kHz wide.

The main task involves scheduling limited resources among users. Subscribers may be in different conditions and require varying levels of service quality. A key characteristic of the scheduling task is that the LTE network standard does not describe a unified scheduling algorithm. Base station manufacturers address the resource scheduling task differently to achieve the primary goal: to meet the diverse needs of cellular network users. Individual components and basic scheduling algorithms can be formulated as an optimization problem.

Description and comparison of existing scheduling algorithms

Existing planning methods can be compared based on the average throughput in the channel. It is worth noting that in the modern world, any radio signal will be subject to channel noise, which can distort transmission. The influence of noise is evaluated by the signal-to-noise ratio (SNR) and the signal-to-interference-plus-noise ratio (SINR).

The average data transfer rate is associated with the SINR value. As the SINR increases, the transmission rate also increases. This relationship is upper-bounded by Shannon's formula:

$$\overline{V} < W \log_2(1 + SINR) \tag{1}$$

Based on the above, let's move on to reviewing existing scheduling algorithms. We will consider a scenario with a fixed number of subscribers connected to the base station. The algorithms will be compared based on the evaluation of the average throughput and the fairness of resource distribution among all subscribers.

Maximizing throughput scheduler (MT)

The goal of this scheduling algorithm is to maximize the data transfer rate in the system [3]. In this case, maximizing throughput means allocating all resource blocks to the subscriber with the best conditions. This algorithm does not guarantee fair resource distribution among subscribers. For each resource block k, the subscriber i_k to whom this block will be allocated is determined by the following formula:

$$i_k = \arg \max_{i=1,M} d_{ik} \tag{2}$$

where d_{ik} is the amount of data that can be transmitted to subscriber i in resource block k, M is the number of subscribers connected to the base station. As seen from the metric, subscribers in worse conditions do not receive any resources at all. This algorithm does not ensure fair resource planning, but it does provide maximum average and total throughput.

Let's assume that the subscriber with the best conditions is labeled as number 1, and their radio conditions do not change. All time and frequency resources will be allocated to this subscriber. Then, the average throughput depends only on the SINR of this subscriber, and it can be estimated using Shannon's formula:

$$\overline{V} < W \log_2(1 + SINR_1) \tag{3}$$

Round robin scheduler (RR)

Provides fair distribution of time resources among users. The scheduler allocates resource blocks to users starting from the first and ending with user M in turn. Each user, regardless of their conditions, receives an equal amount of service time in the channel. In this context, the idea of fairness is associated with the time during which the channel is used by subscribers.

However, such a method is unfair in terms of user throughput, as in wireless systems, it depends not only on occupied resources but also on channel conditions [2]. Since resources are allocated to all us-

ers, the average speed compared to the algorithm maximizing speed will decrease. An expression for estimating the average throughput can be written as follows:

$$\overline{V} < \frac{W}{M} \sum_{i=1}^{M} \log_2(1 + SINR_i) \tag{4}$$

Blind equal throughput scheduler (BET)

This scheduler distributes resources considering the past average throughput. According to this distribution strategy, the user experiencing the lowest throughput undergoes resource displacement: they will be served until they reach the same throughput as other subscribers in the channel.

Let's consider the expression for this scheduler. Let $\overline{R_i(t)}$ represent the average transmission rate of subscriber i over the time interval from the start of operation to time t. $r_i(t)$ is the transmission rate for subscriber i over the time interval with number t. The expression for the average subscriber rate over the observation time t be expressed as follows:

$$\overline{R_i(t)} = \frac{\sum_{x=1}^{t} r_i(x)}{t} \tag{5}$$

According to the scheduling conditions, frequency and time channel resources will be allocated to subscriber i, for whom the average transmission rate $\overline{R_i(t)}$ is minimal over the time interval $(t-1)$. This condition can be expressed as follows:

$$i = \arg\min_{i=1,M} \overline{R_i(t-1)} \tag{6}$$

or

$$i = \arg\max_{i=1,M} \frac{1}{\overline{R_i(t-1)}} \tag{7}$$

Thus, a user with poor channel conditions will be serviced more frequently than others to ensure fairness in throughput. This scheduler does not consider channel conditions when making distribution decisions and belongs to the category of channel-unaware scheduling algorithms. The metric shows that BET has a significant strategy for ensuring fairness in resource allocation. However, if a user consistently encounters poor channel conditions, the scheduler often allocates resources to this user because transmission frequently fails due to the channel state. This ultimately reduces average throughput [3]. Deriving an expression to estimate the average data transfer rate for this scheduler is a non-trivial task since resource allocation among subscribers occurs on more complex principles compared to previous cases.

Proportional fair scheduler (PF)

The PF scheduler finds a compromise between a scheduler that maximizes speed and a scheduler that ensures equal speeds, resulting in a compromise between fairness in resource allocation and spectral efficiency. The scheduler's strategy is as follows: to distribute equal speeds to all users whenever possible, but the user in better conditions will have a higher speed. For each resource block k, the subscriber i_k to whom this block will be allocated is determined by the following formula:

$$i_k = \arg\max_{i=1,M} \frac{d_k^i(t)}{\overline{R_i(t-1)}} \tag{8}$$

where $d_k^i(t)$ is the throughput of subscriber i in subframe t for resource block k, and $\overline{R_i(t-1)}$ is the average transmission rate at subframe $(t-1)$ for user i.

This scheduler provides better fairness in the distribution of frequency-time resources and average throughput compared to the BET scheduler. Expressions for estimating the average speed, similar to those for the BET scheduler, are difficult to derive.

Features of computing estimates for average throughput

Computing $\overline{R_i(t)}$ using formula (5) for BET and PF schedulers is very memory-intensive, as it requires storing large arrays of data for each subscriber for the entire time they are connected to the base station. There is an alternative approach to computing the average throughput, which is more commonly used in practice. Let's introduce a coefficient β ∈ [0, 1]. It indicates how far back in time the average throughput will be analyzed. Now let's rewrite the expression for $\overline{R_i(t)}$ using β:

$$\overline{R_i(t)} = \beta\overline{R_i(t-1)} + (1-\beta)r_i(t) \tag{9}$$

Formula (9) is valid when the values of $r_i(t)$ and $r_i(t-1)$ do not differ significantly from each other, which is typically observed in the LTE wireless communication channel. From formula (9), it can be seen that when β is zero, past average throughputs are not considered at all, and the estimate is based solely on the current throughput. When β is one, the current throughput is not considered. The value of the coefficient β needs to be chosen based on the problem conditions. $\overline{R_i(t)}$ represents a moving average throughput, which is updated every subframe for each subscriber.

Thus, the coefficient β determines the period during which a fairness window will be ensured:

$$T_f = \frac{1}{1-\beta} \tag{10}$$

When β is equal to zero, the results of the average throughput in the past will be equal to the last instantaneous speed, and the fairness window will be equal to one TTI. On the other hand, as β approaches one, the last throughput will never be included in the past throughput calculation, and the fairness window theoretically becomes infinite [2].

Conclusion

In order to develop terahertz beam In conclusion, the presented review of existing methods of scheduling in LTE networks underscores the diversity of approaches to effectively utilize available frequency and time resources. Each method has its own characteristics and advantages tailored to specific network requirements and characteristics. Schedulers can incorporate channel information or build metrics without relying on it. Each of these strategies to some extent ensures fairness in distribution among users and enhances the overall network throughput. The work provides a basic understanding of existing resource planning approaches in the context of LTE networks, which can be applied to optimize data transmission processes in wireless networks.

References

[1] Ndayikunda J. Development and analysis of a model for dynamic resource allocation of wireless access nodes in transmitting heterogeneous IoT traffic.

[2] Capozzi F. et al. "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey." IEEE communications surveys & tutorials 15.2 (2012): 678- 700.

[3] Md. Emdadul H., Faisal T., Muhammad R. A. K., Kai-Kit W. " A Survey of Scheduling in 5G URLLC and Outlook for Emerging 6G Systems." IEEE Access 11 (2023): 34372-34396.

ALGORITHM FOR MATHEMATICAL MODELING OF A RAYLEIGH RANDOM PROCESS WITH A SPECIFIED SPECTRUM WIDTH

Valery Topchiy

State University of Aerospace Instrumentation, Ivangorod Branch
St.Petersburg, Russia,
E-mail: topchy.roma@gmail.com

Abstract. *The study of complex systems, which include automatic control systems, can often be carried out only by mathematical modeling methods. Random processes with the Rayleigh distribution law are often used as mathematical models of information and noise signals in these systems. The paper discusses issues of synthesis of algorithms for modeling Rayleigh random variables with a given correlation coefficient.*
Keywords: *complex system, mathematical modeling, random process, Rayleigh law, information signals.*

When designing and studying automatic control systems, which belong to the class of complex systems, mathematical modeling methods play an extremely important role, since such complex systems usually do not lend themselves to rigorous analytical study [1]. Rayleigh processes are often used as mathematical models of the processes occurring in these systems, for which the width of the power spectral density or, accordingly, the correlation function is usually specified [2]. In this case, they try to limit themselves to Markov processes of a low degree of connectivity. In this paper, various approaches to the synthesis of algorithms for modeling Markov-Rayleigh random processes are considered, analyzed and presented.

The most general type of Rayleigh random process occurs when radio signals are used to transmit information against a background of normal noise. In this case, this distribution describes the two-dimensional distribution density of the sum of the envelope of a stationary Gaussian process $n(t) \sim N\left(0, \sigma_.^2\right)$ and deterministic signal $s(t) = a(t)\cos\left(\omega_0 t\right)$ [3]. This distribution is called the generalized Rayleigh distribution or Rayleigh-Rice distribution. Since in this case the total process will be nonstationary, the distribution of its instantaneous values will also be nonstationary. The functional form of the generalized Rayleigh distribution is written as [4]

$$w\left(\xi_1,\xi_2,t,t+\tau\right) = \frac{\xi_1\xi_2}{\sigma^4\left(1-R^2\left(\tau\right)\right)} \exp\left\{-\frac{\xi_1^2+\xi_2^2+a_1^2+a_2^2-2a_1a_2R\left(\tau\right)}{2\sigma^2\left(1-R^2\left(\tau\right)\right)}\right\} \times$$

$$\times \sum_{m=0}^{\infty} \varepsilon_m I_m\left[\frac{R\left(\tau\right)\xi_1\xi_2}{\sigma^2\left(1-R^2\left(\tau\right)\right)}\right] I_m\left[\frac{a_1-R\left(\tau\right)a_2}{\sigma^2\left(1-R^2\left(\tau\right)\right)}\right] I_m\left[\frac{a_2-R\left(\tau\right)a_1}{\sigma^2\left(1-R^2\left(\tau\right)\right)}r_2\right], \tag{1}$$

where are $\xi_1 > 0$, $\xi_2 > 0$ – the values of the process $\xi(t)$ at points t_1 and t_2, that is $\xi_1 = \xi(t)$, $\xi_2 = \xi(t+\tau)$, $a_1 = a(t)$, $a_2 = a(t+\tau)$, $\varepsilon_0 = 1$, $\varepsilon_m = 2$, $m > 0$, $R\left(\tau\right)$ – ome normalized correlation function, defined through the normalized correlation function of the stationary normal process $n(t)$, $I_m(\ )$ – modified Bessel function of the imaginary argument.

If the deterministic signal is a harmonic oscillation with frequency ω_0 and amplitude U_0, then $a_1 = a_2 = U_0$ and the probability distribution density of a stationary random process $\xi(t)$, after simple algebraic transformations, will take the form [5]

$$w\left(\xi_1,\xi_2,\tau\right) = \frac{\xi_1\xi_2}{\sigma^4\left(1-R^2\left(\tau\right)\right)} \exp\left\{-\frac{\xi_1^2+\xi_2^2}{2\sigma^2\left(1-R^2\left(\tau\right)\right)}\right\} \exp\left\{-\frac{U_0^2}{\sigma^2\left(1+R\left(\tau\right)\right)}\right\} \times$$

$$\times \sum_{m=0}^{\infty} \varepsilon_m I_m\left[\frac{R\left(\tau\right)\xi_1\xi_2}{\sigma^2\left(1-R^2\left(\tau\right)\right)}\right] I_m\left[\frac{U_0\xi_1}{\sigma^2\left(1+R\left(\tau\right)\right)}\right] I_m\left[\frac{U_0\xi_2}{\sigma^2\left(1+R\left(\tau\right)\right)}\right], \quad \xi_1 > 0,\ \xi_2 > 0. \tag{2}$$

In the case when the interference is broadband or when we are interested in transmitting information in digital form only in series of pulses, and the pulses are spaced in time at such a distance that we can assume $R(\tau) = 0$, taking into account the equalities

$$\begin{cases} I_m \left[\dfrac{R(\tau)\xi_1\xi_2}{\sigma^2\left(1 - R^2(\tau)\right)} \right] = I_m[0] = 0, \text{ при } m = 1, 2, \ldots, \\[4mm] I_0 \left[\dfrac{U_0\xi_i}{\sigma^2\left(1 + R(\tau)\right)} \right] = I_0 \left[\dfrac{U_0\xi_i}{\sigma^2} \right], \text{ при } i = 1, 2, \end{cases} \tag{3}$$

in expression (2) only one term of the sum remains at $\varepsilon_m = 0$. Therefore, expression (2) is significantly simplified

$$w(\xi_1, \xi_2) = \frac{\xi_1}{\sigma^2} \exp\left\{ -\frac{\xi_1^2 + a_1^2}{2\sigma^2} \right\} I_0\left(\frac{\xi_1 a_1}{\sigma^2} \right) I_0\left(\frac{\xi_2 a_2}{\sigma^2} \right) \frac{\xi_2}{\sigma^2} \exp\left\{ -\frac{\xi_2^2 + a_2^2}{2\sigma^2} \right\}. \tag{4}$$

Moreover, if we assume that there is no information signal, that is, $a_1 = a_2 = 0$, then from expression (3) we obtain

$$w(\xi_1, \xi_2) = \frac{\xi_1}{\sigma^2} \exp\left\{ -\frac{\xi_1^2}{2\sigma^2} \right\} \cdot \frac{\xi_2}{\sigma^2} \exp\left\{ -\frac{\xi_2^2}{2\sigma^2} \right\}, \tag{5}$$

the most common mathematical model of interference used in mathematical modeling of complex systems.

Expressions (1), (2), (4) represent mathematical models of an additive mixture of transmitted information signals and interference. Although the expressions are very complex, modeling such additive mixtures is not that difficult, since the transmitted information signal and the interference can be separately modeled and then summed together. Modeling a known information signal does not present any difficulties; difficulties arise only when modeling Rayleigh interference, and only in the case when its correlation function is different from zero.

For mathematical modeling of processes with distributions represented by expressions (1), (2), (4), when using additive summation of information and noise signals, it is necessary to take into account that the summation in this case is not algebraic, but vector [6]. It follows that in this case it is necessary to model two quadratures of the normal process, which can be considered independent with zero mathematical expectations, identical variances σ2 and identical correlation functions. In this case, without loss of generality, one of the interference quadratures can be considered in-phase with the information signal, which naturally simplifies the modeling process.

n the case when there is no information signal and only the noise needs to be modeled, the two-dimensional probability density of the modeled process, as follows from expression (1) or (2), is written in the form

$$w(\xi_1, \xi_2, \tau) = \frac{\xi_1\xi_2}{\sigma^4\left(1 - R^2(\tau)\right)} \exp\left\{ -\frac{\xi_1^2 + \xi_2^2}{2\sigma^2\left(1 - R^2(\tau)\right)} \right\} I_0\left[\frac{R(\tau)\xi_1\xi_2}{\sigma^2\left(1 - R^2(\tau)\right)} \right], \quad \xi_1 > 0, \ \xi_2 > 0. \tag{6}$$

The correlation function $R_E(\tau)$, defined as the second mixed moment, is equal to [4]

$$R_E(\tau) = \int\limits_0^\infty \int\limits_0^\infty w(\xi_1, \xi_2, \tau)d\xi_1 d\xi_2 = \frac{\pi}{2} \cdot \sigma^2 \cdot \left[1 + \frac{1}{4}R^2(\tau) + \sum_{i=2}^\infty \frac{\left[(2n-3)!!\right]^2}{2^{2n}(n!)^2} R^{2n}(\tau) \right]. \tag{7}$$

Given the correlation function $R_E(\tau)$, the function $R_E(\tau)$ can be found from expression (7) using numerical methods. However, in practice, the functions of the normal process are specified by some functional form, which is chosen in such a way that the quadratures of the modeled process are Markov processes. Usually, they try to use quadratures with fractional-rational power spectral density and are limited to low-order Markov processes [7]. In this case, the functional form $R(\tau)$, is defined by expression (7), which makes it possible to quite simply numerically determine the width $R(\tau)$, which provides a given

width of the correlation function $R_E(\tau)$, which is inversely proportional to the width of the power spectral density of the simulated Rayleigh process.

In the absence of an information signal, it is often sufficient to model not the fluctuations of the envelope of the Rayleigh process, but the fluctuations of the square of the envelope, that is, power fluctuations. In such a situation, the process of calculating the characteristics of the simulated process is significantly simplified.

The random process $\eta(t)$, simulating power fluctuations of the Rayleigh process $\xi(t)$, is directly proportional to the square of the process $\xi(t)$, that is $\eta(t)=\xi^2(t)$. Therefore, passing to the variables $\eta_1 = \xi_1^2$ and $\eta_2 = \xi_2^2$, from expression (6) we obtain [4,8]

$$w(\eta_1,\eta_2,\tau)=\frac{1}{4\sigma^4\left(1-R^2(\tau)\right)}\exp\left\{-\frac{\eta_1+\eta_2}{2\sigma^2\left(1-R^2(\tau)\right)}\right\}I_0\left[\frac{R(\tau)\sqrt{\eta_1\eta_2}}{\sigma^2\left(1-R^2(\tau)\right)}\right], \quad \eta_1>0,\ \eta_2>0, \quad (8)$$

bivariate exponential distribution. When $R(\tau)=0$, the distribution defined by expression (8) decomposes into the product of widely known and used exponential distributions

$$w(\eta_1,\eta_2,\tau)=\frac{1}{2\sigma^4}\exp\left\{-\frac{\eta_1}{2\sigma^2}\right\}\times\frac{1}{2\sigma^4}\exp\left\{-\frac{\eta_2}{2\sigma^2}\right\}, \quad \eta_1>0,\ \eta_2>0. \quad (9)$$

The correlation function of the exponential process $R_\eta(\tau)$, defined as the second mixed moment, is calculated by the expression (8)

$$R_\eta(\tau)=\int\limits_0^\infty\int\limits_0^\infty\eta_1\eta_2 w(\eta_1,\eta_2,\tau)d\eta_1 d\eta_2=4\sigma^4\left[1+R^2(\tau)\right], \quad (10)$$

and the normalized correlation function of the exponential process $R_\eta(\tau)$ is defined as

$$r_\eta(\tau)=\frac{1}{4\sigma^2}\int\limits_0^\infty\int\limits_0^\infty\left(\eta_1-2\sigma^2\right)\left(\eta_1-2\sigma^2\right)w(\eta_1,\eta_2,\tau)d\eta_1 d\eta_2=R^2(\tau). \quad (11)$$

As normalized correlation functions of quadratures $R(\tau)$, in order to increase the performance of mathematical modeling algorithms, it is advisable to use correlation functions that correspond to Markov processes [9].

