RUSSIAN ACADEMY OF SCIENCE
FEDERAL AGENCY ON EDUCATION OF RUSSIAN FEDERATION
RUSSIAN NATIONAL COMMISSION FOR UNESCO
COMMITTEE ON SCIENCE AND HIGHER EDUCATION OF THE GOVERNMENT OF SAINT PETERSBURG
COUNCIL OF RECTORS OF SAINT PETERSBURG HIGHER EDUCATION ESTABLISHMENTS
SAINT PETERSBURG STATE UNIVERSITY OF AEROSPACE INSTRUMENTATION (SUAI)
UNESCO CHAIR "DISTANCE EDUCATION IN ENGINEERING" OF SUAI
SAINT PETERSBURG SECTION OF THE INTERNATIONAL SOCIETY OF AUTOMATION

ИЗВЕСТИЯ КАФЕДРЫ UNESCO ГУАП «ДИСТАНЦИОННОЕ ИНЖЕНЕРНОЕ ОБРАЗОВАНИЕ»

Сборник статей

Выпуск 10

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

Collection of the papers

Issue 10

РОССИЯ, САНКТ-ПЕТЕРБУРГ 2025 SAINT PETERSBURG, RUSSIA

ББК 378.1
УДК 74.58
И33

И33 Известия кафедры UNESCO ГУАП «Дистанционное инженерное образование» = Bulletin of the UNESCO Chair “Distance education in engineering” of the SUAI: Collection of the papers. Saint Petersburg, Issue 10. – SPb.: SUAI, 2025. – 234 p.
ISBN 978-5-8088-2026-5

The ISA EMEA&Pakistan (The International Society of Automation) and SUAI (Saint Petersburg State University of Aerospace Instrumentation) have organized the Twenty-first ISA Europe, the Middle East, Africa & Pakistan student paper competition (The ISA XXI EMEA&Pakistan SPC-2025). 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, post-graduate students, professors and specialists.

The ISA ESPC-2025 International Jury:

Ovodenko Anatoly (Russia) – chair,
Antokhina Yulia (Russia),
Bobovich Alexander (Russia) – secretary,
Chabanenko Alexander (Russia),
Cockrell Gerald (USA),
Kryachko Alexander (Russia),
Maiorov Nikolay (Russia),
Pau Giovanni (Italy),
Shepeta Alexander (Russia),
Solyonyj Sergey (Russia),
Zamarreno Jesus (Spain).

ISBN 978-5-8088-2026-5

© Saint Petersburg State University of Aerospace Instrumentation, 2025

International Society of Automation
Setting the Standard for Automation™

On behalf of the International Society of Automation (ISA), 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 EMEA&Pakistan Student Paper Competition.

We are grateful that you have taken the time to contribute your time, knowledge, and expertise to prepare for this competition. Thank you also to the volunteers who selected the papers for publication and awards.

Today's students are tomorrow's leaders of the automation community. With so much concern in the broader world about robots, artificial intelligence, and the workforce, it is more important than ever to underscore all the ways that automation depends on people. Technicians, engineers, and professionals who are well trained in automation systems and implementations are fundamental to our future. I am grateful for each of you and for your dedication to our field. I invite you to continue to participate and contribute to ISA so we can work together to make the world a better place through automation.

I wish you all the best as students, lecturers, and committee members in the EMEA&Pakistan Student Paper Competition.

Sincerely,

Scott Reynolds
ISA President 2025

3252 S. Miami Blvd. #102, Durham, NC 27703
PHONE: +1 919-549-8411 • FAX: +1 919-549-8288 • EMAIL: info@isa.org • www.isa.org

Setting the Standard for Automation™

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 Twenty-first ISA Student Paper Competition (The XXI ISA EMEA&Pakistan SPC 2025).
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

PRODUCT QUALITY ASSESSMENT USING VIDEOMETRIC SYSTEMS

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

Abstract. *With increasing demands on product quality and increased competition in the market, companies are increasingly turning to modern technologies to optimize their processes. One of these technologies is videometry, which is a method of assessing quality based on the analysis of images and videos. Videometric systems are used in various industries, such as mechanical engineering, electronics, food and pharmaceuticals. the speed of analysis and the ability to detect defects, making them an important tool for product quality control. In this article, we will consider the basic principles of videometric systems, their advantages and challenges in the implementation of Fig. 1 [1-3].*

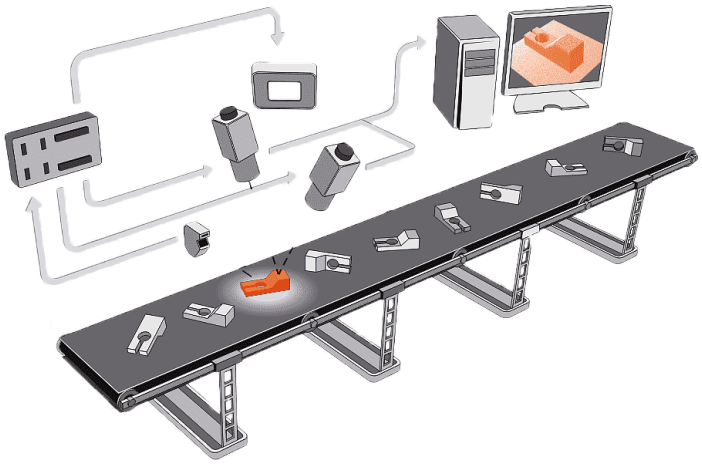


Fig. 1. Structural Model of Neural Network Monitoring Organization

Videometric systems are a set of hardware and software that are used to assess the quality of products based on the analysis of visual data. Their work is based on the use of high-precision cameras, optical sensors, and image processing algorithms. The main stages of their functioning include the following processes:

1. **Data collection.** High-resolution cameras capture the image or video of the object to be monitored. At the same time, image stabilization technologies are used, which ensures the clarity and accuracy of data.
2. **Preprocessing.** At this point, the software performs data filtering, removing noise, and normalizes the images to ensure comparability with reference parameters.
3. **Image analysis.** The system uses machine learning and computer vision algorithms to recognize parts, geometrics, and textural characteristics. For example, cracks, chips, dimensional deviations, or other defects can be detected.
4. **Comparison with the standard.** The data obtained is compared to pre-defined quality standards. Comparisons can be made on several parameters, including the shape, size, color, and texture of the object.
5. **Decision-making.** Based on the analysis, the system makes a conclusion on the compliance of products with the standards. The results are displayed in the form of reports, which may include visualization of the identified defects and recommendations for their elimination.

In addition, videometric systems can be integrated with other elements of the production cycle, such as robotic lines or manufacturing execution systems (MES). This allows you to automate the process of adjusting production parameters in real time [4].

Modern videometric systems are characterized by high data processing speed and versatility. They can be used to inspect both individual elements and finished products. Thanks to their capabilities, these systems not only provide quality control, but also help optimize production processes, reducing equipment downtime and minimizing losses from defects.

Videometric systems are based on the use of high-resolution cameras, sensors, and image analysis software. The main stages of their work include:

1. **Data collection.** Cameras capture an image or video of the object that needs to be analyzed.
2. **Data processing.** The software analyzes the collected images, identifying deviations from the specified standards.
3. **Comparison with the standard.** The system compares the received data with pre-set quality parameters.
4. **Output of results.** The final analysis is presented in the form of reports that allow you to make a decision on the compliance of products with standards.

Table. Benefits of using a videometric system

Advantage	Description
Pinpoint accuracy	Videometric systems provide micrometer measurement accuracy, which is especially important in high-tech industries.
Speed of analysis	The use of videometry allows for real-time quality control, reducing inspection time.
Process automation	Eliminating human error reduces the likelihood of errors and increases the reproducibility of results.
Flexibility of application	Videometric systems can be adapted to a variety of applications, including part geometry inspection, surface defect detection, or packaging inspection.

1. **Pinpoint accuracy.** Videometric systems provide micrometer measurement accuracy, which is especially important in high-tech industries.
2. **Speed of analysis.** The use of videometry allows for real-time quality control, reducing inspection time.
3. **Automation of the process.** Eliminating human error reduces the likelihood of errors and increases the reproducibility of results.
4. **Flexible application.** Videometric systems can be adapted to a variety of applications, including part geometry inspection, surface defect detection, or packaging inspection.

Application examples

1. **Engineering.** Videometric systems are used to verify the accuracy of part assembly, measure dimensions, and detect mechanical damage.
2. **Electronics.** Inspect soldering, chips, and PCBs to detect microcracks or mounting defects.
3. **Food industry.** Checking the packaging and identifying damage or inconsistencies in the labeling.
4. **Pharmaceutics.** Analysis of the integrity of the packaging of medicines, verification of the correctness of dosage and labeling.

Challenges of Implementing Videometric Systems

Despite the many advantages, the implementation of videometric systems is fraught with certain difficulties:

- **High cost.** The upfront costs of hardware and system setup can be significant.
- **Complexity of integration.** Videometric systems require adaptation to existing production processes.
- **Staff training.** To work effectively with such systems, qualified specialists are required.
- **Processing of large amounts of data.** The high resolution of the cameras creates a large data stream that requires powerful computing resources.

Development prospects

With the development of technology, videometric systems open up new horizons for quality control and automation of production processes. The use of artificial intelligence and deep learning allows systems not only to record existing defects, but also to predict possible problems, which significantly reduces the risk of defects. Such systems can analyze large amounts of data integrated from various sources, including production equipment and IoT devices.

One of the key prospects is the introduction of cloud technologies for centralized data management and analysis. This enables the creation of global quality control networks that connect multiple production sites into a single system. Cloud solutions also make it easier to update software and scale systems, making them accessible to small and medium-sized businesses.

The development of quantum computing and neural networks will open up additional opportunities for improving the accuracy and speed of videometric systems. For example, the emergence of new image processing algorithms will improve the quality of defect detection on complex surfaces, such as multilayer coatings or microstructures [5-9].

Another prospect is the integration of videometric systems with additive technologies (3D printing). This will allow not only to control the finished products, but also to make adjustments to the process of their creation in real time. Such solutions are especially relevant for high-precision industries such as aerospace and medical industries.

In addition, the development of virtual and augmented reality technologies will create opportunities for visualizing videometry data in an interactive form. This will simplify the decision-making process and make the quality management system more understandable and accessible to all participants in the production process. technologies, videometric systems are becoming more affordable and efficient. The introduction of artificial intelligence and machine learning opens up new opportunities for quality control automation. Future systems will be able not only to record defects, but also to predict their occurrence based on the analysis of previous data. In addition, the use of cloud technologies will allow the integration of videometric systems into global networks, providing centralized control and analysis [10-17].

Conclusion

Videometric systems are an important tool for improving product quality and reducing production costs. They provide high accuracy and speed of inspection, which allows companies to meet the growing requirements of the market and increase their competitiveness. Despite the challenges associated with implementation, the development of technology is making videometry more and more accessible and universal. The introduction of videometric systems in production is an investment in the sustainable development and success of the company.

Bibliography

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. Assessment of plasticity with combined hardening for the study of deformation processes of structural materials under various modes of low-cycle loads Chabanenko Alexander Valerievich, Rassykhaeva Maria Dmitrievna Certificate of registration of a computer program 2021619545, 11.06.2021. Application No 2021618914 dated 06/11/2021.

GENERALIZED ATOMIC WAVELETS

Michail Kriachko
Saint Petersburg State University of Aerospace Instrumentation
mike_kr@mail.ru

Abstract. *The problem of big data sets processing is considered. Efficiency of algorithms depends mainly on the appropriate mathematical tools. Now there exists a wide variety of different constructive tools for information analysis. Atomic functions are one of them. Theory of atomic functions was developed by V. A. Rvachev and members of his scientific school. A number of results, which prove that application of atomic functions is reasonable, were obtained. In particular, atomic functions are infinitely differentiable. This property is quite useful for smooth data processing (for example, color photos). Also, these functions have a local support, which allows to decrease complexity of numerical algorithms. Besides, it was shown that spaces of atomic functions have good approximation properties, which can reduce the error of computations. Hence, application of atomic functions is perspective. There are different ways to use atomic functions and their generalizations in practice. One such approach is a construction and application of wavelet-like structures. In this paper, generalized atomic wavelets are constructed using generalized Fup-functions and formulas for their evaluation are obtained. Also, the main properties of generalized atomic wavelets are presented.*
Keywords: *data processing, wavelets, atomic functions, V.A. Rvachev up-function, atomic wavelets, generalized atomic wavelet expansion.*

The past few decades have been marked by the rapidly accelerating development of information technology. A huge number of opportunities, which seemed to be completely inaccessible earlier, has been appeared. At the same time new problems have been arisen. The volume of information has been increased and total expenses for its processing has been increased significantly. It is obvious that development of efficient algorithms is the basis for the successful application of new technology. We stress that the complexity of the algorithm and the accuracy of the results are key indicators of quality. These indicators are often highly dependent on the used mathematical tools.

In the last half of the twentieth century some new approximation tools such as wavelets and atomic functions were constructed. The reason was the inability to solve different engineering problems using classic mathematical approaches.

There are many different requirements that can be imposed on systems of functions. However, the most important are the order of smoothness, compactness of the function support (we say that the set $\sup f(x) = \overline{\{x : f(x) \neq 0\}}$ is called a support of the function $f(x)$) and good approximation properties. The first one is important for the case of smooth data processing (for example, color photos). Further, if we use the system of locally supported functions, then it is possible to reduce time and memory complexity of the numerical algorithm. Finally, precision of the data representation and correctness of the results mainly depends on approximation properties. This implies that combination of the above features is necessary for the efficient algorithms of big data sets processing. In this paper we construct the new system of wavelets that have all these convenient properties.

There are many different definitions of the term “wavelet”. In general, wavelet is a function of zero mean that is defined on the real line and decreases sufficiently rapidly at infinity [1]. Various systems of wavelet functions are used in computer graphics [2 – 4], digital data processing and analysis [2,3], economic [4, 5] and so on. Wavelets can be constructed in different ways. One of the methods is the application of solutions of socalled refinement equations

$$y(x) = \sum_k c_k y(ax - k).$$

Note that the equation of this form is a partial case of the linear functional differential equation with a constant coefficients and linear transformations of the argument

$$y^{(n)} + a_1 y^{(n-1)} + \ldots + a_n y = \sum_k c_k y(ax + b_k). \qquad (1)$$

Solutions with a compact support of this equation are called atomic function [6, 7]. Necessary and sufficient conditions of existence of compactly supported solutions of the equation (1) were obtained by V.A.

Rvachev in [7]. For this reason, the authors of the current paper consider it necessary to note that some fundamental principles of wavelet theory were introduced by V.A. Rvachev. One of the most famous atomic functions is wellknown V.A. Rvachev function

$$\mathrm{up}(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(itx)\prod_{k=1}^{\infty}\frac{\sin\left(t\cdot 2^{-k}\right)}{t\cdot 2^{-k}}dt.$$

This function is a solution with a support [-1,1] of the equation

$$y'(x)=2\left[y(2x+1)-y(2x-1)\right].$$

Also, $\mathrm{up}(x)$ is infinitely differentiable. Moreover, it has good approximation properties [8]. Besides, there is a basis of spaces

$$\mathrm{UP}_n=\left\{f(x):f(x)=\sum_k c_k\cdot\mathrm{up}\left(x-\frac{k}{2^n}\right)\right\}$$

that consists of shifts of the locally supported atomic function

$$\mathrm{Fup}_n(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(itx)\left(\frac{\sin\left(t\cdot 2^{-n-1}\right)}{t\cdot 2^{-n-1}}\right)\mathrm{F}\left(\frac{t}{2^n}\right)dt,$$

where $\mathrm{F}(t)$ is the Fourier transform of $\mathrm{up}(x)$.

Atomic functions $\mathrm{up}(x)$ and $\mathrm{Fup}_n(x)$ have a combination of convenient properties. Hence, they have a variety of applications to solution of real world problems [9]. Also, these functions were used in wavelet theory [10].

Some of the results of V.A. Rvachev on the approximation properties of the function $\mathrm{up}(x)$ were generalized for the case of atomic function $\mathrm{mup}_s(x)$, which is a solution of the equation

$$y'(x)=2\sum_{k=1}^{s}\left(y(2sx+2s-2k+1)-y(2sx-2k+1)\right),$$

where $s=2,3,4,\ldots$.

It was shown in [11] that spaces of linear combinations of mup_s-function shifts are asymptotically extremal for approximation of some classes of differentiable functions. Furthermore, the locally supported basis, which consists of shifts of the atomic function

$$\mathrm{Fup}_{s,n}(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(itx)\left(\frac{\sin\dfrac{t}{2(2s)^n}}{\dfrac{t}{2(2s)^n}}\right)\mathrm{F}_s\left(\frac{t}{(2s)^n}\right)dt,$$

where $\mathrm{F}_s(t)$ is the Fourier transform of the function $\mathrm{mup}_s(x)$, was constructed.

There are different ways to use atomic functions in practice. Application of wavelet systems, which are constructed using atomic functions, is one of the approaches to solve real world problems. For this purpose atomic wavelets were introduced in [12]. In addition, it was shown in [13] that these wavelet systems can be effectively used in lossy image compression.

In [12], a generalized Fup -functions were introduced. The function

$$f_{N,m}(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(itx)\left(\frac{\sin\dfrac{t}{N}}{\dfrac{t}{N}}\right)^{m+1}F\left(\frac{t}{N}\right)dt,$$

where $F(t)$ is the Fourier transform of the mother function $f(x) \in L_2(\mathbb{R})$ such that $\sup f(x) = [-1,1]$, $f(-x) = f(x)$, $f(x) \geq 0$ for any $x \in [-1,1]$ and $\int_{-\infty}^{\infty} f(x)dx = 1$, $N \neq 0$ and $m \in \mathbb{N}$ is called a generalized Fup -function.

Wavelet-based data compression. The procedure for the wavelet-based multiresolution analysis

Therefore, the widely used WT-based data compression technique is selected in this paper. In [14], a fundamental multiresolution analysis (MRA) algorithm was utilized for data compression in power systems. The signals were decomposed into scaling coefficients (SCs) and layers of wavelet coefficients data points could be deleted through threshold methods.

The procedure for the wavelet-based MRA is shown in Fig. 1. Through the low-pass filters g_i and high-pass filters h_i, a time series can be decomposed into SCs and WCs, corresponding to the approximations a_i and details d_i of the original signal, respectively. The filters, which are finite impulse response (FIR) filters, are constructed by a scaling function and a wavelet function that depend on the choice of wavelet function. MRA means decomposing the generated SC layer by layer to get SCs and WCs of different scales. SCs and WCs represent the low-frequency and high-frequency components of the signal, respectively. The reconstruction is the inverse procedure of the decomposition.

After the first decomposition, the lengths of the SCs and WCs will be $n_0 + K - 1$, in which n_0 is the number of original sampling points and K is the length of the filters. The total number will be approximately twice as many as before. Other layers are similar. Therefore, two-time downsampling is necessary to avoid information redundancy in decomposition, as indicated by the "$\downarrow 2$" in Fig. 1a. Similarly, two-time up-sampling is obligatory for reconstruction, as indicated by the "$\uparrow 2$" in Fig. 1b. Therefore, assuming that the highest scale is I, the number of a_i and d_i at any scale i can be written as:

$$n_i = \frac{n_0 + K - 1}{2} \approx \frac{n_i - 1}{2} \quad i = 1,2,...,I \tag{2}$$

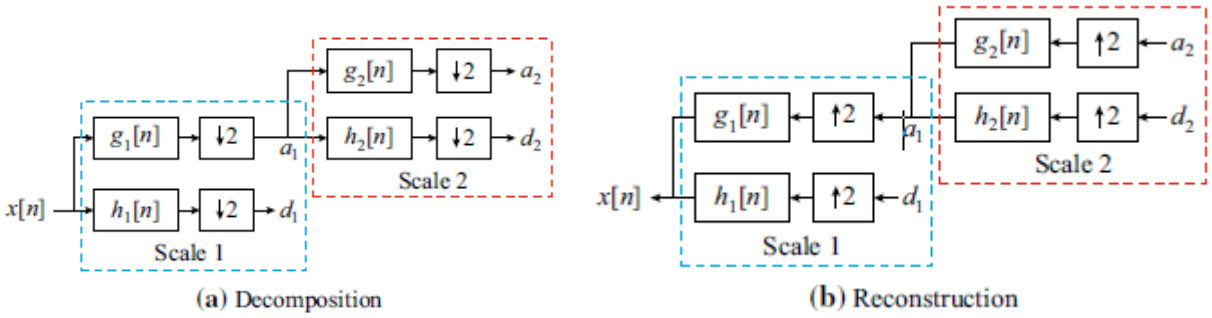


Fig. 1 Procedure of MRA

The general form is:

$$n_i \approx 2^{-i} n_0; \quad i = 1,2,...,I \tag{3}$$

The total number of sampling points before and after decomposition is:

$$n \approx 2^{-I} n_0 + \sum_{i=1}^{I} 2^{-I} n_0 = n_0 \tag{3}$$

As introduced above, SCs represent approximations of the original signal, which is very important. By contrast, WCs represent the details in which the high-value data points experience abrupt changes and the low-value data points are mainly caused by noise. Therefore, a threshold can be applied to WCs to achieve data compression and retain important information of the signal. Many threshold methods have been developed, of which a selection deserves further study in the future. However, no matter what threshold method is chosen, the problem of selecting the best wavelet functions and decomposition scales with

balanced compression performance and reconstruction accuracy always exists. Therefore, a common and widely used fixed-threshold and soft-thresholding method [14] is adopted in this paper. The method is simple but efficient, and the soft thresholding can avoid discontinuities in WCs, making the reconstructed signal smoother. The threshold λ_i and WCs after thresholding $\hat{d}_i$ at scale *i* can be calculated as:

$$\lambda_i = \sqrt{2\ln(n_i)};\ i = 1,2,...,I; \tag{4}$$

$$\hat{d}_i = \begin{cases} \text{sign}(d_i)(|d_i| - \lambda_i), & |d_i| \geq \lambda_i \\ 0, & |d_i| < \lambda_i \end{cases} \tag{5}$$

The main computation of the wavelet-based data compression method is multiplication in convolutions with the wavelet filters. In the decomposition of scale *i*, the number of multiplications is about $2K(n_{i-1} + K - 1)$, so the total number of multiplications in the decomposition of MRA can be calculated as:

$$C(n_0, K, I) = 2K\sum_{i=1}^{I}(n_{i-1} + K - 1) \approx 4Kn_0\left(1 - 2^{-I}\right) \tag{6}$$

Similarly, the number of multiplications in the reconstruction process is the same.

In our previous work [12], the optimal wavelet function and decomposition scale were selected based on the criterion of minimum CDCI for a balanced consideration of compression performance and reconstruction accuracy. The following is an introduction to CDCI. The compression ratio can be used to evaluate compression performance. Assuming that each nonzero data point is of equal size and that zero points can be ignored, the compression ratio λ_{CR} can be calculated as:

$$\lambda_{CR} = \frac{len(\cdot)_len\left(\sum_{i=1}^{I}\hat{d}_i\right)}{len(x)} \tag{7}$$

where x is the original signal; function $len(\cdot)$ means the size of data points.

The distortion rate can be used to evaluate reconstruction accuracy, which can be calculated as a root normalized mean square error. Assuming that the reconstructed signal is x'_n n and *L* is the length of the signal, the distortion rate λ_{DR} can be expressed as:

$$\lambda_{DR} = \frac{\|x'_n - x_n\|}{\|x_n\|} = \frac{\sqrt{\sum_{n=1}^{L}(x'_n - x_n)^2}}{\sqrt{\sum_{n=1}^{L}x_n^2}}. \tag{8}$$

Based on the compression ratio and the distortion rate, the CDCI is constructed as (9) to achieve a compromise between compression performance and reconstruction accuracy.

$$\begin{cases} \xi_{CDCI} = \dfrac{a}{\lambda_{CR}} + b\lambda^*_{DR} = \dfrac{a}{\lambda_{CR}} + \dfrac{\lambda_{DR}}{2\cdot 10^3}, \\ a + b = 1 \end{cases} \tag{9}$$

where *a* and *b* are the weights of the compression performance and reconstruction accuracy, respectively. Owing to the significant difference in magnitude, both $\lambda^{-1}{}_{CR}$ and λ_{DR} should be normalized. $\lambda^{-1}{}_{CR}$ can be seen as a normalized value owing to its value range of (0, 1].

References

1. Novikov, I. Ya. Basic wavelet theory [Text] / I. Ya. Novikov, S. B. Stechkin // Russian Math. Surveys. – 1990. – Vol. 53, No. 6. – P. 1159-1231.

2. Wavelets and multiscale analysis: theory and applications [Text] / J. Cohen, A. I. Zayed (eds). – Springer, 2011. – 353 p.

3. 6. Wavelets in Neuroscience [Text] / A. E. Hramov, A. A. Koronovsky, V. A. Makarov, A. N. Pavlov, E. Sitnikova. – Springer, 2015. – 331 p.

4. Gencay, R. An introduction to wavelets and other filtering methods in finance and economics [Text]/ R. Gencay, F. Selcuk, B. Whitcher. – San Diego : Academic press, 2002. – 359 p.

5. Wavelet applications in economics and finance [Text] / M. Gallegati, W. Semmler (eds.). – Springer, 2014. – 261 p.

6. Рвачёв, В. Л. Неклассические методы теории приближений в краевых задачах [Текст] / В. Л. Рвачёв, В. А. Рвачёв. – К. : Наукова думка, 1979. – 196 с.

7. Rvachev, V. A. Compactly supported solutions of functional-differential equations and their applications [Text] / V. A. Rvachev // Russian Math. Surveys. – 1990. – Vol. 45, No. 1. – P. 87 – 120.

8. Rvachev, V. A. On approximation by means of the function up(x) [Text] / V. A. Rvachev // Sov. Math. Dokl. – 1977. – Vol. 233, No. 2. – P. 295-296.

9. Lazorenko, O. V. The use of atomic functions in the Choi-Williams analysis of ultrawideband signals [Text] / O. V. Lazorenko // Radioelectronics and Communications Systems. – 2009. – Vol. 52. – P. 397-404.

10. Cooklev, T. Wavelets and differentialdilatation equations [Text] / T. Cooklev, G. I. Berbecel, A. N. Venetsanopoulos // IEEE Transactions on signal processing. – 2000. – Vol. 48., No. 8 – P. 670-681.

11. Makarichev, V. A. Approximation of periodic functions by mups(x) [Text] / V. A. Makarichev // Math. Notes. – 2013. – Vol. 93, No. 6. – P. 858-880.

12. Кравченко В.Ф., Чуриков Д.В. Цифровая обработка сигналов атомарными функциями и вейвлетами. – Москва: Техносфера, 2018, 181 с.

13. Крячко М.А., Крячко А.Ф., Дворников С.В. Анализ возможности синтеза спектрально-эффективных сигналов на основе атомарных функций // Вопросы радиоэлектроники. Серия: Техника телевидения. 2022. № 1. С. 94-98.

14. Крячко М.А., Крячко А.Ф., Антонов К.В., Левин Я.Я. Метод повышения спектральной эффективности телекоммуникационных систем на основе аппроксимации огибающей сигналов атомарными функциями // Радиотехника. 2017. № 5. С. 27-31.

AUTOMATING THE DECISION-MAKING PROCESS BY DEVELOPING A MODEL FOR THE EVOLUTION OF A SEA PASSENGER PORT

Angelina Silina, Elena Gaiduk
Saint Petersburg State University of Aerospace Instrumentation
E-mail: angd999@gmail.com

Abstract. *The article analyzes how the relationship between the port and the urban environment has changed over time. The main concepts of development of seaports, as well as the concepts of interaction of the "port-city" system are considered. In addition, a retrospective analysis of the development of sea passenger ports in the Baltic Sea region was carried out, as well as an analysis of existing development projects. Based on the analysis of port development models applied to cargo ports, as well as the analysis of historical data on the development of sea passenger ports and development trends, a model of the evolution of a sea passenger port has been developed, which can serve as a tool for making decisions on port infrastructure modernization and forecasting its further development.*
Keywords: *Sea passenger port, cruise transportation, infrastructure modernization, model of sea passenger port evolution, development forecasting.*

Historically, the cities adjacent to seaports directly depended on the degree of port infrastructure development, both in the economic and social spheres. However, the structure of interaction between the port and the urban environment has constantly changed over time, depending on the degree of influence of sea transportation on the logistics chain of the region, as well as factors affecting the development of the port itself [1-2-2]. For the most accurate representation of the interaction of the "city-port" system, we will consider the key concepts [3] that study the evolution of sea passenger ports, which include historical factors in the development of seaports, as well as theoretical models of development.