In the simplest case, when using the exponential function $R(\tau)=\exp(-\alpha|\tau|)$, $\alpha>0$, determination of α for a given width (at some level, usually $\delta=0.5$) of the normalized correlation function $r_\eta(\tau)$ analytically [10].

Bibliography

1. Izrantsev V.V., Shepeta D.A. Modeling of external signals of on-board instrument systems of fifth-generation aircraft. Scientific instrumentation. 2000. T.10. No. 2, pp. 14-19.

2. Shepeta D.A. Development of mathematical models and synthesis of algorithms for modeling input signals of on-board information processing and control systems. Dissertation for the degree of Candidate of Technical Sciences / Saint Petersburg, 2000.

3. Levin B.R., Schwartz V. Probabilistic models and methods in communication and control systems. – M.: Radio and Communications, 1985. – 312 p.

4. Levin B.R. Theoretical foundations of statistical radio engineering. – 3rd ed., revised. and additional – M.: Radio and Communications, 1989. – 486 p.

5. Shelukhin O.I., Belyakov I.V. Non-Gaussian processes. -SPb.: Politekhnika, 1992. – 312 p.

6. Isakov V. I., Shepeta D. A. Simulation of Radar Signals Reflected from the Edge of Land and Sea // Information and management systems. 2017. № 5 (90), pp. 89-94.

7. Shepeta A.P. Synthesis of nonlinear shaping filters for modeling input signals of loca-tion systems // Proceedings of the international scientific and technical conference (reports), May 1994. – Kiev: Academy of Sciences of Ukraine, NPO Kvant. – Issue 1, pp.81-85.

8. Dech R. Nonlinear transformations of random processes // Transl. from English – M.: Sov. radio, 1965.- 207 p.

9. Blaunshtein N.Sh., Sergeev M.B., Shepeta A.P. Applied aspects of electrodynamics // Saint Petersburg: Agraf+, 2016. – 272 p.

10.Shepeta D.A., Bozhenko V.V., Dolgov E.N. Algorithm for modeling correlated numeri-cal se-quences distributed according to the Weibull law // In the book: Wave electronics and in-focommunication systems. Collection of articles of the XXV International Scientific Confer-ence. In 3 parts. Saint Petersburg, 2022. pp. 130-134.

THE BOTTLE FACTORY

Gabriele Trovato, Gabriele Zaffora

Computer Engineering and Networks Laboratory – Kore University of Enna – Italy
trovato.gabriele10@gmail.com,
gabriele.zaffora93@gmail.com

Abstract. *Our case study focuses on the implementation of a bottle factory with the aim of determining the time required to meet the demand based on production costs and the quantity of bottles requested by the client.*

Introduction

The project simulates an industrial scenario, which, thanks to **fuzzy logic**, is able to determine the optimal number of bottles to produce in response to a client's request, taking into account the production cost at that specific moment. The scenario consists of three networks:

1. The first one is tasked with managing the produc- tion of components necessary for assembling the bottles (empty), such as caps, bottles, and labels.

2. The second one is responsible for filling and assem- bling the filled bottles.

3. The third and final network is comprised of a wire- less sensor tasked with detecting when the empty bottles are positioned below a solenoid valve, activating it accordingly.

The three networks communicate with each other through a gateway that distributes information among the various sections of production, detection/filling, and assembly.

Related Works

Before starting the project in Matlab Simulink, we con- ducted research in order to learn about the techniques for manufacturing a plastic bottle. Among the various techniques used globally, we chose the Plastic Extrusion Molding Process, since the amount of plastic available in the warehouse is measured in units, namely in plastic granules. Furthermore, to determine the most suitable material for the production of plastic bottles, we con- ducted research and concluded that the best material for this technique is PET. Finally, to assign truthful pa- rameters to the project regarding the proximity sensor, we consulted the technical data sheet of a manufacturer specialized in the produc- tion of proximity sensors.

Proposed Approach

Within the first network, there is a fuzzy logic system that calculates, based on the client's re- quested quantity and production cost, the actual number of bottles that should be produced. This infor- mation is then sent to the machinery responsible for producing the caps, bottles, and labels needed to meet the demand. Each machinery consists of an **MBSD** that simulates the production of the desired object. In fact, at the output of each chart, there are bottles, caps, and labels that have been produced based on a production time. These data are sent via a wired connection from the machinery kernel, which manages the machines, to the warehouse kernel, which is responsible for updating the available plastic, decre- menting the consumed plastic based on the number of items produced. In the event that the plastic falls below a certain threshold, which in the scenario is 300 units, the warehouse sends a warning signal (indi- cating the re- maining quantity of plastic) to the machines, which temporarily halt production until a new plastic restocking is performed.

When all the machinery has completed production, the number of caps, bottles, and labels pro- duced is sent via the **gateway** to both the second wired network and the third wireless network. This en- sures the activation of the proximity sensor and the subsequent filling of the empty bottles. Specifically, each time the sensor detects a bottle below the solenoid valve, a signal is sent to the valve, causing it to open and facilitate the filling of the bottle. If the water level in the tank drops below 300 liters, similar to the plastic scenario, a refill of the tank will be carried out.

Once the filling is completed, the actual assembly of the bottle takes place. This is achieved through an MBSD that simulates the machinery responsible for adding a cap and a label to each filled bottle.

Scenario

In the first network, at the input of the fuzzy block, there are two **Uniform Random Numbers**, each with a randomized seed that changes with each execution, deter- mining the number of bottles requested by the client and the production cost. The output from the fuzzy block provides the number of bottles to produce. This data is fed into a **MATLAB function**, responsible for round- ing the result and then sent through the Goto tool to the machinery kernel and individual charts, which follow the logic of the MBSD.
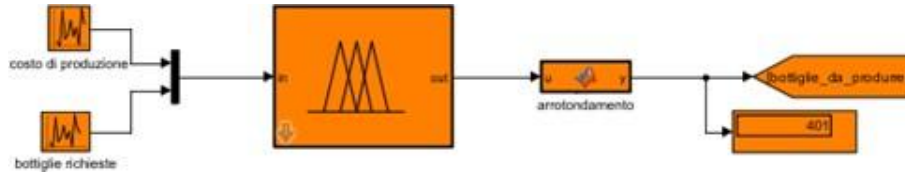
Fig. 1. Fuzzy Controller

Each chart is tasked with simulating the operation of a machinery and consists of four states:
4. **Off**
5. **On**
6. **Processing**
7. **Error**

The state, initially off, will transition to on when it receives the *ONOFF* flag set to one as input; otherwise, it will remain off. It will transition to the processing state only when it receives plastic from the warehouse kernel, and the plastic quantity is greater than three hundred. After a constant, received as input, representing the processing time, it will increment the produced object count by one.

It will transition to the off state once all necessary objects are produced, when the number of bot- tles to produce matches the produced objects, setting the *ONOFF* flag to zero. This behavior is managed within a MATLAB function.
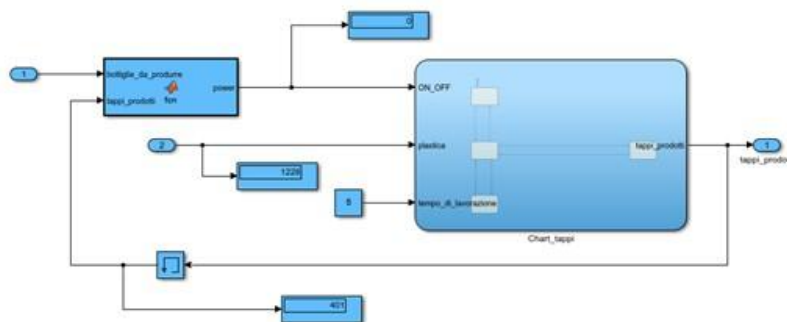
Fig. 2. Cap machine.

The warehouse kernel receives the produced items from the machines and sends them to an out- put where a MATLAB function is present. This function is responsible for managing the quantity of plas- tic.

Specifically, when the plastic quantity falls below 300 units, it pauses for three seconds to simu- late the replenishment of plastic.
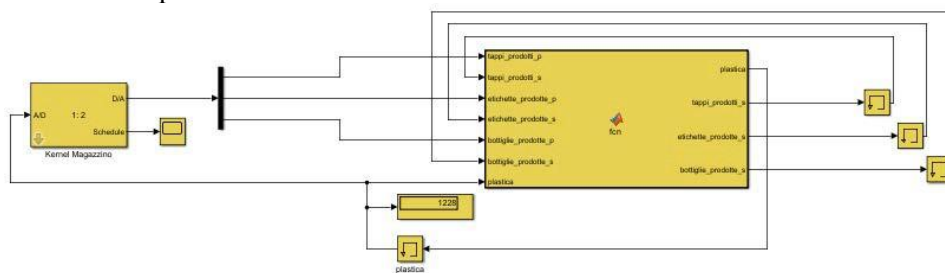
Fig. 3. Warehouse

Once the production of the necessary items is completed, the chart will send the actual number of caps, bottles, and labels produced to the machinery kernel, and this data will travel through a wired connection to the gateway.
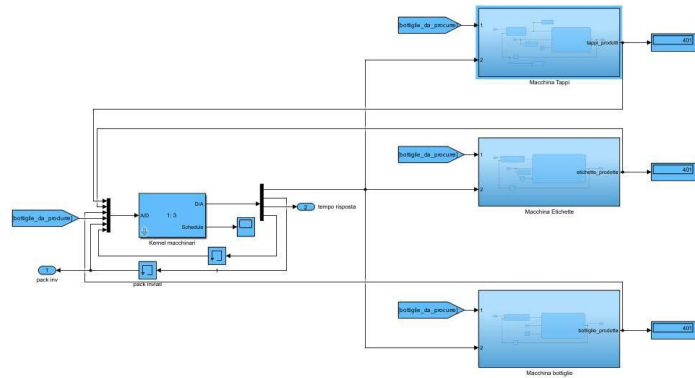
Fig. 4. Machinery

The gateway is responsible for sending the data received from the first network to the third network, ensuring the activation of the proximity sensor. This, facilitated by a **Pulse Generator**, simulates the detection of bottles for filling through the solenoid valve. As the sensor is wireless, it requires a battery, in our case simulated by the **TrueTime Battery**, with a capacity of 1500 mAh.
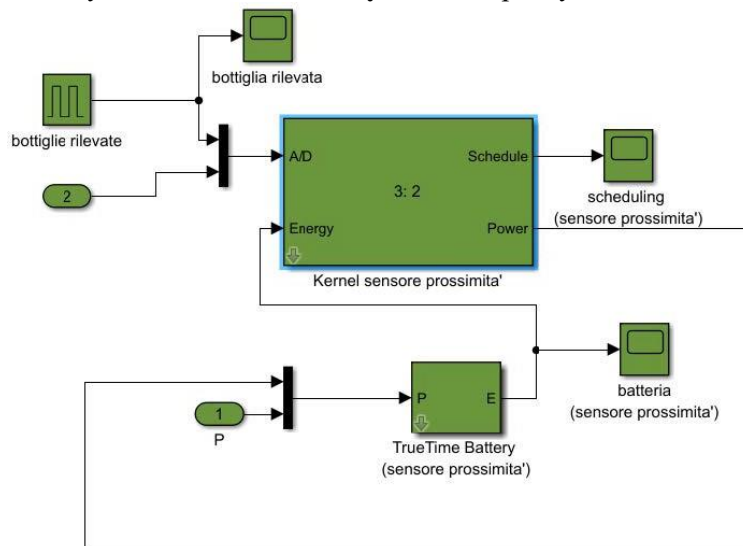
Fig. 5. Proximity Sensor

The sensor kernel, once it detects the bottle, sends this data through a wireless connection to the gateway, which then forwards it through a wired connection to the tank. This process aims to decrement the water inside the tank, consequently activating the solenoid valve. The valve, on the other hand, is responsible for increasing the count of filled bottles only when the detected flag is set to one.
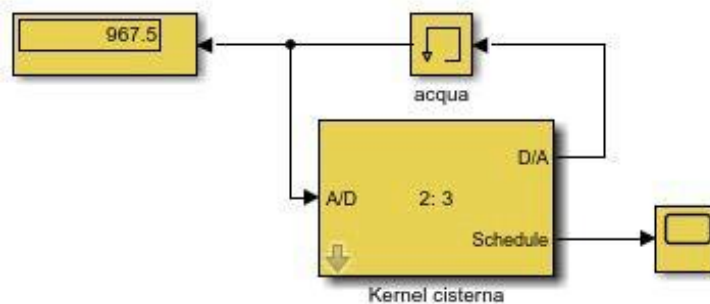
Fig. 6. Tank
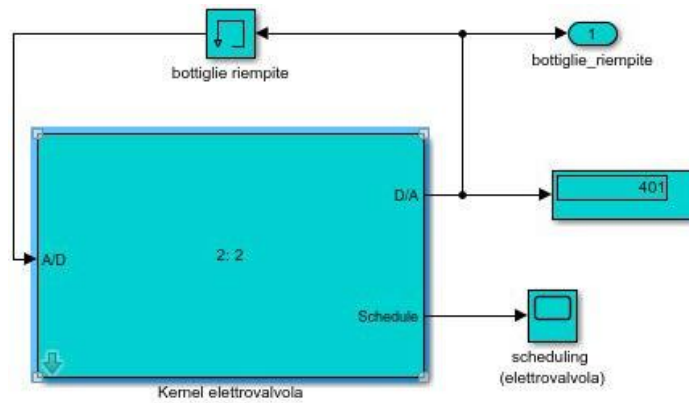
Fig. 7. Solenoid valve

Inside the tank kernel, there is a control that simulates the tank refill by using a 3-second pause and a random increment once the water level drops below 300 liters. It sends an alarm flag to the solenoid valve kernel to stop the bottle filling process.

When the number of filled bottles equals the number of produced bottles, the solenoid valve kernel sends a message to the packaging kernel containing all the necessary data. Within this node, there is also a MATLAB function responsible for decrementing the count of filled bottles, used caps, and labels consumed in production. It then sends the *ONOFF* flag to a chart simulating the bottling process.

The chart takes as input the produced caps, produced labels, filled bottles, the activation flag, and a constant indicating the processing time. It outputs the number of bottles produced.
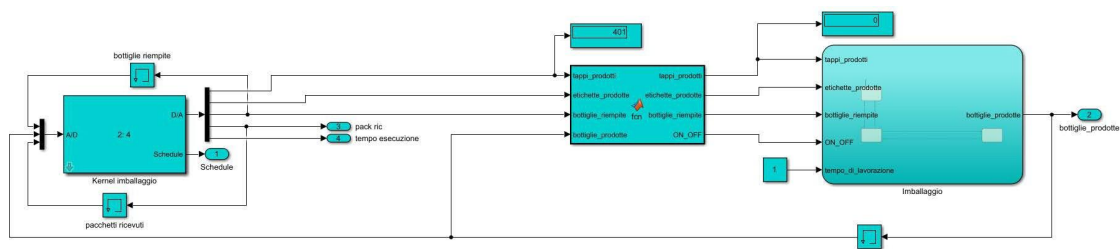
Fig. 8. Bottling

Performance Metrics

As for the performance evaluation, we have decided to calculate the **packet loss** and **response time** of the system in the following manner:

• We calculated packet loss by taking the ratio of sent packets to received packets. The sent packets are calculated when the machinery completes its production of objects and sends data from the machinery kernel to the packaging kernel. The received packets are calculated at the end of the assembly line, inside the packaging kernel. To ensure that the packets received in the packaging kernel are indeed those sent by the machinery kernel, we added a "sequence" field within each sent message, so that the increment occurs only when this field has a value of ten. The packet loss value is calculated within a MATLAB function and multiplied by one hundred to obtain a percentage result. Note that following 1959 sent packets, the packet loss is only 0.05%.
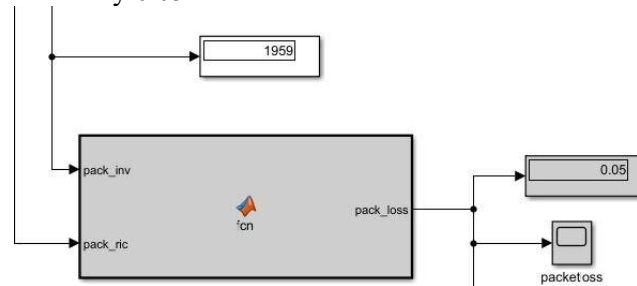
Fig. 9. Packet Loss's Example
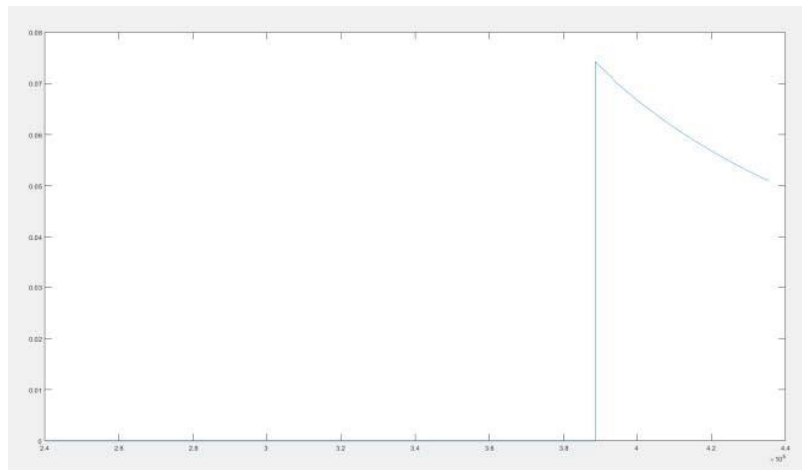
Fig. 10. Packet Loss Plot

• The calculation of the response time was performed considering the initial moment as the activation of the machinery kernel, once it received the number of bottles to be produced from the fuzzy logic. The final moment is defined as the time when the last bottle is produced in the final packaging phase. Taking these two time instances using the ttCurrentTime() tool within a MATLAB function, we subtracted the initial moment from the final moment. Note that, for example, to produce 620 bottles, it took 8235 seconds, which is approximately 2 hours.
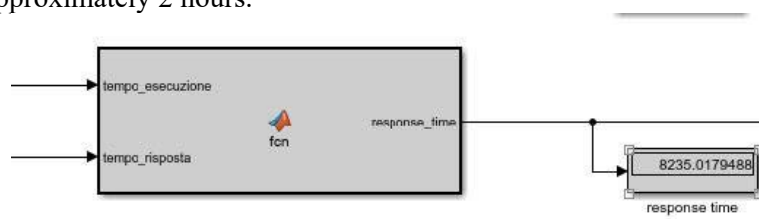
Fig. 11. Response Time

Conclusions

In summary, following various simulations, we have observed that the packet loss ranges from approximately 0.02% to 0.20% (this variation is due to the number of bottles to be produced). Certainly, this is an acceptable result, as even in the worst-case scenario, the achieved outcome is still considerable.