The "Anyport" model. The evolution of British ports in the J. Bird model [4] is a five-stage model that reflects the development of a seaport from a small port with berths in the city center to more specialized cargo terminals located far from the city center (Fig. 1). The process of creating new cargo berths and terminals is driven by the development of marine technologies and improved cargo handling, this also changes the spatial relationship between the port and the city. The main goal of the model was to offer a universal approach to the evolution of port development, but it soon became clear that the model does not fully cover all ports, since they do not always follow the same sequence of development stages.

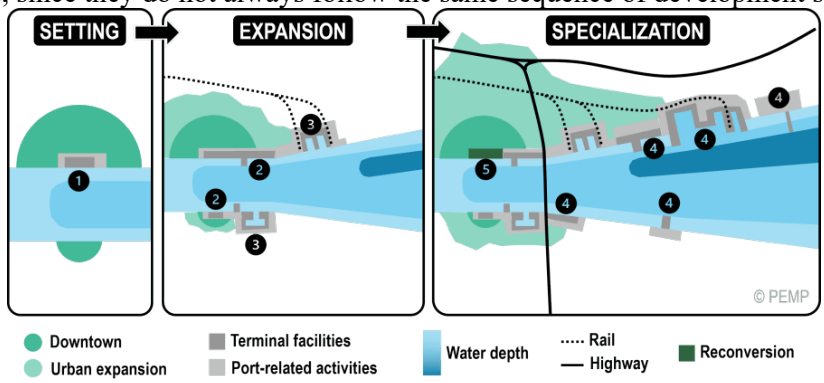


Fig. 1. Stages of spatial development of the seaport of the Anyport model[5]

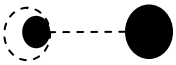
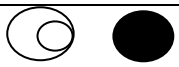
Model of the port city interface, Y. Hayuth. The model was created based on the analysis of the Port of San Francisco [6-7], which previously operated exclusively with finger berths with several branches, but with the growth of containerization, there was a need for large cargo handling spaces and direct access to transport hubs. Given the specifics of the urban environment, the available territories could not meet these needs, which affected the decline in the port's competitiveness. Thus, in search of more adequate locations for cargo handling, port structures were forced to move from the central part of the city to its outskirts, which led to a weakening of the spatial and functional links between the port and the urban structure.

Port-city interface model, B. S. Hoyle. The model [8] was developed at the end of the 20th century under the influence of globalization trends and the growth of intermodal transport. The visual scheme developed by B.S. Hoyle to explain the interaction between the city and the port is presented in Table 1. The

evolution of the seaport begins with the stage of territorial and functional dependence of the port and the city, then its territories are expanded beyond the city limits due to rapid industrial growth. The continued growth of the industry is leading to containerization and requirements for new technologies for bulk and rolling cargo. After that, there are changes in transportation technologies, which contributes to the growth of the port's specialization. The final stage is the renewal of urban cargo terminals in response to the growing demand for territories.

Table 1

Port-city interface model

Development Stage	Designation: City Port	City Period
Primitive port and city		until the 19th century
The expanding port and city		of the XIX century-the beginning of the XX century.
A modernized port and city		in the mid-20th century.
Moving away from the urban freight front		in the 1960s and 1970s.
Revival of the urban sea freight front 1970-1990		1970-1990

Port-city competition model G Norcliffe et al. G Norcliffe, K. Bassett, and T. Hoare [9], in contrast to previous researchers, focus on the socio-cultural aspects of interaction between ports and cities: with the distribution of the population by class, the middle class sought to avoid living in portside industrial zones and moved closer to the outskirts. Fig. 2 shows changes in relations between major cities and ports over time. According to the proposed model, there are three stages in the development of these relations: the period of symbiotic existence, the period of active separation of cities from ports, and the period when ports are actively expanding beyond the boundaries of the urban territory.

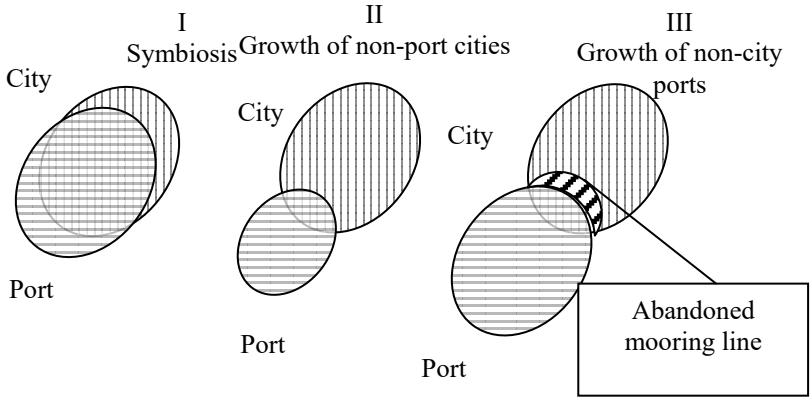


Fig.2. Port-city competition model

The «three generation» model, UNCTAD. The model developed by UNCTAD [10] is an alternative to more outdated models, representing the evolution of ports in three time generations, taking into account the development of information systems, monitoring systems for tracking ship movements, without taking into account the specifics of the geographical location of the port and the impact of the external environment on the port

Workport model, A. K. C. Beresford et al. . A.. K.. C.. Beresford et al. [11] expressed criticism of the model, since most often ports do not develop according to the selected stages, but go through constant

evolution, adapting to changes in the external environment. The development of the infrastructure of a particular seaport may not be linked to a specific period of time, since the geographical location of the port plays a crucial role in determining its potential for further development. The model covers issues of ownership and diversity of cargo handled at the port, methods of organizing transshipment processes, management of auxiliary processes and information communications, and others.

It is worth noting that the presented models were developed **exclusively for cargo terminals**, taking into account the peculiarities of their interaction with the city. These models can also be used to analyze the development of passenger ports, but their application will be very limited. In the case of sea passenger ports, the key development factors influencing its development strategy include such factors as geographical location, cultural characteristics, the volume of passenger traffic in the navigation region and the conditions of its formation. Let's consider the historical factors of development of sea passenger ports on the example of the Baltic Sea region.

Strategic projects for the development of passenger terminals in Saint Petersburg

From the 1920s, the Port of Leningrad began to establish regular routes with many ports in Western Europe, including London, where passenger ships docked in the city center. However, since 1980s, cruises and ferries began to use the «Morskoy Vokzal» for mooring [12], which includes 5 berths with a total length of 720 meters, and maximum capacity of 1 million passengers per year. Due to the limited channel of the Sea Station, in 2008 a specialized passenger complex «Morskoy fasad» was opened [13], which includes seven berths with a total length of 2.2 km. The port is capable of receiving vessels up to 340 m long, 42 m wide and up to 8.8 m draught.

Given the impact of trends on the ever-increasing size of cruise ships, ports need to adapt to changing conditions for continuous access to the region's maritime passenger lines. In this regard**, a project was developed for the construction of outrigger bollards** [14], which provided for the reconstruction of berth No. 7. As part of the project, three remote bollards were installed, which made it possible to simultaneously moor vessels up to 333.3 meters long to berths No. 6 and No. 7. In addition, within the framework of the Russia – South-Eastern Finland cooperation project 2014-2020 [15] **a project was developed for additional equipment of the checkpoint**, in order to provide year-round ferry service and use the checkpoint for receiving cargo vehicles (Fig. 3).

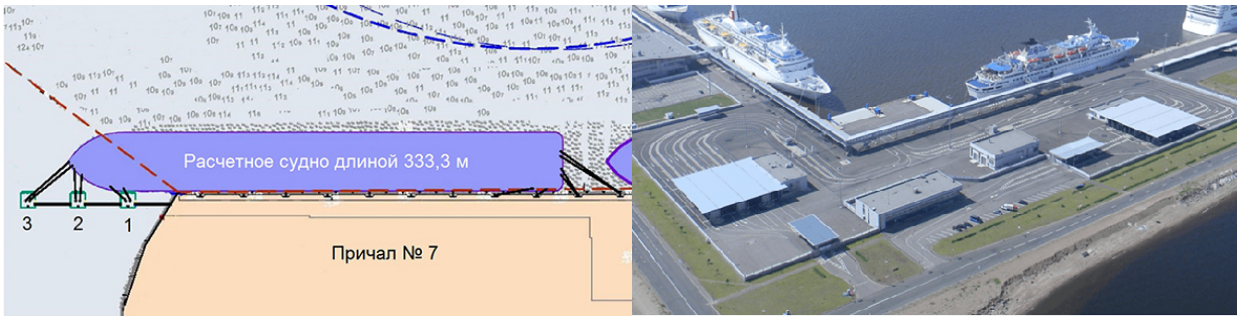


Fig. 3. a) project for the construction of remote bollards; b) project for additional equipment of the checkpoint

Strategic development projects of the Port of Tallinn

Sea passenger traffic gradually developed with the emergence of regular routes in the 1850s, such as St. Petersburg-Tallinn-Riga, Helsinki-Tallinn-Lubeck, etc., and the connection with the Baltic Railway also contributed to the development. An important step was the construction of a separate berth for receiving cruise ships in 2004, as the number of passengers arriving at the port by cruise service, was constantly increasing. Also in 2013, a new cruise ship marina was built next to the existing Old Town marina. To date, the port has 3 passenger terminals, the length of which is 4.2 km. The maximum length of vessels entering the port, is 340 m with a draft of 10.7 m [16].

In 2017**, a comprehensive plan for the development of the Old Town Harbor area until 2030 was developed** [17], which provides for the unification of the port territories and public areas into a single functional system. The purpose of creating the plan was to solve the problem of separating the port territories from the city center, which creates inconvenience for passengers arriving at the port. The changes are aimed at creating infrastructure that connects passengers with the urban environment (Fig. 4).

Fig. 4. a) Port of Tallinn; b) Integrated development strategy of the Port of Tallinn

Strategic development projects for the Port of Helsinki

The port of Helsinki has historically developed as a hub for cargo transportation, and the development of passenger services began in the 1820s with the development of connections between Turku, Helsinki and St. Petersburg. The construction of railways to Saint Petersburg and investments in the development of the Southern Harbor of the port were also significant steps in the development of the sea passenger port. Today, the port of Helsinki consists of three active harbors – South, West, for receiving passengers from various destinations, as well as the port of Vuosari, which is mainly used for container transportation.

One of the nearest directions of the port's development is **the project on centralization and** movement from Tallinn to the Western Port and from Stockholm to Katajanokka. The Port of Helsinki development program is designed for 10 years, and ticking traffic to Stockholm from South Harbor will be transferred to Katajanokka, which will transform the territory of South Harbor into a public space for the city (Fig. 5).

Fig. 5. a) Helsinki Port; b) Helsinki Port Development Project

In the considered port area development projects, there is a desire to integrate port areas with the urban environment, as well as a response to trends towards increasing the size of cruise ships and the need for quick access to the historical part of the city. These requirements lead to an increase in berthing fronts, the creation of specialized passenger terminals and specialized cruise berths for certain destinations. Based on a retrospective analysis of the development of passenger ports in the Baltic Sea region and their infrastructure projects, as well as the considered concepts for the development of the "port city", it is possible to identify the model of evolution of the sea passenger port presented in Table 2.

Table 2

Model of evolution of a sea passenger port

№	Evolution Stage	Description
1.	Cargo and passenger port	A cargo and passenger port without regular passenger services. The port is actively developing infrastructure for receiving passenger ships.
2.	Establishment of passenger services	Regular passenger routes are gradually beginning to be established, in addition, there is a need for integration with other modes of transport
3.	Infrastructure development	The workload of berths is gradually increasing, as well as the size of passenger ships, which leads to the desire of ports to keep incoming passenger traffic and investments in the

№	Evolution Stage	Description
		development of new infrastructure. In addition, there are dedicated berths for receiving cruise ships.
4.	Separation	The rapid growth of passenger traffic leads to the creation of specialized terminal projects, and joint cooperation projects are being implemented to connect passengers with other modes of transport. In addition, there is a need to predict the development of passenger traffic in order to meet demand and quickly respond to external factors
.5	Integration of ports into the urban environment	The trend towards continuous interaction of cruise and ferry lines with the port infrastructure and the urban environment requires port managers to ensure connectivity with the city and quick access to its historical part. The port becomes part of the urban infrastructure, which simplifies access for passengers. In addition, terminals are beginning to adapt to serve certain passenger flows, focusing on specific routes

Based on a retrospective analysis of individual ports in the Baltic region, a model for the development of a sea passenger port was developed. Taking into account modern infrastructure projects, the main trends are identified, one of which is to improve the connection of the sea passenger port with the regional route lines and land transport, increasing the accessibility of the central part of the city.

Separating the different stages of the sea passenger port evolution is essential for decision makers to improve the port infrastructure, increasing the ability to adapt flexibly and in a timely manner to the changing needs of the sea passenger transportation market and grow as the demand for sea transportation increases. It is also worth noting the common features in the evolution of sea passenger and cargo ports, where the passenger port develops from a small berth that accepts mainly cargo ships, to a modern marine complex that provides passengers with a connection to the historical center and other vehicles. For example, as with cargo, terminals are beginning to specialize in receiving passengers in certain directions. It is important to take rapid measures to develop infrastructure to minimize the gap between passenger flows and the region, as well as to restore routes in the future.

References

1. Recent developments and trends in international maritime transport affecting trade of developing countries URL: https://unctad.org/system/files/official-document/cid30_en.pdf (date of reference: 29.01.2025).

2. Yangyang, H. Spatial correlation network structure of port performance and its drivers: A case study of Chinese coastal ports / H. Yangyang, X. Xiaofeng // Ocean & Coastal Management. – 2023. – Vol. 244. – art. no. 106780.

3. Ouariti, O. O. Z. Ports and spatial planning: an exploratory study in the Moroccan context / O. O. Z. Ouariti, E. Mehdi // 13ème Conference internationale de modelisation, optimisation et simulation (MOSIM2020). -2020.

4. Bird, J. The Major Seaports of the United Kingdom / J. Bird // London: Hutchinson. – 1963.

5. The Evolution of a Port URL: https://porteconomicsmanagement.org/pemp/contents/part3/changing-geography-of-seaports/spatial-evolution-port/ (date of reference: 29.01.2025).

6. Hayuth, Y. (1982). The Port-Urban Interface: An Area in Transition. *Area*, *14*(3), 219–224.

7. Ziyue, Li. Conflicts and Reconciliation at the Port-City Interface in Contemporary European Cities / Li. Ziyue // Engineering Management Research. – 2018.

8.Hoyle, B.S. The port—City interface: Trends, problems and examples, Geoforum, Vol. 20, Issue 4, 1989, P. 429-435.

9. Norcliffe, G. The emergence of postmodernism on the urban waterfront: Geographical perspectives on changing relationships / G Norcliffe, K. Bassett, T. Hoare // Journal of Transport Geography. – 1996. – P. 123-134.

10. Port Marketing and the Challenge of the Third Generation Port URL: https://unctad.org/system/files/officialdocument/tdc4ac7_d14_en.pdf (date of reference: 29.01.2025).

11. Beresford, A. The UNCTAD and WORKPORT models of port development: evolution or revolution? / A. Beresford, B. Gardner, S. Pettit, A. Naniopoulos, C. Wooldridge // Maritime Policy & Management. – 2004. – Vol. 31. – P. 93-107.

12. «Morskoy vokzal» 1983-1989 URL: https://pastvu.com/p/522341 (date of reference: 29.01.2025).

13. General information port «Morskoy fasad» URL: https://www.portspb.ru/O_porte/about (date of reference: 29.01.2025).

14. Investment projects Port Morskoy Fasad URL: https://www.portspb.ru/O_porte/about/invest_programm (date of reference: 29.01.2025).

15. Russia–Southeastern Finland Cross-Border Cooperation Program URL: https://docs.cntd.ru/document/542633724 (date of reference: 29.01.2025).

16. Statistics. Port of Tallinn URL: https://www. ts. ee/en/statistics/ (date of reference: 29.01.2025).

17. Tallinn Masterplan 2030 URL: https://www.ts.ee/wp-content/uploads/2020/01/Tallinn-Masterplan_2030_Report_ENG_.pdf (date of reference: 29.01.2025).

18. Port of Helsinki. Annual report URL: https://vuosikertomus2022.portofhelsinki.fi / en/port-of-helsinki/ (date of reference: 29.01.2025).

AUTOMATION IN GAMEDEV: TOOLS FOR AUTOMATIC TRANSLATION AND LOCALIZATION

Kirill Alekseev
Saint Petersburg State University of Aerospace Instrumentation
E-mail: lordkaoron@gmail.com

Abstract. *Automation in GameDev is an interesting and voluminous topic. I believe that this topic may include a sub-topic, namely Procedural Content Generation(PGC): A detailed description of various PGC techniques, such as generating landscapes, cities, characters, weapons, and other game elements based on algorithms. Review the advantages and disadvantages of PGC, as well as examples of successful use of PGC in popular games.*
Automated Testing: An overview of tools for automating game testing, including automatic creation of test scenarios, bug detection, performance and compatibility analysis. Emphasizing the importance of automated testing to ensure high-quality games.
Automating animation creation: Review techniques for automating animation creation, such as motion capture, procedural animation, and AI-driven animation. Analyze the benefits of these techniques for creating realistic and dynamic characters.
Performance Optimization: Use automated tools to identify and fix problems with game performance, such as code bottlenecks, inefficient use of resources, etc.
Automated Project Management Systems: An overview of game development project management tools such as Jira, Trello, and Asana that allow you to automate tasks, track progress, and coordinate team work.
Automatic Translation and Localization Tools: An overview of automatic translation systems that help you quickly and efficiently adapt games for different markets.

For this article, I decided to take the subtopic Tools for automatic translation and localization. In my opinion, it is the most feasible one listed by me.

This sub-topic examines software solutions that can automate or significantly simplify the process of translating and adapting games for different languages and cultures. This includes not only translating text, but also adapting graphics, sound, and even game code.

Technical aspects:

Machine Translation (MT) systems:

Basis: Using machine learning algorithms (neural networks, statistical models) to automatically translate text from one language to another.[1]

MT Types:

- Rule-based MT: Based on linguistic rules and dictionaries. They require manual creation and updating of rules, which makes them time-consuming.
- Statistical MT (SMT): They are trained on large volumes of parallel texts (human-translated texts). They require less manual labor, but the quality of translation can be unpredictable.
- Neural MT (NMT): A state-of-the-art approach that uses neural networks for translation. They provide the most natural and high-quality translation, but require large amounts of training data and computational resources.

Examples: OpenAI Translate API, Microsoft Translator API, DeepL API.

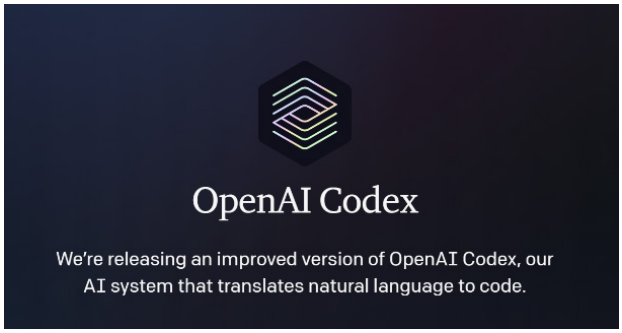


Fig. 1. OpenAI Codex

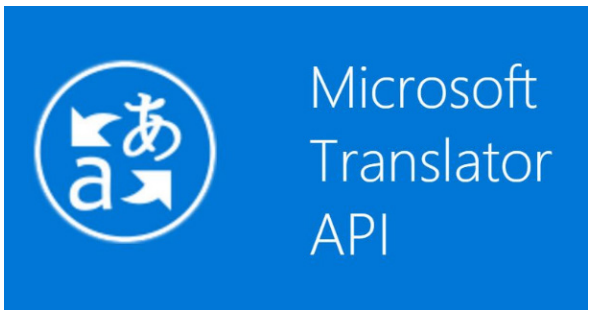


Fig. 2. Microsoft Translator API

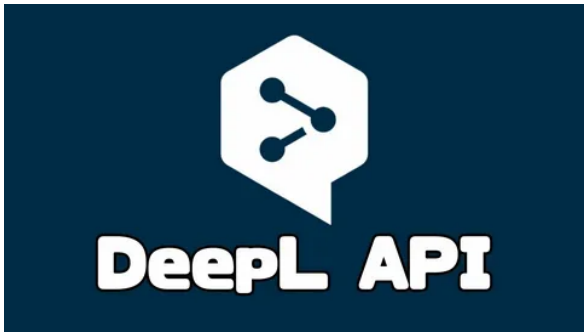


Fig. 3. DeepL API

Интеграция вGameDev Integration: APIs are integrated directly into движок the game engine (Unity, Unreal Engine) or в инструменты для управления content Management Tools (CMS).

Системы управления Translation Management Systems (TMS):

Core: Platforms designed to manage the entire translation and localization process.

Functionality:

- Centralized storage: Stores all texts, graphics, and other resources that require localization.
- Project management: Tracking translation progress, assigning tasks to translators, and managing versions.
- Translation Memory (TM): A database containing already translated text segments. TMS automatically offers re-use of existing transfers, which reduces time and costs by offering re-use of existing transfers.
- Terminology Management: Manage a glossary of terms to ensure consistency of translation.
- Quality Assurance(QA): Automated translation quality checks, such as spelling errors, grammatical errors, and inconsistencies in terminology.
- Integration with MT: Support for automatic translation to speed up the process. [3]

Examples: Crowdin, Lokalise, Smartling, Phrase.

Technical implementation: They use web interfaces and APIs for integration with various systems.

File formats for localization:

- JSON, XML, and YAML: Standard formats for storing text data.
- Gettext (.po and. mo): A common format for localization of software, especially open-source.
- Custom formats: Some games use custom formats for storing text. In this case, you need to develop special tools for importing and exporting data to TMS.

Localization of graphics and audio:

- Image editing tools: Photoshop, GIMP for adapting text elements in images.
- Audio editing tools: Audio editors (Audacity, Adobe Audition) for recording and processing new audio tracks in different languages.
- Asset versioning systems: (for example, Perforce, Git) for tracking changes in image and audio files.
- Localization Testing Automation:
- Basis: Creation of automatic tests to check the correctness of translation, text display, correct formatting of numbers and dates, and compliance with cultural norms. [2]

Technologies: Using Selenium, Appium, and other tools to automate interface testing.

Integration example:

- Texts from the game (character dialogs, item descriptions, interface messages) are stored in JSON format.
- JSON files are imported to TMS (for example, Crowdin).
- Translators are working on translating texts in Crowdin.
- TMS uses Translation Memory and machine translation to speed up the process.
- After the translation is completed, the JSON files are exported from Crowdin and imported back into the game engine.
- The testing system checks whether the text in the game is displayed correctly in different languages.

Key benefits of automation:

- Reduce translation time and costs.
- Improve translation quality by using Translation Memory and terminology databases.
- Simplify localization project management.
- Accelerate the release of games in different languages.

Conclusion

Automation of translation and localization is a critical aspect of modern game development, allowing you to reach a wide audience and increase the commercial success of the project.

The development of the translation automation concept opens up wide opportunities for improving efficiency and reducing costs in the gaming industry. However, to realize this potential, you need specialists who have deep knowledge in the field of machine translation, linguistics, programming and, of course, understand the specifics of game localization. Traditional educational approaches do not always keep up with the rapid development of automation technologies. Therefore, for effective training of new generation personnel, it is worth paying attention to the gamification of the educational process. Creating interactive simulators of working with TMS, tasks for developing scripts for automating routine operations, as well as gamified machine learning courses will not only speed up the learning process, but also significantly increase the involvement and motivation of future specialists. Investments in translation automation should be accompanied by investments in gamified education to ensure that qualified personnel are available to effectively use and develop this new paradigm.

References

[1] “Statistical Machine Translation” Philipp Koehn (2006)

[2] “The Guide to Translation and Localization: Preparing for the Global Marketplace” by Heather Maxwell Chandler(2004)

[3] “Effective Software Testing: A Developer’s Guide” by Maurício Aniche(2022)

CONSTRUCTION OF A SMALL-SIZED AIRCRAFT IMPLEMENTED ON AN ARDUINO BOARD WITH REMOTE CONTROL FUNCTIONALITY VIA A MOBILE DEVICE INTERFACE

Anastasia Bagaeva
Saint Petersburg State University of Aerospace Instrumentation, Ivangorod Branch, Saint Petersburg, Russia, E-mail: 1998anastasia19982401@mail.ru

Abstract. *The article discusses the technological features of building a small-sized unmanned aerial vehicle implemented on an arduino board with remote control functionality via a mobile device interface. During the development of the project, a frame for a quadcopter was developed and printed, the appropriate range of parts and modules was selected, the device (drone) was assembled and tested, and a graphical control interface was implemented for a mobile device running on the Android operating system.*
Keywords: *Graphical interface, Arduino, DIY, electronic components, Automation, Wireless communication, Android.*

In the tasks of this project track, research and design work was carried out to create a small-sized unmanned aerial vehicle with remote control. The type of device is a quadcopter with symmetrical placement of screws at the corners of the square

The following stages were attributed and completed to the basic tasks performed as part of the work on the project: setting a task and forming a prototype, forming a range of parts for assembly, modeling and printing a frame, assembling and configuring, writing code, testing software and hardware, identifying weaknesses and opportunities for modernization.

When modeling a frame, regardless of the type of plastic chosen, the main thing is to take into account all the strength characteristics of the structure with minimal weight. For this, it is necessary to carefully study the model at the project stage and remove excessive use of plastic while maintaining rigidity. In this case, stiffeners, perforations and other ways to facilitate structures can be used.

In fact, the development of an unmanned aerial vehicle can be divided into two parts. These are the "Physical part" and "Working with the controller". The physical part is, in fact, everything that has to do with mechanics, what makes the drone fly. These are elements such as the engine, frame, battery, propellers and everything else that physically provides the drone with the ability to lift and move in the air.

A controller is a device or set of modules for flight control. This set of electronic components controls the physical process so that the drone can fly as a whole without falling. In fact, it is a microcontroller, software on it and sensors that help it determine the direction of movement. So, to create a drone, we selected a controller and a set of various compatible physical parts (blocks) that the controller could control.

As for the physical parts of the body (the UAV platform), it must usually include: a frame, propellers, a battery and an electronic control unit.

Components for the project:

- The frame is made of ABS plastic, printed on a 3D printer;
- blades;
- Battery pack\accumulator;
- ESC (Electronic speed controls);
- motors;
- arduino nano;
- arduino USB cable;
- Bluetooth module;
- LED indicators;
- resistors;
- wires;
- accelerometer and gyroscope;
- prototyping board;
- wire crimping tips (mom\dad);
- ABS plastic.

During assembly, it is important to check whether the propellers are properly installed (inverted or not). When they rotate in the direction indicated by the engines (most engines have arrows showing how they

should rotate), you can feel the breeze blowing under the propellers, not above them. You also need to make sure that the location of the installed motors is appropriate so that they rotate exactly as shown in Fig. 1. To change the direction of rotation of the motor, you just need to swap the wires at opposite ends. This will reverse the direction of rotation of the engine.

Fig. 1. The rotation scheme of the blades

In addition to assembling the platform from parts, there is a nuance that concerns the propellers, or rather their quality of execution and balancing. Quality and balance make it possible to improve the flight performance of the vehicle, reduce noise from a small-sized unmanned aerial vehicle (UAV), increase aerodynamic efficiency, reduce vibration and load on nodes, and improve the image quality from the camera on the suspension.

Accordingly, based on the balancing requirements, the manufacture of blades using 3D printing on most printers cannot be effective due to different quality, different density of plastic in the areas of the blades, and surges. The smoothness of the surface without additional processing is also questionable.

Before completing the assembly of the device on the printed circuit board and assembling the electronics on the frame, it is necessary to use the breadboard for testing. This makes it easy to make adjustments and check the entire circuit for operability.

The Arduino Nano board is chosen as the basis of the controller system. The pin designation diagram of the Arduino Nano board used is shown in Fig. 2 below. It allows you to provide the necessary level of functionality and has small overall dimensions, which is an important factor in terms of balancing the placement of cargo on a drone. Accordingly, the center of mass should have a minimum displacement from the center point between the propellers. Also, large dimensions create excessive resistance to the air environment "windage", which naturally negatively affects the flight characteristics and stability of the device as a whole.

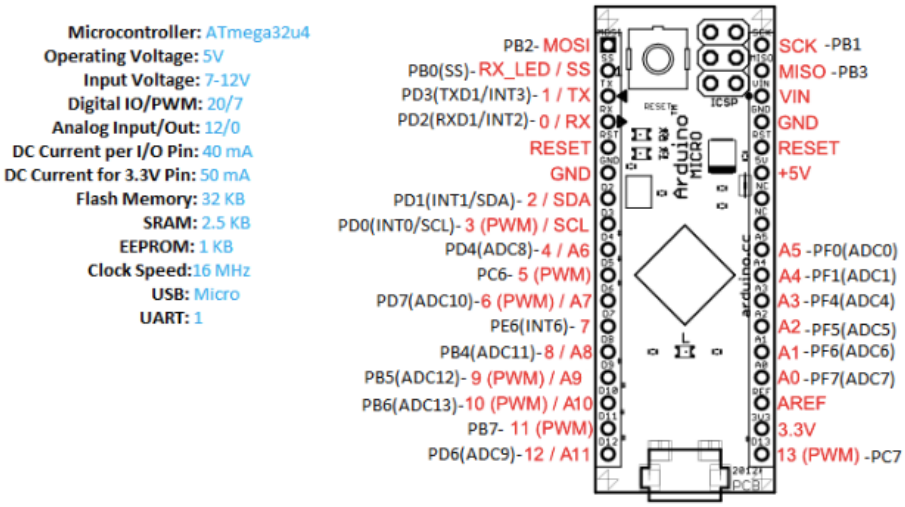


Fig. 2. Arduino Nano

For proper control and orientation in space, the device must be equipped with a gyroscope and an accelerometer. This combined module tells us the acceleration in a certain direction (X, Y, Z) and the angular acceleration along these directions.

The diagram of this module with detailed ports is shown in Fig. 3.

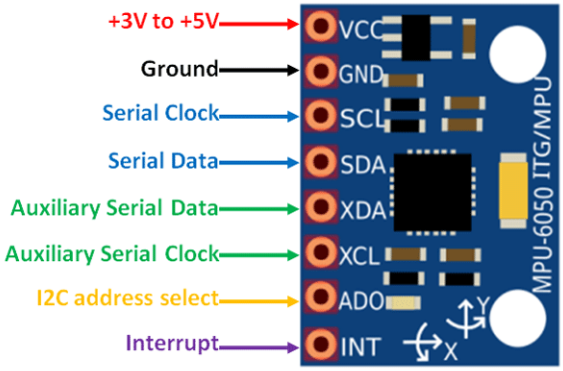


Fig. 3. The gyroscope/accelerator module

The Arduino controller does not support wireless communication out of the box. And accordingly, it needs a compatible expansion module. The connection diagram of the wireless communication module (Bluetooth) to the Arduino Nano board is shown in Fig. 4.

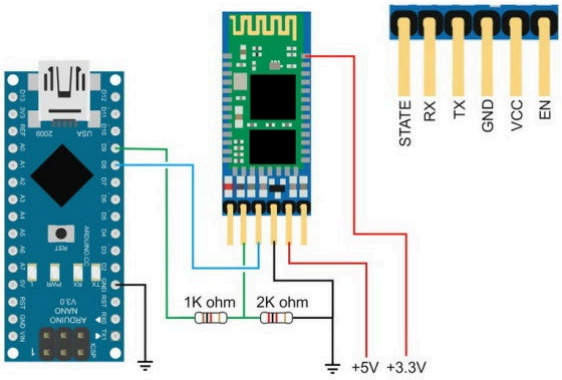


Fig. 4. Connection diagram of the Bluetooth module to the board

Prototyping a system with connectivity details in the Arduino ecosystem is usually performed in the Fritzing package. This package is a simplified CAD system that contains a collection of Arduino elements for modeling, taking into account the wiring from board to board (module, device).

For this project, the corresponding scheme was assembled in the Fritzing package. The switching scheme of modules and components of the small-sized unmanned aerial vehicle under development is shown in Fig. 5.

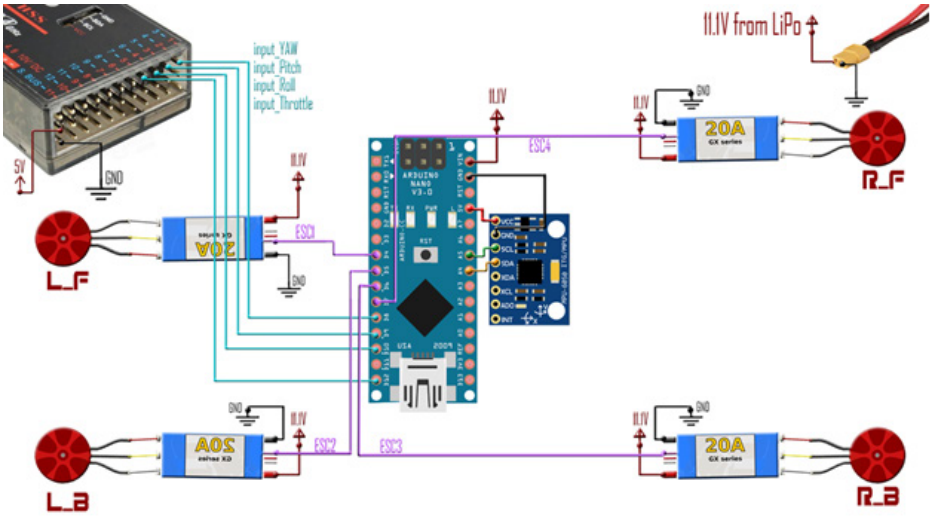


Fig. 5. Connection diagram of the components in the Fritzing system

The main objective of the project is to control the drone from a mobile device. This task can be solved in various ways. In a specific case, the option of using the online tool "Remote XY" was chosen.

Using this service, you can get an interface for Android or iOS and the corresponding source code for uploading to the Arduino controller.

Accordingly, the interface was modeled in the RemoteXY service environment and the necessary set of files was obtained for use in the device under development. The necessary files are embedded in the microcontroller of the Arduino Nano board.

An example of an interface under development in the Remote XY service is shown in Fig. 6.

Fig. 6. Interface development in the Remote XY service

As part of the project, comprehensive work was carried out on the design and construction of a Quadcopter-type UAV with remote control via a mobile device (Android OS phone). The nomenclature of Arduino modules and electronic components for the implementation of the device has been selected. The copter frame was modeled for subsequent printing on a 3D printer, taking into account the necessary stiffeners to reduce the total weight. Modules and components were assembled on a breadboard and their operability as a system as a whole was checked accordingly. At the moment, the type of wireless connection used and the components used do not allow for real long-distance flights.

В общей концепции для более утилитарного использования планируется замена типа беспроводного соединения, замена компонентов беспроводной связи и оснащение радарной установкой для возможности получения картины рельефа земной поверхности.

In the general concept, for more utilitarian use, it is planned to replace the type of wireless connection, replace wireless communication components and equip a radar installation to obtain a picture of the relief of the earth's surface.

References

1. Blaunstein, N. S. Applied aspects of electrodynamics / N. S. Blaunstein, M. B. Sergeev, A. P. Shepeta. – St. Petersburg : Agraf+, 2016. – 272 p.

2. Isakov, V. I. Modeling of location signals reflected from the edge of the earth-sea / V. I. Isakov, D. A. Shepeta // Information and control systems. 2017. No.5(90). pp. 89-94.

3. Kulemin, G. P. Radar interference from sea and land radar of centimeter and millimeter ranges / G. P. Kulemin // Modern radar. 1994. pp. 23-29.

4. Sorokin, A. A. Classification of types of the Earth's surface by airborne small-sized radars / A. A. Sorokin, A. P. Shepeta // Innovations in the field of information technology: from idea to implementation (IvSIT-2024). 2024. pp. 12-17.

THE ROLE OF AUTOMATION IN MODERN METROLOGY

Pavel Banukov
Saint Petersburg State University of Aerospace Instrumentation
E-mail: pool.jump.com@gmail.com

Abstract. *Modern metrology, as the science of measurements, their unity, and accuracy, plays a key role in ensuring the quality of products and technological processes. With the rapid development of industry and the digitalization of production, the need for automation of measurement processes is increasing. Automation not only improves the accuracy and speed of measurements but also minimizes the influence of the human factor, which is especially important in high-precision industries such as aerospace, mechanical engineering, and microelectronics. This paper discusses the main aspects of automation in metrology, its advantages and challenges, as well as prospects for further development.*

Automation of measurement processes involves the use of specialized equipment, software, and data processing algorithms that minimize human involvement in the measurement process. One of the key tools of automation in metrology is high-precision measuring instruments, such as roundness testers, coordinate measuring machines (CMM), and laser scanners [1]. These devices are equipped with sensors capable of detecting microscopic deviations in the shape and size of parts, making them indispensable in precision manufacturing (Fig. 1).

Fig. 1. Example of a signal change tracking system

The main components of automated measurement systems include:

1. High-precision sensors – devices that record deviations in the shape and size of parts with nanometer accuracy. Modern sensors can detect even the smallest surface changes, allowing defects to be identified at early stages of production.
2. Data management and analysis systems – software that processes signals received from sensors and generates measurement reports. Modern systems are equipped with digital filtering algorithms and automatic shape parameter determination, which improves measurement accuracy and reliability.
3. Robotic manipulators – devices that automatically position and secure parts for measurement. This eliminates the influence of the human factor and reduces preparation time for measurements.

The principle of operation of automated measurement systems is based on comparing the actual parameters of parts with reference values. During the measurement, sensors record deviations in shape and size, allowing the construction of a part profile and the calculation of deviation characteristics. To improve measurement accuracy, methods of multi-angle analysis and adaptive data interpolation are used.

Automation of measurement processes has several significant advantages that make it indispensable in modern production conditions. One of the main advantages is the improvement in measurement accuracy and repeatability [2]. Automated systems minimize errors associated with the human factor and ensure high reproducibility of results.

Moreover, automation significantly reduces the time required for quality control. In mass production, where it is necessary to quickly identify defects and deviations from design parameters, automated measurement systems are integrated into production lines, providing non-contact control and instant analysis of the obtained data. This significantly reduces the likelihood of defective products and increases product reliability.

Another important advantage of automation is the ability to integrate measurement data with production management systems, such as ERP (Enterprise Resource Planning) and MES (Manufacturing Execution Systems). This allows enterprises to manage product quality in a unified digital space, which contributes to increased efficiency and reduced costs for corrective operations [3].

Automation also opens new opportunities for predicting and optimizing production processes. Modern measurement systems are equipped with intelligent data analysis algorithms that not only record shape deviations but also predict potential defects. The use of artificial intelligence and machine learning in measurement data processing allows for the identification of hidden patterns and increases process predictability, which is critically important in high-precision manufacturing.

Despite numerous advantages, the automation of measurement processes faces several challenges. One of the main challenges is ensuring the high accuracy and reliability of measurement systems [4]. In mass production, even minor errors can lead to serious consequences, such as equipment failure or reduced product quality. Therefore, the development and implementation of high-precision sensors and data processing algorithms remain key tasks for metrology.

Another challenge is the integration of automated measurement systems into existing production processes. Successful integration requires ensuring compatibility of measurement systems with other components of the production line, such as robotic manipulators and control systems. This requires significant investments in the development and implementation of new technologies, which can be challenging for small enterprises.

Additionally, the automation of measurement processes requires highly qualified specialists capable of developing, configuring, and maintaining complex measurement systems. This creates the need for training and retraining personnel, which also requires additional resources.

However, despite these challenges, the prospects for automation in metrology remain very optimistic. The development of artificial intelligence and machine learning technologies opens new opportunities for improving the accuracy and efficiency of measurement processes. The implementation of intelligent data analysis algorithms not only records shape deviations but also predicts potential defects, contributing to the optimization of production processes [5].

Furthermore, the further development of automated measurement technologies will contribute to increased efficiency and reliability of production processes. The implementation of comprehensive automated geometry control systems will allow enterprises to reduce costs, increase production efficiency, and guarantee high quality of manufactured products.

Conclusion

Automation of measurement processes plays a key role in modern metrology, ensuring increased accuracy, speed, and reliability of measurements. The implementation of automated systems minimizes the influence of the human factor, reduces time spent on quality control, and integrates measurement data into the overall enterprise management system. Despite the challenges associated with ensuring high accuracy and integrating complex systems, the prospects for automation in metrology remain very optimistic. The further development of artificial intelligence and machine learning technologies opens new opportunities for improving the efficiency and reliability of production processes, making automation an integral part of modern metrology.

References

1. Gostev V. N., & Kozlov A. I. (2020). Methods and means of measurements in mechanical engineering. Moscow: MSTU Publishing House.

2. Smith, G. T. (2016). Industrial Metrology: Surfaces and Roundness. Springer.

3. .GOST 24642-81. Control of geometric accuracy of parts. Methods for measuring roundness and cylindricity.

4. ISO 1101:2017. Geometrical product specifications (GPS) – Geometrical tolerancing – Tolerances of form, orientation, location and run-out.

5. Brignol P., Petty R. "Psychology of Decision Making: Rational and Irrational Strategies." – St. Petersburg: Piter, 2020.

AUTOMATION OF PROCESSES OF TECHNOLOGICAL PREPARATION OF PRODUCTION BY APPLICATION OF NEURAL NETWORK TECHNOLOGIES

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

Abstract. *The article explores the possibilities of automation of processes of technological preparation of production with the use of neural network technologies. The article analyses neural network algorithms that demonstrate high efficiency in processing large amounts of data, forecasting production performance and optimising resource allocation.*

Intensive growth of technological processes productivity in mechanical engineering is combined with relative stability of design and technological production preparation methods. As a result, the gap between production preparation and the production cycle itself is increasing, which negatively affects the efficiency of the totality of design and manufacturing processes of the final product. The application of neural network technologies to automate production preparation is a promising solution to this problem, allowing not only to speed up data processing, but also to provide adaptive control of technological parameters [1,2].

Traditional methods of production preparation are often based on fixed algorithms and manual analysis, which makes it difficult to integrate the stages of product design, technological preparation and production directly. With the need to develop aggregated technological processes that include various processing methods, the requirements for adaptability and speed of decision-making increase. The use of neural network models makes it possible to integrate data from multiple sources, including experimental results, historical data and equipment performance, which facilitates the accurate identification of technological parameters necessary for the formation of target properties of the product.

Thus, **the purpose of this study** is to develop and implement neural network technologies to automate the processes of technological preparation of production in mechanical engineering.

The need to satisfy end-user requirements is the main motivating factor in product development. The target properties of the product Z_i represent a set of characteristics $Z_i = \{z_{i,1}, z_{i,2}, \ldots, z_{i,m}\}$, and the state of the product satisfying the requirements of the consumer can be expressed by the inequality:

$$\sum_{i=1}^{n} z_i \geq F,$$

where F is a set of functional requirements. At the same time, the actual values of workpiece characteristics may differ from the required values. If we consider the target characteristics as values with a defined tolerance field δ, the criterion for defining the deviation can be given by the formula:

$$\Delta = z - \left(z_F + \delta \right),$$

where z is the current value of the characteristic, and z_F is the required value. The main task of the automated system based on neural network technologies is to ensure the possibility of correction of technological parameters in such a way that the final values of product characteristics fall within the established tolerance limits.

Realisation of the method of automation of production preparation with the use of neural network technologies is based on the construction of the system of identification of the influence of target properties of the product and initial parameters of components on technological parameters of processing (Fig.1). Neural network algorithms trained on extensive data sets are able to identify complex dependencies between input characteristics and final quality indicators, which allows optimising the technological process in real time. This approach facilitates the integration of the design, pre-production and manufacturing phases, ensuring a high degree of synergy between them.

Thus, the introduction of neural network technologies into the processes of technological preparation of production opens up new opportunities for improving the efficiency and adaptability of machine-building production. Automated systems based on intelligent data analysis can not only reduce time and resource

costs, but also significantly improve the quality of final products, meeting the dynamically changing requirements of the market and consumer.

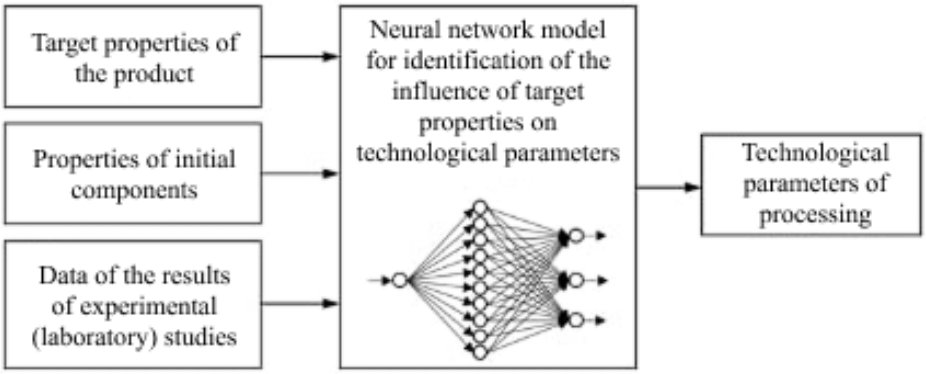


Fig. 1. Structure of interrelation of key elements of the system of identification of the influence of target properties on technological parameters of processing

Optimisation of the efficiency of the processes of technological preparation of production directly depends on the availability of a control system capable of adequately reflecting the ongoing technological processes. In this context, the use of neural network technologies to build a system for identifying the influence of target properties on technological parameters becomes a key element of automation. For correct modelling of the processes of product properties formation it is necessary to create an extensive database based on the results of experimental research, which provides an accurate representation of all stages of production preparation.

Formation of neural network model architecture for automation of production preparation includes determination of the network structure, namely, the number of layers and neurons in each of them [3]. The input layer, the number of neurons of which is determined by the dimensionality of the input signal, serves for the separation and initial processing of input data without performing computational transformations. Similarly, the choice of the number of neurons for the output layer is determined by the dimensionality of the output signal, which guarantees the correct transformation of the input information into the required technological parameters [4,5].

Since the tasks of identifying technological parameters of processing and determining the target properties of the product are related to prediction tasks, it is reasonable to use the architecture of a feed-forward neural network. In such a model, the input signal passes from the input layer to the output layer without feedback, which simplifies the interpretation of the results and increases the adaptability of the system [6-8].

The key stage of development is the optimisation of neural network parameters, in particular, the determination of the number of hidden layers and the number of neurons in them. Despite the absence of strict criteria for selecting these parameters, empirical evidence suggests that a three-layer perseptron can approximate functions defined on a finite set of points or compact areas, and a model with four layers (including two hidden layers) may be required for more complex dependencies. However, in real-life conditions, a more complex architecture can be used when solving problems of automation of technological preparation of production.

In this case, optimisation of the number of neurons in the hidden layers plays a decisive role in the quality of approximation. An insufficient number of neurons limits the training capabilities of the network, while their excessive number can lead to increased training time and the problem of overtraining, when the network demonstrates high convergence on training data, but is unable to adequately process the test sample. Thus, the balancing of neural network parameters is a critical factor for successful automation of manufacturing process preparation processes (Fig. 2).

This is due to the fact that when the number of neurons is excessive, the training of the network increases the weights of the factors characteristic only for the training sample in equal priority with the factors that are significant for the given task.

There are heuristic rules for selecting the number of neurons in hidden layers, one of which is the geometric pyramid rule [9], according to which the number of hidden layer neurons in a three-layer perseptron in the form:

$$k = \sqrt{n \cdot m}$$

where *k* is the number of neurons in the hidden layer, *n* is the number of neurons in the input layer, and m is the number of neurons in the output layer.

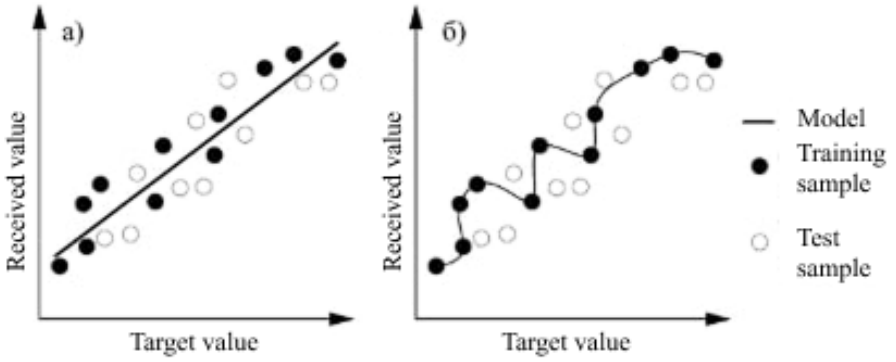


Fig. 2. Convergence of optimally trained (a) and retrained (b) network

Training a neural network is the process of adjusting its parameters by modelling the environment in which it operates (Fig. 3). Modern approaches to learning can be divided into three main categories: learning with a teacher, learning without a teacher, and blended learning. In learning with a teacher, the network is presented with a set of training examples where each input signal corresponds to a reference response. The obtained result is compared with the correct value, after which the error is calculated and the network parameters are adjusted according to the chosen algorithm. When learning without a teacher, there is no required reference, so the network independently groups the input data into clusters based on the internal structure or the degree of correlation between samples. In the mixed approach, a part of neurons is tuned according to the method of learning with a teacher, and the other part is corrected using self-learning methods.

The neural network training process can be represented as a multiextremal non-convex (since the function can have an arbitrary type of training) multidimensional optimisation problem with a large dimensionality, for the solution of which iterative local optimisation algorithms with the calculation of first and second order partial derivatives, stochastic algorithms and global optimisation algorithms are applicable, solved by means of a sequential search for the values of variables defining the control function.

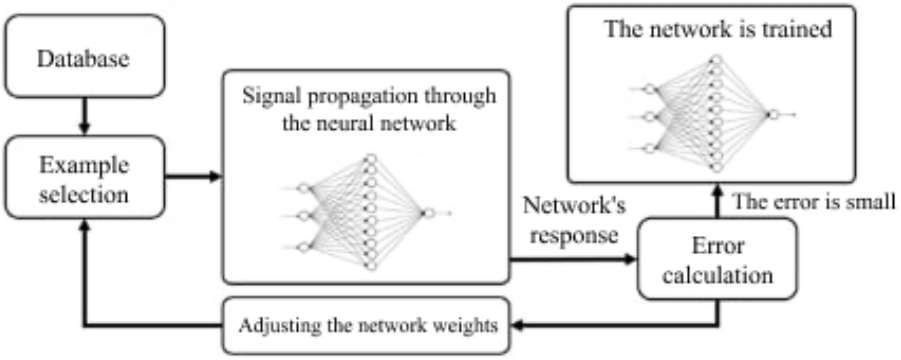


Fig. 3. Structure of the neural network training process

Thus, the process of training a neural network for the automation of technological preparation of production can be considered as a multidimensional optimisation problem with a large number of extrema. The applied iterative local optimisation algorithms based on the calculation of first and second order partial derivatives, as well as stochastic and global search methods, allow to adjust the model parameters to achieve optimal approximation of technological dependencies. However, the successful application of these algorithms requires not only high-quality training data, but also precise tuning of the network architecture, which directly affects the ability of the model to generalise the results beyond the training sample [1-3].

Optimisation of training processes of neural network models for automation of technological preparation of production involves solving a number of complex problems. Training of such models can be considered as a multidimensional optimisation problem with a large number of extrema, which requires the

application of both iterative local optimisation algorithms (using first and second derivatives) and stochastic methods or global optimisation algorithms. The choice of the optimal algorithm depends on the specific problem, the amount and quality of training data, and the computational resources available to solve the problem.

The key aspect is the formation of a high-quality database reflecting all stages of technological preparation of production. Availability of such a database allows neural network algorithms not only to learn, but also to identify complex correlations between input data (characteristics of initial components, processing parameters and target properties of products) and final quality indicators. However, the quality of training directly depends on the completeness and reliability of input data, which requires careful preparation and periodic updating of information sources.

Developing the architecture of a neural network model involves selecting the number of layers and neurons, which is critical to the network's ability to adequately approximate the relationships between process parameters and desired product characteristics. The application of heuristic rules, such as the geometric pyramid rule, can serve as a starting point for determining the optimal model structure. However, empirical tuning of the parameters remains necessary because an excess of neurons leads to overtraining and their deficiency leads to the inability of the network to capture important features of the production process.

When implementing forward propagation models, special attention is paid to the choice of the learning method. In the case of learning with a teacher, the network receives a set of reference data, which allows to correct its parameters through error calculation, while the methods of learning without a teacher contribute to revealing the internal structure of the data by clustering. Mixed approaches, which combine the advantages of both methods, are of particular interest for integration into manufacturing process engineering systems, as they combine accurate prediction with real-time adaptive correction of parameters.

A promising area for further research is the development of hybrid systems combining neural network algorithms with digital twin technologies and the Internet of Things. Such integration will allow not only modelling the technological process with a high degree of detail, but also promptly responding to changes in production conditions, ensuring continuous improvement of product quality and reduction of operating costs.

Thus, the application of neural network technologies to automate the processes of technological preparation of production is a powerful tool capable of increasing the efficiency and adaptability of machine-building enterprises. The solution of the set tasks requires a comprehensive approach, including optimisation of neural network architecture, development of effective training algorithms and integration with existing production management systems. The implementation of such systems opens the way to the creation of intelligent production lines capable of real-time adaptation to changing market conditions and technological challenges of our time [4-7].

Conclusion

The introduction of neural network technologies into the processes of technological preparation of production is a promising direction that can significantly increase the efficiency and adaptability of machine-building enterprises. The use of intelligent algorithms for identification and correction of technological parameters allows to integrate the stages of design, preparation and direct production into a single synergetic system, which leads to a reduction in time and resource costs, as well as to an increase in the quality of the final product. Despite the difficulties associated with optimising the neural network architecture and ensuring the quality of training data, modern training methods and optimisation algorithms offer great opportunities for the practical implementation of these systems.

Future research in this area should focus on the development of hybrid models that combine neural network algorithms with digital twins and Internet of Things technologies to create self-regulating production lines capable of real-time adaptation to changing market conditions and technological challenges. Thus, the application of neural network technologies is becoming a key factor in shaping the competitive and highly adaptive manufacturing of the future [9].

List of used literature:

1. Silaev N.S. Kuznechno-shtampovochnoe proizvodstvo. 1990. No1. pp. 2-5.
2. Semenov A.B. et al. Advanced materials & technologies. 2019. No. 1. doi: 10.17277/amt.2019.01.pp.003-011
3. Bryukhanov V.N., Kosov M.G., Protopopov S.P., Solomentsev Yu.M. Teoriya avtomaticheskogo upravleniya [Automatic Control Theory]. Edited by Solomentsev Yu.M. 3-e, ster. izd. M.: Vysshaya shkola, 2000. 268 p.

4. Sannikov A.S., Nakhushev R.S., Glashev R.M. Elektrotekhnicheskie i informatsionnye kompleksy i sistemy. 2018. T. 14. No 4. pp. 48-53.

5. Kruglov V.V., Borisov V.V. Iskusstvennye neyronnye seti. Teoriya i praktika [Artificial neural networks. Theory and practice]. M.: Goryachaya liniya- Telekom, 2002. 382 p.

6. Nguyen D., Widrow B. IJCNN International Joint Conference on Neural Networks. IEEE, 1990. pp. 21-26. doi: 10.1109/ijcnn.1990.137819

7. Kulakovich A.Yu. Inženernyj vestnik Dona (Rus), 2018, No3. URL: ivdon.ru/ru/magazine/archive/n3y2018/5119.

8. Polyakov S.V., Koroleva I.Yu., Avdeyuk D.N., Pavlova E.S., LemeshkinaI.G. Inženernyj vestnik Dona (Rus), 2019, No1. URL: ivdon.ru/ru/magazine/archive/n1y2019/5561.