Furthermore, by calculating the hourly bottle production, we approach a value of around 300 units. Therefore, in the event that this project is realized and becomes a small business, the result is certainly modest. It may not be suitable for a large-scale supplier. However, to make this feasible, the implementation of multiple machines for the production of caps, bottles, and labels, along with an additional machine for bottling, would double the hourly production, making it significantly better.

References

1. Technical Specifications of the Proximity Sensor: https://it.rs-online.com/web/p/sensori-di-prossimita/1243769?matchtype=e&pla- 4575411501126251

2. Process of Creation and Bottling of Water Bottles:
https://www.thomasnet.com/articles/materials- handling/plastic-bottle-manufacturing/
https://sciencing.com/plastic-bottle-making- process-5249914.html

ESTIMATION OF THE DEGREE OF CONNECTIVITY OF A NORMAL MARKOV PROCESS WITH A FRACTIONAL-RATIONAL POWER SPECTRAL DENSITY

Violetta Tyurinova

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: viletteee@yanbdex.ru

Abstract. *The paper proposes a criterion for evaluating the connectivity of a normal Markov process with a fractional rational spectral power density. The criterion is based on calculating the rank of a matrix composed of the correlation coefficients of the process used to calculate the parameters of a linear recursive discrete shaping filter with an infinite impulse response.*

Keywords: *Markov process, spectral power density, discrete shaping filter, correlation coefficient, matrix rank, impulse response.*

Automation systems belong to the class of complex systems, in the design and research of which mathematical modeling methods play a key role, often being the only methods of designing and researching these systems. At the same time, simulation modeling, which allows to evaluate the characteristics of the system, often acts as the main research method. One of the tasks when using simulation modeling is the task of synthesizing algorithms for modeling input processes of the designed systems, which are random processes with predetermined statistical characteristics [1]. In turn, these statistical characteristics are evaluated based on some empirical data, which makes it possible to bring the simulated input processes closer to the real processes operating at the inputs of the designed and studied systems [2].

When using empirical data to build modeling algorithms in order to obtain efficient fast algorithms, it is necessary to limit oneself to modeling Markov processes [3]. In this case, the problem arises of estimating the order of connectivity of this process. In practice, they try to limit themselves to processes of a low order – the first or second. However, such simplification can lead to simulation results that are far from real processes. Therefore, this paper considers the solution of the problem of determining the order of the Markov process for normal processes in which the spectral power density is approximated by fractional rational functions of a general form, for which the most used moving average and autoregression processes are only special cases [4]. The modeling algorithm uses the method of forming filters, which are understood as discrete forming filters (DFF), which make it possible to write the modeling algorithm explicitly – in the form of a stochastic difference equation [2,5].

Let's imagine a discrete transfer function of a recursive DFF of the nth order, in the most general form

$$H(z) = \frac{\sum_{i=0}^{N-1} b_i Z^{-i}}{1 - \sum_{j=1}^{N} a_i Z^{-j}} = \frac{U_k}{\xi_k}, \tag{1}$$

where $\mathbf{a}_N = (a_1, a_2, a_N)$ and $\mathbf{b}_N = (b_1, b_2, b_N)$ – the vectors of the DFF parameters, U_k – is its output signal, and ξ_k – is discrete white normal noise with zero mean and unit variance, $\xi_k \sim N(0,1)$, Z^{-j} – is the delay operator for j clock cycles, that is $Z^{-j} U_k = U_{k-j}$. Hence the filter output signal U_k is

$$U_k = \sum_{j=1}^{N} a_j Z^{-j} U_k + \sum_{i=0}^{N-1} b_i Z^{-i} \xi_k = \sum_{j=1}^{N} a_j U_{k-j} + g_k, \tag{2}$$

where g_k – is the colored noise

$$g_k = \sum_{i=1}^{N-1} b_i \xi_{k-i}. \tag{3}$$

There are many different ways to synthesize DFF according to the specified correlation and spectral characteristics of the output signal. The main of these methods is the synthesis of DFF by spectral density of the output signal by factorization [5], however, this method can be successfully applied in practice only for second-order filters (theoretically for 4th-order filters) due to the difficulties encountered in factorizing the spectral density as a function of pseudo-frequency [5]. Synthesis methods have also been developed that require setting the correlation function at several points, however, these methods, firstly, lead to a very high-order DFF, and secondly, do not guarantee the behavior of the correlation function of the modeled process outside the interval of setting its correlation function.

So, let the lattice function of the DFF output signal be equal to $r_{(i-n(}$ corresponds to the Markov process of the nth order. We use the DFF synthesis method developed in [6], in which the DFF, in accordance with expression (2), is represented as a sequential connection of two filters – a non-recursive filter with a finite impulse response determined by a vector $\mathbf{b}_N$, forming colored noise $g_k, k = 1, 2, \ldots$, with a correlation function $R_0, R_1, \ldots, R_{N-1}$ and $R_n = 0$ by (moving average model), and a recursive filter with an infinite impulse response defined by a vector $\mathbf{a}_N$, at the output of which the required random process is formed U_k [2].

The elements of the vector $\mathbf{b}_N$ are determined from a system of nonlinear equations

$$
\begin{cases}
b_0^2 & + & b_1^2 & + & \cdots & + & b_{N-1}^2 & & = & \sum\limits_{j=0}^{N-1} b_j^2 & & = & R_0, \\
b_0 b_1 & + & b_1 b_2 & + & \cdots & + & b_{N-2} b_{N-1} & & = & \sum\limits_{j=0}^{N-2} b_j b_{j+1} & & = & R_1, \\
\ldots & & \ldots & & \ldots & & \ldots & & & \ldots & & & \ldots \\
b_0 b_i & + & b_1 b_{i+1} & + & \cdots & + & b_{N-1-i} b_{N-1} & & = & \sum\limits_{j=0}^{N-i-1} b_j b_{j+i} & & = & R_i, \\
\ldots & & \ldots & & \ldots & & \ldots & & & \ldots & & & \ldots \\
& & & & & & b_0 b_{N-1} & & = & R_{N-1},
\end{cases}
\tag{4}
$$

where $R_0, R_1, R_{N=1}$ - are the samples of the lattice correlation function of colored normal noise $g_k, k = 1, 2, \ldots$,

$$
R_n = \begin{cases}
r_n - \sum\limits_{l=1}^{N} a_l \left(r_{|n-l|} + r_{|n+l|} \right) + \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} a_i a_j r_{|n+i-j|}, & n = 0, 1, 2, \ldots N-1, \\
0, & n \geq N.
\end{cases}
\tag{5}
$$

The elements of the vector $\mathbf{a}_N$ are determined from a system of linear equations

$$
\begin{cases}
r_0 a_N & + & r_1 a_{N-1} & + & \ldots & + & r_{N-1} a_1 & = & r_N, \\
r_1 a_N & + & r_2 a_{N-1} & + & \cdots & + & r_N a_1 & = & r_{N+1}, \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
r_j a_N & + & r_{J+1} a_{N-1} & + & \cdots & + & r_{N-1+j} a_1 & = & r_{N+j}, \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots \\
r_{N-1} a_N & + & r_N a_{N-1} & + & \cdots & + & r_{2(N-1)} a_1 & = & r_{2N-1}.
\end{cases}
\tag{6}
$$

Expressions (4), (5), (6) they solve the problem of synthesizing an *N*-order DFF filter. Note that expressions (6) can be considered as independent expressions – recurrent expressions that allow calculating the correlation functions of the output signals of the DFF of an arbitrary order. Therefore, as shown in [2], the parameter vector can be determined not from solving a system of linear equations (6), but from a modified system (7)

$$
\begin{cases}
r_1 a_N & + & r_2 a_{N-1} & + & \dots & + & r_N a_1 & = & r_{N+1}, \\
r_2 a_N & + & r_3 a_{N-1} & + & \cdots & + & r_{N+1} a_1 & = & r_{N+2}, \\
\dots & & \dots & & \dots & & \dots & & \dots \\
r_j a_N & + & r_{J+1} a_{N-1} & + & \cdots & + & r_{N-1+j} a_1 & = & r_{N+j}, \\
\dots & & \dots & & \dots & & \dots & & \dots \\
r_N a_N & + & r_{N+1} a_{N-1} & + & \cdots & + & r_{2N-1} a_1 & = & r_{2N}.
\end{cases}
\tag{7}
$$

Let's introduce another DFF parameter $a_0 = -1$ and add another equation to the system (6)

$$
r_N a_N + r_{N+1} a_{N-1} + \cdots + r_{2N-1} a_1 - r_{2N} = r_N a_N + r_{N+1} a_{N-1} + \cdots + r_{2N-1} a_1 + r_{2N} a_0 = 0.
\tag{8}
$$

Let's rewrite the system of equations (6) in the following form

$$
\begin{cases}
r_0 a_N & + & r_1 a_{N-1} & + & \dots & + & r_{N-1} a_1 & + & r_N a_0 & = 0, \\
r_1 a_N & + & r_2 a_{N-1} & + & \cdots & + & r_N a_1 & + & r_{N+1} a_0 & = 0, \\
\dots & & \dots & & \dots & & \dots & & \dots \\
r_j a_N & + & r_{J+1} a_{N-1} & + & \cdots & + & r_{N-1+j} a_1 & = & r_{N+j} a_0 & = 0, \\
\dots & & \dots & & \dots & & \dots & & \dots \\
r_N a_N & + & r_{N+1} a_{N-1} & + & \cdots & + & r_{2N-1} a_1 & = & r_{2N} a_0 & = 0.
\end{cases}
\tag{9}
$$

This homogeneous system of linear equations has a solution $(a_0, \mathbf{a}_N)$ only if the determinant of the matrix of coefficients $\mathbf{r}_{N+1,N+1}$ of the system of equations is zero

$$
\det \mathbf{r}_{N+1,N+1} =
\begin{vmatrix}
r_0 & r_1 & r_2 & \cdots & r_N \\
r_1 & r_2 & r_3 & \cdots & r_{N+1} \\
r_2 & r_3 & r_4 & \cdots & r_{N+2} \\
\dots & \dots & \dots & \dots & \dots \\
r_N & r_{N+1} & r_{N+2} & \cdots & r_{2N}
\end{vmatrix}
= 0.
\tag{10}
$$

The last expression is expression (10) and can be used as a criterion for determining the order of connectivity of a Markov process, namely: the order of the filter N is equal to the rank of a matrix of the form (10) composed of the correlation coefficients of the random process U_K. Naturally, in practice, it is not the values of the correlation coefficients of the process that are used, but their estimates determined from experimental data [7], therefore it is necessary to set a threshold below which the matrix composed of estimates of correlation coefficients is considered poorly conditioned, since its determinant is close to zero.

Naturally, this criterion can be directly used to determine the degree of connectivity of Markov normal processes. However, if the DFF is used to model non-Gaussian processes formed by a nonlinear inertia-free transformation of the generating normal process [8], then the criterion can be modified, since the nonlinear inertia-free transformation does not change the order of connectivity of the Markov process.

Reference

1.   Izrantsev V.V., Shepeta D.A. Modeling of external signals of on-board instrument complexes of fifth-generation aircraft. Scientific instrumentation. 2000. Vol.10. No.2. pp. 14-19.

2.   Shepeta D.A. Development of mathematical models and synthesis of algorithms for modeling input signals of on-board information processing and control systems. Dissertation for the degree of Candidate of Technical Sciences / Saint Petersburg, 2000.

3.   Levin B.R., Schwartz V. Probabilistic models and methods in communication and control systems. – M.: Radio and Communications, 1985. – 312 p., ill.

4.   Marpl. – ml. S.L. Digital spectral analysis and its applications / Trans. from English – M.: Mir, 1990. – 584 p

5.   Besekersky V.A. Digital automatic systems. -M.: Nauka, 1976. – 576 p., ill.

6. Shepeta A.P. Synthesis of nonlinear shaping filters for modeling input signals of location systems // Proceedings of the International Scientific and technical conference / Academy of Sciences of Ukraine; NPO Kvant. Kiev, – May, 1994, Issue 1.

7. Tyurinova, V. A. Algorithms for input signals simulation of complex systems synthesized based on empirical data/ V. A. Tyurinova, G. M. Wattimena, D. A. Shepeta // Wave Electronics and Its Application in Information and Telecommunication Systems. – 2022. – Vol. 5. – No 1. – P. 506-510. – EDN OOXSVU.

8. Shepeta, D. A. Direct method of modeling logarithmic-normal distribution / D. A. Shepeta, V. I. Isakov, V. A. Tyurinova // Wave electronics and infocommunication systems : Collection of articles of the XXV International Scientific Conference. In 3 parts, Saint Petersburg, May 30 – 03 2022. Volume Part 3. – Saint Petersburg: Saint Petersburg State University of Aerospace Instrumentation, 2022. – pp. 135-139. – EDN JHRTGC.

PROGRAMMING OF COMPUTER GRAPHICS ELEMENTS IN THE OPENGL ENVIRONMENT

Dmitry Vinogradov

Saint Petersburg State University of Aerospace Instrumentation
E-mail: vinmitya@gmail.com

Abstract. *The paper presents methods for using the capabilities of a graphical software interface and provides examples of the implementation of dynamic scenes and projective transformations – the main elements for creating 3D images, elements of computer graphics and video games.*
Keywords: *OpenGL, computer graphics, 3D images, projective transformations.*

OpenGL (Open Graphics Library) is an open standard graphical programming interface (API) designed to create high-performance applications capable of displaying 2D and 3D graphics. This cross-platform interface has applications in a variety of fields, including gaming, data visualization, science and engineering, and virtual and augmented reality development.

Theoretical Foundations

OpenGL uses a frame buffer to form an image, and also supports shaders to create lighting, shadow, and texturing effects. This standard allows for multithreading and parallelization, allowing multiple CPU cores to be used to improve performance.

One of the key advantages of OpenGL is its open and free nature, which allows free use and distribution by developers without any restrictions or licenses.

Let's take a look at some of the main tasks that arise when programming 3D graphics. The first task will be to create a dynamic scene in a high-level programming language.

Dynamic scene development using OpenGL.
Software Implementation

To create a dynamic 3D scene, the Python 3 programming language was chosen. The program uses **the Pygame** and OpenGL libraries to render 3D graphics. At startup, the program creates an 800x600 pixel window and configures OpenGL to display 3D graphics using the **gluPerspective()** and **glTranslatef() functions**.

The main loop of the program handles **Pygame** events, responding to the **QUIT** event (pressing the window close button) to shut down. The render loop uses the glRotatef() **and** glClear() **functions to rotate objects and clear the screen buffer, respectively.**
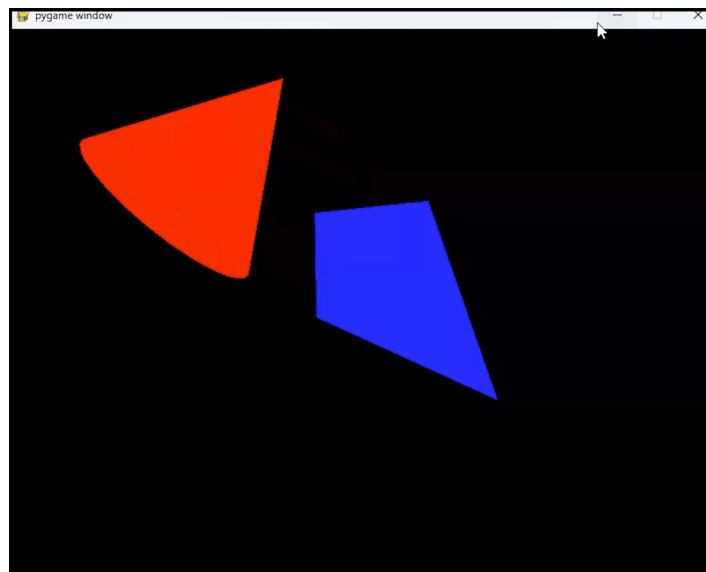
Fig. 1. An example of program execution, the cone is located above the pyramid

OpenGL functions such as **glBegin(),** glColor4f**(), and glVertex3fv**()) are used to render objects. In this case, **the cube is drawn in blue and the cone is rendered in red. The cube and cone are also rotated using** glPushMatrix() and glPopMatrix()**.**

Finally, the program refreshes the screen and waits for some time between frames to create the illusion that objects are moving. Examples of how the program works can be found in Figures 1 – 2.
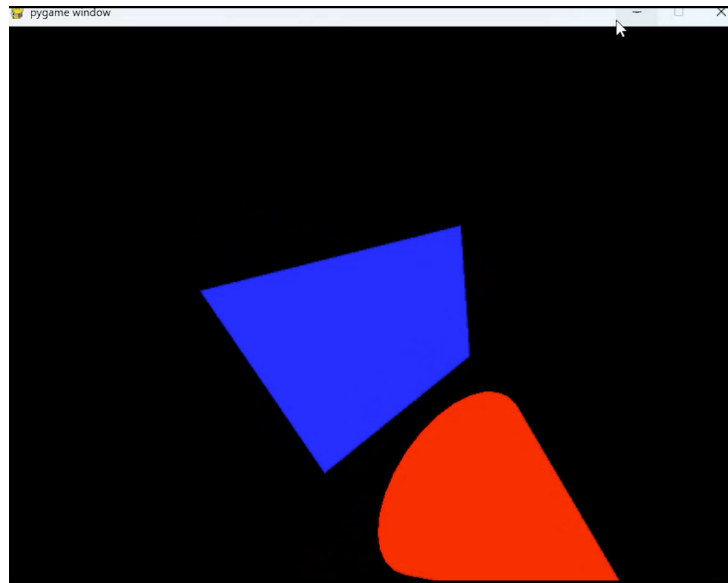
Fig. 2. an example of program execution, the cone is located under the pyramid

The program code is shown below.

```
import pygame
from pygame.locals import *

from OpenGL.GL import *
from OpenGL.GLU import *
import math

vertices = (
    (0, 0, 1),
    (1, 0, -1),
    (-1, 0, -1),
    (0, 1, 0)
)

triangles = (
    (0, 1, 3),
    (1, 2, 3),
    (2, 0, 3),
    (0, 2, 1)
)

cone_vertices = []
cone_triangles = []
cone_radius = 1
cone_height = 2
cone_slices = 30

for i in range(cone_slices):
    angle = i * 2 * math.pi / cone_slices
```

```
    x = cone_radius * math.cos(angle)
    y = 0
    z = cone_radius * math.sin(angle)
    cone_vertices.append((x, y, z))
cone_vertices.append((0, cone_height, 0))

for i in range(cone_slices):
    triangle = (i, i+1, cone_slices)
    cone_triangles.append(triangle)
cone_triangles.append((cone_slices-1, 0, cone_slices))

def Cube():
    glBegin(GL_TRIANGLES)
    for triangle in triangles:
        for vertex in triangle:
            glColor3f(0, 0, 1)
            glVertex3fv(vertices[vertex])
    glEnd()

def Cone():
    glBegin(GL_TRIANGLES)
    for triangle in cone_triangles:
        for vertex in triangle:
            glColor3f(1, 0, 0)
            glVertex3fv(cone_vertices[vertex])
    glEnd()

def main():
    pygame.init()
    display = (800, 600)
    pygame.display.set_mode(display, DOUBLEBUF|OPENGL)

    gluPerspective(45, (display[0]/display[1]), 0.1, 50.0)

    glTranslatef(0.0,0.0,-5)

    while True:
        for event in pygame.event.get():
            if event.type == pygame.QUIT:
                pygame.quit()
                quit()

        glRotatef(1, 1, 1, 1)
        glClear(GL_COLOR_BUFFER_BIT|GL_DEPTH_BUFFER_BIT)

        # Draw the cube
        glPushMatrix()
        Cube()
        glPopMatrix()

        # Draw the cone on the left of the cube
        glPushMatrix()
        glTranslatef(-2.5, 0, 0)
        glRotatef(90, 0, 1, 0)
        Cone()
        glPopMatrix()
```

```
        pygame.display.flip()
        pygame.time.wait(10)

if __name__ == "__main__":
    main()
```

As a second task, let's look at performing projective transformations using OpenGL.

Perform projective transformations using OpenGL.
Software Implementation

Projective transformations are an important concept in geometry and computer graphics. They allow you to change the shape and position of objects in three-dimensional space and project them onto a two-dimensional plane. Projective transformations are widely used in computer graphics to create three-dimensional scenes, perspective distortion, camera mapping, and other effects.

In the main loop of the program, events are processed, in this case only the **QUIT** event (closing the window). If a window closing event occurs, the program exits.

Inside the render loop, objects are rotated and rendered. First, the glRotatef() function is called to specify the rotation angle of the objects. The **glClear()** function is then called to clear the screen buffer.

Next, the object rendering functions are called: **HexagonalPyramid()** to render a hexagonal pyramid. function use OpenGL functions to determine the vertices and colors of objects.

The hexagonal pyramid also rotates on the screen. To do this, use the glPushMatrix() **and** glPopMatrix() **functions.**

At the end of each iteration of the loop, **pygame.display.flip() is called to refresh the screen and** pygame.time.wait() **to delay between frames. This allows graphics to be displayed on the screen at a specific frame rate.**

Thus, the program creates a window, configures OpenGL for 3D graphics, renders and rotates objects on the screen using **Pygame** and OpenGL. Examples of how the program works with comments can be seen in Figures 3 – 5.
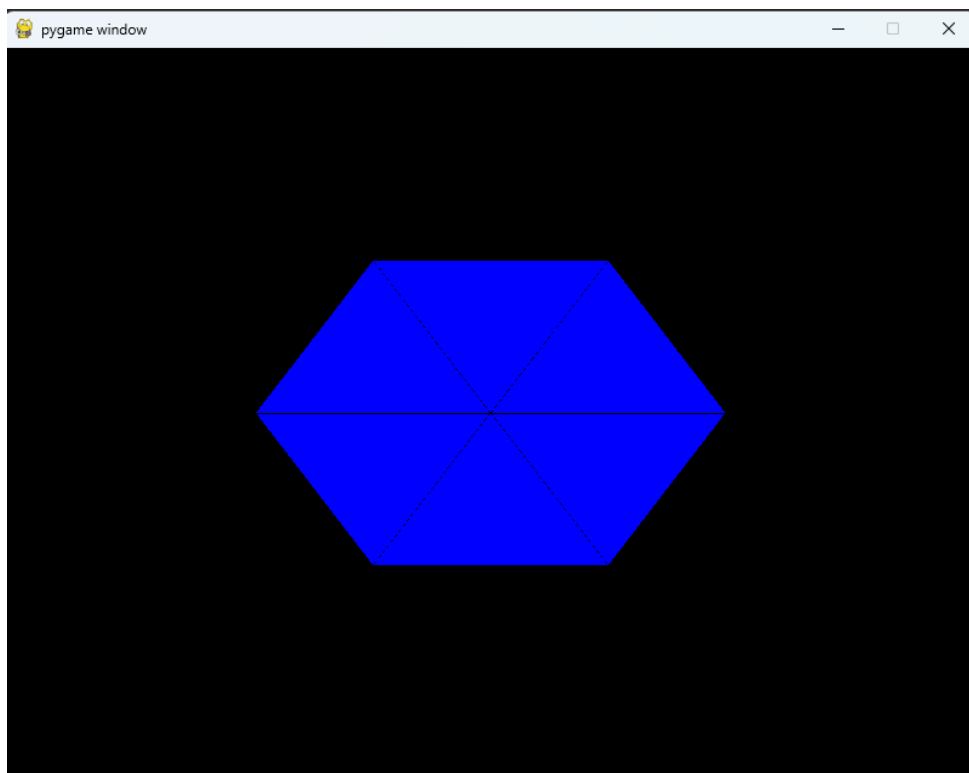
Fig. 3. a projection of the image on the xy plane
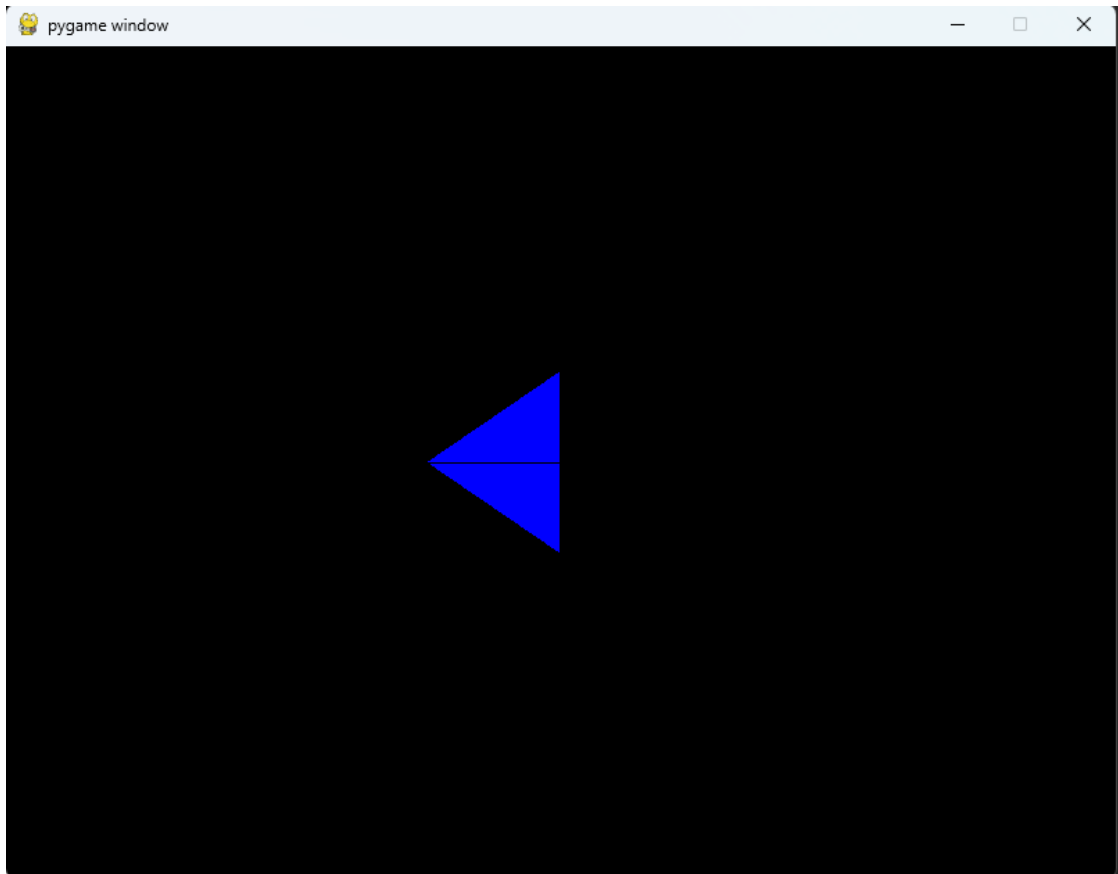
Fig. 4. a projection of the image on the xz plane
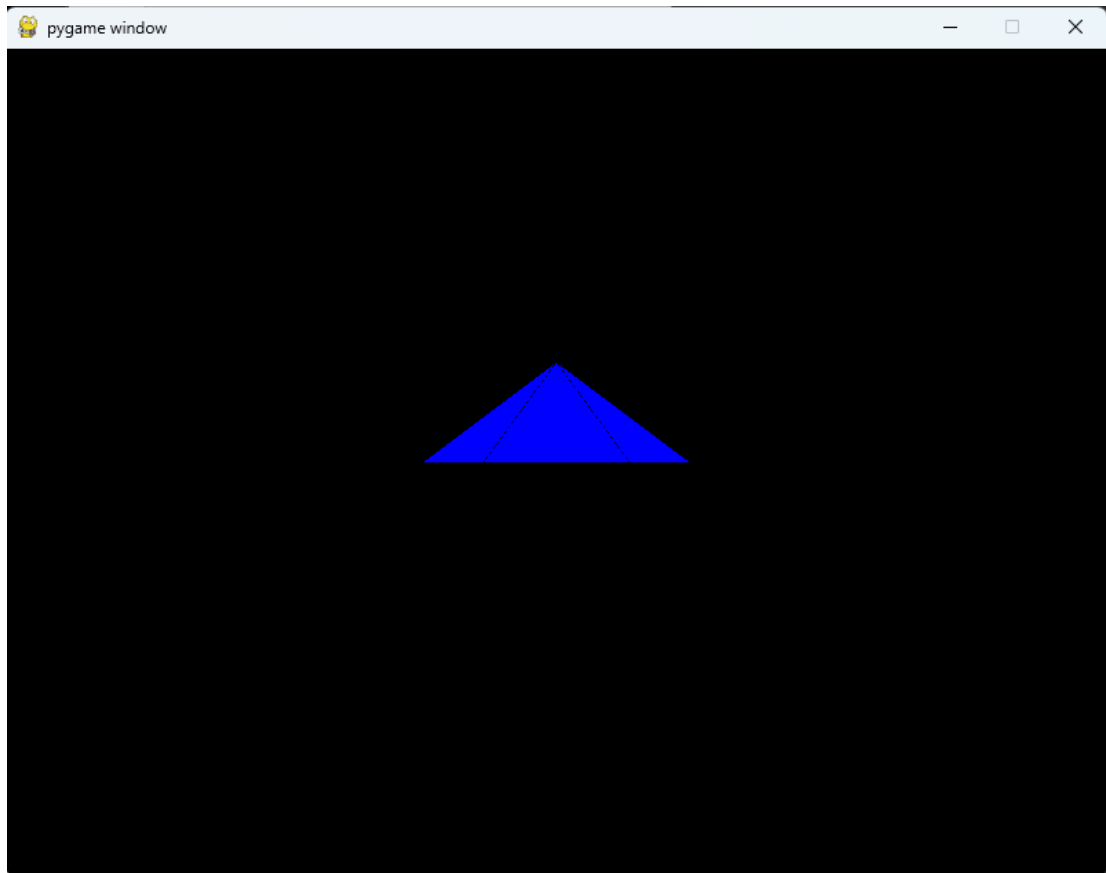
Fig. 5. a projection of the image on the yz plane

The code with comments is shown below.

```python
import pygame
from pygame.locals import *
from OpenGL.GL import *
from OpenGL.GLU import *
from math import cos, sin, pi

def HexagonalPyramid():
    # Draw hexagonal base
    glColor4f(0, 0, 1, 1)
    glBegin(GL_POLYGON)
    for i in range(6):
        angle = i * pi / 3
        glVertex3f(cos(angle), sin(angle), 0)
    glEnd()

    # Draw top triangle faces
    glBegin(GL_TRIANGLES)
    for i in range(6):
        angle1 = i * pi / 3
        angle2 = (i + 1) * pi / 3
        glVertex3f(0, 0, 1)
        glVertex3f(cos(angle1), sin(angle1), 0)
        glVertex3f(cos(angle2), sin(angle2), 0)
    glEnd()

    # Draw edges
    glColor4f(0, 0, 0, 1)
    glBegin(GL_LINE_LOOP)
    for i in range(6):
        angle = i * pi / 3
        glVertex3f(cos(angle), sin(angle), 0)
    glEnd()
    glBegin(GL_LINES)
    for i in range(6):
        angle = i * pi / 3
        glVertex3f(cos(angle), sin(angle), 0)
        glVertex3f(0, 0, 1)
    glEnd()

def main():
    pygame.init()
    display = (800, 600)
    pygame.display.set_mode(display, DOUBLEBUF|OPENGL)
    gluPerspective(45, (display[0]/display[0]), 0.1, 50.0)
    glTranslatef(0.0, 0.0, -5.0)
    glEnable(GL_DEPTH_TEST)
    glEnable(GL_BLEND)
    glBlendFunc(GL_SRC_ALPHA, GL_ONE_MINUS_SRC_ALPHA)

    while True:
        for event in pygame.event.get():
            if event.type == pygame.QUIT:
                pygame.quit()
                quit()

        # Keystroke Handling
```

```
        if event.type == pygame.KEYDOWN:
            if event.key == pygame.K_UP:
                # y-axis view
                gluLookAt(5, 0, 0, 0, 0, 0, 0, 1, 0)
            elif event.key == pygame.K_RIGHT:
                # x-axis view
                gluLookAt(0, 0, -5, 0, 1, 0, 0, 0, 1)
            elif event.key == pygame.K_LEFT:
                # z-axis view
                gluLookAt(0, 5, 0, 0, 0, 0, 0, 0, 1)


    glClear(GL_COLOR_BUFFER_BIT | GL_DEPTH_BUFFER_BIT)

    HexagonalPyramid()

    pygame.display.flip()
    pygame.time.wait(10)

main()
```

Findings

In the course of this publication, the principles of the OpenGL open graphics library were studied. The created dynamic 3D scene in the Python 3 programming language using the **Pygame and** OpenGL libraries demonstrates the use of various library functions, such as working with matrices, lighting, and rendering primitives.

These skills can be successfully applied in the development of graphical applications for data analysis, virtual reality, and other areas. As a powerful tool, OpenGL provides developers with the flexibility and freedom to create a variety of visual effects and 3D models.

References

1. Python: documentation; [Electronic resource]. URL: https://www.python.org/ (accessed: 21.01.2024)

2. PyOpenGl: documentation; [Electronic resource]. – URL: https://pyopengl.sourceforge.net/documentation/index.html (accessed: 21.01.2024)

MONITORING AGRICULTURAL OBJECTS ON THE BASIS OF VISION METHODS

R. M. Voronov, A. V. Berezin, A. D. Matveev
Saint Petersburg State University of Aerospace
instrument engineering, Saint Petersburg, Russia
E-mail: co3.14paste@gmail.com

Abstract. *The article is devoted to the development of a method of automatic recognition of rural objects based on the processing of a stream of frames taken from a small aircraft. The article deals with the actual task of monitoring and control of agricultural objects. The possibilities of computer vision are demonstrated, which allows recognizing haystacks and counting their number on the frames taken from a small aircraft. The results of the study show high accuracy of recognizing and counting agricultural objects when using vision methods.*
Keywords: *technical vision, monitoring, video frame streaming, object counting, neural network.*

Introduction

Realization of the process of automatic recognition of rural objects (livestock, crops, etc.) is an urgent task of agriculture. The most important aspect of this development is the training of a neural network (NN) and subsequent recognition of objects located on the frames of the video stream.

Modern agriculture is a dynamic field where conditions can change daily. Factors such as weather conditions, soil conditions, plant and animal health affect the yield and performance of farmland. However, regardless of these changes, farmers and landowners have the opportunity to take advantage of innovative methods and technologies to optimize their operations and improve efficiency [1-3]. One such approach is the application of vision techniques to monitor, control and analyze agricultural land, for example when controlling harvested crops or controlling livestock numbers.

PROBLEM STATEMENT

One of the major challenges in modern agriculture is the need for constant operational monitoring and control of the condition of agricultural facilities. Farmers and landowners must constantly monitor the condition of their land, plants and animals, as well as assess the volume of crop yields after it has been harvested. This can be quite a time-consuming task, especially when monitoring large areas.

In order to overcome this problem, a vision-based approach has been implemented.

As a result of this approach, the counting of the number of diverse ground objects on the video stream frames collected from small aircraft is realized [4, 5].

PREPARATION OF DATA FOR TRAINING THE NEURAL NETWORK MODEL

Data preparation is a key step in the NS training process when implementing vision-guided systems. This stage includes data collection, preprocessing and conversion into a format that can be used by the NS model for training. Subsequently, the data is divided into training (70%) and test (30%) samples. In this case, the data for recognizing haystacks [6] on video stream frames were trained. Below we present the methodology of video data preparation for NS training, which consists of 5 steps.

Step 1: Data collection and processing.
The original video material was split into a separate number of frames using ffmpeg tool [7]. This process is called dividing the video into frames. This step is necessary to convert the video into a format that can be used to train the NS model.

Step 2: Data partitioning
Further selected at the previous step frames on the basis of using the software resource roboflow [8], it is required to carry out data partitioning on this set of frames, only those of their areas where there are objects of interest to form a mask of the object in order to further recognize it.

This data partitioning process performs the assignment of labels or categories to objects on the frames, which are then used to train the NS model.

Step 3: Data granularity
Next, each individual frame is divided into 9 3x3 cells to increase the accuracy of small object recognition. This process is referred to as splitting the training sample frames into individual cells, which

allows the model to analyze and process images at a more detailed level, which subsequently improves the accuracy and efficiency of recognizing agricultural ground objects.

Step 4: Data Augmentation

Also, the training frames were supplemented with similar frames with blur effect. This was done to enable the NS model to recognize haystacks [6] in poor quality images. This process is called data augmentation and involves creating new data by applying various transformations to existing data such as rotation, scaling, flipping, etc.

Step 5: Splitting the data

Finally, the original video data was divided into two parts: 70% of the data was used for training the NS model and 30% was used for testing. This process is called splitting the data into training and test samples. Splitting the data into training and test samples allows us to evaluate how accurately the trained NS model can extend its knowledge to previously unused test data.

OBJECT SEARCH AND COUNTING

Having trained the neural network, it is further required to test it on the data that did not participate in the process of its training. These test data also need to be converted from video format into separate frames – images. This is done by splitting the video into separate frames. For example, a video with a rate of 60 frames per second and a duration of 30 seconds will be split into 1800 individual frames.

Each individual frame is then passed through an NS model that highlights the haystacks in the current image. This can be accomplished using a variety of methods, including the use of Ultralytics library tools [7].

After that, each haystack is selected in the frame. In this case, each haystack [6] is assigned a unique identification number (UIN), which allows identifying and counting these agricultural objects in the frame.

Thus, counting of the number of haystacks on the current frame at the moment is realized, and also the UINs of haystacks are used to count their total number of haystacks. This allows to avoid counting when counting the same object several times.

Once all the frames are processed, they are merged back into the video and saved for future use.

It should be noted that the objects to be recognized (Fig. 1) are haystacks and represent a characteristic structure, easily recognizable due to their conical or cylindrical shape. Their color is usually yellowish-brown, which is due to the drying and maturing process of hay. This color allows haystacks to be distinguished from other objects in the field, such as animals or farm machinery.