9. Grabusts P., Zorins A. Proceedings of the 10th International Scientific and Practical Conference. Volume III. 2015. Vol. 76. P. 81. doi: 10.17770/etr2015vol3.213

WHY LINEAR ACTIVATION FUNCTIONS SHOULDN'T BE USED IN MACHINE LEARNING

Alena Bobovich
Technical University of Munich
E-Mail: alenabobovich08@gmail.com

Abstract. *The paper explains how a neural network is built and what role activation functions in neural networks play. The main focus is on the reason why a pure linear activation function is generally not used as an activation function in hidden layers of a neural network. The limitations of linear activation functions are presented and the mathematical formulas of the output of a neural network are given.*
Keywords: *neural networks, linear activation functions, matrices and vectors.*

In machine learning, neural networks are universal function approximators inspired by the structure of neural networks of animal brains [1]. An artificial neural network consists of multiple layers: an input layer, one or several hidden layers and an output layer. Fig. 1 shows a simple example of an artificial neural network with a single hidden layer. Each layer consists of a number of nodes that are fully connected to the nodes of the subsequent layer. Within a neural network, each layer is responsible for applying specific transformations to its inputs.

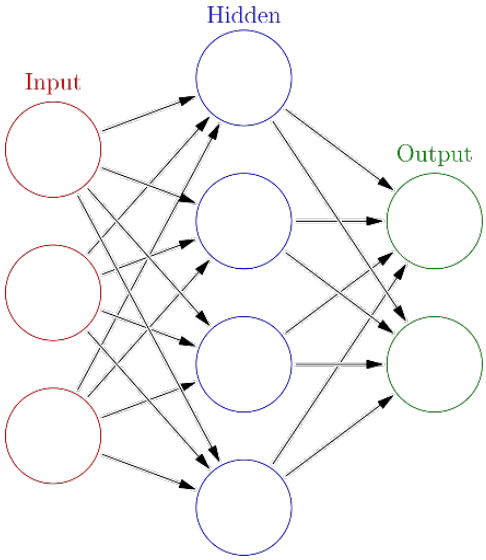


Fig. 1. Simple example of a neural network

Each neuron in the hidden and output layers is associated with a bias and an activation function, while each connection between neurons is assigned a weight. The neurons process their inputs by computing a weighted sum of the incoming signals, adding the bias, and then applying the activation function to produce an output as shown in Fig. 2 [3]. By connecting the neurons with appropriate weights and using appropriate activation functions, a sufficiently large neural network can approximate complex functions with high accuracy [4].

In this paper, the main attention is paid to the activation functions of a neural network. Without the activation functions, a neural network would behave like a simple linear model, which limits its ability to capture more complex patterns in data. Regardless of the number of layers, it would be equivalent to a single-layer linear model, as the composition of several linear transformations is itself a linear transformation and can be collapsed into a single matrix multiplication. This can be illustrated using a simple neural network consisting of two inputs and one hidden layer with two neurons, all employing linear activation functions. While this is a simplified example for illustration, the concept can be generalized to any number of neurons and layers.

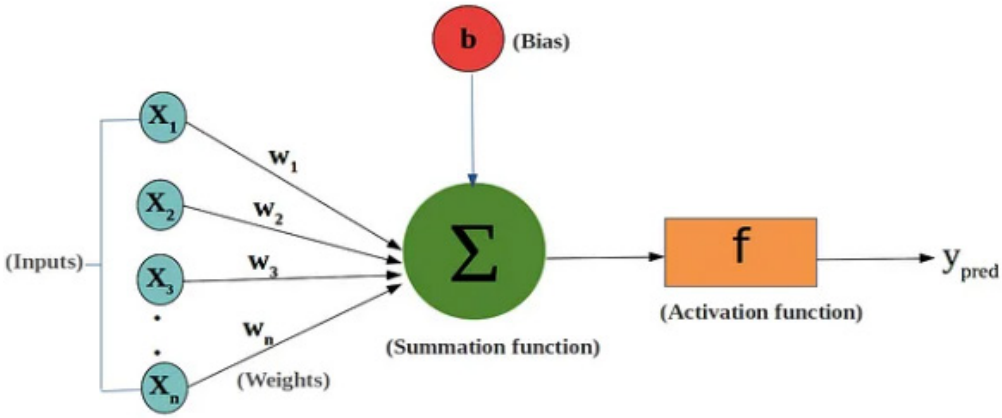


Fig. 2. An input and an output of a single neuron [2]

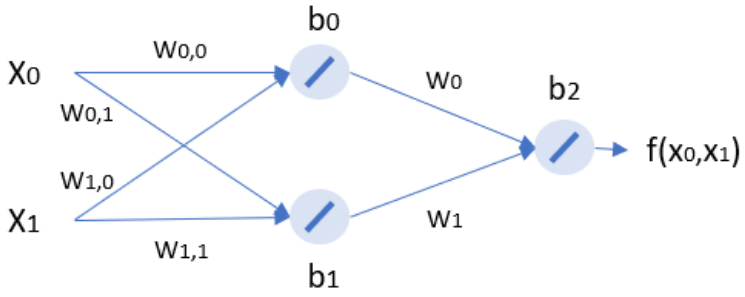


Fig. 3. A simple neural network with 2 inputs, a single hidden layer with 2 neurons and an output layer. All of the activation functions are linear

Consider a simple neural network shown in Fig. 3. The linear activation function is denoted by g with g(x) = x. The output of the 0^{th} neuron in the hidden layer is given by

$$Z_0 = g(x_0 w_{0,0} + x_1 w_{1,0} + b_0) = x_0 w_{0,0} + x_1 w_{1,0} + b_0$$

And the output of the other neuron is given by:

$$Z_1 = g(x_0 w_{0,1} + x_1 w_{1,1} + b_1) = x_0 w_{0,1} + x_1 w_{1,1} + b_1$$

The output of the output layer is then computed as:

$$f(x_0, x_1) = g(Z_0 + Z_1 + b_2) = Z_0 + Z_1 + b_2$$

$$= (x_0 w_{0,0} + x_1 w_{1,0} + b_0) + (x_0 w_{0,1} + x_1 w_{1,1} + b_1) + b_2$$

$$= x_0(w_{0,0} + w_{0,1}) + x_1(w_{1,0} + w_{1,1}) + (b_0 + b_1 + b_2)$$

$$= x_0 W_1' + x_1 W_2' + b'$$

The given neural network consisting of an input, a hidden and an output layer reduced to a neural network with an input and an output layer, with W' being a new weight matrix and b' being the new bias of the output neuron. This demonstrates that using a linear activation function in the hidden layer prevents the neural network from capturing complex, non-linear relationships in data, regardless of its depth.

Thus, a linear activation function could only be useful in the output layer of a neural network for regression problems, where we want to predict a numerical value. To approximate complex functions, nonlinearity must be assured within the hidden layers.

References

1. Hardesty L (14 April 2017). "Explained: Neural networks". MIT News Office
2. Kuruva Satya Ganesh (Jul 24, 2020), “What’s The Role Of Weights And Bias In a Neural Network?”: https://medium.com/data-science/whats-the-role-of-weights-and-bias-in-a-neural-network-4cf7e9888a0f
3. S. Sharma, S. Sharma, A. Athaiya. “Activation functions in neural networks”. International Journal of Engineering Applied Sciences and Technology, 2020 Vol. 4, Issue 12, ISSN No. 2455-2143, Pages 310-316
4. A.D.Dongare, R.R.Kharde, Amit D.Kachare. “Introduction to Artificial Neural Network”. International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 1, July 2012.

DEVELOPMENT OF THE SYSTEM OF STRESS-TESTING OF ELECTRICAL CONTROL AND MEASURING DEVICES FOR DETERMINATION OF ESTIMATED FUNCTIONS OF DEGRADATION OF ACCURACY OF DEVICES

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

Abstract. *The accuracy of electrical instrumentation is a key technical parameter that determines the quality of this device. It is this parameter that determines the correctness and reliability of the obtained measurement results, which can be used to calculate and indicate critical modes, the exceeding of which can lead to the failure of expensive equipment or even become a cause of potential danger to human life. It should be noted that long-term operation of devices is inadmissible due to the dynamics of the accuracy parameter, namely its deterioration due to changes in the parameters of the elements included in the design of instrumentation. It is for these reasons that such a concept as the verification interval is introduced, which establishes a period of time for assessing the technical condition of the electrical device [1]. These values are regulated by legislation and metrological services, where depending on the design and the value of passport accuracy determines the period of maximum operation before re-certification of devices. Such intervals can range from a year to several years, provided that the device operates under normal conditions.*

Regulatory and design features of electrical test and measurement instruments

Understanding the dynamics of accuracy can determine the critical, boundary and even unacceptable values of instrument accuracy in absolutely different time intervals. To determine this dependence of time on the manifestation of errors, it is possible to use the scheme of stress testing of electrical instrumentation [2]. This method of reliability assessment and duration of preservation of permissible values of accuracy consists in modeling of operational loads that accelerate the reduction of parameters of the elements included in the design of devices.

At the moment, the task of developing a methodology and scheme of instrumentation stress testing is a rather interesting technical task. The obtained dependence parameters and measurement data in the process of aggressive testing of the instrument will allow to determine more accurately the values of the testing interval. Of course, in an ideal case there is a need to simulate accelerated aging in different operating modes, for example, under the action of high temperature or pressure or two or more parameters at once. But due to the complexity and multiplicity of these quantities, it is possible to identify a key basic aging that is mathematically possible to model the effects of other quantities. Thanks to the basic understanding, the relationship of which and the resulting dependence of the accuracy dynamics, it will be possible to estimate how much and in what way the accuracy parameter of electrical instrumentation will change under different operating conditions.

It is also important to understand the type of current, the type of device mechanism, the difference in design, because the functioning of such devices differs significantly due to the differences in design, for example, for digital devices the important elements are: a reference voltage source, comparators in the form of operational amplifiers, comparators and sets of resistors. In turn, analog-arrow devices differ in the type of mechanism, namely: electromagnetic, magnetoelectric, ferrodynamic, etc. At the same time, the filling of such devices is also different: coil, indicating hand, magnets, diodes, spring, resistors, etc. and other types of devices [3]. In terms of design, each type of devices is characterized by its own features of accuracy degradation, which should be taken into account when planning stress tests.

The result of such tests is a functional dependence that determines the dynamics of the accuracy of the devices, under normal but long term operation. Another result of the model is the determination of critical points and parameters of devices, which are important to consider during the operation of devices. On the basis of the obtained data it is possible to form and develop recommendations for determining and adjusting the verification interval for different classes of electrical instrumentation.

Important standards in this matter are basic standards and laws, for example, the order of the Ministry of Industry and Trade of the Russian Federation from 31.07.2020 № 2510 “On approval of the procedure for verification of measuring instruments, requirements for the verification mark and the content of the verification certificate”, GOST R 8.883-2015 “State System for Ensuring Uniformity of Measurements.

Software for measuring instruments. Algorithms for processing, storage, protection and transmission of measurement information. Test methods", IEC 60051 2016 "Direct-acting analog electrical measuring instruments and parts thereof. Part 1. Definitions and basic requirements common to all parts", and GOST IEC 61010-1-2014" Safety of electrical test and measurement instruments and laboratory equipment. Part 1. General requirements" [4, 5, 6, 7]. These standards and the order regulate the procedures for assessing the accuracy, reliability, quality and resistance of instrumentation to various operational factors.

Considering the procedure of stress testing from the point of view of qualities, we can determine that this process of conducting controlled and predictable tests is necessary to assess the technical quality, namely the reliability and accuracy of performance of the stated functionality. For quality, the concept of accuracy in metrology can be considered synonymous. The development of the scheme and correct measures for stress testing will allow to ensure the prediction of accuracy of devices and increase the reliability and durability of operation of devices.

Causes of changes in the accuracy of electrical test and measurement instruments during normal operation of the instruments

The normal operating conditions for electrical instrumentation, as described earlier, are regulated by standards and orders. The conditions for operation are selected and given considering the impact and mutual influence of these values on the design and as a consequence the operation of the devices. From the variety of conditions such parameters as: ambient temperature, relative humidity, magnetic field induction, vibration, atmospheric pressure, voltage pulsations, frequency of the supply network are selected. Considering the different design of instrumentation, manufacturers set different value intervals for different parameters. For example, for temperature, the range of operating conditions can be from -10 °C to +25 °C.

Standards for the norms of verification and normal operating conditions set the basic parameters and points so, for example, for temperature the point of normal operation is set at +25 °C, it is assumed that at this value the impact of temperature on the elements of devices is reduced to the maximum minimum, because the mentioned components are designed to work in this temperature regime. The permissible fluctuations of ± 5 °C during measurements are also assumed. The relative humidity parameter is taken at the level of 40 to 80% relative humidity, from the parameters of dew point of moisture formation on the elements, as well as the air dryness affecting the short circuit between two current-carrying elements of the device. The parameter of magnetic field determines the electromagnetic compatibility of devices, namely its resistance to work under the external influence of interference, which includes the influence of the magnetic field, for normal conditions is taken point of 0.5 mTl external source. Mechanical effects for electrical instrumentation is determined in the form of vibration and is set in the range of 5-55 Hz with an amplitude of 0.35 mm. For atmospheric pressure is determined by the norms of 84-106.7 kPa. The parameters of the electrical network are characterized by a large list of values such as amplitude, harmonics, overvoltage, etc. For electrical instrumentation, voltage ripple is regulated to not more than 0.5% of the nominal value, with a supply frequency of 5 Hz, with a tolerance of ± 0.2 Hz [8].

It is worth noting that even when operating under nominal conditions, over time the ratings of the elements, as well as the entire device, are subject to change. This is due to physical and chemical aging of materials, changes in normal operating points for the elements, and mechanical wear and tear.

Phenomena of influence on the operation of electrical control and measuring devices expressed in functional dependencies

For different types of instrumentation, the effects of external conditions, their values also vary. As described earlier, the devices are divided into two main categories, namely digital and pointer-analog instrumentation. The former, in turn, are differentiated into magnetoelectric, electromagnetic, electrodynamic, ferrodynamic, etc. systems. Devices of this type are more susceptible to wear and tear from mechanical wear of the elements. So, it is possible to distinguish several reasons that worsen the accuracy of devices.

The change of resistance parameter for resistors and coils is one of the common phenomena, besides it is not always possible to ensure the established temperature standards, that is why this reason is important. So, the function of resistance dependence on temperature has the following form:

$$R(T) = R_0 \cdot (1 + \alpha \cdot (T - T_0)),$$

where R_0 is the initial resistance, α is the temperature coefficient of resistance (TCR), T_0 is the nominal temperature, T is the modified temperature.

The temperature coefficient of resistance parameter is to a large extent an experimental value, which is written out in the normative documentation and the passport of the device, element or material. It should be noted that the relationship between temperature and resistance is more linear and is clearly subject to prediction and analysis. Of course, this parameter does not refer to the cause of normal operation, but it has a place [9].

Normal operating conditions are more suitable by reducing the stiffness of the return spring. At long enough operation of the device and critical, and frequent, in terms of changing values, the movement of the pointer, the spring will change its stiffness parameters, which will lead to the manifestation of deviations. Zero shifts, changes in the stiffness curve, all of these elements will result in errors that require the determination of the spring parameter curve, which is a unique and difficult task in terms of detection.

The change of magnetic field refers to electromagnetic compatibility and is defined as the possibility of interaction of electrical apparatus in the complex between each other. Sources of electromagnetic interference can be electromagnetic fields in electrical machines, power lines, other products generating electromagnetic pulses. For pointer instrumentation, this parameter has less impact on the design, because a sufficiently strong field is needed to distort the measurement values.

A different picture of the effects of the operating conditions for digital measuring instruments, in view of the specific design and functioning of the instruments. Of course, due to the general functionality, the conditions to be kept constant, because they also affect digital instruments, but in a different way and can reveal a different kind of errors. These devices have more stable characteristics, but individual influences can cause errors to varying degrees.

The main and key structural element is the reference voltage source and comparison elements. Violation of the values and ratings of these elements distorts the measurement results. Thus, the drift of the reference voltage is affected by temperature, exactly and according to the same formula as the dependence of resistance on ambient temperature [10].

Variation of parameters and general characteristics of semiconductor elements is also an important aspect. The triggering of comparators in the form of comparators, operational amplifiers or changes in the resistance ratings have a sufficiently large effect on the measurement results, namely the manifestation of electrical measurement errors. They can also generally change the ratings and mode of operation of the devices permanently, which is a displacement of the normal zero point. Displacement of this point and further progressive influence of other conditions can lead to the manifestation of errors not according to a linear law, but at all in a chaotic format, which is reflected by the error in ratings as plus minus.

Separately it is worth to dwell on the growth and instability of analog-to-digital converters. Noise, manifested when operating at different temperatures, is a nonlinear influence on the operation of devices, more from refers to the quadratic or chaotic manifestation. This value is described by the following relationship:

$$V_n = \sqrt{4 \cdot k \cdot T \cdot R \cdot B},$$

where k is Boltzmann constant, T is temperature, K is resistance, V is bandwidth.

Modeling of stress testing scheme

To determine the parameters and dependencies describing how much the instrumentation loses its ability to maintain the declared accuracy during long-term operation, it is necessary to reproduce the conditions of long-term operation of the device. Thus, when the temperature is exceeded, the drift of component parameters accelerates, when high humidity changes, the corrosion of contacts increases and there is a risk of short-circuit formation, which is also characteristic of low humidity. Strong influence of electromagnetic field affects the electrical value itself or structural elements of the instrumentation, vibration and mechanical effects can cause damage to components, wedge or provide greater impact for other conditions.

To simulate the effects of temperature, humidity, pressure and mechanical influences it is possible to use climatic chambers, vibration test chambers, heat, humidity and cold chambers. Of course, to reproduce accelerated effects, these parameters must be changed dynamically, but such dynamics may not be acceptable for e.g. climatic chambers. A vibration chamber can be affected mechanically dynamically, which will undoubtedly distort the results of analog pointer instruments, but for digital instrumentation this test will have no effect. Changing electrical parameters, dynamically and with varying amplitude, in some cases exceeding ratings or rapidly changing the results of the applied electrical quantity, is a solution to the problem of stress testing of electrical instrumentation of both pointer-analog systems and digital instruments.

It is possible to design such systems for devices of different types of current and measured electrical quantity. It is possible to use one scheme of electric quantity change, for example, reversible electric energy converter.

There are many circuits that can be used for these tasks, e.g. inverter, rectifier, step-up and step-down dc-dc converter. The use of set transformers or regulated transformer is excluded, because of magnetization and filtering of measured values by the device. Such schematic solutions will be able to simulate measurement results dynamically, programmatically and unpredictably, which is extremely important for the purity of the experiment.

In the functional dynamic simulation program SimInTech the circuits of electric energy conversion are built. Figures 1-2 show constant voltage converters that change both voltage and current parameters. These circuits are controlled by changing the duty ratio, which regulates the opening and closing times of the transistor. Fig. 3 shows the control system assembled in the program, simulating different values of the coefficient, which affects the final characteristic of the power parameters. Fig. 4 shows an example of the resulting graph with the constructed inverters.

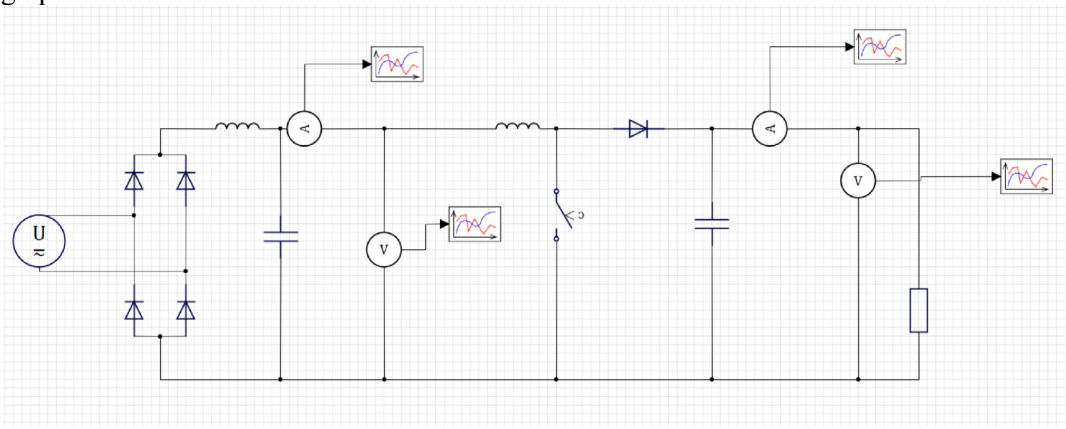


Fig. 1. DC-DC step-up converter

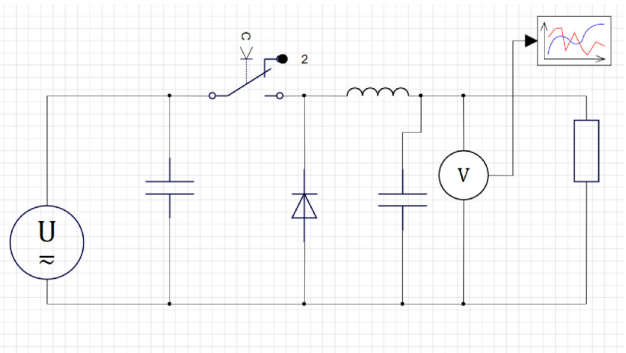


Fig. 2. DC-DC step-down converter

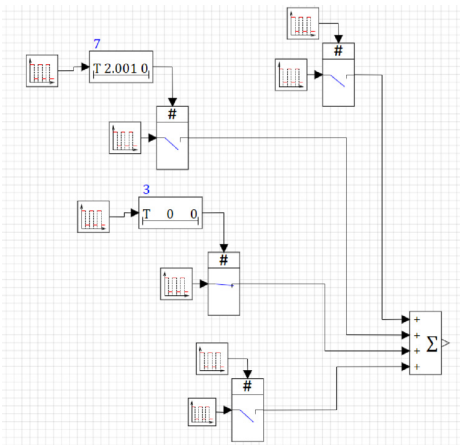


Fig. 3. Control system of transistors of dc-voltage converters built in SimInTech

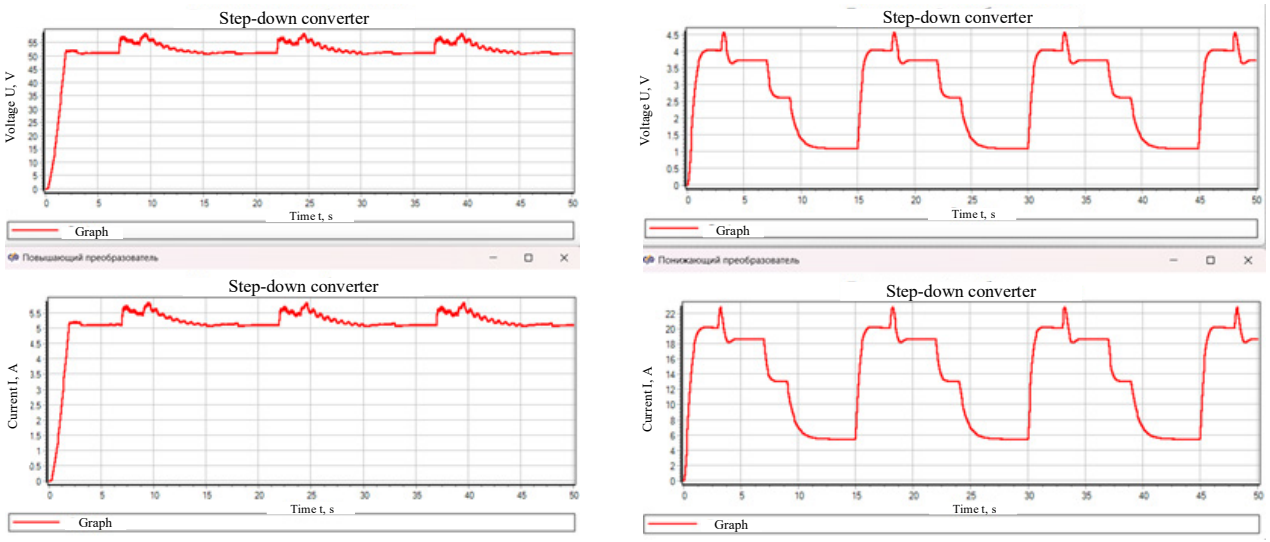


Fig. 4. Graph of energy parameters of current and voltage-built converters in SimInTech program

The graphs in Fig. 4 show an example of the operation of the circuit at different duty ratios, and their operation times. Since the functionality of the constructed control system is limited to changing the pulse filling factor by a given queue with equal time sections, the graphs are periodically equal. But it should be noted that these dependencies confirm the possibility of operation of this scheme and dynamic change of parameters. It is more preferable from the point of view of universality of change of energy parameters, to use serial connection of these converters for the organization of the scheme of stress testing of electrical instrumentation of direct current.

For AC electrical devices it is possible to use either a set of transformers with the ability to change the transformation ratio, or a more accurate schematic solution in the form of an inverter. The constructed circuit is shown in Fig. 5 as a schematic of a reversible DC voltage inverter in symmetrical mode of operation. This mode implies inverting voltage and current, in the asymmetrical mode (Fig. 6), to a greater extent, this converter is a reversible step-down voltage converter, because it is possible to reduce energy parameters and at the same time change the direction of flow of electric current in the circuit..

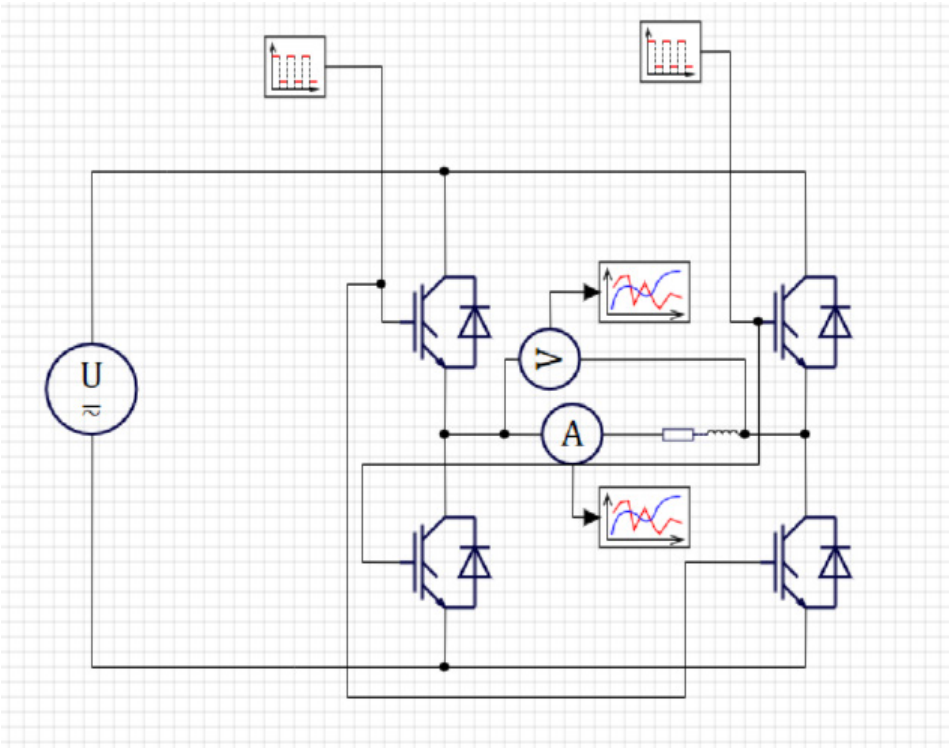


Fig. 5. Schematic diagram of a reversible dc-voltage converter in symmetrical operation mode

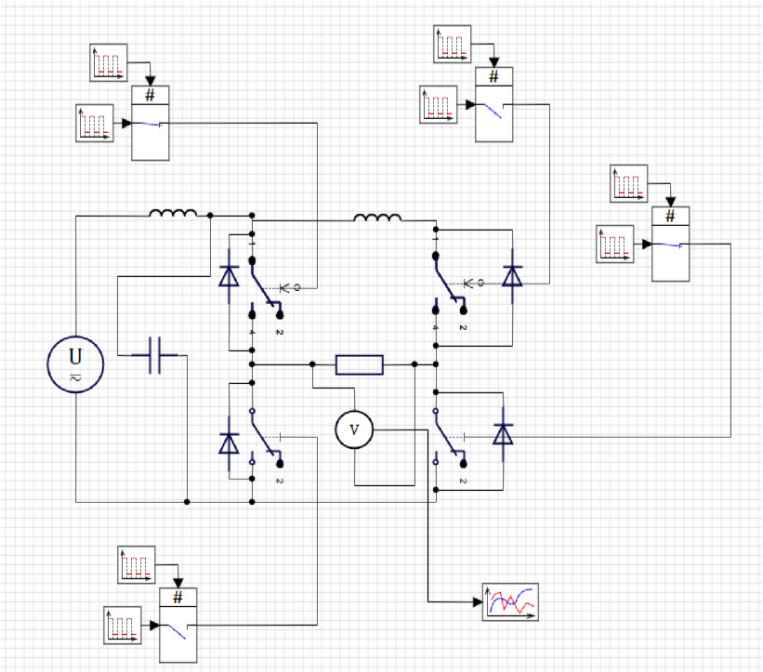


Fig. 6. Schematic diagram of a reversible dc-voltage converter in unbalanced operation mode

This control system is less complex than the previous one, but it also shows the possibility of changing the energy parameters by adjusting the duty cycle. Fig. 7 shows the energy parameters obtained at different duty cycle ratios. Analyzing the obtained results, we can say that the obtained spikes and the graph as a whole demonstrate the necessary values within the framework of accelerated stress testing for AC electrical instrumentation.