Fig. 1. Object to be recognized – haystack

Thus, the distinguishing features of haystacks [6], such as shape and color, allow the use of vision and object recognition techniques to identify and distinguish them in the field, which can be useful in applications such as farmland monitoring, harvest planning and yield estimation in general.

RESULTS OF COMPUTER EXPERIMENTS

On the video stream frame, the NS successfully detected haystack objects and performed counting of their total number as well as the current number in the frame (Fig. 2).
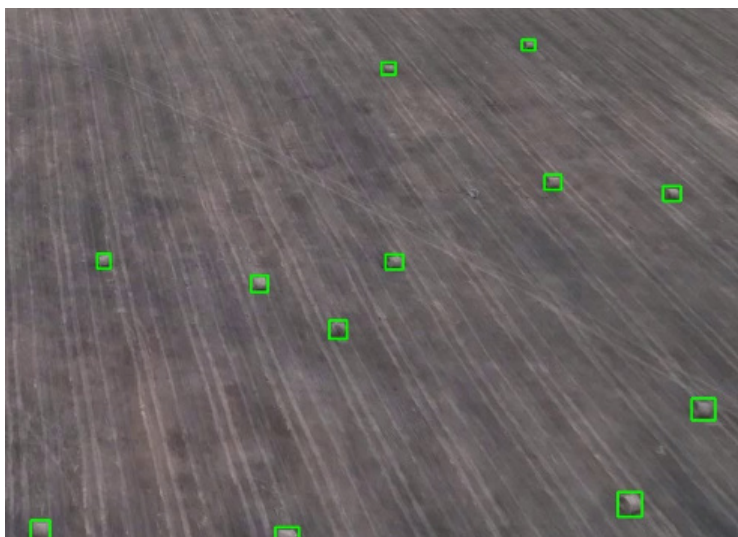
Fig. 2. Program results

Green borders around the objects indicate correct recognition and assignment of UINs to them. It should be noted that the size of the stacks is quite small, which makes their visual detection difficult, but the model coped with this task successfully.

From the computer experiment, it follows that the NS model correctly recognized 192 haystacks out of 193 possible haystacks, this indicates a high recognition accuracy, which is 99%.

Conclusion

Within the framework of this research, a process is implemented that allows to automatically recognize haystacks on video stream frames and count their number. This together allows to automate the process of accounting of agricultural objects.

Much attention is paid to the process of preparing the data selected for training the NS model. The trained model is implemented on the basis of YOLOv8 [9-11], which demonstrated high accuracy in haystack recognition, reaching a level of about 99%.

As a result, the problem of automatic haystack detection in wheat fields is successfully solved. The developed model is able to recognize haystacks in the frames of the video stream, track their total number, as well as the number in the current frame. This makes the model an effective tool for automation and optimization of agricultural processes.

LIST OF SOURCES USED

1. Rybakov A. V. V. V., Vybornov N. A., Rybakov I. A. Analysis of computer vision methods promising for application in agroindustrial complex // Caspian Journal: Management and High Technologies. 2022. №1. – pp. 128-138.

2. Rybakov A. V. Khodarova Z.M. Prospects for the use of mobile robots with vision systems in agriculture // Symmetries: theoretical and methodical aspects: a collection of scientific papers of the VII International Scientific and Practical Conference. 2018. C. 39-46.

3. Hongkun, Tian, Tianhai Wang, Yadong Liu, Xi Qiao, Yanzhou Li Computer vision technology in agricultural automation // Information Processing in Agriculture. 2020. Vol. 7, iss. 1. pp. 1-19.

4. Nenashev V.A., Shepeta A.P., Kryachko A.F.. Fusion radar and optical information in multiposition on-board location systems // 2020 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2020, Saint Petersburg, June 01-05, 2020. P. 9131451.

5. Nenashev V.A., Khanykov I.G. Formation of fused images of the land surface from radar and optical images in spatially distributed on-board operational monitoring systems // Journal of Imaging. 2021. Vol. 7, No. 12.

6. URL: https://ru.wikipedia.org/wiki/Сено (accessed on 20.02.2024)

7. URL:https://zenway.ru/uploads/knigi/ffmpeg-ru-man.pdf?ysclid=ls4g6owhdw995815842 (accessed 20.02.2024)

8. URL:https://roboflow.github.io/roboflow-python (accessed 20.02.2024)

9. Borislava Toleva. The Proportion for Splitting Data into Training and Test Set for the Bootstrap in Classification Problems Article in Business Systems Research Journal 2021.

10. Juan R. Terven, Diana M. Cordova-Esparaza, a comprehensive review of yolo: from yolov1 to yolov8 and beyond, 2023. Pp. 18-27.

11. Nenashev, V.A.; Khanykov, I.G. Formation of the complex image of the Earth surface on the basis of the pixel clustering of the location images in the multi-position airborne system // Informatics and Automation. 2021. Т. 20, № 2. С. 302-340.

STABILIZATION OF FALSE ALARMS OF THE DETECTOR OF INFORMATION SIGNALS OBSERVED AGAINST THE BACKGROUND OF INTERFERENCE FROM UNDERLYING SURFACES

Ivan Yudin

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia,
E-mail: ivan-yudin@mail.ru

Abstract. *When detecting physical objects in a given search area against the background of reflections of location signals from underlying surfaces, it is necessary to ensure a specified level of false alarms of detectors. An algorithm is proposed to ensure the level of false alarms when searching for objects against the background of reflections from the sea surface and from the ground, while stabilizing false alarms occurs regardless of the type of underlying surface.*

Keywords: *mathematical model, underlying surface, location pulses, digital detectors, false alarms, information signal,*

When receiving information signals observed against the background of the internal noise of receiving devices and reflections from underlying surfaces, the problem of stabilizing false alarms of input signal detectors is always solved [1]. This task is complicated by the fact that reflections from underlying surfaces, in particular the surface of the earth and the sea surface, differ quite strongly in their statistical characteristics [2]. For reflections from the sea surface, the most common model of fluctuations in signal amplitudes and powers is the log-normal distribution law [3]. For reflections from the earth's surface, when approximating the amplitude distribution law, the Rayleigh distribution law is often used [4], and for power fluctuations, the exponential distribution law is often used [5]. The intrinsic noise of receiving devices is always assumed to be normal, so the Rayleigh distribution law is used to describe the envelope of this noise.

Since the arrival time of information pulses is unknown, detection of the presence of an information signal at the inputs of receiving devices occurs over a certain time interval. In this case, this interval is divided into several subintervals, each of which contains a separate detector. Thus, detection occurs in parallel on all subintervals simultaneously [6]. The most commonly used detectors are well-known digital detectors of the "k out of n" type, which are quite simple to implement and are actually quasi-optimal [7], since they lose to optimal detectors of deterministic signals against the background of normal noise by no more than 1.5-2.0 decibels.

A generalized block diagram of the "k out of n" detector is shown in Fig. 1.
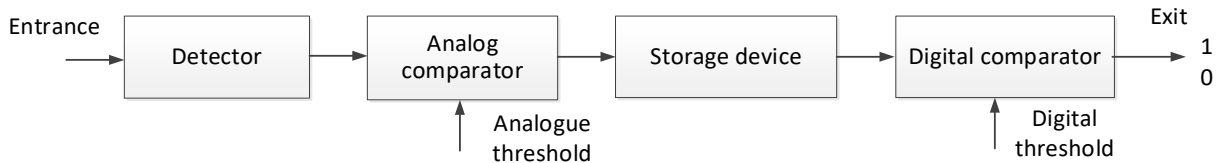
Fig. 1. Generalized block diagram of the "k out of n" detector

A packet of n radio pulses is supplied to the detector input. The detector contains two thresholds: an analog threshold for video pulses after the detector equal to U_0 and a digital threshold k_0 for the storage sum. An analog comparator converts a sequence of analog video pulses into a sequence of binary signals "1" and "0", accumulated in the drive. The digital threshold k_0 is the detection threshold; in the case when the total signal of the drive exceeds this threshold, a detection signal "**1**" is issued. If during the reception of a packet of n pulses the threshold k_0 is not exceeded, then a "**0**" signal is issued – no detection, that is, the absence of an information signal at the input of the receiving devices.

The probabilities of correct detection P_{po} and false alarm P_{lt}, with a fixed burst length n, are determined by the expressions

Вероятности правильного обнаружения $P_{по}$ и ложной тревоги $P_{лт}$, при фиксированной длине пачки n, определяются выражениями

$$\begin{cases} P_{\text{po}} = \sum_{k=k_0+1}^{n} C_n^k P_s^k (1 - P_c s)^{n-k}, \\ P_{\text{lt}} = \sum_{k=k_0+1}^{n} C_n^k P_i^k (1 - P_i)^{n-k}, \end{cases} \qquad (1)$$

where P_s and P_i are the probabilities of generating a "1" signal at the output of the analog comparator, that is, the probability of exceeding the analog threshold U_0 in the presence of an additive mixture of information signal and interference or only interference at the input of the analog comparator (detector output), respectively.

We are interested in the lower expression of formula (1), which determines the probability P_{lt} the probability of operation of an analog comparator (formation of a binary signal "1") in the presence of only noise at the input of the comparator. This probability is given by the expression

$$P_i = \int_{U_0}^{\infty} f_i(U) \cdot dU, \qquad (2)$$

where $f_i(U)$ is the noise envelope distribution density.

By interference we mean reflections from the underlying surface of the earth or sea and the inherent noise of the receiving devices of automation systems. The envelope of reflections from the earth's surface and its own noise are described by the Rayleigh distribution

$$f_R(x) = \frac{x}{\sigma_R^2} \cdot \exp\left(-\frac{x^2}{2\sigma_R^2}\right), \; x > 0. \qquad (3)$$

where $\sigma_R^2 = \tilde{P}_R$ is the average interference power. The envelope of reflections from the sea surface is described by a log-normal distribution

$$f_L(x) = \frac{1}{\sqrt{2\pi}\sigma_L x} \cdot \exp\left(-\frac{\left(\ln(x) - \ln(\overline{A}_L)\right)^2}{2\sigma_L^2}\right), \; x > 0. \qquad (4)$$

where $\overline{A}_L$ and σ_L are the distribution parameters through which the average power of radio interference pulses is expressed as $\tilde{P}_L = 0.5\overline{A}_L^2 \exp(2\sigma_L^2)$ [8].

When observing input signals in receiving devices, automatic gain control systems are always used to expand the dynamic range, maintaining the power of the input signal due to reflections from underlying surfaces and self-noise at a certain fixed level. Due to this, the powers of the input signals from the underlying surfaces and our own can be considered the same. In our case, this leads to the equality of the second initial moments of the distributions defined by expressions (1) and (2), that is, to the equality

$$\int_{0}^{\infty} x^2 f_R(x) dx = \int_{0}^{\infty} x^2 f_L(x) dx. \qquad (5)$$

In equality (5) it is necessary to take into account the following feature. The Rayleigh distribution, represented by expression (3), is one-parameter, and the log-normal distribution, represented by expression (4), contains two distribution parameters. Therefore, the right side of expression (5) depends on only one parameter σ_R^2, and the left side of expression (5) depends on two parameters $\overline{A}_L$ and σ_L^2. In this case, equality (5) can be achieved for different ratios of these two parameters. It is known from empirical observations that the underlying surfaces have almost identical coefficients of variation, which makes it possible to select parameters $\overline{A}_L$ and σ_L^2 in such a way that not only the second initial moments of distributions (3) and (4) are equal, but that their average values and variances are equal.

As follows from expression (1), stabilization of false alarms of the P_{lt} is achieved due to stabilization – the probability of generating a "1" signal at the output of the analog comparator in the absence of an information signal at its input. It follows that the analog comparator thresholds for Rayleigh noise U_{0R} and log-normal noise U_{0L} must be such that ensure equality of P_i.

Calculating the corresponding dependences $\varphi_R(P_i)$ and $\varphi_L(P_i)$ of the thresholds U_{0R} and U_{0L} on P_i, we get

$$\begin{cases} U_{0R} = \varphi_R\left(P_i\right) = \sigma_R\sqrt{-2\ln P_i}, \\ U_{0L} = \varphi_L\left(P_i\right) = \sqrt{2}\sigma_R\exp\left(\sqrt{\dfrac{4-\pi}{2}}\cdot\left(\Phi^{-1}\left(1-P_i\right)-\sqrt{\dfrac{4-\pi}{2}}\right)\right), \end{cases} \tag{6}$$

where $\Phi^{-1}(.)$ is the inverse function of the Laplace integral. When calculating expression (6), we additionally used the condition of equality of mathematical expectations and variances of the Rayleigh and log-normal probability distribution densities. The dependence graphs $U_{0R}=\varphi_R(P_i)$ and $U_{0L}=\varphi_L(P_i)$ are shown in Fig. 2.
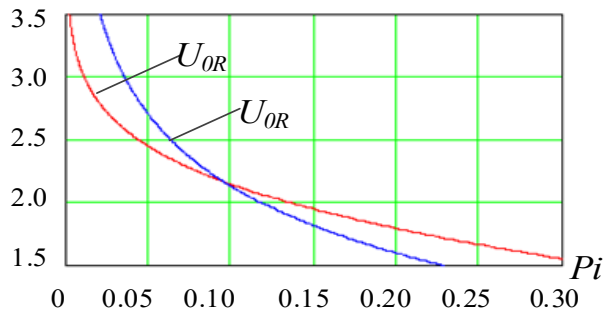
Fig. 2. Dependence of analog comparator thresholds U_{0R} and U_{0L} on the value of P_i.

The curves intersect at $P_i = \bar{P}_i$, which is the root of the equation

$$\exp\left(2\sigma_L\left(\Phi^{-1}\left(1-P_i\right)-\sigma_L\right)\right)+\ln P_i = 0,, \tag{7}$$

the numerical solution of which gives the value $\bar{P}_i = 0.096$, and the threshold is equal to $\bar{U}_{0R} = \bar{U}_{0L} = \bar{U}_0 = 2.166$. The values of the thresholds U_{0R} and U_{0L} are expressed in units σ_R, which must be taken into account in practical calculations; here we set it σ_R equal to unity.

Moreover, if we introduce a generalized mathematical model of reflections from land and sea, using the Huber model [9] in the form

$$f_{RL}\left(x\right) = \gamma f_R\left(x\right)+\left(1-\gamma\right)f_L\left(x\right),, \tag{8}$$

where γ is a certain coefficient $0 \le \gamma \le 1$, in the Huber model is the clogging coefficient, but in this case it is simply a parameter of the generalized distribution, which at $\gamma=1$ goes into the Rayleigh distribution, and at $\gamma=0$ goes over into the lognormal distribution. This distribution is also convenient in that sometimes interference from the sea surface is described using the Rayleigh distribution, and fluctuations in interference from the earth's surface are described by a lognormal distribution [10]. The obtained conclusions regarding the stabilization of false positives by detectors apply to the entire presented family of distributions, of which the Rayleigh distribution and the lognormal distribution are particular extreme cases.

Thus, when choosing the digital threshold of the analog comparator equal $\bar{U}_0 = 2.166\sigma_R$, we ensure, in the presence of an automatic gain control system, that the probability of false alarms of $P_{лт}$ detectors is independent of the type of underlying surfaces – from the underlying surface of the earth and the sea surface. This, of course, ensures support for the $P_{лт}$ value in the absence of interference signals caused by reflections of location signals from underlying surfaces, that is, the probability of false alarms of the $P_{лт}$ in the presence of only the inherent noise of the receiving devices.

231

References

1. Izrantsev V. V., Shepeta D. A. Simulation of External Signals on-Board Instrument Sys-tems for Aircraft of the Fifth Generation. Izvestiia vuzov. Priborostroenie, 2000, no. 2, pp. 14-19 (In Russian).

2. Shepeta D.A. Development of mathematical models and synthesis of algorithms for model-ing input signals of on-board information processing and control systems. Dissertation for the degree of Candidate of Technical Sciences / Saint Petersburg, 2000.

3. Sesin A.E., Shepeta D.A. Mathematical model of echo signals of the sea surface ob-served by on-board locators of aircraft // Information and control systems. 2010. No. 2. pp. 21-25.

4. Isakov V. I., Shepeta D. A. Simulation of Radar Signals Reflected from the Edge of Land and Sea // Information and management systems. 2017. № 5 (90), pp. 89–94.

5. Isakov V.I., Shepeta D.A. Power distribution density of the envelope of location signals re-flected from the land-sea edge // In the book: Processing, transmission and protection of information in computer systems '21. International scientific conference: collection of reports. Saint Petersburg, 2021. pp. 25–28.

6. Blaunstein N. S., Sergeev M. B., Shepeta A. P. Prikladnye aspekty elektrodinamiki [Ap-plied Aspects of Electrodynamics]. Saint Petersburg, Agraf+ Publ., 2016. 272 p. (In Russian).

7. Modern radar (analysis, calculation and design of systems). Per. from English edited by Kobzareva Yu.V. Publishing house "Soviet Radio", 1969.

8. Shepeta A.P. Synthesis of nonlinear shaping filters for modeling input signals of loca-tion systems // Proceedings of the international scientific and technical conference (reports), May 1994. – Ki-ev: Academy of Sciences of Ukraine, NPO Kvant. – Issue 1, pp.81-85.

9. Huber J. P. Robustness in statistics / Transl. from English-M.: Mir, 1984.- 304 p.

10. Kulemin G. P. Radar Clutter from Land and Sea Rdar in the Centimeter and Millimeter. Trudy Mezhdunarodnoi nauchno-tekhnicheskoi konferentsii (doklady) [Proc. of Intern. Scien-tific-Technical Conf. (reports)]. Kiev, AN Ukrainy, NPO Kvant Publ., 1994, vol. 1, pp. 23–29 (In Russian).

RESEARCH OF NEURAL NETWORK METHODS
FOR CLASSIFICATION OF THE EARTH'S SURFACE

Alexander Zalishchuk

Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
E-mail: sacha1501@yandex.ru

Abstract: *The article is devoted to the study of the use of machine learning methods for classifying the earth's surface based on satellite images as earth remote sensing data. In this work, some of the most popular and accurate neural network models were selected for consideration. The process of data preparation, feature selection, and model tuning is described. The results of the study show high classification accuracy when using the considered methods. The resulting technique can form the basis of a comprehensive system for high-precision analysis of remote sensing data not only from satellite images, but also from all optical sources.*

Keywords: *Machine learning, convolutional networks, image classification, remote sensing.*

Introduction

Classification of satellite images from a satellite or from an unmanned aerial vehicle is one of the key tasks in the field of processing and analysis of remote sensing data. As the number of remote sensing images available increases, the need for automatic processing and analysis of these data increases [1].

The purpose of this research is to develop an automatic algorithm for processing remote sensing images. To do this, it is proposed to solve the problem of classifying sectors in satellite images into various classes.

Formulation of the problem

To extract meaningful information from huge volumes of earth remote sensing data, it is necessary to divide large data into small parts and only then perform detailed processing of satellite images for the resulting parts or sectors. However, manual processing of large volumes of data is time-consuming, subjective, and often impractical for large-scale analysis. Therefore, there is a need to develop an automated solution that can quickly and accurately process images based on their content [2].

This research addresses the challenge of creating an efficient and robust image classification model that can cope with the complexities and variations present in remote sensing imagery.

In solving this problem, emphasis should be placed on the best solution for analyzing earth remote sensing images and ensuring timely decision making in various fields [3].
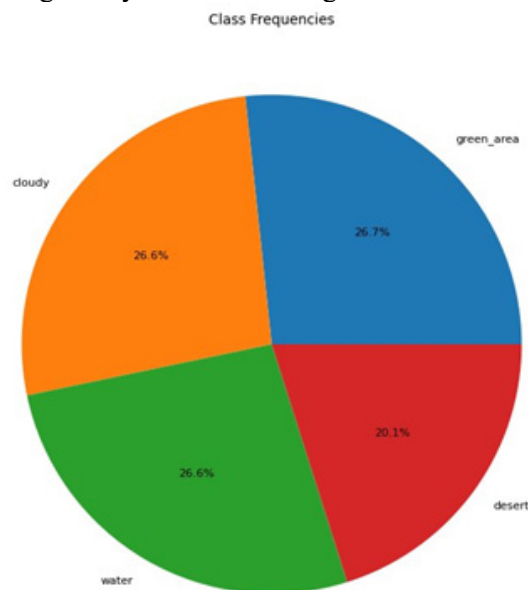
Fig. 1. Graphic description of data in the form of a chart

The selected dataset provides a valuable resource for training, validating and testing machine learning models for satellite image classification tasks in solving earth remote sensing problems.

Description of the dataset

The RSI-CB256 satellite image classification input parameters used in the study [4] represent an extensive collection of satellite images specifically dedicated to the analysis of remote sensing images. This dataset contains imagery from a variety of sources, including sensors and Google Maps snapshots, providing a rich representation of different landscapes and scenarios. In addition, the dataset includes four separate classes (Fig. 1), each of which corresponds to a unique category of satellite imagery.

Data processing

Before we start training our models, we need to examine the selected dataset. To do this, the number of images in each class is displayed using a bar chart (Fig. 2).
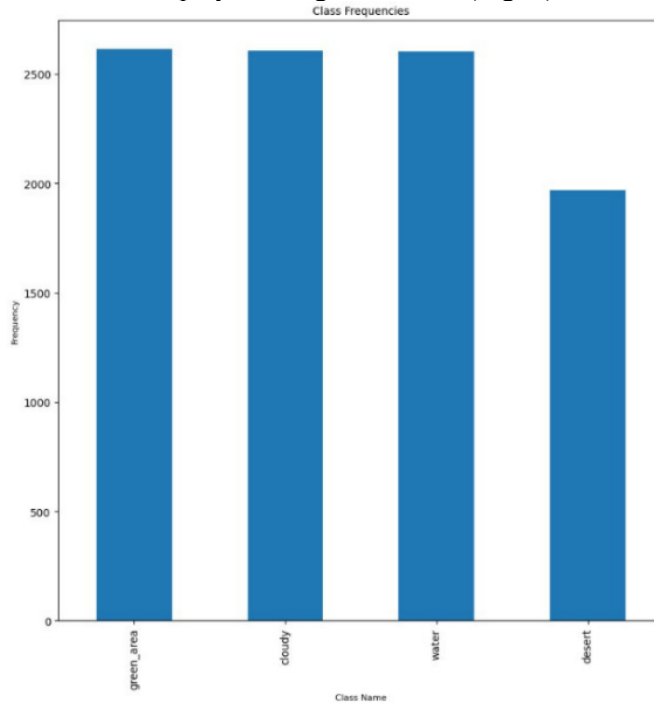
Fig. 2. Graphical determination of the number of images in each class

To get an idea of the data, 4 random images are selected in each class. This will help you visualize the content of the images and better understand what data they contain (Fig. 3).
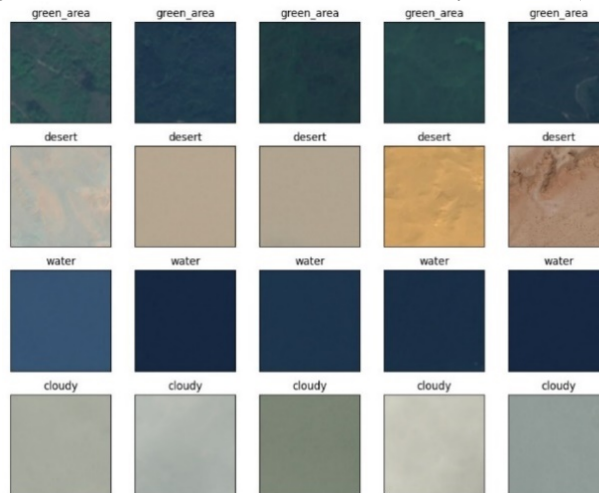
Fig. 3. Selecting a random sample of each class for clarity of data presentation

Having examined the sample, we can conclude that the data does indeed contain remote sensing images of the earth.

After this, the entire data set is mixed and divided into 2 samples: training and test [5].

To expand the training set and bring the model "retraining" to different conditions for obtaining images, the "ImageDataGenerator" class is used with modification methods, such as: rotation by a certain number of degrees, shift, scaling and offset. While the test sample with the "ImageDataGenerator" class uses only scaling to avoid loss of classification accuracy when checking models during "retraining" [6].

Results

In the process of "retraining" the models, those that showed the best results were selected. The classification efficiency of each model exceeds 90 percent. The resulting loss and accuracy achieved by each model are presented in Table 1.

Table 1

Comparative table of the results of assessing the accuracy of retrained models on the test sample

Model name	Minimal loss of accuracy	Maximum model accuracy
ResNet	1.037233e-07	1.000000
VGG19	8.473640e-07	1.000000
MobileNet	1.545546e-02	0.995438
InceptionV3	1.602876e-02	0.993998
CNN	1.958975e-01	0.927491

ResNet and VGG19 models show the best results, achieving more than 99% accuracy in satellite image classification. These models successfully extract relevant features from images, resulting in accurate predictions. The InceptionV3 and MobileNetV2 models also perform well with accuracy close to 100%. Of all of them, only the classical convolutional neural network (CNN) is inferior in accuracy, achieving approximately 92% accuracy.

Conclusion

The high accuracy of the retrained ResNet and VGG19 models demonstrates their effectiveness in classifying satellite images. These models use pre-trained architectures and have rich feature extraction capabilities, allowing them to capture important patterns and structures in images. InceptionV3 and MobileNetV2 also perform well, but with slightly lower accuracy than ResNet and VGG19, making them suitable choices for this task.

Only the classical convolutional neural network (CNN) with the recreated architecture shows lower accuracy, indicating limitations in its ability to capture complex patterns and features compared to models that have pre-trained weights.

The results highlight the importance of using pre-trained models and knowledge transfer in satellite image classification. By leveraging knowledge gained from large sampled images, pre-trained models can effectively extract meaningful features from satellite images, improving classification accuracy. The higher accuracy achieved by the ResNet and VGG19 models indicates that they learned generalized features that are most relevant to the task at hand. However, specific requirements and constraints must be taken into account when choosing a model, since the "ResNet" and "VGG19" models may require higher computational resources due to their deep architecture. In contrast, the convolutional model, although less accurate, may be more suitable for resource-constrained environments. But it is worth considering that CNN is still inferior in accuracy to models "heavier in computational weight," which indicates potential limitations in its ability to identify complex patterns.

In conclusion, satellite image classification research successfully develops and evaluates machine learning models for satellite image classification.

The results of the study have significant implications for remote sensing image analysis and for various fields such as environmental monitoring, land cover mapping, urban planning and disaster response [10].

Accurate classification of satellite images allows you to make informed decisions and allocate resources. The use of pre-trained models and knowledge transfer provides a valuable approach to solving problems of classification of remote sensing images [11].

References

1. Ageev S.V., Mikhailov I.S., Shiryaeva G.N. Analysis of machine learning methods for classifying the earth's surface using satellite images. Geoinformatics, 2016, pp. 30–49.

2. Bakunov D. V., Plyushkin S. A., Solovyov D. A. A machine learning model for classifying the earth's surface based on satellite images. Izv. RAS, 2017. T. 72. No. 3. P. 425–438.

3. Vinogradov S. A., Kuznetsov D. V., Perelygin V. P. Assessing the quality of machine learning models for satellite image analysis problems. Izv. RAS, 2017. T. 72. No. 3. P. 439–452.

4. Li, H.; Tao, C.; Wu, Z.; Chen, J.; Gong, J.; Deng, M. RSI-CB: Large Scale Remote. Sens. Image Classif. Benchmark Via Crowdsource Data. ArXiv 2017, arXiv:1705.10450.

5. «Assessing the potential of machine learning algorithms for land cover classification using remote sensing data» by A. El-Sayed et al., Remote Sensing, vol. 11, no. 9, pp. 1153–1176, Sep. 2019.

6. "A Comprehensive Review of Machine Learning Techniques for Land Cover Classification" by S. Alzahabi, A. Alzahabi, and A. AlZahabi, 2020.

7. "Machine Learning for Land Cover Classification Using Remote Sensing Data: A Review" by S. K. Shrivastava, S. K. Shukla, and S. K. Mishra, 2017.

8. D. D. Dandin, S. K. Shrivastava, and S. K. Shukla. «A comparative study of land cover classification techniques in remote sensing: A review». Journal of Remote Sensing, 42(3): 1450–1476, 2018.

9. J. Gong, X. Liu, and X. Zhang. «A review of land cover classification methods based on remote sensing images using machine learning algorithms». Journal of Remote Sensing, 42(3): 1350–1373, 2018.

10. "A Survey of Methods for Land Cover Classification Using Remote Sensing Data" by A. El-Sayed and M. El-Bakry, 2018.

11. «Accurate land cover classification using deep learning and its application in remote sensing» by H. Zhang et al., Remote Sensing, vol. 11, no. 8, pp. 1075–1092, Aug. 2019.

STUDY AND ANALYSIS OF LOAD FORECASTING METHODS IN THE ELECTRIC NETWORK WITH DISTRIBUTED ENERGY RESOURCES

Roman Zenin

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: romazenin@gmail.com

Abstract. *In the process of design and operation of urban power grids there arises, due to technological and economic reasons, the problem of forecasting electrical loads for a certain time period. To date, the forecasting methods used are limited by the accuracy of forecasts of network parameters. To increase the accuracy and improve the process of predicting the parameters of electrical systems, today machine learning, neural networks and parameter estimation models in terms of statistics and probability theory are actively used. In this research paper, the types of load forecasting methods in power grids based on neural networks will be discussed and their advantages and disadvantages will be discussed.*

The need for load forecasting in power systems

It is known that the use of high-power electrical equipment and pulse sources has a detrimental effect on the public power grid in terms of the quality of electrical energy. Starting, stopping and operation of such devices distorts the basic electrical characteristics of the network, subsequently affecting other less powerful and more energy-sensitive appliances. Today, the most active electrification is taking place, more and more electrical equipment is being used. The requirements for power quality indicators remain unchanged, while the parameters themselves are deteriorating due to the growth of electrical appliances that actively affect the network indicators by their work, polluting and distorting it. This is why compliance with the established parameters of the power grid is one of the most important tasks in the electric power industry today. In order to avoid overloading of the circuit due to unstable voltage values and other parameters, forecasting tools are used to build a strategy for supplying energy to consumers. Based on the obtained forecast, the energy supplier company can adjust the power generation strategy. In order to obtain a more accurate forecast of power consumption, there is a need to form a model describing the changes in the system parameters considering the dependencies of changes in power consumption. The forecasting task means predicting the next values considering the previous values l of some time series x(t):

$$x(l + d) = q(x(t)...,(x - l),(\eta_1, \eta_2 ..., \eta_q),$$

where d – the prediction step, q – the number of independent variables, η – the independent variables [1].

Forecasting can be done in various ways, which can be divided into two main groups: qualitative and quantitative. Qualitative methods use a mathematical apparatus based on expert estimates of electricity consumption. It includes a generalized prediction method and two heuristic methods. On the other hand, quantitative methods utilize time sequence analysis, statistical analysis, Bayesian probability based forecasting and neural networks. The selection of an appropriate forecasting method should consider various factors such as the required form of forecast, forecasting time interval, availability of statistical data, accuracy and behavior of the system to be forecasted. Depending on the forecasting time interval, forecasts can be operational, within the current day, short-term, up to ten days, medium-term, up to several months, long-term, up to five years and forward-looking for several years [2]. Traditional forecasting methods are based on the analysis of stationary random processes using the correlation function.

Analytical forecasting method

Statistical methods are based on analyzing historical load data and extrapolating the data to a future period. Examples of such methods are regression analysis and exponential smoothing. Statistical methods are based on analyzing past load data. The least squares method is used to find the best straight line that approximates past data, the analytical expression is given by the following formula:

$$i(x.t) = \sum_{r=1}^{R} \eta_r x_r (q, t),$$

where: $x_r(q,t)$ – expressions forming the basis of the formula of the forecasting algorithm; η_r – additional load correction factor depending on the season.

Statistical methods have a number of advantages, such as simplicity in use; no requirement for a large amount of data, the ability to adapt to different time series and consider seasonal fluctuations of the power grid. The disadvantages include low accuracy, especially with non-linear data, and lack of consideration of individual external factors affecting the data.

Statistical analysis

The statistical analysis method involves collecting data about the system and constructing an algorithm that calculates a moving average of the predicted load. This type of analysis is often used in working with neural network and fuzzy neural network forecasting methods. This method allows you to work with new data without resorting to a complete recalculation of the model, but most often anomalous values beyond the value of statistics will be ignored by it. The main advantages of this method include: consideration of many factors affecting the load, and high accuracy of the forecast. The disadvantages are: the need for a large volume to build dependencies and probabilities of occurrence of certain events, the presence of a clear mathematical description of the process and the model of the system of occurrence of events to obtain accurate forecast results.

Regression analysis

The regression method is most often used mainly with tasks related to the construction of a mathematical model of the object and its identification. In order to obtain a forecast, it is required to build a model of the object behavior and identify external factors influencing the system. Due to the fact that this method considers many parameters affecting the system, this method is suitable for solving the problem of predicting the output value of voltage in the power grid. The advantages of regression analysis are: simple implementation both in mathematical and software form, construction of the result of predicted values on the basis of previous values and consideration of external factors affecting the system. The disadvantages of the method include: the dependence of the result on the hypothesis on the basis of which the model of object behavior is built, in case of choosing an unrealistic scenario, the results obtained after forecasting will be unsuitable for working with them, the need to select the type of linearity of the system, as well as the calculation of independent variables affecting the behavior of the system, as well as the need for a large number of interdependent calculations to obtain forecasting data.

Forecasting methods using neural networks

Neural networks refer to a computational structure, a model. For ease of understanding, the principle of networks is associated with the work of a human neuron. A neuron consists of three types of elements adder, multipliers and nonlinear converter. The structure of the above networks consists of trained parallel systems by analyzing positive and negative results. Training of neural networks, according to its principle, is divided into two types of training with a teacher and without a teacher. The list of tasks solved by neural networks in modern power systems includes such tasks as load prediction, control of power flows in networks, diagnostics of systems in order to determine faults, management of power system operation modes and many other tasks.

Learning with a teacher implies training a model on pre-labeled data. This data implies input data and corresponding target responses. The training data and the training algorithm are prepared before training the neural network. Unsupervised learning is a type of training a model on an unlabeled data set. The essence of this approach is to allow the neural network to identify patterns, structures or relationships without prior information about the results. This feature automates the data partitioning process and speeds up the neural network development process. More training iterations are required to obtain high accuracy. The time spent on the process of training the model is more than that of training with a teacher, but reduces the human workload.

Examples of using the neural network method in grid load forecasting

In the regression method using neural network, the regression equation describes the growth of daily energy consumption on the time interval from 0:00 to 24:00 of one day [2]. The learning model, however, is learning with a teacher, the structure of the neuron is the simplest with two inputs and one output is shown in Fig. 1.
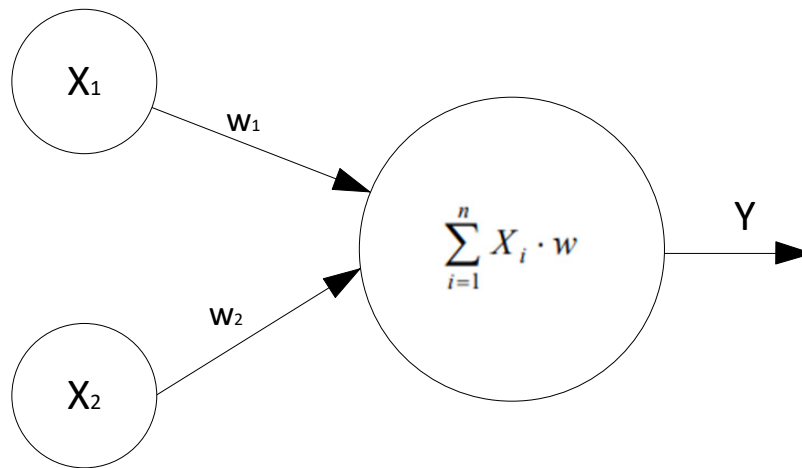
Fig. 1. Schematic of the neuron used in the study

The accuracy of the described prediction is 96.94%, but it should be understood that to achieve such a result we need ready-made, marked-up data on the basis of which the network will be trained. In case of changing the step of forecasting depth, it will be necessary to completely redesign the whole sample of training data and retrain the model. For a particular case the use of such a forecasting method is redundant, as more traditional forecasting methods with less waste of resources could cope with such a task.

Multilayer neural networks are a special case of neural network realization using several neurons forming a layered structure [3]. A multilayer neural network consists of an input layer, which is the first layer of neurons in the network and takes as input the data to be processed, a hidden-intermediate layer between the input and output layers, for data processing, and an output layer responsible for the generation of output values of the entire network, the number of neurons in the output layer depends on the task condition. The number of neuron cells in the first layer is equal to the number of characteristics or dimensionality of the input data. Each artificial element in this area takes as input one value from the input data. The neurons of the output layer are connected with the neurons of the hidden layer, the results are generated and calculated using the activation function, and then interpreted into the final result.

This example considers energy consumption forecasting using a multilayer network consisting of three layers. The main information for forecast generation is average data on network loads, as well as indirect data affecting the load. Due to its structure, namely the use of a large number of neurons, the network is able to process and analyze a huge amount of data, allowing them to perform complex tasks such as decoding and image recognition, etc. Due to the ability to use different combinations of neuron species in the hidden layer, the network is able to recognize many more connections and patterns in the data than conventional single neuron designs. Due to the layered design of the network, adapted systems for other tasks do not require a complete redesign of the network, but only a change in one of the layers.

Due to the complex and voluminous structure of the model, a large amount of training data is required to obtain accurate modeling results, otherwise one may encounter undertraining of the model or overtraining. The process of training and using a multilayer neural network involves a large amount of computational resources. This can be an obstacle when working with large amounts of data or on equipment with limited computational capabilities. Due to the complex structure of the neural network, the algorithm to compute the results may be sensitive to anomalous data, which may affect the accuracy of the method.