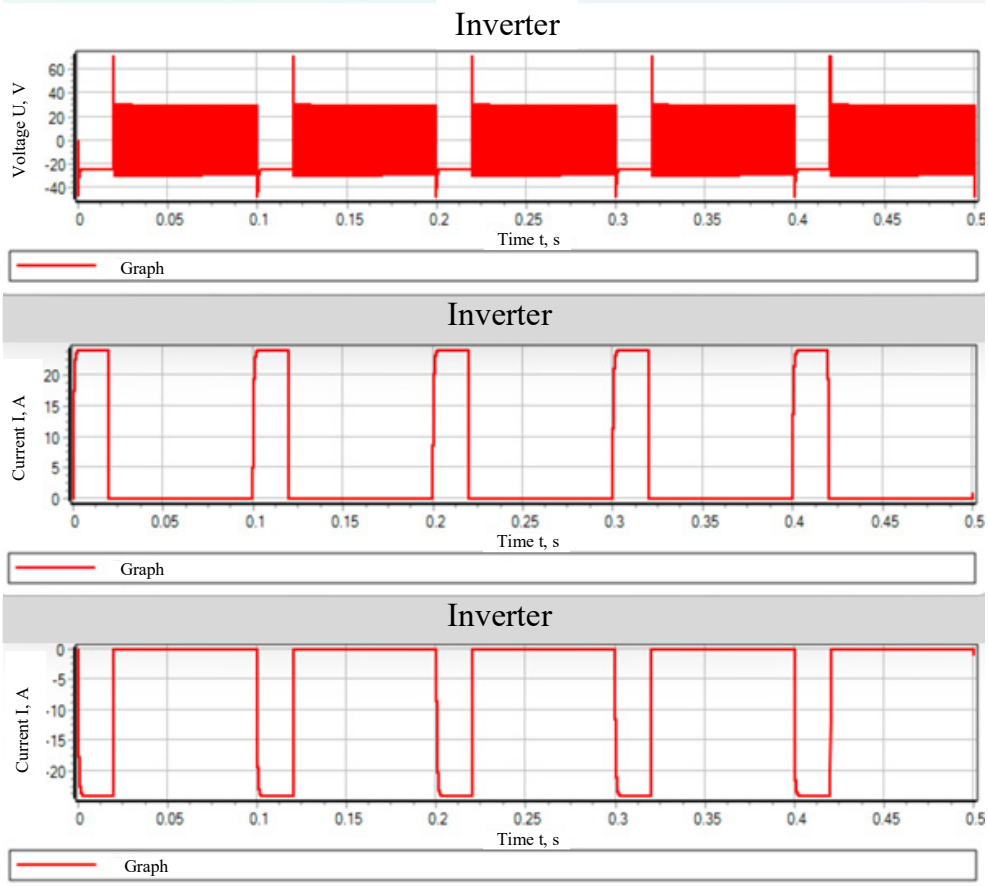


Fig. 7. Graph of current and voltage energy parameters of the constructed converter in the symmetrical mode of SimInTech operation

Carrying out the general analysis it is possible to note, that taking into account the received results of schemes, focusing on the most widespread instrumentation operated in networks 220/380 AC, to the decision of designing the scheme it is offered to use the combined scheme consisting of uncontrolled rectifier with a filter, with serial connection of step-down and step-up converters of direct current, and after the conclusion of the final scheme by the block of inverter. After the step-up and step-down converters, it is necessary to install a controlled transistor to turn off the inverter and turn on the load block, to test the dc devices. The final circuit is shown in figure 8.

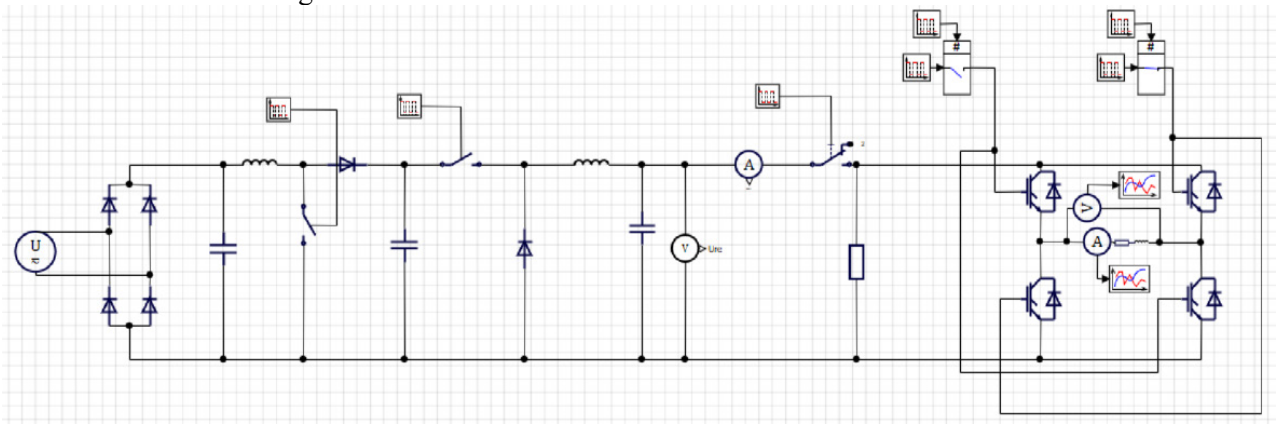


Fig. 8. Formed converter of AC and DC electrical instrumentation stress testing scheme for 220/380-volt network

Conclusion

Stress testing of electrical instrumentation will make it possible to study the peculiarities of instrument quality reduction during operation, and on the basis of the obtained dependencies to form and correct more accurately the results of the verification interval, calibration or instrument operation rules. Prediction of reliability is an important aspect in technical quality. Variation of instrumentation accuracy is a natural process due to physical laws and instrument design. Operation under normal established conditions will certainly provide the stated level of accuracy, but over time the device will in any case go beyond the established errors.

As part of this research study, the performance characteristics of electrical energy conversion circuits were determined and modeled. As a general result, a final circuit suitable for both AC and DC devices was formed.

References

1. GOST 8.565-99 State system for ensuring the uniformity of measurements. Procedure for Establishing and Adjusting the Interval of Verification Intervals of Standards;

2. Bobryshov A.P., Solyoniy S.V. Development of the system of stress-testing of the electrical control and measuring devices / Modern Technologies in Problems of Control, Automation and Information Processing XXXIII International Scientific and Technical Conference, Sochi, 13-19.09.2024 – p.110;

3. Shishkina Y.M. Specificity of the device of electrodynamic and electromagnetic measuring devices and the basic principles of their work / Actual researches ¹2, 2019 p. 5 – 6;

4. Ministry of Industry and Trade of the Russian Federation order of July 31, 2020 n 2510 on approval of the procedure for verification of measuring instruments, requirements for the verification mark and the content of the verification certificate URL: https://normativ.kontur.ru/document?moduleId=1&documentId=376362 (date of circulation 20.01.2025);

5. GOST R 8.883-2015 State System for Ensuring Uniformity of Measurements. Software of measuring instruments. Algorithms for processing, storage, protection and transmission of measurement information. Test methods;

6. IEC 60051-1(2016) Direct-acting analog electrical measuring instruments and parts thereof. Part 1. Definitions and basic requirements common to all parts;

7. GOST IEC 61010-1-2014 Safety of electrical test and measurement instruments and laboratory equipment. Part 1. General requirements;

8. GOST R 8.656-2009 State System for Ensuring Uniformity of Measurements. Means of measurement of quality indicators of electric energy. Verification methods;

9. Borodina E.A., Semenova L.L. Dependence of resistance on temperature of different materials and determination of their temperature coefficient / Vestnik cybernetiki No. 1, 2019, p. 55-59;

10. Bobryshov A.P., Solyoniy S.V., Kuzmenko V.P. Analysis and evaluation of the key design features that determine the quality of electrical control and measuring devices / Innovatsionnoe priborostroenie vol. 3 No. 4, 2024, pp. 5 – 11.

SEMANTIC SEGMENTATION OF AIRBORNE LIDAR DATA USING RANDLANET NERUAL NETWORK MODEL

Polina Bukhvalova, Ekaterina Rachkovskaya
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: tim.kirp@mail.ru

Abstract. *The method of semantic segmentation of airborne lidar data for forest monitoring is considered. Several neural network models for the task of semantic segmentation of lidar data are analyzed. The Randlanet neural network model is selected in the analysis. Experiments with the selected neural network model are carried out.*
Keywords: *semantic segmentation, lidar data, neural network model, airborne laser scanning, digital model.*

Introduction

Nowadays, airborne laser scanning technologies are actively used to build volumetric and highly accurate digital models. Analysis of these models makes it possible to automate the control over the state of natural resources, in particular, the forest area, which reaches millions of square kilometers in area. To analyze such a large territory, it is necessary to use means of automatic analysis of the earth surface. Therefore, to ensure visual informativeness of the operator it is necessary to divide a large digital array into segments. In a volumetric digital terrain map, it is necessary that each pixel (point in the lidar cloud) be assigned to a separate group, i.e. the task of semantic segmentation in digital three-dimensional terrain models. [1].

A digital terrain map obtained by airborne laser scanning technology is volumetric and contains spatial information including elevation characteristics of terrain and objects. To effectively utilize this data, several neural network models exist for the task of semantic segmentation of the land surface.

Analysis of neural network models of semantic segmentation of the earth surface

A digital terrain map is a three-dimensional model of the earth's surface consisting of a cloud of points, where each point carries information about elevation, position and, if a camera is available, color [2-3]. This data format allows not only to visualize the relief, but also to analyze in detail the structure of objects on the surface due to the coloring of each pixel in a color scheme.

Semantic segmentation of airborne lidar data is a method for automatic segmentation of terrain objects in a point cloud acquired by airborne lidar. This method consists of several steps:

1. Lidar scans the terrain, forming a point cloud in 3D space.
2. Noise filtering and normalization of the lidar data is performed to improve segmentation quality.
3. Split into training and test sample of lidar data.
4. Neural network processing of the point cloud.
5. A local feature aggregation mechanism that analyzes neighboring points while preserving spatial dependencies.

There are different approaches to semantic segmentation such as PointNet, PointNet++, PointCNN, SPG and others.

The advantage of the RandLA-Net neural architecture [4] is that a random sample is selected and processed from a large point cloud. This reduces the total computation time, which helps to process a larger number of points in a shorter time. The total time taken by the network to sequence the scanned point clouds and the maximum number of points to be output is evaluated. Table 1, shows the results of comparative analysis of RandLA-Net with other semantic segmentation approaches.

All experiments were conducted on an electronic computer (ECM) with an «AMD 3700X 3,6 GHz» processor and an «NVIDIA RTX2080Ti».

RandLANet is a neural architecture (Fig. 1) for semantic segmentation of lidar point clouds. This network utilizes an encoder-decoder architecture. Input data is fed to a common layer of a multilayer perceptron (MLP), followed by four layers of encoding and decoding to train the model with features for each point. At the end, three fully connected layers and a “thinning” layer associate each point in the 3D point cloud with a class label, such as car, truck, land, or vegetation.

Table 1

Results of comparative analysis

Neural network models	Total processing time, seconds	Number of trained network parameters, million	Maximum number of output points, million
PointNet	192	0.8	0.49
PointNet++	9831	0.97	0.98
PointCNN	8142	11	0.05
SPG	43584	0.25	-
KPConv	717	14.9	0.54
RandLA-Net	185	1.24	1.03

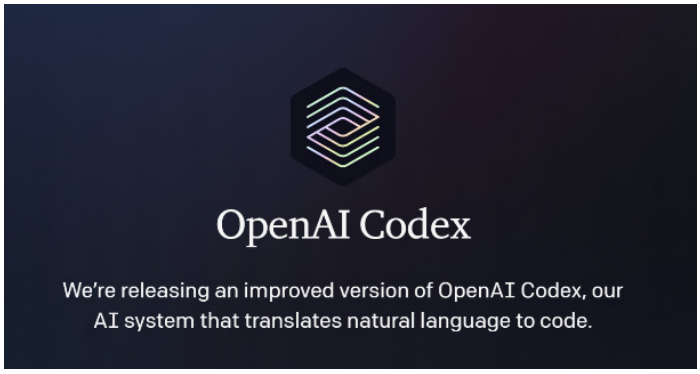


Fig. 1. RandLA-Net architecture

The RandLA network architecture shown in Fig.1 consists of: N – number of points in the input data; dm – feature dimensionality in the input data; layer FC – Fully Connected Layer, fully connected layer; layer LFA – features Local Feature Aggregation, local aggregation; layer RS – random sampling (Random Sampling); MLP – shared Multi-Layer Perceptron, shared Multi-Layer Perceptron; layer US – Up-sampling, increasing the sampling rate of the signal; layer DP – Dropout, thinning [5].

Since RandLA-Net uses random sampling, there is a possibility of losing important geometric features. The local aggregation module helps to avoid this (Fig. 2).

Fig. 2. Architecture of the local aggregation module

The local aggregation module consists of three neural blocks: local spatial encoding (LocSE), group aggregation (pooling), and extended residual block.

The local spatial coding block receives a point cloud as input. This block assigns x-y-z coordinates to each point so that each corresponding point has a different relative location in space. This allows the LocSE block to notice geometric repeating elements, which further helps the entire network to handle complex structures.

For each point, this block uses the k-nearest neighbors metric algorithm. Then points with similar characteristics are combined into groups. As a result, the output is a new set of points characterizing the local geometric structure with a central point.

The grouping (pooling) block combines sets of neighboring points according to certain features. These features can be considered as a mask that automatically selects important features.

To summarize, from the original input set of points, LocSE and grouping blocks search for and combine geometric patterns and features of the closest points, and generate an informative feature vector (Fig. 3).

Advanced Residual Block is designed to handle large point clouds. It increases the perceived space of each feature, which helps preserve detail while discarding unnecessary detail. This allows more points to be considered with each step. The more such blocks, the more computational resources are required. The RandLa-Net model uses two such blocks, which provides a good balance between efficiency and performance.

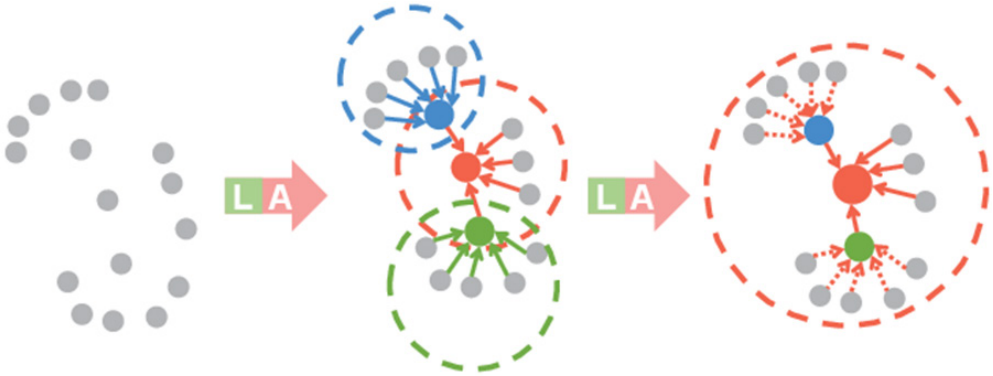
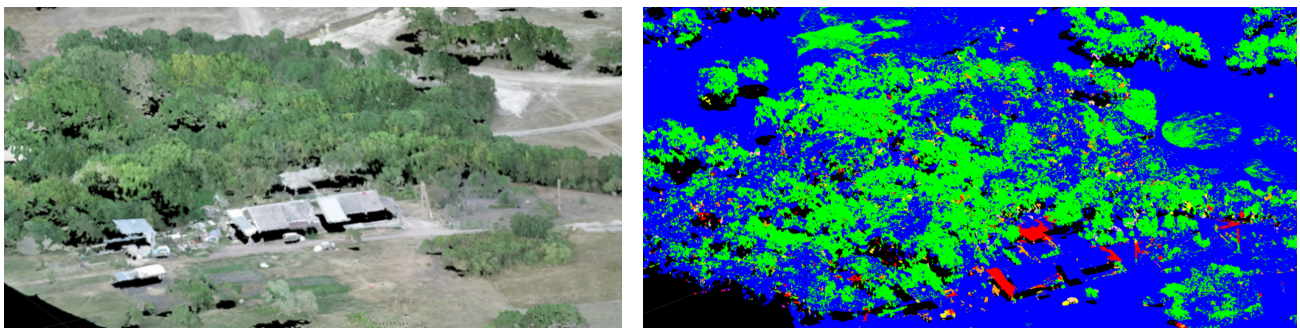


Fig. 3. Operation of the extended residual block,
where colored dots are aggregated features, and symbol L is the "LocSE" block,
symbol A is the "pooling" block.

This local aggregation module efficiently copes with preserving complex local structures taking into account their peculiarities. Also, this module uses only direct transfers. Thanks to this, all computations are faster than with other approaches.

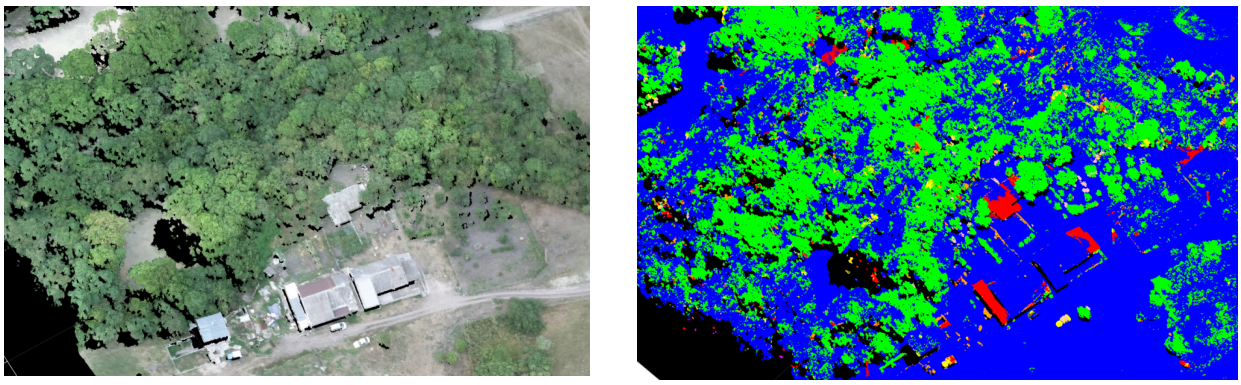
Experiment of the RandLA-Net neural network model under study

To analyze the RandLA-Net neural network model under study [6], an experiment was conducted on the obtained data from airborne lidar. For the experiment, raw point clouds were loaded to the input of the developed software (Figs. 4a, 5a, 6a). As a result of application of the RandLA-Net neural network model, the terrain zones (Fig. 4b, 5b, 6b) such as anthropogenic landscape (transmission line poles, fences), roads, vegetation, and cars were identified.



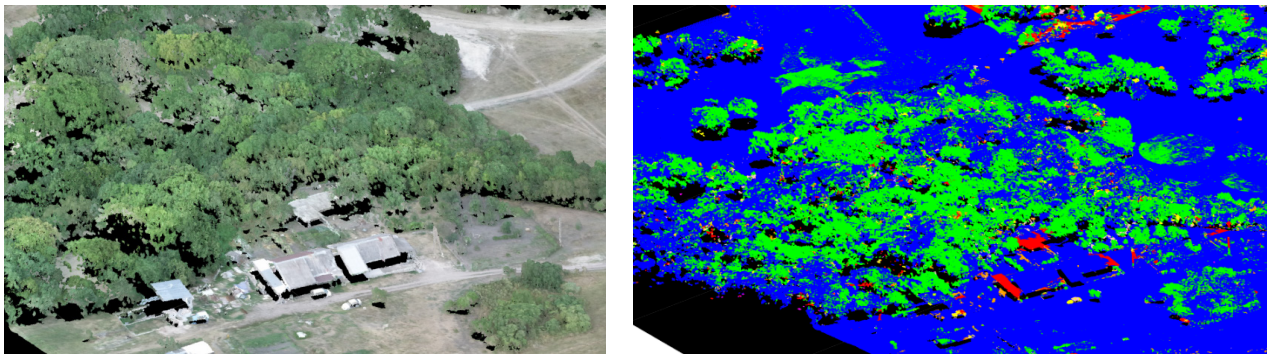
a) input point cloud *b) result of semantic segmentation*

Fig. 4. Experimental result on the test lidar dataset foreshortening 1



a) input point cloud *b) result of semantic segmentation*

Fig. 5. Experimental result on the test lidar dataset foreshortening 2



a) input point cloud *b) result of semantic segmentation*

Fig. 6. Experimental result on the test lidar dataset foreshortening 3

As can be seen from the results, areas with forested areas were colored green, cars were highlighted in yellow-orange, buildings and other structures in red, and the rest of the land surface in blue.

The experiments were performed on a computer with an “AMD 3700X 3.6 GHz” processor and an “NVIDIA RTX2080Ti” graphics processor. The total time spent on software processing processes in the semantic segmentation task was 621 seconds.

Conclusions

As a result of this work, several neural network models of semantic segmentation for 3D terrain models were considered and analyzed. In particular, the RandLA-Net neural network architecture was considered.

Studies have confirmed the high efficiency of RandLa-Net compared to other methods such as PointNet, PointNet++, PointCNN, SPG and KPConv. Experiments have shown that using RandLa-Net can achieve high processing speed, which makes this method most suitable for analyzing large amounts of data.

To summarize, the RandLA-Net neural network architecture is a high-performance neural model for semantic segmentation for pattern analysis and state control automation.

References

1. Forest stand segmentation using airborne lidar data and very high-resolution multispectral imagery / C. Dechesne, C. Mallet, A. Le Bris, V. Gouet, A. Hervieu // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2016. – Vol. XLI-B3. – P. 1-2. – Proc. of XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.

2. Airborne Lidar: A Tutorial for 2025 Part I: Lidar basics // Lidarmag URL: https://lidarmag.com/2024/12/30/airborne-lidar-a-tutorial-for-2025/ (date of address: 23.02.2025).

3. Lidar Zenmuse L1 // Djimsk URL: https://www.djimsk.ru/catalog/products/industrial/zenmuse/lidar_dji_zenmuse_l1.html?srsltid=AfmBOoqKL-WbN1A8lLUXIEIQryDHM5yRH7l8L9cB-ypXwKpkrIk6Etrc (date of address: 04.03.2025).

4. Aerial Lidar Semantic Segmentation Using RandLANet Deep Learning // MathWorks. – [Electronic resource]. – URL: https://www.mathworks.com/help/lidar/ug/aerial-lidar-semantic-segmentation-using-randlanet-deep-learning.html (date of address: 23.02.2025).

5. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds / Q. Hu, B. Yang, L. Xie, S. Rosa, Y. Guo, Z. Wang, N. Trigoni, A. Markham // Proc. of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – P. 3-5.

6. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds (CVPR 2020) // GitHub. – [Electronic resource]. – URL: https://github.com/QingyongHu/RandLA-Net?tab=readme-ov-file (date of address: 5.03.2025)

THE ETHICS OF USING ARTIFICIAL INTELLIGENCE IN DECISION-MAKING: BALANCING AUTOMATION AND HUMAN CONTROL

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

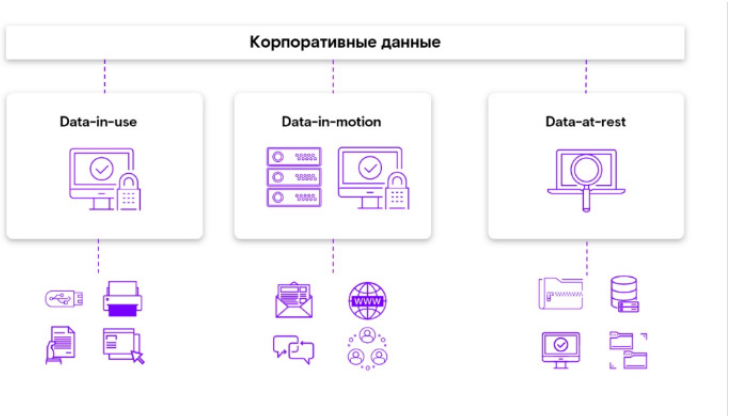
Abstract. *In today's world, artificial intelligence (AI) is playing an increasingly important role in decision-making processes in various fields, from medicine and law to business management and public administration. However, the widespread adoption of AI raises ethical questions related to the balance between automation and human control. This paper examines the key moral and ethical aspects of the use of AI, including the transparency of algorithms, the problem of bias, responsibility for decisions made by the system, and the need to preserve human involvement in critical processes. The need to develop a regulatory and ethical framework that ensures fairness, safety and trust in automated systems is analyzed. It is emphasized that effective interaction between AI and humans should be based on the principles of ethical responsibility, which will minimize risks and maximize the benefits of using smart technologies.*

One of the key ethical issues in the use of AI is the transparency of algorithms. In complex machine learning systems based on neural networks, the decision-making process is often opaque even to the developers themselves. This can lead to unpredictable consequences, especially if algorithms make decisions that affect the fate of people (for example, in lending, litigation or medicine). The lack of transparency also makes it difficult to detect possible bias related to the quality of training data. If the data used to train the model contains hidden social or economic biases, the system can reproduce and even amplify these biases, compromising the principles of fairness.

Various methods are proposed to increase the transparency of algorithms, such as interpretable machine learning models, visualization of decision-making, as well as the implementation of explainable artificial intelligence (XAI) mechanisms. These approaches allow users and professionals to better understand the grounds behind which AI decisions were made, as well as identify potential sources of bias. Algorithm audits and peer reviews also play an important role to help ensure the objectivity and fairness of automated decisions.

One of the most controversial issues in the use of artificial intelligence is the distribution of responsibility for its decisions. In traditional decision-making systems, a specific person or organization is responsible for mistakes. However, in the case of automated systems, this aspect becomes more complicated, as decisions can be made without direct human involvement.

One of the key aspects is legal responsibility. Currently, there is no universal legislation that clearly regulates who should be held responsible for AI errors. Possible models include the responsibility of developers, system owners, or operators who use algorithms in their work. It is important to develop mechanisms to identify the source of error and to provide compensatory measures in the event of damage (pic.1) [1-3].



Picture 1. An example of a corporate data protection system.

In addition, there is the question of moral and ethical responsibility. AI is not conscious and cannot be aware of the consequences of its decisions, so it is necessary to implement control systems, such as

mandatory human intervention in critical situations. This will minimize the risks associated with uncontrollable algorithm errors and their impact on society.

Addressing these issues requires the development of a multidisciplinary approach that includes both legal and technological aspects, as well as the introduction of peer review mechanisms to monitor AI decisions and prevent possible negative consequences.

Another important aspect is the issue of responsibility for decisions made by AI. In traditional decision-making systems, a specific person or organization is responsible for mistakes and their consequences. However, in the case of automated systems, liability is blurred between developers, data providers, system operators and end-users. This creates legal uncertainty, especially in cases where AI errors result in property damage or life-threatening. Developing clear mechanisms for assigning responsibilities is a necessary step to integrate AI into critical areas.

Despite the desire for automation, the complete rejection of human participation in decision-making processes carries serious risks. Artificial intelligence should be seen as a tool to support decision-making, not as the ultimate arbiter. This is especially important in areas where a moral assessment of the situation is necessary, such as medicine or judicial practice.

A key aspect of human control is to provide feedback between the AI and the user. This includes mechanisms for monitoring, analyzing decisions and correcting them in case of errors or deviations from regulatory requirements. Ethical norms governing the interaction between AI and humans should provide for mechanisms for verifying decisions, the possibility of human intervention, and a system for appealing decisions made by algorithms.

In addition, it is important to consider the degree of trust in AI and its recommendations. This requires the development of transparent standards for certification of artificial intelligence systems that would guarantee their reliability, safety and compliance with moral and ethical principles. Human control must be built into all phases of the AI lifecycle, from design and training to operations and upgrades.

Despite the desire for automation, the complete rejection of human participation in decision-making processes carries serious risks. Artificial intelligence should be seen as a tool to support decision-making, not as the ultimate arbiter. This is especially important in areas where a moral assessment of the situation is necessary, such as medicine or judicial practice. Ethical standards governing the interaction between AI and humans should include mechanisms for verifying decisions, the possibility of human intervention, and a system for appealing decisions made by algorithms [3-5].

To ensure the safe and equitable use of AI, it is necessary to create regulatory and ethical frameworks that govern the development and use of intelligent systems. Such a framework should include the principles of transparency, impartiality, protection of user data, as well as the responsibility of developers for the consequences of their algorithms.

A key element is the introduction of international standards that ensure uniformity of approaches to AI regulation. For example, the development of codes of ethics for developers and users of AI will minimize the risks associated with uncontrolled decisions of automated systems.

It is also important to establish independent expert committees that will assess the compliance of algorithms with ethical standards before implementing them in critical areas. The introduction of AI audit and certification mechanisms will provide additional transparency and trust from society [5-8].

International organizations, including the UN and the European Union, are already developing regulatory documents aimed at ethical regulation of AI. However, the question of the implementation of these norms in national legislation and their adaptation to the specific conditions of different countries remains.

To ensure the safe and equitable use of AI, it is necessary to create regulatory and ethical frameworks that govern the development and use of intelligent systems. Such a framework should include the principles of transparency, impartiality, protection of user data, as well as the responsibility of developers for the consequences of their algorithms. International organizations, including the UN and the European Union, are already developing regulatory documents aimed at ethical regulation of AI, but the question of their implementation remains open.

Conclusion

The ethics of using artificial intelligence in decision-making is the most important problem of modern society, which requires careful consideration and development of regulatory mechanisms. The balance between automation and human control must be based on the principles of transparency, fairness and accountability. The complete transfer of decision-making to AI carries risks associated with the possible bias

of algorithms, the lack of moral values in machines, and the difficulty of determining responsibility for the consequences of such decisions.

The best approach is to develop hybrid systems in which AI plays a supporting role and humans retain control over critical decisions. This will minimize the risks associated with the incorrect operation of algorithms, and at the same time use the advantages of AI in data processing and improving the efficiency of processes.

Thus, the future application of AI in decision-making must be based on a judicious combination of technological capabilities and ethical principles that ensure trust, security, and fairness in society.

Bibliography

1. A. Asmolov, G. Soldatova, Man in the Digital World: Artificial Intelligence and Ethics. – Moscow: Kogito-Center, 2021.
2. N. Bostrom, Artificial Intelligence: Stages, Threats, Strategies. – Moscow: Alpina Publisher, 2019.
3. Brignolle P., Petty R. "Psychology of Decision Making: Rational and Irrational Strategies". – St. Petersburg: Piter, 2020.
4. Wendland R. "Ethics of Digital Technologies and Artificial Intelligence". – Moscow: Yurayt, 2022.
5. E. Gurova, Artificial Intelligence and Morality: Opportunities and Limitations. – St. Petersburg: Saint Petersburg State University Publishing House, 2020.
6. A. Deryugina, Social Ethics and Technology: Challenges of the Digital Age. – Moscow: Nauka, 2021.
7. Kaplan J. "Artificial Intelligence: What Everyone Needs to Know". – Moscow: Alpina Publisher, 2020.
8. N. Lukyanchenko, Digital Ethics and Artificial Intelligence. – Kazan: Kazan Federal University, 2022.

USING A CHATBOT SYSTEM TO AUTOMATE MARKETING PROCESSES

Georgiy Chabanenko
Saint Petersburg State University of Aerospace Instrumentation
E-mail: Chagosha@gmail.com

Abstract. *This article describes the chatbot technology. A description of the existing types of chatbots is given, as well as their possible functionality is listed. The main systems for managing chatbots are named.*

An example of human work and interaction with a chatbot system is Fig. 1.

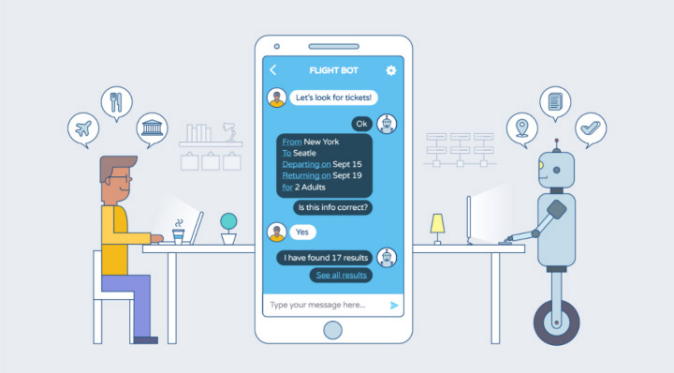


Fig. 1. Example of a chatbot

The result of scientific and technological progress has been the evolution of human behavior in the digital environment, which has expanded the subject and number of digital marketing tools. The continued growth in the number of "digital" and "social" users has contributed to the development of new patterns of user behavior. Many UXs researchers come to the conclusion that the current trend in the digital world offers a person to get any information without contact with a real user. Thanks to new technologies, users are able to control several parallel tasks and quickly get answers to their questions. In turn, the increasing demand for services capable of filtering the huge amount of data present on the web has led to the rapid proliferation of chatbots that can provide relevant information according to the context. Despite the rapid development of chatbots and the digital marketing industry as a whole, this technology does not yet have an unambiguous definition. Some researchers and practitioners believe that our ideas about the functionality of chatbots come from our approach to defining the concept of "context", which is the subject of their analysis. The concept of "context" involves the analysis of any information that can be used to characterize a particular subject. Therefore, the functioning of chatbots is based on the mechanism for receiving feedback on request in the form of some entities. An entity is a person, place, or object that is considered relevant to the interaction between the user and the application. Thus, in the broadest sense, a chatbot is a virtual assistant, a program that is created to imitate human behavior when communicating with one or more interlocutors. In the most relevant literature, chatbots are divided into simple and complex. Simple ones are characterized by the fact that they act according to a certain algorithm, which includes processes from the list, depending on the planned behavior of the user. Sophisticated bots are based on artificial intelligence. Their programs are able to learn independently in the process of communication. Therefore, complex bots can perform more tasks related not only to recognizing the meaning of text, but also photos, videos, and audio files.

Chatbots are widely used in social networks and instant messengers, they are on Facebook, WhatsApp, Twitter, Vkontakte, Odnoklassniki, Telegram, Viber, etc. It is important to emphasize that the effectiveness of the use of this technology depends on the scale of the company. If the company has few customers, then chatbots will not be able to bring the expected effect, and the management will only spend resources on development. 39 The main functionality of chatbots is as follows: maintaining a dialogue, integrating with the customer's systems, analyzing dialogues, connecting an operator and voice recognition. One of the main tasks of bots is the initial and subsequent consultation of customers, which allows you to more effectively integrate this tool into the company's CRM system. Chatbots also allow you to increase sales and expand your customer base [1].

On many sales sites and mobile applications, bots are also used to reduce the load on technical support. A similar function is performed by information bots, which, at the request of the user, can aggregate data in an easy-to-understand format. In general, the practical benefit of chatbots for digital marketing is the automation of marketing processes. There are a large number of chatbots on the chatbot market, from the simplest constructor in form to the most complex with the ability to modify third-party code. Among the current chatbot management systems, the most popular are Flow XO, ManyChat, Meya.ai, Chatfuel, leeloo.ai. The main advantage of Flow XO is the ability to create and host chatbots for various messengers and social networks. This control system has a fairly simple interface, so no additional programming knowledge is required to use it. Flow OX can be integrated with more than 90 services. ManyChat is a more highly specialized solution, as it only allows you to work with Facebook chats to automate them. Leeloo allows you to work simultaneously with four messengers: Facebook, Vkontakte, telegram and viber. It has an advanced analytics system and setting up the entrance to the automatic sales funnel. To create a chatbot structure, you need an entry point, that is, the user must start the process, after which he will receive a notification about a successful attempt and further feedback on the request. Since bots function on the basis of algorithms, it is necessary to think through all the user's steps and emerging cases in CJM. That is, write all the messages that the user will receive when performing certain actions. It is best to think over the entire structure of the chatbot in the form of a scheme, for example, with the help of services for generating mental maps (navigation). To form the structure, the following data are required: information about the product and categories of consumers (their portraits); frequently asked questions from users for; current promotions and bonuses; audience targets; repulsive parameters; Generation of feedback messages.

In the era of digitalization, business process automation is becoming a key tool for increasing efficiency. One of the most affordable and convenient ways to automate is a chatbot in Telegram. This tool allows you to streamline operations, reduce costs, and improve customer engagement.

Why do you need business process automation?

Using a chatbot in Telegram helps businesses:• Reduce personnel costs – the bot takes care of routine tasks.• Speed up customer service – instant responses without waiting.• Eliminate the human factor – reduce the number of errors.• Work 24/7 – the bot does not require rest and is available at any time.• Optimize sales and logistics processes – automatic ordering, delivery calculation, etc.

How does a chatbot work in Telegram?

Its functionality may include:1. Answers to frequently asked questions – the bot helps customers without the participation of a manager.2. Accepting orders and applications – users can place an order directly in the chat.3. Sending notifications – automatically informing customers about promotions and new products.4. Integration with CRM – the bot transmits customer data to the accounting system.5. Payment processing – accepting payments directly in Telegram.6. Collecting feedback – reviews and evaluations of the company's work.

Advantages of Telegram for business automation

Telegram is a convenient platform for automating business processes, because:• Popular among users – millions of people use the messenger every day.• Easy to develop – you can create a bot without deep programming knowledge.• Secure – data transmission to Telegram is encrypted.• Flexible in integration – you can connect CRM, payment systems, databases.

Stages of chatbot implementationIn order for the automation of business processes using a chatbot in Telegram to be successful, it is important to implement it correctly:1. Determine the goals – what tasks the bot should solve.2. Develop communication scenarios – think over the logic of dialogues.3. Create a bot – using programming or a constructor (ManyBot, BotFather, Chatfuel).4. Integrate with systems – connect CRM, databases, payment systems.5. Test and launch – check the performance before 6. Analyze and optimize – collect statistics, improve algorithms.

Examples of using chatbots in various fields

1. Online storesExample: "SmartTech electronics store "The chatbot helps customers choose products, find out about availability, place an order and pay for the purchase. The bot also sends notifications about the status of the order and offers personal discounts.

2. Scope of servicesExample: "Beauty salon BeautyPro"The bot allows customers to sign up for procedures, choose a master, receive reminders about the visit and leave a review.

3. Financial sectorExample: "FinBot Bank "The chatbot advises customers on tariffs, accepts loan applications, sends payment notifications and helps to calculate the loan burden.

4. HR and recruitmentExample: "HR agency WorkFinder"The bot conducts a survey of candidates, tests their skills and automatically sends suitable vacancies.

5. EducationExample: "EduBot Online School" The bot sends a class schedule, links to lessons, tests, and reminders for upcoming tests.

ConclusionAutomating business processes with a Telegram chatbot is a convenient way to increase business efficiency, reduce costs, and improve customer service. Bots help automate orders, customer service, marketing, and other processes.

Bibliography

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

FORMATION AND PROCESSING OF SIGNAL-CODE STRUCTURES IN SMALL-SIZED AIRBORNE RADAR STATIONS

Renata Chembarisova
Saint Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia
E-mail: magna7746@gmail.com

Abstract. *Modern small-sized airborne radar stations require increased resolution to obtain more detailed images. To do this, it is necessary to correctly select the signal modulation form, its duration and spectral characteristics. The use of complex probing signals with a wide radiation spectrum allows for increased accuracy in determining the coordinates of objects, which is especially important in conditions of limited resources and high accuracy requirements.*
Keywords: *small-sized airborne radar systems, probing signals, aviation control, M-sequences, correlation characteristics, phase modulation of signals.*

Introduction

In modern small-sized airborne radar systems, improving the resolution is important for obtaining detailed images of the earth's surface [1]. For this purpose, complex probing signals with a wide emission spectrum are used, which increases the accuracy of estimating the coordinates of objects.

The choice of the signal modulation method and its spectral characteristics plays an important role in resolving and detecting ground objects. Particular attention is paid to phase modulation methods, which allow signals to be distinguished from the background of internal receiver noise.

The development of modern aviation control radar systems requires the integration of technologies to improve the accuracy and reliability of signal processing. It is necessary to ensure high compression rates of phase-modulated signals by reducing the level of autocorrelation characteristics of modified M-sequences. A high level of side lobes of the autocorrelation function (ACF) can reduce the accuracy and reliability of aviation control systems, so this parameter must be considered when designing them.

Phase modulation of signal

In small-sized airborne radar systems for monitoring areas of the earth's surface, phase-shift keyed signals are often used, which are a sequence of elementary pulses with different initial phases changing from one element to another. The law of change of the manipulated parameter is completely determined by the code sequence used. In this regard, an urgent task arises of studying the properties of code sequences and choosing the most suitable one for solving a specific problem.

One example of such sequences is M-sequences, which are widely used in communication and radar systems due to their correlation characteristics [2-5], as well as the ability to minimize mutual interference and improve signal detection.

At present, several methods are known for modifying M-sequences using real values [6, 7]. In this article, to solve the problem under consideration, the found complex M-sequences and the corresponding values of the initial phases φ for negative elementary pulses, presented in Table 1, are used.

Table 1

Phase values φ for new complex M-sequences

Length of M-sequence, *N*	Obtained values of initial phases, φ
7	±41.41°
15	±28.955°
31	±20.36°
63	±14.354°

The methods of generation and analysis of correlation properties of signals modulated by phase using new complex M-sequences are considered. As an example, an M-sequence of length N = 31 is considered, the generating polynomial of which is: $x^5 + x^4 + x^3 + x^2 + 1$ [8-10]. For M-sequences of this length, the initial phase is φ = ±20.36°. All elements «0» of the classical sequence are replaced by the complex value «-0.9375-0.3479i». Thus, the new complex M-sequence has the form presented in Table 2.

Table 2

New complex M-sequence of length N = 31

Length, N	Generating polynomial	New complex M-sequence
31	[1 1 1 1 0 1]	-0.9375-0.3479i -0.9375-0.3479i -0.9375-0.3479i -0.9375-0.3479i 1 1 -0.9375-0.3479i -0.9375-0.3479i 1 -0.9375-0.3479i -0.9375-0.3479i 1 1 1 1 1 -0.9375-0.3479i 1 1 1 -0.9375-0.3479i -0.9375-0.3479i -0.9375-0.3479i 1 -0.9375-0.3479i 1 -0.9375-0.3479i 1 1 -0.9375-0.3479i 1

Fig. 1 shows the normalized ACF (NACF) of a modified M-sequence of length N = 31, as well as a signal modulated [11, 12] by this sequence.

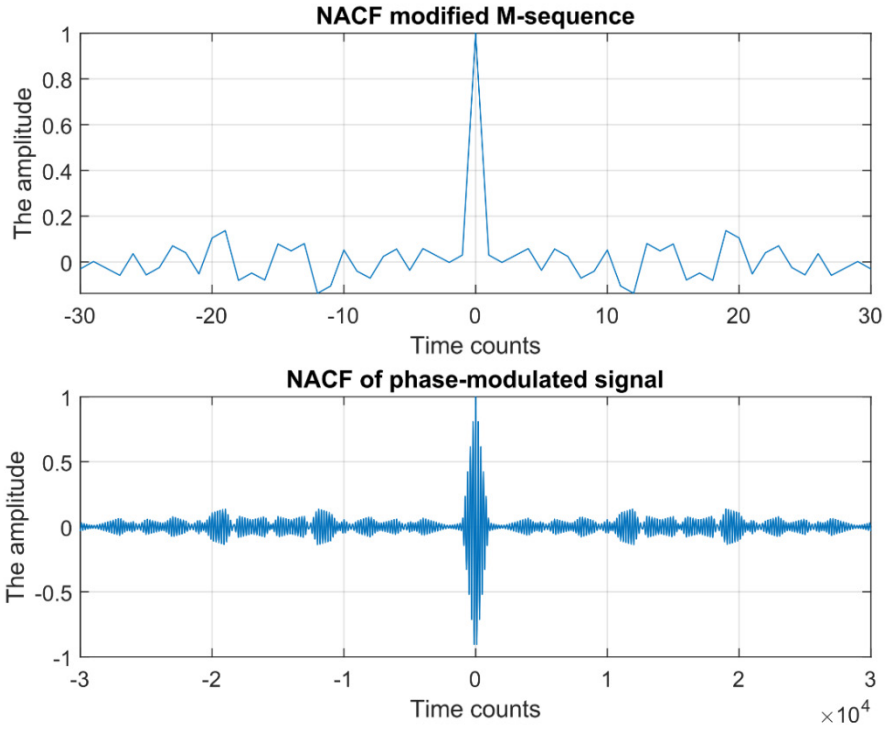


Fig. 1. NACF of the modified M-sequence and phase-modulated signal

After analyzing Fig. 1, we can conclude that the NACF of the modified M-sequence and the NACF of the signal phase-modulated by it correspond to each other, from which we can conclude that the signal is modeled correctly.

Conclusion

In the course of the study, an example of compression of a probing signal with phase modulation was considered, and a mechanism for its formation based on modified M-sequences was proposed. The analysis showed that the use of phase-shift keyed signals with new complex M-sequences ensures a low level of side lobes of the autocorrelation function.

Such signals can be used in aviation control radar systems to obtain detailed images of the earth's surface. The results obtained emphasize the need for further research in the field of searching for new code structures to improve the capabilities of radar technology.

References

1. I. Kudrya, E. F. Tolstov, and V. N. Chetverikov, "Expanding the capabilities of using M-sequences in RSA," in Proc. II All-Russian Armand Readings on Radio-Physical Methods in Remote Sensing, Murom, Russia, 26–28 June 2012, Murom: IPC MI VLGU, 2012, pp. 518–531.

2. T. D. Sokolov, N. A. Askerova, and A. A. Askerova, "Modeling pseudo-random sequences," Polytechnical Youth Journal, vol. 2022, no. 02(67), 2022. [Online]. Available: http://dx.doi.org/10.18698/2541-8009-2022-02-771

3. D. Zakharov and A. A. Ozhiganov, "Using generator polynomials of M-sequences for constructing pseudo-random code scales," Izvestiya Vuzov. Instrument Engineering, vol. 54, no. 6, pp. 49-55, 2011.

4. O. V. Ivantsov, D. E. Gorokhov, and M. N. Mishustin, "Correlation properties of symbol M-sequences in divergent decoding," Vestnik VGU. Series: System Analysis and Information Technologies, vol. 3, pp. 38-46, 2021. [Online]. Available: https://doi.org/10.17308/sait.2021.3/3734

5. A. Ozhiganov and Z. D. Zhuan, "Using pseudo-random sequences in constructing code scales for linear displacement transducers," Izvestiya Vuzov. Instrument Engineering, vol. 51, no. 7, pp. 28-33, 2008.

6. T. D. Bhatt, "Construction of perfect periodic binary sequences for radar applications," in 2020 IEEE International Radar Conference, 2020, pp. 662-667. [Online]. Available: https://www.researchgate.net/publication/343524590

7. V. A. Nenashev and S. A. Nenashev, "Search and study of marked code structures for a spatially distributed system of small-sized airborne radars," Sensors, vol. 23, no. 15, p. 6835, 2023. [Online]. Available: https://doi.org/10.3390/s23156835

8. V. A. Nenashev and S. A. Nenashev, "Search and study of marked code structures for a spatially distributed system of small-sized airborne radars," Sensors, vol. 23, no. 15, p. 6835, 2023. [Online]. Available: https://doi.org/10.3390/s23156835

9. Z. R. Garifullina, M. A. Ivanov, V. E. Ryabkov, and I. V. Chugunkov, "Method for generating nonlinear M-sequences," Information Technology Security, vol. 18, no. 2, pp. 31-36, 2001.

10. L. E. Varakin, Communication Systems with Noise-Like Signals, Moscow: Radio i svyaz, 1985, 384 p.

11. U. Khazikaramov and A. R. Sagitova, "Methods of modulation," Vestnik Magisterium, no. 9-1 (96), pp. 5-6, 2019.

12. V. E. Vakhtin, E. S. Lebedev, and D. A. Bobrov, "Signal modulation and its types: a comparison," Young Scientist, no. 33 (480), pp. 66-68, 2023. [Online]. Available: https://moluch.ru/archive/480/105521

COMPARISON OF CONVOLUTIONAL NEURAL NETWORK MODELS FOR THE POSSIBILITY OF HUMAN DETECTION IN THERMAL IMAGING DATA USING KERAS AND TENSORFLOW

Vlas Dianov
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: vlasd2863@gmail.com

Abstract. *Thermal imaging devices are widely used in various fields such as medicine, industry, and agriculture. This article discusses the comparison of pre-trained convolutional neural network models implemented in Python using the Keras library based on TensorFlow for the analysis and processing of thermal imaging data.*
Keywords: *neural networks, thermal imaging, Keras, TensorFlow, Python, model comparison.*

Introduction

Thermal imaging is a technology designed to obtain information about the temperature characteristics of objects, which is essential for automated monitoring of specific tasks. The main challenge when using thermal imaging data is the need for effective integration of image processing algorithms with software for data analysis. To address this issue, convolutional neural networks can be used, which can increase accuracy, reduce computational costs, and minimize human error.

To improve the quality of human detection in thermal images, a study of convolutional neural networks was conducted. For this purpose, the following tasks were set::

- Conduct a review of the Keras library to analyze existing neural networks capable of processing thermal imaging data;
- Test the selected convolutional neural network models;
- Compare the obtained data to identify the best model for detecting humans in thermal images.

The neural network should ensure high efficiency in processing thermal images, demonstrate ease of use for implementing such systems on devices, and confirm the feasibility of using this technology in real-world applications.

Neural networks for object detection in thermal images

Keras, serving as a high-level interface for TensorFlow, provides tools for building and adapting complex neural network architectures capable of effectively processing the specific characteristics of thermal images. The use of pre-trained models significantly reduces computational costs and accelerates the deployment of detection systems. Thermal images are characterized by low contrast, which leads to a significant level of noise and complicates their processing using traditional methods [1].

Thermal images are characterized by low contrast, high dependence on external conditions, and pronounced noise components, which complicates their processing using traditional methods. The use of deep convolutional neural networks in the Keras environment ensures robustness to distortions and variability in input data. The ResNet50 architecture, based on residual connections, demonstrates a high ability to generalize information, which is critically important for analyzing infrared images. VGG-19, with its deep sequential structure, allows for the extraction of detailed spatial features, improving the accuracy of object classification and segmentation in thermal images. YOLO v8, belonging to the family of single-stage detection models, provides a balance between speed and detection accuracy, making it particularly promising for real-time tasks [2].

One of the key advantages of using Keras is the ability to fine-tune pre-trained models by retraining them on specialized datasets of thermal images. Transfer learning mechanisms allow for effective adaptation of ResNet50, VGG-19, and YOLO v8 architectures to the specifics of the infrared spectrum, which improves the quality of object recognition in low-light conditions and complex meteorological environments.

Experiment: Comparison of Convolutional Neural Network Models for Human Detection Capability

An experiment was conducted using various neural network models. The experiment was carried out in a Python environment. The implementation of the experiment was done on the Google Colab Research

platform—a remote Jupyter Notebook server. This service provides allocated resources of 4 GB GPU, 12 GB RAM, and CPU. Such resources enable the implementation of a wide range of neural networks.

The first part of the experiment involved the ResNet50, VGG-19 models, and a custom convolutional neural network architecture. The dataset used for training was LLVIP, containing 16,000 thermal images. To expand the dataset, data augmentation methods were applied, which increased the training dataset threefold [3-4].

At the first stage of the experiment, during the testing of the ResNet50 model (Table 1), a decrease in errors for object coordinate regression (box_loss) and classification (cls_loss) is observed, indicating inefficient adaptation of the model to the specifics of thermal images. The low recall value suggests that the model misses a large number of objects [5]. Meanwhile, precision fluctuates within a narrow range but at a low value, indicating the model's stability but low prediction accuracy. These results point to the low effectiveness of the model.

Table 1

Experimental Results for the ResNet50 Model

epoch	box_loss	cls_loss	dfl_loss	precision(B)	recall(B)
1	2.5037	2.7521	1.5991	0.2714	0.0451
2	2.3104	2.534	1.2553	0.323	0.0523
3	2.1267	2.2722	1.2012	0.4571	0.0576
4	2.0876	2.142	1.1456	0.5211	0.0598
5	1.9765	1.8992	1.0541	0.6137	0.0611

Next, testing of the VGG-19 model was conducted. It is observed (Table 2) that there is a high level of errors in object coordinate regression (box_loss) and classification (cls_loss), indicating several shortcomings of the model, such as its inability to adapt to such a complex task and insufficient training data. The low recall value suggests a low percentage of correctly processed images relative to the total number. At the same time, precision is too low, which may indicate the model's low accuracy.

Table 2

Experimental Results for the VGG-19 Model

epoch	box_loss	cls_loss	dfl_loss	precision(B)	recall(B)
1	110.9294	93.8754	41.8158	0.1512	0.0251
2	97.1213	91.8224	35.8158	0.2143	0.0523
3	76.8991	55.4152	17.1089	0.3414	0.0376
4	56.5453	44.5732	10.1206	0.4809	0.0598
5	55.4272	37.4221	8.3298	0.5201	0.0411

During the testing phase of the model with a custom architecture, a decrease in errors for object coordinate regression (box_loss) and classification (cls_loss) is observed (Table 3), indicating the model's adaptation to thermal imaging data. An increase in the recall value is also noted, suggesting that the model improves the percentage of correctly detected objects. At the same time, precision systematically increases, indicating high-quality model training. Based on the results of testing the custom architecture model, it can be concluded that it demonstrates relatively higher training efficiency compared to the other models.

Table 3

Experimental Results for the Custom Architecture

epoch	box_loss	cls_loss	dfl_loss	precision(B)	recall(B)
1	1.9012	2.5678	4.9101	0.2345	0.1789
2	1.7890	2.2890	4.1234	0.3567	0.1901
3	1.5678	1.9012	3.3456	0.3789	0.2123
4	1.3456	1.6234	2.5678	0.4901	0.2345
5	1.1234	1.3456	1.7890	0.5123	0.2567

In the second part of the experiment, the YOLO model and a custom dataset containing a sample of over 15,000 infrared spectrum images were used.