Fuzzy neural networks (NNF) are a hybrid approach that combines the advantages of fuzzy logic and artificial neural networks. They are applied in various fields, including the electric power industry, to solve prediction and control problems. This method is used when it is impossible to create a mathematical model of the process using standard methods. For realization of neuro-fuzzy forecasting the interaction of fuzzy algorithms and neural network approach is used, and sometimes additional description of the logic of the investigated process is used. Fuzzy neural networks are able to produce smaller error in forecasting, due to smooth approximation of threshold functions by fuzzy way of data representation in the network.

This approach is based on a system of fuzzy inferences, elements of fuzzy sets, rules and reasoning that form fuzzy inferences [3]. The technical realization of such an approach is provided by the interaction of two types of data for the communication of which a fuzzy system is introduced. It has a phasifier, which converts the input data set into a fuzzy set at the input, and a dephasifier, which converts the

fuzzy sets into a specific value of the output variable at the output. The phasificator converts a crisp set of input data into a fuzzy set defined through a membership function. The defuzzifier performs the inverse task, creating a definite solution with respect to the input variable based on many fuzzy outputs that the executive module of the fuzzy system creates. The output of the output module can be represented as M fuzzy sets, which define the measurement range of the output variable. The defuzzifier then converts the range into a specific value which is taken as the output of the whole system [4]. The most commonly used membership functions for phasification are Gaussian type, triangular and trapezoidal functions.

The features of this method include: the possibility of obtaining a model of a complex system in the form of fuzzy rules, and because of the similarity of fuzzy systems, it is possible to use the experience of previous systems and experts in the creation of new ones, NNFs are able to adapt to changing conditions and non-stationary data, which makes them useful for predicting electricity consumption and other variables in the power system, NNFs are based on fuzzy logic, which provides a clearer and more transparent view of the prediction process, NNFs have a high level of predictive power, NNFs are able to use the data from the model to predict electricity consumption.

At the same time in this methodology there is no model training and it is not possible to use model data to modify the model. Model building is due to the combination of fuzzy mathematics and logic, which leads to the limitation of the model structure and input data, and also the model will have to be rebuilt completely when the problem changes, as it is not capable of adaptation.

Conclusion

Load forecasting in an electrical network with distributed energy resources is an important task that requires the application of effective analysis and forecasting methods. The methods considered in this paper, such as linear regression, statistical method, neural network method and fuzzy neural networks, allow obtaining sufficiently accurate results and can be used depending on the specifics of the problem and available data.

However, it should be noted that none of the considered methods is universal and ideal for all cases. Therefore, in order to obtain the most accurate and reliable results, it is recommended to combine different methods and approaches, as well as to evaluate their effectiveness based on different metrics and criteria.

Thus, to solve the problem of load forecasting in power grids with distributed generation, it is necessary to develop a flexible forecasting system that will be able to adapt to different conditions and consider multiple factors affecting the load.

References

1. Smirnov V.V., Chernova M.A., Starchenko M.A., Maksimenko T.S. Models of forecasting of electric loads. Energeticheskie installatsii i tekhnologii 2018. C. 103-109. Date of circulation (20.12.23).

2. Ageev V.A., Rep'ev D.S., Kazakov D.V. Review of traditional and neural network methods of electric load forecasting. Technical Sciences 2023. C. 1-5. Date of circulation (13.01.24).

3. Alferova T.V., Trokhova T.A.. Computer forecasting of electric loads by methods of neural networks. Agrotechnics and energy supply 2019. C. 166-173 Date of circulation (20.01.24).

4. Sakhno E.P., Dyachenko R.A., Reshetnyak M.G., Kapustin K.Y. Toward the issue of short-term forecasting of electric loads using fuzzy neural steys. Modern problems of science and education 2013. C. 162. Date of circulation (12.02.24).

5. Ovsyannikov N.B., Pichuev A.V. Methods of forecasting the electric load at mining enterprises. Mining information and analytical bulletin 2015. C. 292-296. Date of circulation. (14.02.24).

SIMULATION OF A CORRELATION NAVIGATION SYSTEM FOR CONTROLLING THE TRAJECTORY OF AN AIR OBJECT WHEN DEVIATING FROM THE ROUTE

Alexandr Zhmurin

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: qqwwww2006@gmail.com

Abstract. *The simulation of a navigation system in the Matlab environment based on the principle of correlation of two trajectories is considered. The model simulates an aerial object (VO) with a vertically mounted altimeter and a pre-loaded reference altitude map of the area being flown with a known flight route.*

Keywords: *Matlab, height matrix, sampling of values, calculation of correlation coefficients, algorithm, index search*

Introduction

Correlation navigation systems are a promising tool for flight control and orientation on the ground. This scientific article presents a simulation of a correlation navigation system in a development environment Matlab.

Creating an application

We will simulate the system in the Matlab App Design package, which allows you to create complete applications and presents all the necessary tools for this. An illustration of the appearance of the application and its operation is shown in figures 1 and 2.

The application works according to the following principle. There is a reference set of heights of the area being flown in the form of a matrix and a flight route of this area in the form of a set of indices of this matrix. A two-dimensional array simulating a hilly terrain map was taken as a reference matrix [1].

The distance between the route points, as well as the distance between the cells of the reference matrix, depends on the speed of the shooting device. For simplicity of construction, we assume that all points of the route were taken after the time Δt [2].

At the initial moment of time and after Δt, the obtained height value from the altimeter is compared with the reference height value from the matrix. If the values match, or if the received value does not exceed the acceptable limits determined by the deviation value, the application considers that the AO is moving correctly and compares the point of our location with the route point at a given time.
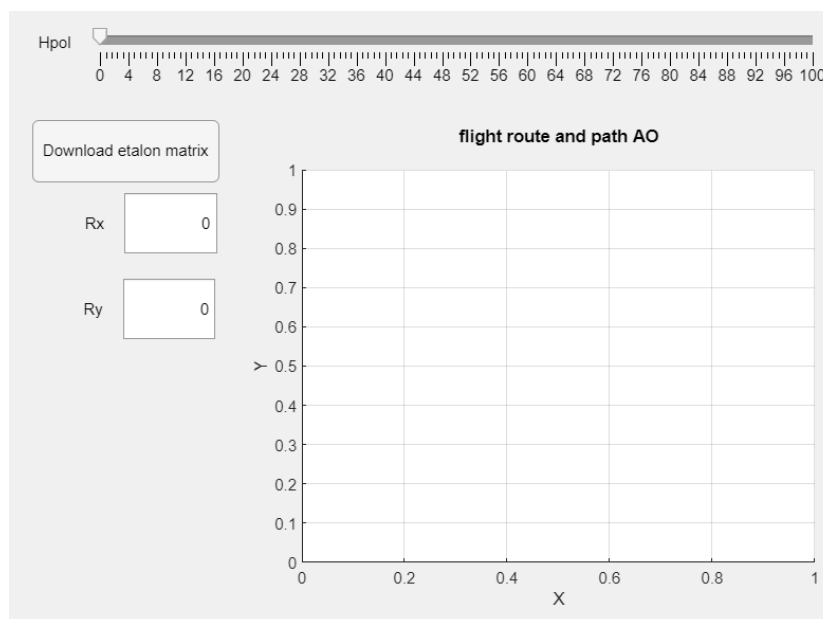
Fig. 1. Application appearance
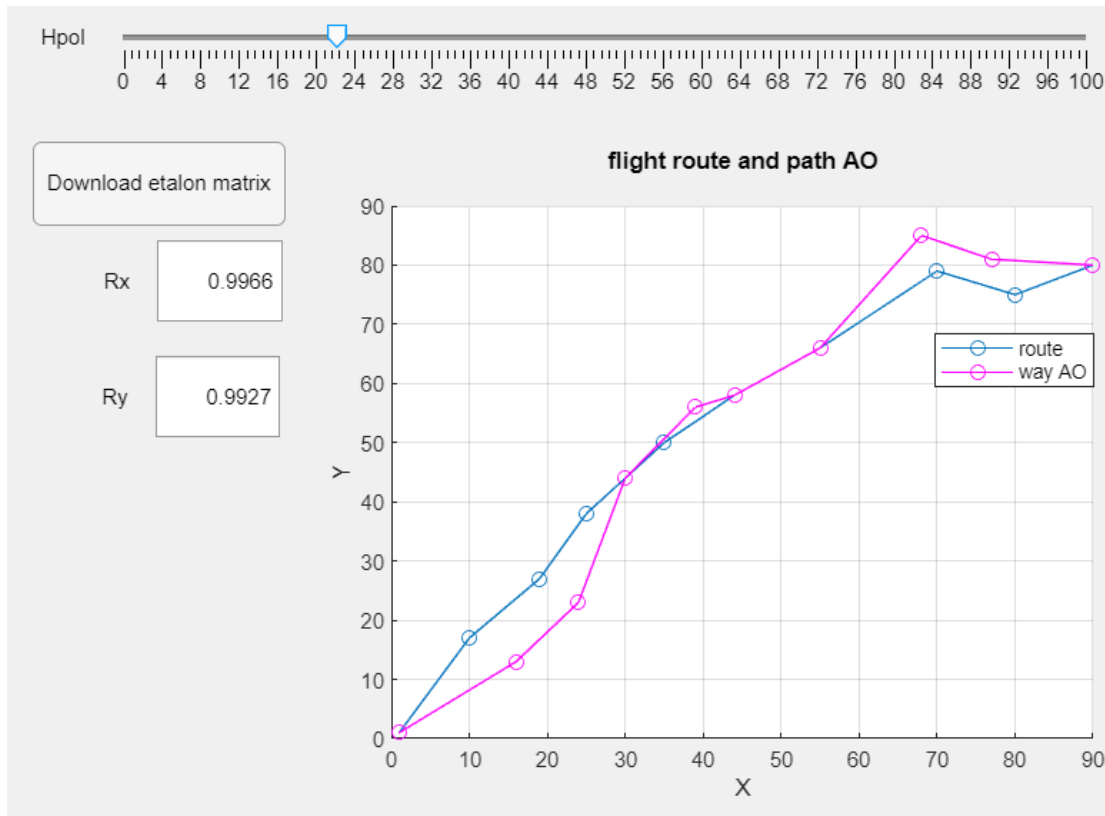
Fig. 2. An example of how the application works

If the obtained altitude value from the altimeter goes beyond the boundaries set by the error, that is, we deviated from the route, then an algorithm for determining the location of the AO is launched, during which a matrix with altitude values in the vicinity of the route point is formed, which can also be adjusted further depending on the correlation coefficients [3]. In this "cut-out" matrix, the height value obtained from the altimeter is searched, if the result is negative, the cell that is as close as possible to the obtained height value is taken for the current location.

The indices of this cell are translated into the indices of the reference matrix and the point of our location is placed on the graph. Also, based on the indices found, it is possible to correct the movement AO [4]. If there are more than three marks (points) of our flight, then the calculation of the correlation coefficients according to the formula (1) for the x and y coordinates will begin [5].

$$r = \frac{\sum_{i=1}^{n}(x_i - x_{av})(x_i^{'} - x_{av}^{'})}{\sqrt{\sum_{i=1}^{n}(x_i - x_{av})^2 * \sum_{i=1}^{n}(x_i^{'} - x_{av}^{'})^2}} \quad (1)$$

Where x_i is the number of the column (line) at the i-th moment of the route, x_{av} – the arithmetic mean of n numbers of columns (lines) of the route, $x_i^{'}$ – the number of the column (row) at the i-th moment of flight time AO, $x_{av}^{'}$ – the arithmetic mean of the n numbers of columns (rows) of the flight AO.

An illustration of the principle of operation of the algorithm is shown in figure 3.

By pressing the "Download etalon matrix" button, the reference matrix is loaded, variables, matrices, arrays are declared and initialized, and a flight path already pre-recorded in the code is also built upon pressing. The block diagram is shown in figure 4.

The "Hpol" slider simulates the altimeter readings taken after a time of Δt. Also, by moving the slider, a subroutine is called that sets the point of our location and calculates the correlation coefficients [5]. The block diagram is shown in figure 5.
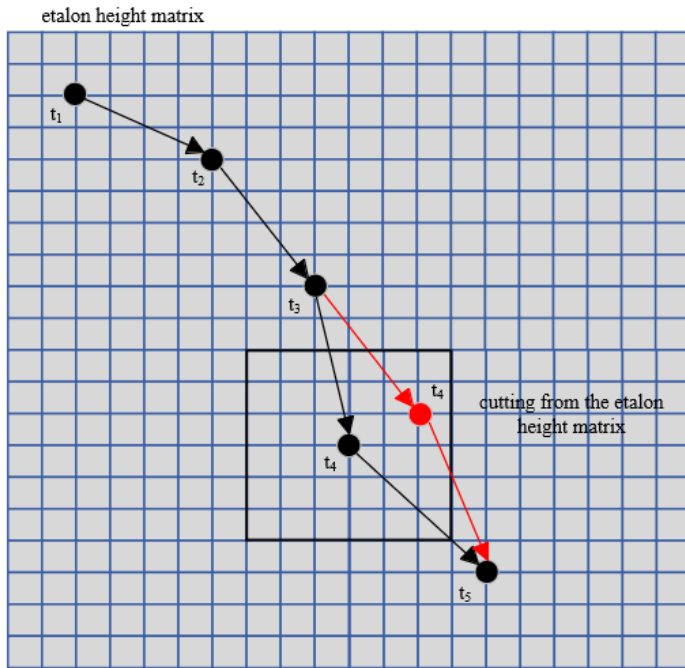
Fig. 3. An illustration of the principle
of operation of the algorithm
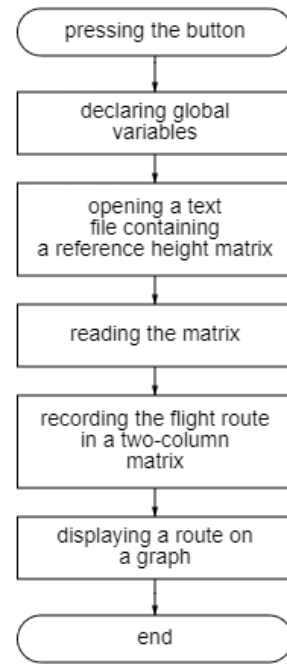
Fig. 4. Block diagram
of the button press processing routine
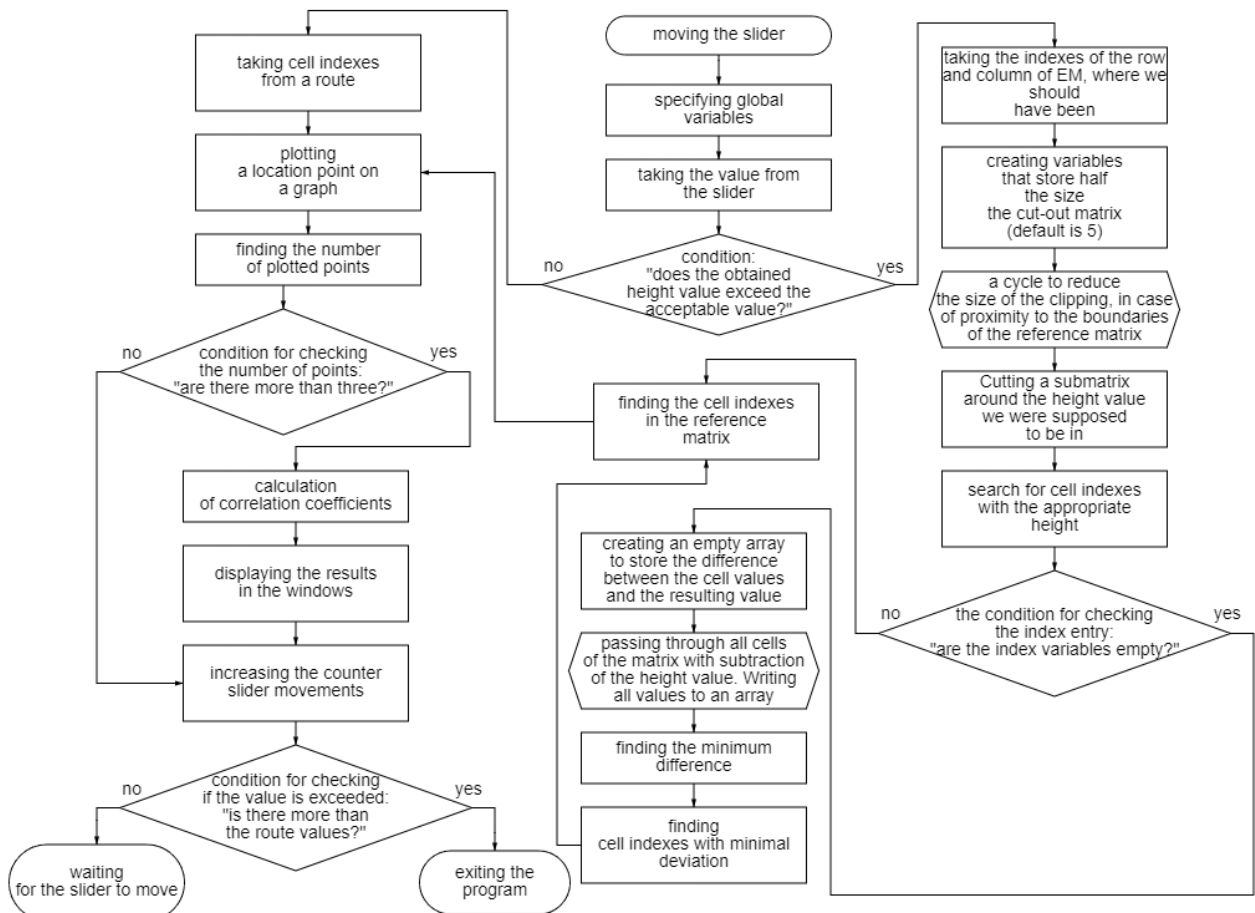
Fig. 5. Block diagram of the slider movement processing routine

Conclusion

This article discusses the simulation of a correlation navigation system in the Matlab environment. As a result, an application was created that simulates the operation of this system. Using such an application allows you to visually display the results of research in the field of autonomous navigation, as well as serve as a prototype in the creation of advanced programs for a similar purpose.

References

1. The method of forming the matrix of errors of the navigation field of radio engineering systems of short-range navigation and landing using GNSS, V. I. Baburov, N. V. Ivancevich, O. I. Sauta // Anniversary XXV Saint Petersburg International Conference on Integrated Navigation Systems. Collection of materials. Editor-in-Chief V. G. Peshehonov. – Saint Petersburg.: Publishing house: "Concern "Central Research Institute "Electropribor ", 2018.

2. Digital model of an on-board radio altimeter for a complex correlation and extreme navigation system, O. I. Sauta, E. P. Vinogradova. // Wave electronics and infocommunication systems. Proceedings of the XXVI International Scientific Conference. In 3 parts. Saint Petersburg, 2023. P. 339-349.

3. The limits of applicability of the mathematical model of the return of an air object to a given trajectory when deviating from the route, O. I. Sauta, E. P. Vinogradova, A. V. Zhmurin// III International Forum « Mathematical methods and models in high-tech production». Collection of abstracts and reports P.1 Saint Petersburg, 2023 P. 256-261.