In this part of the experiment, a systematic decrease in errors for object coordinate regression (box_loss) and classification (cls_loss) is observed (Table 4), indicating the model's adaptation to the specifics of thermal images. At the same time, there is a trend of increasing recall, suggesting a growing proportion of correctly detected objects as the number of training epochs increases. Precision fluctuates within a relatively narrow range, reflecting the model's stable ability to correctly identify detected objects. The combination of positive metric trends and reduced losses confirms the high effectiveness of the model.

Table 4

Experimental Results for the YOLO Model.

epoch	box_loss	cls_loss	dfl_loss	precision(B)	recall(B)
1	1.8031	3.5419	1.0723	0.4366	0.1038
2	1.4469	1.4486	0.9968	0.4719	0.1172
3	1.3285	1.1694	0.9461	0.601	0.1206
4	1.2817	1.0698	0.9345	0.6431	0.1258
5	1.2103	0.9998	0.9221	0.7547	0.1223

Fig. 1 shows the result of human detection using the YOLO model. As a result of the YOLO model's operation, an image is produced with a highlighted region identified as an area containing people.

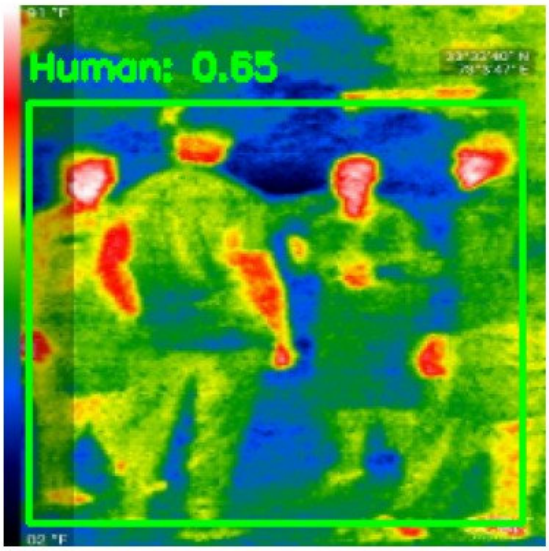


Fig. 1. Recognition Result

Conclusion

During the experiment, four convolutional neural network models were tested: VGG-19, ResNet50, YOLO, and a custom architecture. Each of these models has unique characteristics that determine their applicability in thermal image processing tasks.

During the experiment, VGG-19 performed well; however, due to its depth and large number of parameters, VGG-19 requires significant computational resources. Without these resources, the model cannot be sufficiently trained.

The ResNet50 neural network proved to be more efficient. However, like VGG-19, ResNet50 also requires significant resources for data processing.

The custom architecture, specifically developed for the experiment, was designed with the task and data specifics in mind, allowing it to become a sufficiently effective model. However, despite being tailored to specific conditions, the custom model fell short of YOLO v8 in terms of processing speed and accuracy.

Based on the experiment, YOLO v8 demonstrated the best results among all tested models. Its high speed, accuracy, and versatility make it the optimal choice for tasks involving real-time thermal image

processing. While VGG-19 and ResNet50 remain powerful tools for tasks requiring high detail, their application is limited due to high computational resource demands. The custom architecture, although it performed well, requires further optimization to achieve results comparable to YOLO v8.

References

1. Nikolenko S., Kadurin A., Arkhangelskaya E. Deep Learning. Immersion in the World of Neural Networks. – 3rd ed. – Moscow, St. Petersburg : Peter, 2018. – 480 p.

2. Fisenko V.T. Computer Processing and Image Recognition / V.T. Fisenko, T.Yu. Fisenko: textbook. – St. Petersburg: ITMO University, 2008. – 192 p.

3. Grigoriadis G.M., Kaliberda I.V. Application of Neural Network Technologies for Analysis and Pattern Recognition in Thermal Imaging // Control Systems, Complex Systems: Modeling, Stability, Stabilization, Intelligent Technologies. – Yelets: Yelets State University named after I.A. Bunin, 2023. – 182 p.

4. Logachevsky, I. A. Analysis of Thermal Imaging / I. A. Logachevsky // Soft Measurements and Computing. – 2022. – Vol. 57, No. 8. – Pp. 18-30. – DOI 10.36871/2618-9976.2022.08.002.

5. Borodin, G. D. A Brief Review and Classification of Artificial Neural Networks / G. D. Borodin // News of Tula State University. Technical Sciences. – 2021. – No. 11. – Pp. 45-53.

GENERAL ARTIFICIAL INTELLIGENCE: OPPORTUNITIES, RISKS AND THE FUTURE OF DIFFERENT INDUSTRIES

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

Abstract. *The article is devoted to analyzing the opportunities, risks, and prospects of implementing general artificial intelligence (AGI) in various fields of activity. The potential benefits of AGI are discussed, including increased efficiency, personalization of processes, optimization of workflows, and improved quality of decision-making. Special attention is paid to the impact of AGI on education, medicine, industry, and business, where its application can lead to significant changes. At the same time, the author emphasizes the importance of developing control and security mechanisms and preserving the human role in key processes. The potential threats associated with AGI autonomy and the need for an ethical approach to its use are discussed. The conclusion is that successful implementation of AGI requires a balanced approach that combines innovation and precautions.*
Keywords: *general artificial intelligence, AGI, automation, optimization, data analytics, education, medicine, industry, business, risk, security, ethics, technology adoption.*

AGI (Artificial General Intelligence, “general artificial intelligence”) is one concept for the theoretical study of artificial intelligence, which has general intellectual abilities comparable or superior to human intelligence. [1]

The capabilities of this artificial intelligence are vast and can be utilized in various fields. For example, if AGI is trained with the necessary knowledge, it can become a useful assistant in professional activities. It can be used to develop individualized development plans and strategies based on everyone's needs, abilities, and pace, which will help to increase efficiency and engagement. AGI is also capable of analyzing large amounts of data, identifying hidden patterns, and providing recommendations to help optimize workflows and make more informed decisions. This intelligence can be a valuable tool for increasing productivity, improving work quality and reducing time costs, which opens new horizons for various industries.

Also, despite the enormous potential of general AI in education, its implementation will require careful planning and the development of robust controls and security mechanisms. It is important that the use of AI is not just about automating processes, but also about preserving the role of humans in learning, their ability to mentor and personally interact with students.

In the medical field, AGI has the potential to make a significant difference. AI can improve diagnosis by analyzing large amounts of data, detecting diseases in their early stages, and recommending personalized treatments. Moreover, AI can be used to create robotic solutions, increasing the accuracy of operations and minimizing human error. However, despite the high level of accuracy, AI cannot replace the doctor's personal relationship with the patient, which remains an important aspect in medicine. Empathy and attentiveness to the patient is something that cannot be replicated by machines.

Industry will also benefit greatly from the adoption of general artificial intelligence, especially in the area of process automation and optimization. AI can predict equipment breakdowns and optimize production lines, which helps increase productivity and reduce costs. In addition, AI can manage risk and improve safety in manufacturing facilities. However, despite all these benefits, it is important to remember that you can't rely entirely on machines to solve complex problems, and there is always a need for human oversight.

In business, AI has the potential to change the way decisions are made, improve analytical capabilities and speed up decision-making processes. AI will help companies accurately predict market needs, build effective strategies and automate many processes, such as accounting or recruitment.

The introduction of general artificial intelligence, despite all its prospects, involves certain risks. One of the main challenges is ensuring security and control over AI actions. So far, there are no universal methods that can ensure that artificial intelligence always acts in the interests of users and society, rather than pursuing its own goals. Cybersecurity experts continue to look for solutions to prevent AI from being used for harm, but the problem remains.

In summary, general artificial intelligence has huge potential to improve various sectors, from medicine and education to industry and business. However, its implementation must be careful and

thoughtful. It is important to develop control and security mechanisms to minimize possible risks. And while AI can significantly improve the efficiency and quality of work in different areas, it should not replace human communication and personal involvement in the process.

References

1. Jaspreet. Artificial general intelligence research: fact or fiction? URL: https://aimojo.io/ru/artificial-general-intelligence/

THE PROCESS OF COLLECTING A DATASET FOR TRAINING A NEURAL NETWORK FOR CLASSIFICATION BASIC ALGORITHMS AND DATA STRUCTURES WRITTEN IN THE PYTHON PROGRAMMING LANGUAGE

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

Abstract. *This paper presents the development of a neural network classifier for recognizing algorithms and data structures in the Python language, aimed at automating the verification of laboratory works in the subject "Algorithms and Data Structures". One of the key stages of the research was the formation of a specialized dataset including more than 2000 algorithms and data structures. To improve the quality of model training, a method of representing each algorithm in two versions: with meaningful variables and with variables devoid of semantic information was proposed. This approach improved the generalization ability of the model and increased robustness to code changes. The final dataset was divided into training and validation samples, and supplemented with a separate dataset for pre-training. The methods used significantly improved the accuracy of algorithm classification, which is an important step in the development of automated code analysis systems.*
Keywords: *neural network classifier, algorithms, data structures, Python, machine learning, dataset, automated validation, code analysis, classification, model training*

To optimize the verification of student's laboratory work in the subject "Algorithms and Data Structures" the task was set: to create a neural network – a classifier for recognizing algorithms and data structures in the Python programming language. The task is especially relevant for sorting algorithms, search and other categories often encountered in programming [1].

One of the key stages of the work was the formation of a specialized dataset containing more than 2000 different algorithms and data structures that reflect the characteristic features of their class [2]. This process proved to be complex and involved numerous experiments aimed at finding optimal approaches to dataset generation. The initially collected dataset consisted of standard algorithm implementations, but it had significant limitations: it contained frequently repeated variables and some algorithms lacked sufficient uniqueness. This had a negative impact on the quality of the model training, as it could not effectively generalize the data and classify algorithms correctly.

To minimize the influence of variables on code structure, the concept of creating a dataset consisting solely of control constructs such as conditional statements, loops, and other flow controls was proposed. However, this method did not yield significant improvements because the algorithms remained structurally similar and some of them contained borrowed elements, which reduced the classification accuracy.

The next stage of the research was to investigate methods that allow variables in the code to be accounted for, but minimize their impact on the model predictions. Two types of data were developed:

- Algorithms with meaningful variables, where variables had logical and understandable names (e.g. count, index, result). An example based on the Bubble Sort algorithm can be seen in Fig. 1.
- Renamed algorithms where variables were replaced by random character sets that carried no semantic information (e.g., vqwer1, xzy123). An example based on the "Stirred Sorting" algorithm can be seen in Fig. 2.

```
def bubble_sort(numbers):
    outer = 0
    while outer < len(numbers) - 1:
        inner = 0
        while inner < len(numbers) - outer - 1:
            if numbers[inner] > numbers[inner + 1]:
                numbers[inner], numbers[inner + 1] = numbers[inner + 1], numbers[inner]
            inner += 1
        outer += 1
```

Fig. 1. «Bubble sorting» with meaningful variables

```
def fdsazxcv(qwr):
    rfe = len(qwr)
    ght = 0
    bvc = rfe - 1

    while ght < bvc:
        lkj = ght
        for mkj in range(ght, bvc):
            if qwr[mkj] > qwr[mkj + 1]:
                qwr[mkj], qwr[mkj + 1] = qwr[mkj + 1], qwr[mkj]
                lkj = mkj
        bvc = lkj
        plm = bvc
        for mkj in range(bvc - 1, ght - 1, -1):
            if qwr[mkj] > qwr[mkj + 1]:
                qwr[mkj], qwr[mkj + 1] = qwr[mkj + 1], qwr[mkj]
                plm = mkj
        ght = plm
```

Fig. 2. «Mixing sorting» with nonsensical variables

This approach allowed us to create a training dataset in which each function was represented by two versions: with meaningful variables and with variables devoid of meaning. Thanks to this, the model learned to analyze the structure of the algorithm independently of variable names, which increased its ability to generalize and its robustness to code changes.

Data assembly was performed in a semi-manual format and was time-consuming. Approximately 100 unique implementations with variations in variable and function names were created for each category of algorithms. This process involved writing a large amount of code preserving the logic of the algorithms, but with modified variable names. Subsequently, new classes of algorithms were added to the dataset, also including about 100 unique implementations.

To simplify the process, a Python program was developed that automatically selected algorithms for the dataset, creating their codes and allocating them to the appropriate directories. In cases where an algorithm did not match any of the known classes, it was assigned a special class “-”, which allowed algorithms outside the standard categories to be considered.

The final dataset was divided into training (80%) and validation (20%) samples. The training sample was dominated by algorithms with meaningful variables, while the validation sample was dominated by versions with renamed variables. This method allowed us to check how robust the model is to changes in the structure and content of the code.

Additionally, a separate dataset was assembled for pre-training, which excluded keywords such as “def” and function names. This was done in order to minimize the risk of misclassifying algorithms based on the same function names. Excluding these elements allowed the model to focus on analyzing the structure of the algorithm and improve classification accuracy.

The approaches used significantly improved the accuracy of the model and its ability to generalize. They made the model less sensitive to changes in variable and function names, which is an important step towards creating a universal algorithm recognition system. In the future, we plan to further expand the dataset by adding new classes of algorithms, as well as to investigate the possibilities of automated code generation to increase the diversity of data.

References

1. Golovkin M.M., Tatarnikova T.M. Neural network approaches to identification of classical algorithms in Python. In Collection: Applied Artificial Intelligence: Prospects and Risks. Collection of reports of the International Scientific Conference. St. Petersburg, 2024. C. 232-233.

2. Tatarnikova T.M., Bimbetov F., Bogdanov P.Y. Выявление аномалий сетевого трафика методом глубокого обучения // Известия СПбГЭТУ ЛЭТИ. 2021. № 4. С. 36-41.

MODELING OF A TRAFFIC LIGHT OBJECT

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

Abstract. *This work discusses the process of modeling a traffic light object based on a microcontroller. It justifies the selection of specific tools, modules, and components for development and includes testing of the developed model.*
Keywords: *modeling, microcontroller programming, Keil µVision5, C programming language, STM32.*

Introduction

A traffic light is a light-signaling device used to regulate the sequence of passage for vehicles and pedestrians [1].

A traffic light object is a group of road traffic lights installed on a section of the street-road network, where the sequence of movement for conflicting vehicles or vehicles and pedestrians is regulated by traffic light signaling operating in a unified cycle [1]. An example of a traffic light object is a regulated intersection.

Traffic lights play a key role in organizing road traffic, ensuring safety and efficiency of traffic flow both within cities and beyond. Advances in microelectronics and related technologies have significantly simplified and improved the management of such systems, opening new possibilities for their automation and integration with modern information networks.

This paper describes the process of modeling a traffic light object installed at an intersection, equipped with four three-section traffic lights to regulate the sequence of vehicle passage and eight three-section traffic lights to ensure safe pedestrian crossing.

Thus, the goal of this work is to create a microcontroller-based model of a traffic light object.

Problem statement and component selection

The primary function of the designed model is to regulate the sequence of passage for vehicles and pedestrians. The device's task is to timely activate and deactivate the corresponding traffic light signals. Additionally, the model should provide the capability for manual signal control to forcibly activate a specific phase regardless of the current state.

In the designed model, the coordinating and controlling component is the 32-bit microcontroller (MCU) STM32H743VIT6 [2]. This MCU was chosen due to its high performance (ARM Cortex-M7 core with a clock speed of up to 480 MHz) and the presence of multiple SPI interfaces for connecting four displays.

As traffic lights for vehicle passage, an OLED display based on the ST7789 controller [3] was selected. This display operates via a standard SPI interface, simplifying its integration with the MCU. It features a fast response time, ensuring timely image updates, which is particularly important for simulating dynamic traffic light phase changes. OLED technology provides high contrast, deep black levels, and vibrant colors, allowing for clear distinction of displayed signals.

Additionally, it offers convenient software support for creating displayed images and their corresponding code. The designed model will include four such displays, one for each direction of the road.

For pedestrian passage, eight three-color LEDs, BL-L51SRGBW-CC, are used. These LEDs are easy to control, have small dimensions, low power consumption, and can display three colors: red, green, and yellow.

To implement the manual control function, a 1x3 matrix keyboard is used.

The connection diagram design

Based on the stated task and the aforementioned components, an electrical connection diagram was designed.

The connection diagram (assembly diagram) illustrates the connections between the components of the device (installation) and specifies the wires, harnesses, cables, or pipelines used for these connections, as

well as the points of attachment and entry [4]. It should depict all devices and elements included in the product, their input and output elements, and the connections between these devices and elements.

The main function of the device is performed by the microcontroller (MCU) M3. The MCU is powered via the XS1 connector – USB Micro B.

The matrix keyboard M2 is connected to the MCU through the XS2 connector.

Displays M4 and M7 are connected to the MCU via the SPI1 interface through connectors XS7 and XS10, while displays M5 and M6 are connected to the MCU via the SPI2 interface through connectors XS8 and XS9. Corresponding pairs of adjacent displays show the same light signals, meaning they are connected in parallel to each other.

The LEDs HL1 – HL8 are connected in such a way that the MCU controls only four of them, with the remaining LEDs connected in parallel to the corresponding adjacent LEDs, as shown in the diagram.

Current-limiting resistors R1 – R16 are installed between the anodes of each LED and the input/output ports to protect the MCU ports. When a logic high is applied to the line leading to the red crystal (R), the red color lights up. If a logic high is applied to the line leading to the green crystal (G), the green color lights up. When both outputs of the MCU are activated simultaneously, both crystals light up, and the LED emits a mixed color (yellowish-orange).

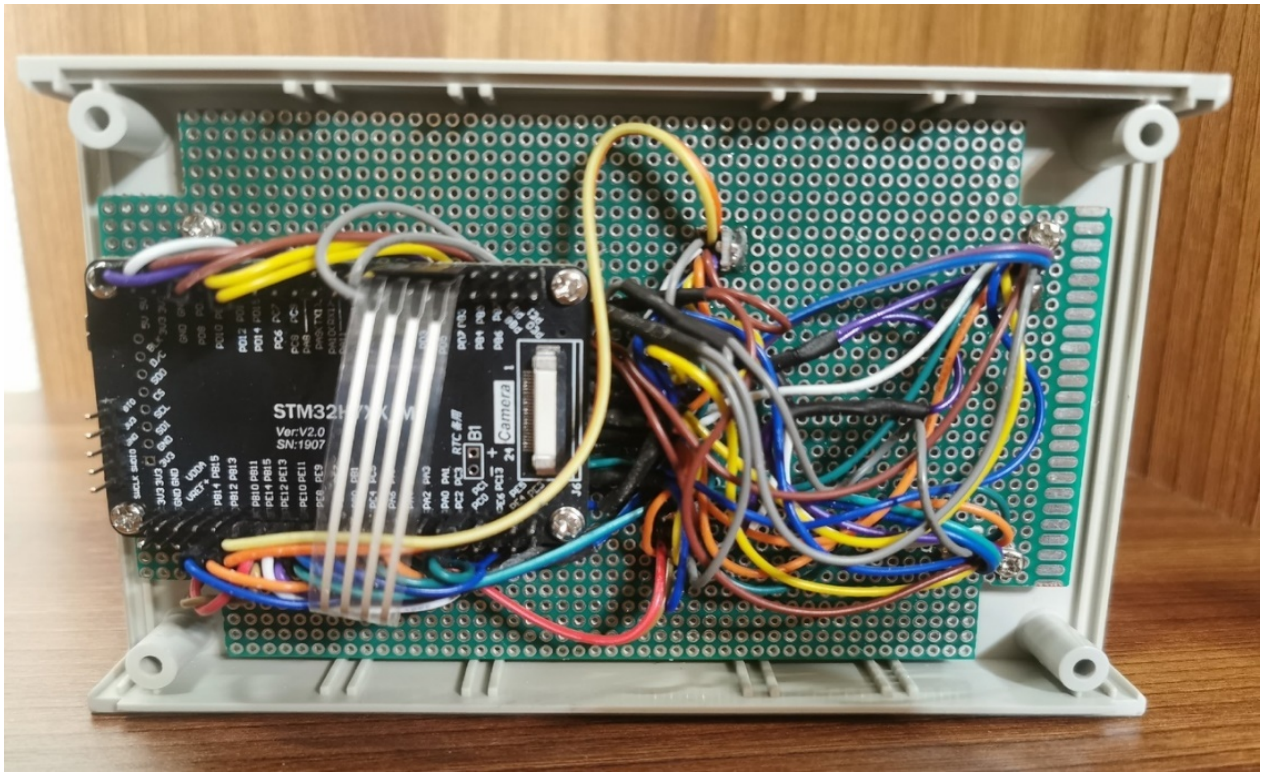


Fig. 1. Model in disassembled form

Development of the software model and debugging

Based on the constructed electrical connection diagram of the designed model, the development of the software model for the device is carried out.

To develop the software model, an algorithm for the device's operation is created, which will be used to write the program.

The core of the algorithm is based on the idea of switching traffic light phases using interrupts from timer TIM2, which overflows every 0.5 seconds. Additionally, timer TIM3 is used to poll the matrix keyboard every 273 microseconds.

The C programming language [5] is used to write the program executed by the MCU, as its design closely aligns with typical machine instructions and assembly language.

The STM32CubeMX tool is used for configuring and initializing the MCU, simplifying and automating the process of generating C code for use in various development environments.

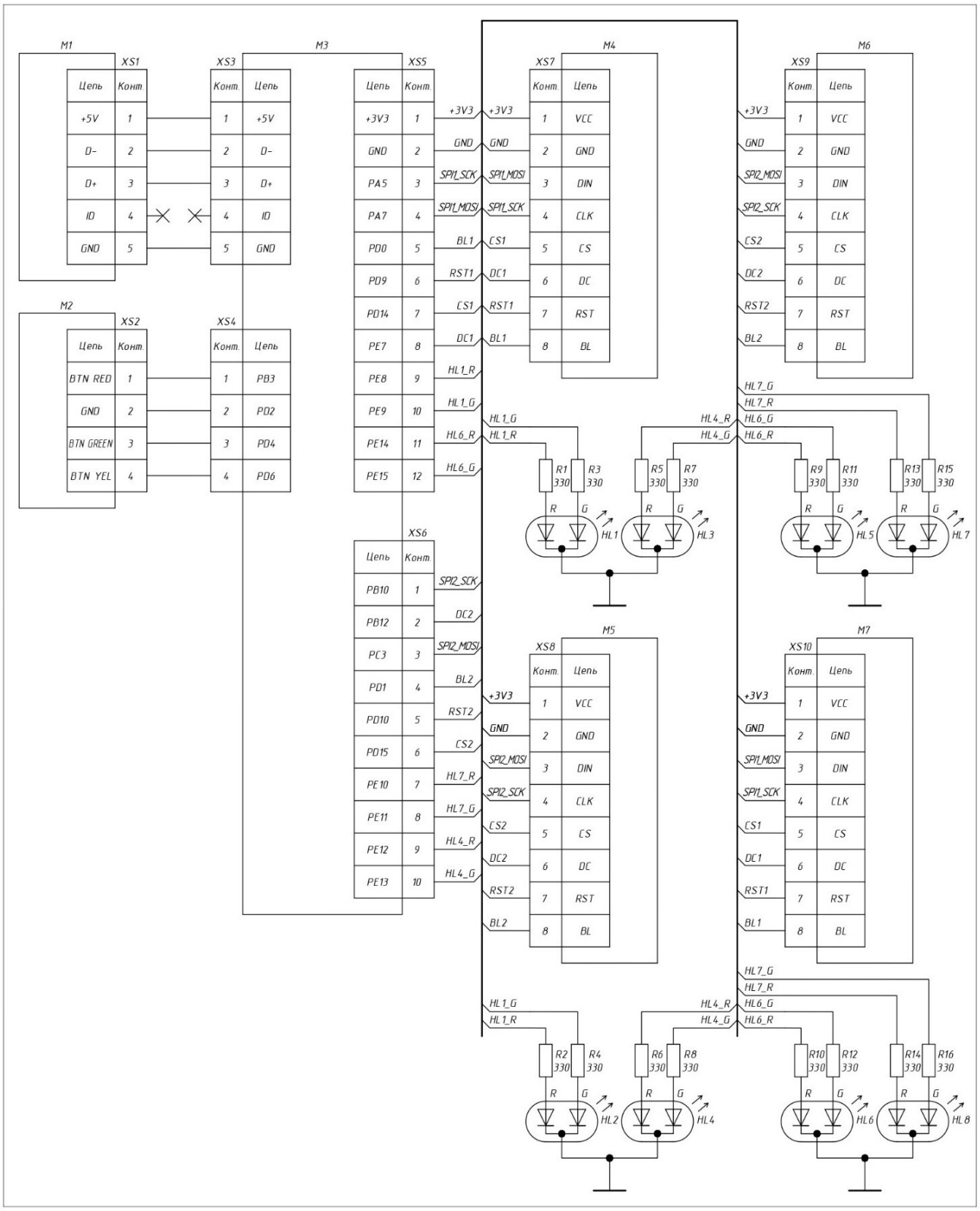


Fig. 2. A connection diagram

The Keil µVision5 environment is used for writing the program, as it supports the entire STM MCU family, allows for writing the program into the MCU's memory, and interfaces with many popular programmers.

```
245   case 1: //phase1 LCD_UD GREEN, LCD_LR RED
246
247     if (!draw_flag) {
248       draw_flag = 1;
249       DEV_Digital_Write(DEV_BL1_PIN, 0);
250       DEV_Digital_Write(DEV_BL2_PIN, 0);
251
252       LCDN=LCD_UD;
253       Paint_DrawCircle(LIGHT_XPOS, RED_YPOS, LIGHT_SIZE, BLACK, 2, DRAW_FILL_FULL);
254       Paint_DrawCircle(LIGHT_XPOS, YLW_YPOS, LIGHT_SIZE, BLACK, 2, DRAW_FILL_FULL);
255       Paint_DrawCircle(LIGHT_XPOS, GRN_YPOS, LIGHT_SIZE, GREEN, 2, DRAW_FILL_FULL);
256       LCDN=LCD_LR;
257       Paint_DrawCircle(LIGHT_XPOS, RED_YPOS, LIGHT_SIZE, RED, 2,   DRAW_FILL_FULL);
258       Paint_DrawCircle(LIGHT_XPOS, YLW_YPOS, LIGHT_SIZE, BLACK, 2, DRAW_FILL_FULL);
259       Paint_DrawCircle(LIGHT_XPOS, GRN_YPOS, LIGHT_SIZE, BLACK, 2, DRAW_FILL_FULL);
260       HAL_Delay(BL_DELAY);
261
262       DEV_Digital_Write(DEV_BL1_PIN, 1);
263       DEV_Digital_Write(DEV_BL2_PIN, 1);
264
265       LED_UL_RG_off; LED_UL_GR_on;   LED_UR_RG_off; LED_UR_GR_on;
266       LED_DL_RG_on;  LED_DL_GR_off;  LED_DR_RG_on;  LED_DR_GR_off;
267     }
268
269       sprintf(str, "%02d", phase_counter);
270
271       LCDN = LCD_UD;
272       Paint_DrawString_EN(STR_XPOS, STR_YPOS, str, &Font24, BLACK, GREEN);
273       LCDN = LCD_LR;
274       Paint_DrawString_EN(STR_XPOS, STR_YPOS, str, &Font24, BLACK, RED);
```

Fig. 3. Keil µVision5 development environment. Listing of the program code. The fragment

To verify the correct operation of the program, it is necessary to conduct testing of the model.