4. The concept of creating an applied geographic information system for modeling search correlation and extreme autonomous navigation systems, A. I. Alchinov, I. N. Gorohovskiy // Management issues, 2022, № 1, 54–66; Control Sciences, 2022, no. 1, 43–54.

5. Metrological aspects of the construction of a complex correlation-extreme navigation system using pseudo-radar maps, O. I. Sauta, E.P. Vinogradova // Metrological support of innovative technologies. Collection of articles of the V International Forum, Saint Petersburg, 2023. P. 127-128.

CENTRALIZED SYSTEM FOR BUILDING DYNAMICALLY RESCHEDULED ROUTES OF ROBOTIC COMPLEXES

Dmitry Zyryanov

Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: zyrikdima@gmail.com

Abstract. *The navigation system is an integral part of the functioning of the mobile robotic complex. The success of the tasks performed will depend on the effectiveness of this system. The purpose of building an intelligent centralized navigation system is to collect more information for the operation of the complex, simplify the process of its processing and transformation, accelerate the process of forming and creating tasks for mobile robots. This system also provides for the creation of several security circuits to ensure fault tolerance. The advantages of the resulting system are considered, as well as the conditions for its effective use.*

In general, a mobile robotics complex includes: mobile robots, a navigation system, communication channels, and external security systems. In this case, the robot is the executive body. The navigation system converts incoming tasks for the robot to complete them. Communication channels are necessary for the exchange of information between all parts of the complex. Safety systems ensure proper trouble-free operation. Accordingly, a navigation system is necessary for the proper functioning of the robotic complex, which includes mobile platforms. These systems are created for certain robots and for specific working conditions, which is why there are a huge number of them.

Task statement: it is necessary to develop a system for the centralized construction of dynamically tunable routes for a production complex with a pre-known layout and with an unknown constantly updated environment of obstacles, people, other robots, loaders. At this stage, a simplified simulation model of the room is used to test the system (Fig. 1) [1].
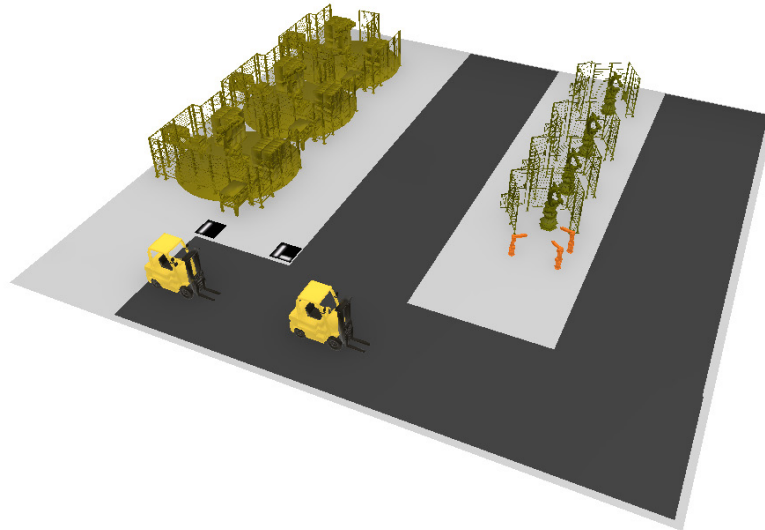
Fig. 1. Simplified simulation model

Navigation systems are divided into 3 types [2]:

- global: the mobile robot is able to determine its absolute coordinates; it is most often used over long distances;

- local: coordinates are calculated relative to a specific point, applied to an already known territory;

- Personal: Robot positioning is used to perform small localized tasks.

For a production facility, the choice of a local system prevails. The mobile robot will transmit its coordinates to the centralized control system at each time interval. This is necessary for proper positioning and route planning. Movement tasks will be issued in small portions, following the points of move-

ment. This is necessary to quickly change the movement of the mobile robot in case of rescheduling the final route.

The main problem is to identify dynamic obstacles and choose the right method to bypass them. In most robotic complexes (RTC), this task will be performed locally on mobile robots. However, the sensor field of view is not enough for a comprehensive analysis of the situation. To solve the described problem, the use of computer vision is proposed.

Cameras installed in the production room will allow you to get information immediately from a large area of the terrain, even outside the current range of robots. The analysis of objects will provide comprehensive information about the dimensions, speed and nature of the movement of a potential obstacle.

Object detection is performed by two parallel systems:

1) The neural network allows you to accurately determine the type of object and its approximate parameters: the speed, the nature of the movement, the potential purpose of the movement. However, image processing can take a long time, and it also requires significant computing power.

2) The system based on image subtraction allows you to accurately determine the discrepancy with the original and identify an obstacle. For the system to work, it is necessary to image at the initial moment of time with the absence of obstacles and unnecessary objects. The images are converted to a black and white spectrum. In the future, the color value of each pixel of the new and original image is subtracted. An image is obtained in which only objects that were not in the primary image are visible. For more accurate work, it is necessary to convert brightness, contrast and blur to compensate for changes in lighting and visibility in the production room. It is also necessary to filter out information by the boundary brightness coefficient of the pixel. This system can get the parameters of an object only when it is actually moving. The definition of an object in time occurs by combining the area of the object's location at the moment with the area of its location at the last moment. This method is not able to accurately determine the type of obstacle, but its reliability allows you to guarantee the reliability of the system as a whole.

At the next stage, it is necessary to make a perspective transformation to combine the found obstacle with the layout. To do this, it is necessary to obtain the homography of the image [3]. Image homography is the basic transformation of planes in a two-dimensional image. The transformation is carried out by obtaining a homography matrix connecting the positions of the points of the plane of the source image with the points of the plane of the final image. The minimum number of required points is 4 (Fig. 2).
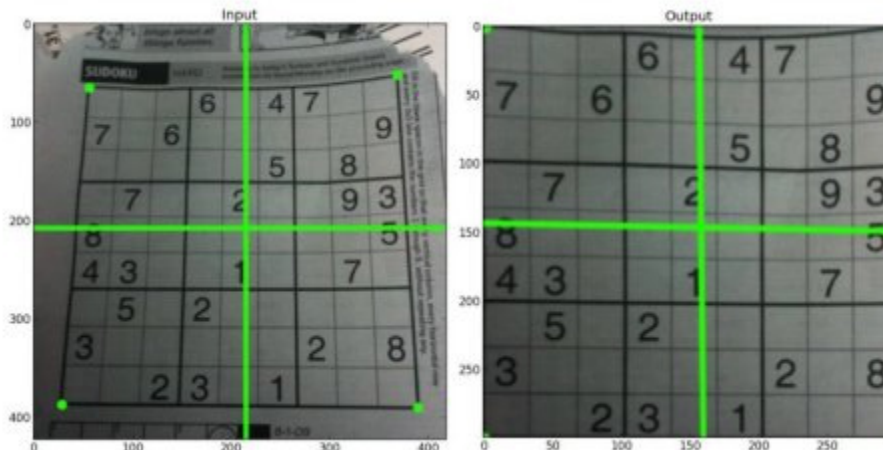
Fig. 2. Image homography

Obtaining these points is possible by defining markers. It is also possible to obtain homography using the already known homography matrix, but this approach implies complex configuration, low flexibility of changes, as well as labor-intensive implementation and modification of the system. Any uniquely identifiable object can be a marker, and there are 2 approaches to detecting them:

1) neural network: It is fast, insensitive to distortion, and uniquely identifies an object, but this method does not allow us to calculate homography.;

2) Highlight key points: allows you to calculate the tomography.

To simplify and speed up the process, a special type of markers is used: ArUco [4]. To combine, a simplified enterprise layout is used to plan routes with marked markers (Fig. 3). This layout is an image in a two–color format, where white is the area of movement allowed for mobile robots, black is a prohibited area, obstacles.
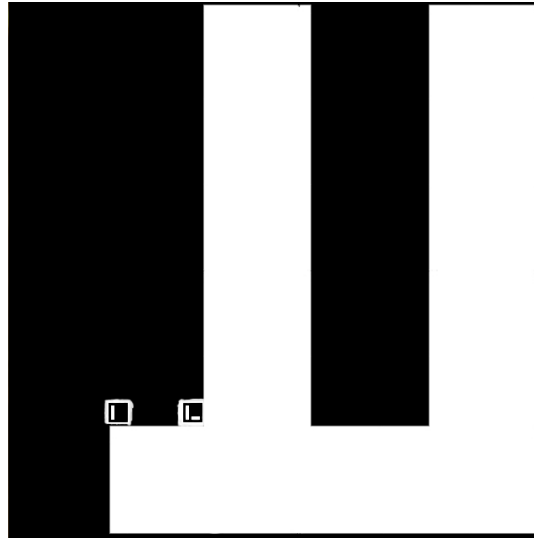
Fig. 3. Simplified layout

One marker in the image can provide enough information to obtain the homography of the image, however, with a small relative size of the marker in the image, perspective lines will vary greatly in angle of inclination, because of this it is advisable to combine several ArUco markers into a group. This reduces the image overlay error. Fig. 4 and 5 – comparison of the transformation of 1 marker and 2 with a small change in the camera angle, while the markers themselves are visually the same square shape.
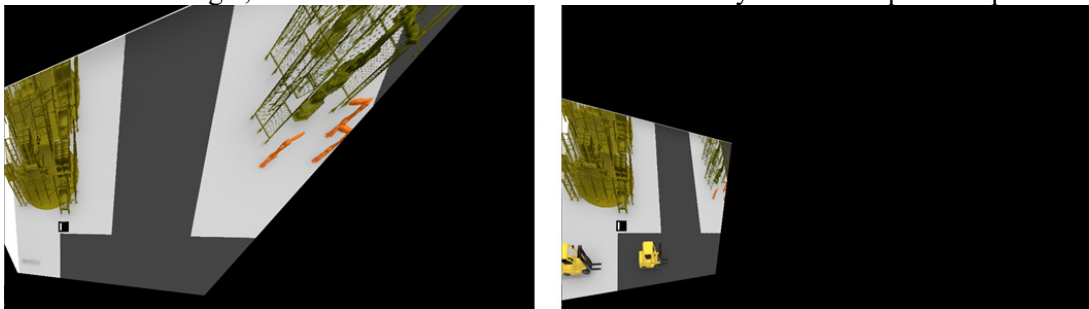
Fig. 4. Homography by 1 marker
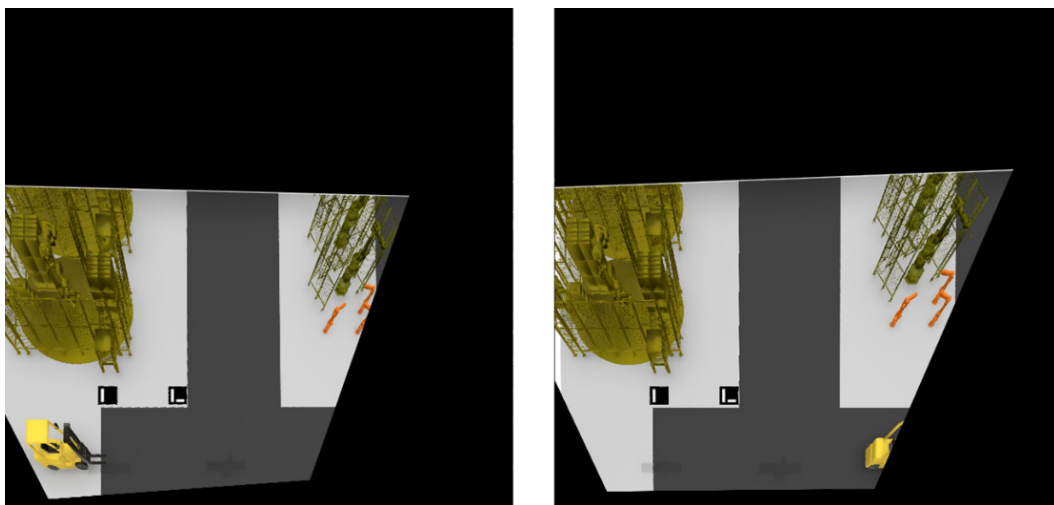
Fig. 5. Homography by 2 markers
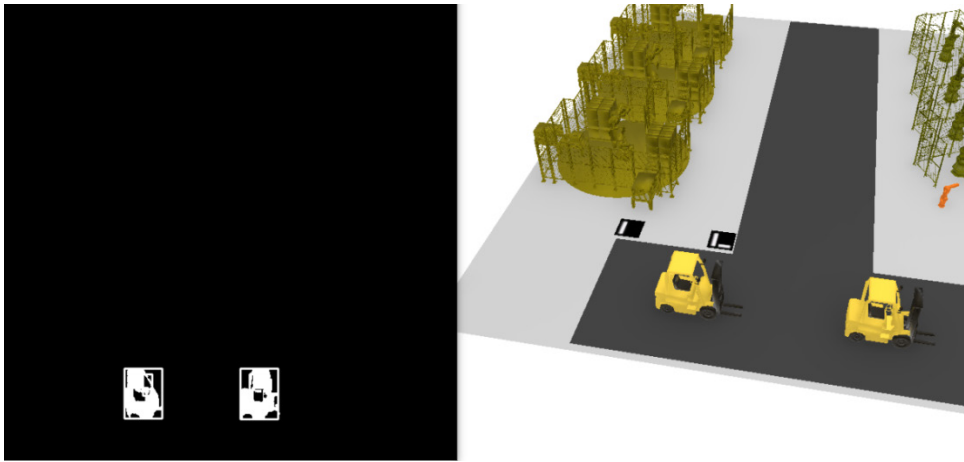
Fig. 6. Detected objects

The resulting final image is superimposed on a simplified layout. In this way, the relative coordinates of the obstacle are unambiguously determined. In the future, it is also possible to transform the obstacle map depending on the nature, direction and speed of movement.
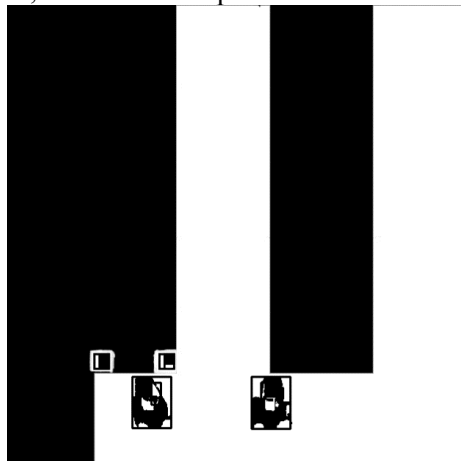
Fig. 7. The resulting map of the permitted traffic area

At the next stage, the resulting map of the permitted traffic area for mobile robots is transferred to the route building module. In this case, a module based on the A* algorithm is used [5, 6]. This module in production receives the current position of mobile robots, a blue marker in the example (Figures 8 and 9), and a target position, a red marker in the example, which comes as a task from MES, another production management system or from the operator.

а)　　　　　　　б)
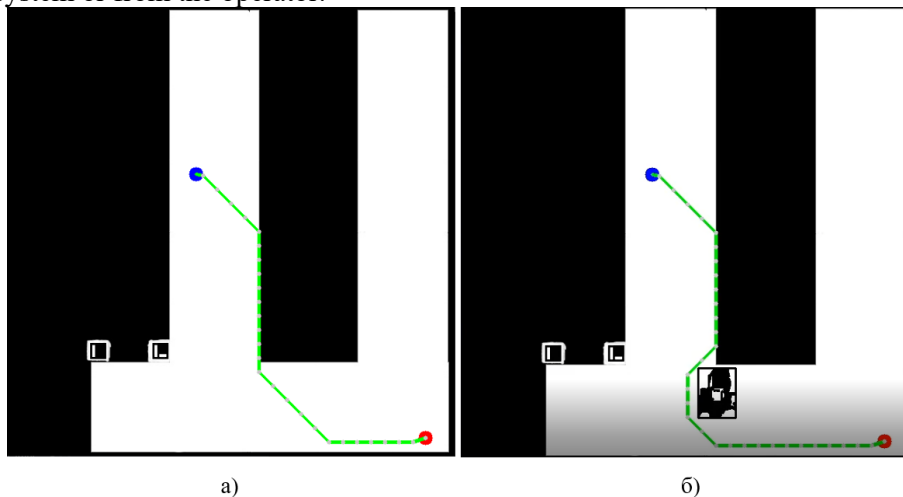
Fig. 8. RTC route: a – initial, b – rebuilt

To improve the reliability of the system, it is also necessary to transmit information from the local security systems of mobile robots to create re-planned routes in case of failure of the computer vision system.

Thus, the resulting system of centralized construction of dynamic RTC routes differs:

1) Good reliability due to additional information processing circuits;

2) Accuracy due to the operation of several systems for detecting and confirming the situation in the production complex;

3) Speed, however, this parameter strongly depends on the scale of the production complex, on the parameters and quality of the neural network, on the selected transformations of the camera image in brightness, contrast, and blur. In the example presented, the speed of information processing and building new routes for 1 mobile robot is less than 0.002 s.

The advantages of the system in question, in comparison with others, are:

1) Comprehensive analysis of the situation and information about the movement of objects in the enterprise;

2) A wider overview as opposed to local robot security systems;

3) Optimal route realignment due to the information received earlier.

However, it is advisable to use this system in industries with a frequently changing environment, with loaders and people moving along the trajectories of mobile robots. For other cases, this system is overloaded and difficult to set up.

References

1. Zyryanov D.A., Badika E.M., Istomina Yu.P., Tspilev V.S. Advantages of using digital production twins on the example of a model in Delfoi robotics // "Actual issues of scientific research ". Saratov: NOP "Digital Science", 2023. – pp. 153-163.

2. Vlasov S. M., Boikov V. I., Bystrov S. V., Grigoriev V. V. Contactless means of local orientation of robots. — Saint Petersburg: ITMO University, 2017. — 169c.

3. Geometric Transformations of Images // OpenCV URL: https://docs.opencv.org/3.4/da/d6e/tutorial_py_geometric_transformations.html (date of application: 01/15/2024).

4. Kasoyan K. F. Development of the augmented reality marker recognition system //Physics. Technologies. Innovations (FTI-2019).—Yekaterinburg, 2019. – 2019. – pp. 123-133.

5. Piskorskii, D. S. Optimization of a-star search algorithm / D. S. Piskorskii, F. K. Abdullin, A. R. Nikolaeva // Bulletin of the South Ural State University. Series: Computer Technologies, Automatic Control, Radio Electronics. – 2020. – Vol. 20, No. 1. – P. 154-160. – DOI 10.14529/ctcr200115. – EDN PRWNKB.

6. Belkov, D. N. Methods and algorithms for constructing an optimal path / D. N. Belkov, I. V. Ponomarev // Proceedings of the seminar on geometry and mathematical modeling. – 2020. – No. 6. – pp. 78-81. – EDN GBWHGN.

CONTENTS

The scientific edition

ИЗВЕСТИЯ КАФЕДРЫ UNESCO ГУАП «ДИСТАНЦИОННОЕ ИНЖЕНЕРНОЕ ОБРАЗОВАНИЕ»

Сборник статей

Выпуск 9

BULLETIN OF THE UNESCO CHAIR "DISTANCE EDUCATION IN ENGINEERING" OF THE SUAI

Collection of the papers

Issue 9

For notes