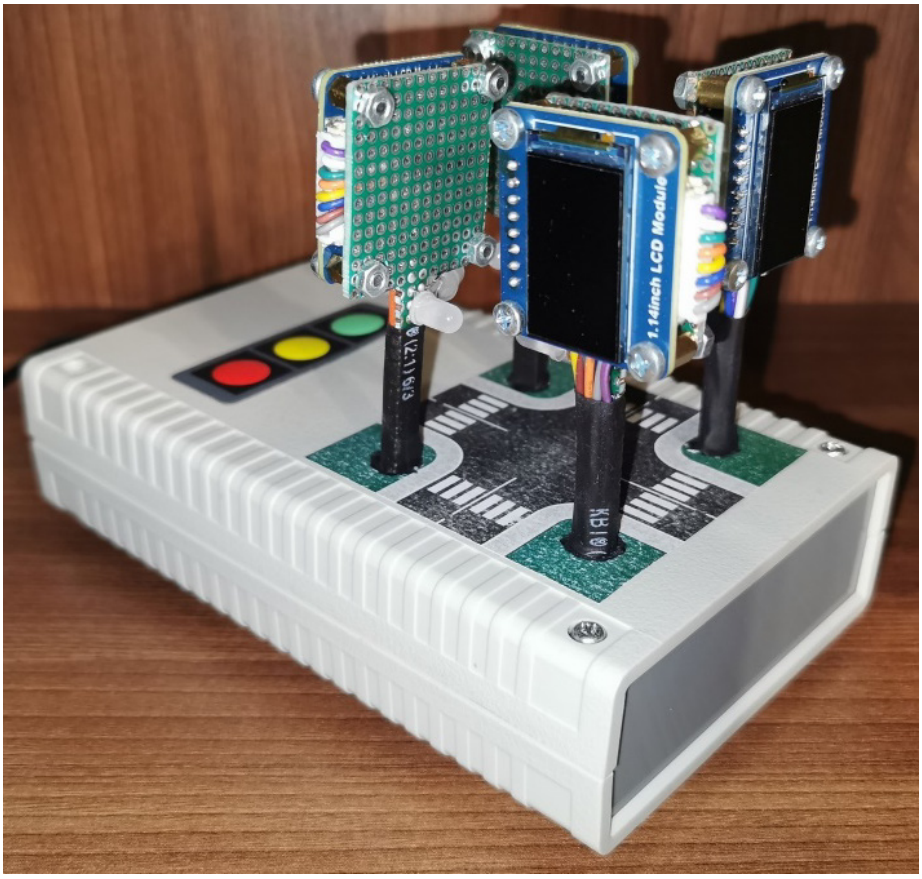
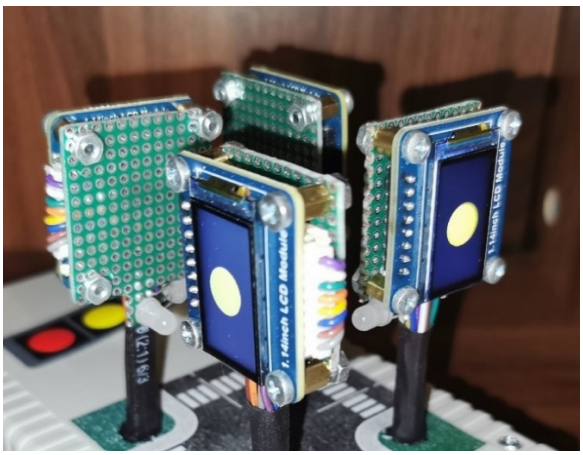
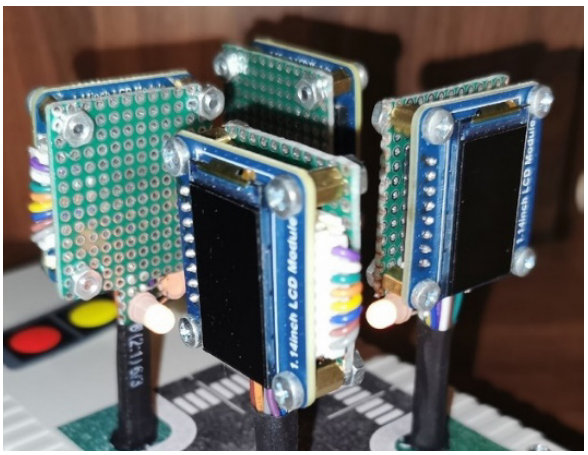


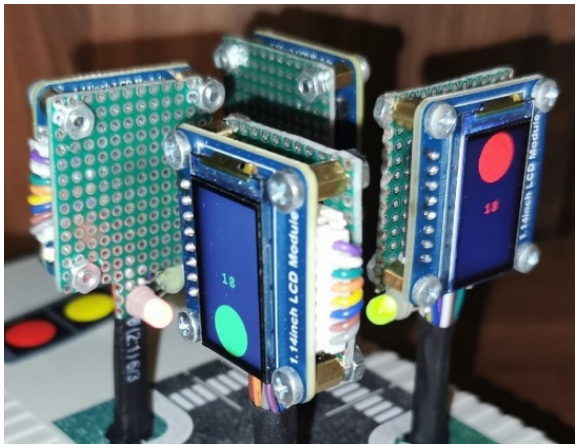
Fig. 4. Model in assembled form



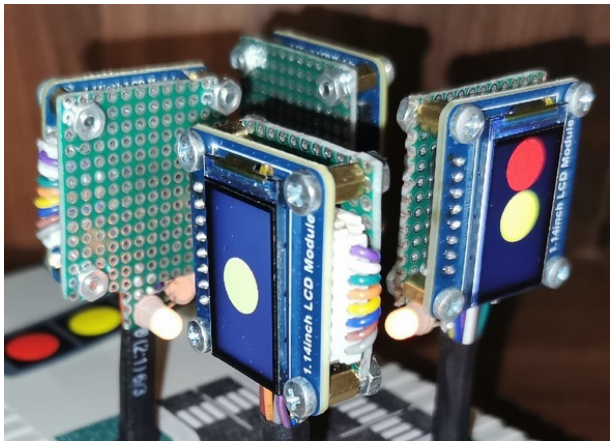
a) Yellow signal phase (transport)



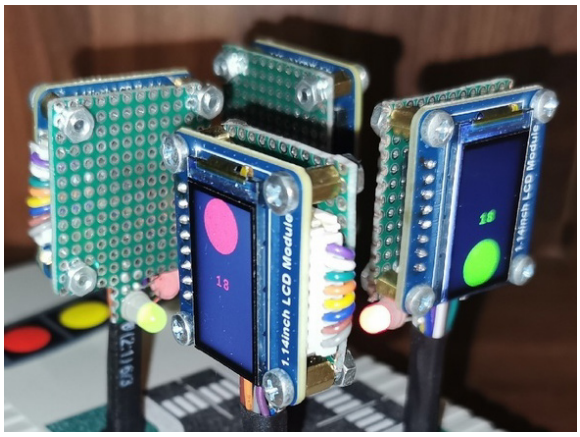
b) Yellow signal phase (pedestrians)



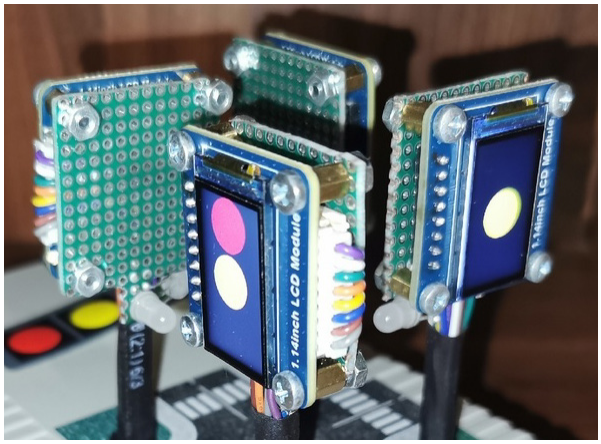
c) Movement phase I



d) Preparation phase I



e) Movement phase II



f) Preparation phase II

Fig. 5. Model testing results

To clearly demonstrate the operation of the traffic light within the computer-aided design (CAD) system AutoCAD 2024, an intersection diagram was created where the traffic light is planned to be installed. The image was printed and attached to the model's housing to visualize the traffic flow management at the traffic light location.

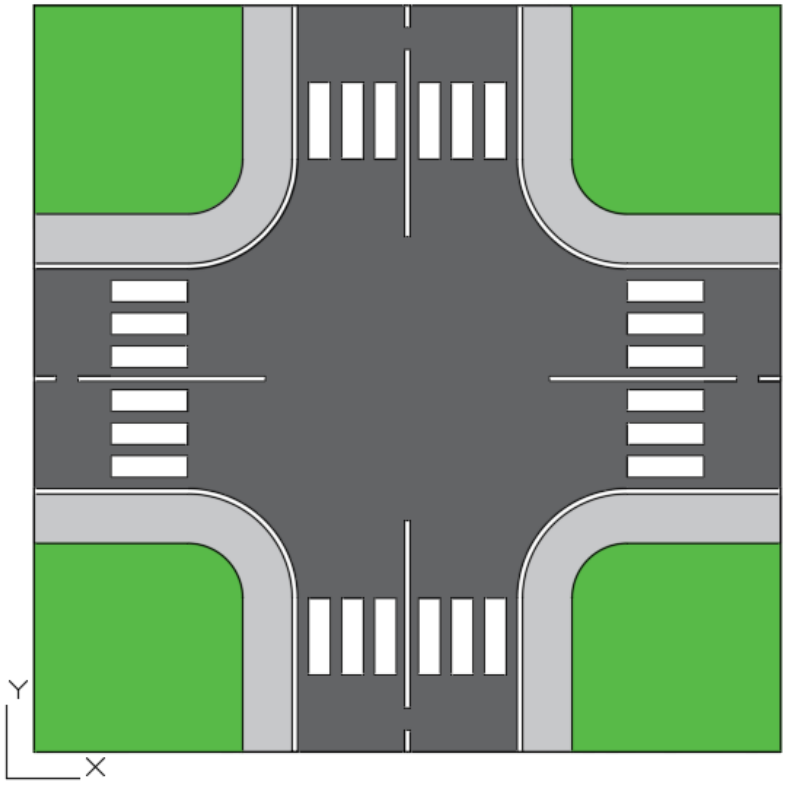


Fig. 6.. Drawing of the intersection for placement on the corpus

Conclusion

This work explored the method of modeling a traffic light object using a microcontroller. Based on the analysis of this method, the necessary components were selected, and an electrical connection diagram was designed.

A software model of the prototype was developed, including the creation of an algorithm for the program's execution and the writing of its code in the C language within the Keil µVision5 environment. Testing of the device confirmed that it operates correctly according to the designed algorithm.

References

1. Government standard 52289-2019 Traffic control devices. Rules of application of traffic signs, markings, traffic lights, guardrails and delineators. URL: https://www.mos.ru/upload/documents/files/9432/GOSTR52289-2019.pdf?ysclid=m7cykeh7pm830754891 (date of request: 02.01.2025).

2. STM32H743VIT6. Datasheet.

URL: https://www.st.com/resource/en/datasheet/stm32h743vi.pdf (date of request: 05.01.2025).

3. ST7789V. Datasheet. URL: https://newhavendisplay.com/content/datasheets/ST7789V.pdf (date of request: 05.02.2025).

4. 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: 10.01.2025).

5. Deitel Harvey. How to program C. URL: https://vk.com/doc7608079_448866001?hash=wXGAxIlNo8Eq1ErMEhVINVavyMaFpoohYREav9BSfzD&dl=Zw8yjOyoriF8oT25M9z2yomSMLXvMjZCx5mkogPrAgL (date of request: 21.01.2025).

RESEARCH OF THE NOISE IMMUNITY OF IMAGES MASKED BY HADAMARD MATRICES

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

Abstract. *This article describes an experiment aimed at evaluating the noise immunity of images masked by Hadamard matrices. As image transformation methods in this experiment, one-sided and two-sided masking were used. To simulate data loss in communication channel, a function was implemented. The evaluation of the restored image was conducted using metrics such as SSIM, PSNR, MS-SSIM, MSE, BRISQUE, and Pearson correlation coefficient. To illustrate the results, graphs were created to visualize the quality assessment of the restored image. The results of the work allow us to conclude that in the context of solving the problem of noise immunity, image masking is best implemented in a one-sided version, with a matrix of small sizes. The results of the work complement the known theoretical works devoted to matrix masking of images.*
Keywords: *matrix masking, quasi-orthogonal matrices, Hadamard matrices, noise immunity.*

Introduction

In modern science and technology, a significant portion of tasks related to information transformation and data analysis is directly associated with image processing and transmission. This is because visual data is one of the most informative and universal ways of presenting information, making them indispensable in various fields of knowledge.

Examples of such areas include remote sensing of the Earth, where the surface is scanned and analyzed using radar systems [1]. The resulting images allow for monitoring the environmental situation, studying climate changes, and addressing tasks related to cartography and geodesy. Equally important are images in medicine [2-3], where methods such as X-ray imaging, computed tomography, and magnetic resonance imaging are used for diagnosing diseases, planning surgeries, and monitoring patient conditions. The quality of internal organ and tissue visualization directly affects the accuracy of diagnosis and the effectiveness of treatment. Furthermore, image processing is widely used in research on biological and chemical processes [4-5]. For example, microscopy enables the study of cell and tissue structure, as well as the analysis of chemical reaction dynamics at the micro level. In these cases, the quality of images determines the ability to obtain reliable data and their subsequent interpretation.

Various image transformation methods can be used for noise-resistant image transmission over communication channels, particularly using orthogonal and quasi-orthogonal matrices [6-9]. An original solution in this field is the strip method of image transformation, proposed by Professors L.A. Mironovsky and V.A. Slaev [6]. This method divides the image into blocks and multiplies each block by a special matrix. For example, in the case of impulse noise, the use of Hadamard matrices allows for the perfect uniform distribution of a single impulse noise over the entire image area [6]. A disadvantage of this method is the need for Kronecker matrix multiplication. Further development of this solution is known as the "matrix masking method" [7-9], which addresses two tasks simultaneously: ensuring the confidentiality of information with a short lifespan and providing noise-resistant image transmission.

It should be noted that when using IP networks for data transmission, the occurrence of impulse noise, which is effectively suppressed by the strip method, is unlikely. Since data is transmitted in blocks, it is more likely that a row of the matrix representing the transformed image transmitted over the communication channel will be lost. Preliminary studies on the noise immunity of Hadamard-matrix-masked images in the case of transmission losses over a communication channel were conducted in the dissertation [7]. The analysis of the restored images showed that visually, the original image can be recognized, but characteristic stripes appear on it.

The aim of this work is to refine the results obtained in [6-7] using objective metrics for image restoration quality.

Metrics overview

1. SSIM – an image quality assessment metric based on three criteria: brightness, contrast, and structure. It takes values from 0 to 1, where a higher value indicates lower image distortion and higher quality.

$$SSIM = \left(\frac{\sigma_{XY}}{\sigma_X \sigma_Y}\right)\left(\frac{2\overline{X}\,\overline{Y}}{\overline{(X)}^2 + \overline{(Y)}^2}\right)\left(\frac{2\sigma_X \sigma_Y}{\sigma_X^2 + \sigma_Y^2}\right);$$

$$\overline{X} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} x_{ij}, \overline{Y} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} y_{ij};$$

$$\sigma_X^2 = \frac{1}{(M-1)(N-1)}\sum_{i=1}^{M}\sum_{j=1}^{N}(x_{ij} - \overline{X})^2;$$

$$\sigma_Y^2 = \frac{1}{(M-1)(N-1)}\sum_{i=1}^{M}\sum_{j=1}^{N}(y_{ij} - \overline{Y})^2;$$

$$\sigma_{XY} = \frac{1}{(M-1)(N-1)}\sum_{i=1}^{M}\sum_{j=1}^{N}(x_{ij} - \overline{X})(y_{ij} - \overline{Y}),$$

where *SSIM* is the value of the similarity (quality) measure of images; $X=\{x_{ij}\}$ and $Y=\{y_{ij}\}$ are the compared images; M and N are the image dimensions. The first component of the expression is the correlation coefficient between images X and Y. The second component characterizes the similarity of the mean brightness values of the two compared images. The third component characterizes the similarity of the contrasts of the two compared images.

2. MSE (Mean Squared Error) measures the average squared difference between the actual and ideal pixel values. The higher the metric value, the worse the image quality.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - y_i\right)^2,$$

where x_i and y_i are the pixel values of the reference and original images.

3. PSNR is a metric that measures the difference between the maximum possible signal level and the noise level in an image. The higher the PSNR value, the better the image quality.

$$PSNR = 20\log_{10}\left(\frac{255}{\sqrt{MSE}}\right),$$

where *MSE* – Mean Squared Error.

4. MS-SSIM is a metric that analyzes an image at different levels (scales) using successive blurring. This allows for the assessment of structural similarity at various levels of detail. The higher the value of this metric, the better the quality.

$$MS - SSIM(x, y) = \left[l_M(x, y)\right]^{\alpha_M} \cdot \prod_{j=1}^{M}\left[c_j(x, y)\right]^{\beta_j} \cdot \left[s_j(x, y)\right]^{\gamma_j},$$

where x and y are the original and processed images; M is the number of scales; $l_M(x,y)$ is the luminance similarity at the lowest level; $c_j(x,y)$ is the contrast similarity at each level; $s_j(x,y)$ is the structural similarity at each level; $\square\square\square\square\square\square$ – represents the weights for each component.

5. BRISQUE is a no-reference metric, meaning there is only one image whose quality needs to be measured. The BRISQUE model is trained on a database of images with known distortions and is limited to assessing the quality of images with the same type of distortions. The lower the metric value, the better the image quality. The metric algorithm works as follows: first, pixel normalization is performed, then structures are analyzed for naturalness, feature vectors are generated, and finally, they are fed into a trained SVM. (Support Vector Machine)

6. The Pearson correlation coefficient measures the degree of linear dependence between two images (original and processed). The value ranges from -1 to 1. The higher the value, the stronger the dependence.

$$r = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X_{ij} - \overline{X}\right)\left(Y_{ij} - \overline{Y}\right)}{\sqrt{\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X_{ij} - \overline{X}\right)^2\right)\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(Y_{ij} - \overline{Y}\right)^2\right)}},$$

where X and Y are the original and masked images, respectively; $\overline{X}$ and $\overline{Y}$ are the mean values of X and Y, respectively; M and N are the vertical and horizontal dimensions of the original and masked images.

Description of the experiment

Masking is a procedure for converting images using matrix operations. In this experiment, we used two types of masking: one-sided and two-sided. With one-sided masking, the image on the transmitting side is transformed and it looks like this:

$$Y = XH,$$

where H is a Hadamard matrix. It should be noted that the matrix can be any in principle, the main condition is the presence of an inverse matrix. Hadamard matrices, due to their quasi-orthogonality, always have an inverse matrix, which is equal to the transposed one.

With two-sided masking, additional multiplication by the transposed special matrix is performed:

$$Y = H^T XH.$$

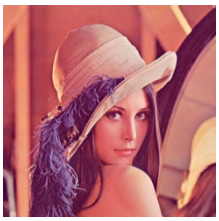
Demasking is the procedure of inverse transformation using matrix operations. Demasking is performed on the receiving side, and, for one-sided masking, it looks as follows:

$$X = YH^{-1},$$

with two-sided masking, the inverse transformation will be represented as follows:

$$X = (H^T)^{-1} YH^{-1}.$$

For the experiment, several images from different classes were selected, namely: distinct objects, aerial photographs, and textures. The images are shown in Fig. .



а) Lena

б) San-diego

в) 1.5.07

Fig. 1. Images used in the experiment.

A program was written that performs image masking and unmasking using the Hadamard matrix and then evaluates the quality of the restored image using six metrics. The program begins by specifying the path to our images. Then, we load the image under investigation, convert it to grayscale, change it to the required format for processing, and save it. Next, we define the experimental parameters: two types of masking – one-sided and two-sided, the Hadamard matrix used, and matrix sizes – 8, 16, 32, 64, 128, 256, and 512, as well as the 10-packet loss function defined at the end of the program. In the main cycle of the experiment, we iterate through all combinations of the previously set parameters, apply the masking to the image, use the loss function (in this case, the 10-packet loss), and unmask the image after the loss. We evaluate the restored images using built-in MATLAB metrics for quality assessment – SSIM, PSNR, MSE, BRISQUE, Pearson correlation coefficient, and MS-SSIM. To handle the data efficiently, we save the evaluation results in an Excel spreadsheet. This involves first saving the data to an array and converting it into a table. Then, we can create graphs to analyze the results.

For the sake of the experiment, we must fix the row indices that we reset during the 10-packet loss, so we define a fixed list of rows in the program itself – these are the ones we will reset during the experiment. All pixels in these rows will be black on the image, and at the end, the function returns the modified image.

To conduct the experiment, we needed to select three images from different classes, then specify the required image in the program and evaluate the results of the metrics in the table created by the program.

Results of simulation

The simulation results are presented in Figures 2–13. In all figures, the blue color corresponds to the SSIM metric, orange to MSE, purple to the correlation coefficient, light green to MS-SSIM, cyan to PSNR, and dark green to BRISQUE.

Based on the simulation results, it can be concluded that with one-sided masking, the matrix order does not affect the image quality assessment results. The metric values remain the same across all orders. It was also observed that the no-reference quality assessment metric BRISQUE behaves differently in the case of the San-Diego and 1.5.07 images with one-sided masking, does not correspond to human visual perception and significantly differs from the results of other metrics used. This can be observed in Figures 4 and 6.

At order 512, with two-sided masking, there is a sharp jump in metric values. Strangely, the correlation coefficient shows values corresponding to good or even excellent image quality (close to one), which suggests that using this metric at order 512 is not a reliable method for quality assessment. For the 1.5.07 image, we can observe an increase in the metric starting from order 256.

For two-sided masking, it was observed that for the Lena and 1.5.07 images, the SSIM, MS-SSIM, and BRISQUE metrics at order 16 demonstrate better image quality than at order 8, while other metrics show a gradual deterioration in the output image quality as the matrix order increases. For the San Diego image, the correlation coefficient is also added to this list of metrics. This observation is illustrated in Figures 8–13. For the SSIM, MS-SSIM, and BRISQUE metrics, this may be due to better preservation of the overall image structure compared to order 8. As for the correlation coefficient, the reason may be slight local changes that do not significantly affect the image itself but influence the dependency between the original and output images.

According to the graphs shown in Figures 11 and 13, it is evident that the BRISQUE metric produces results that do not depend on the matrix order and differ from the results of other metrics. This may be due to several reasons: most often, this metric shows unexpected results due to artifacts that are difficult for a human to notice, for example, after compression. Additionally, the results of the metric are influenced by texture, smoothing and noise reduction, brightness, and contrast. In my opinion, this is a drawback of no-reference metrics—since the algorithm is trained on natural images, when working with images that have, for instance, blurry details or an unexpected brightness range, we may obtain results that are incorrect compared to other metrics.

For the 1.5.07 image, for one-sided masking, the MSE metric is not presented on the graph, as it was 0.003 and was not visible on the graph due to the range differences.

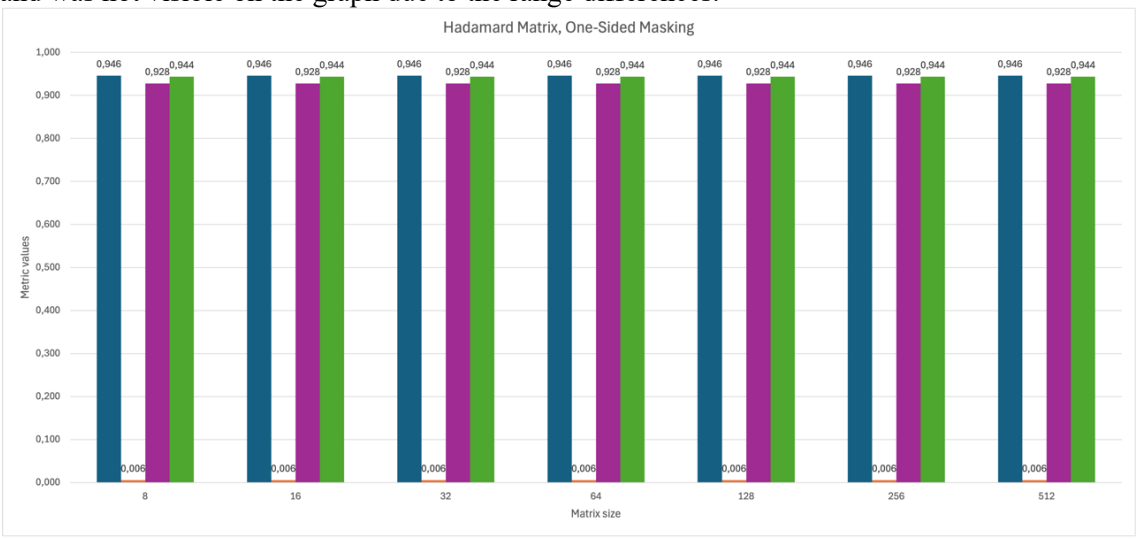


Fig. 2. Lena. SSIM, MSE, correlation coefficient, MS-SSIM.

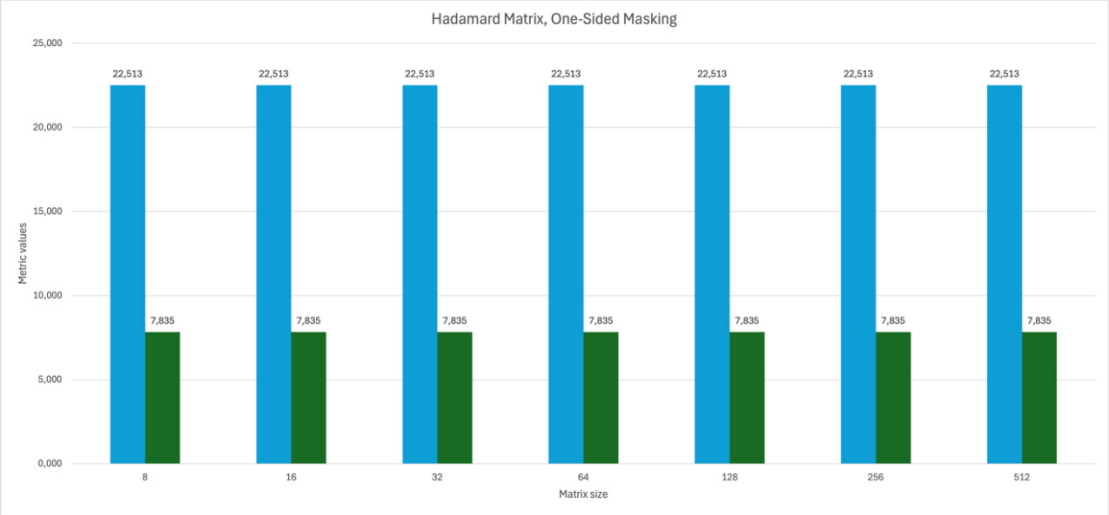


Fig. 3. Lena. PSNR, BRISQUE

Fig. 4. San-Diego. SSIM, MSE, correlation coefficient, MS-SSIM

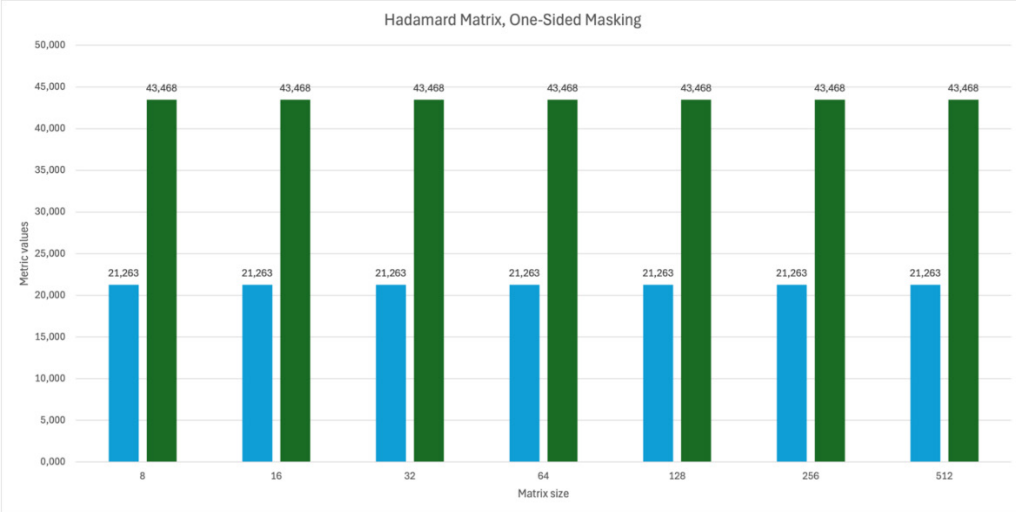


Fig. 5. San-Diego. PSNR, BRISQUE

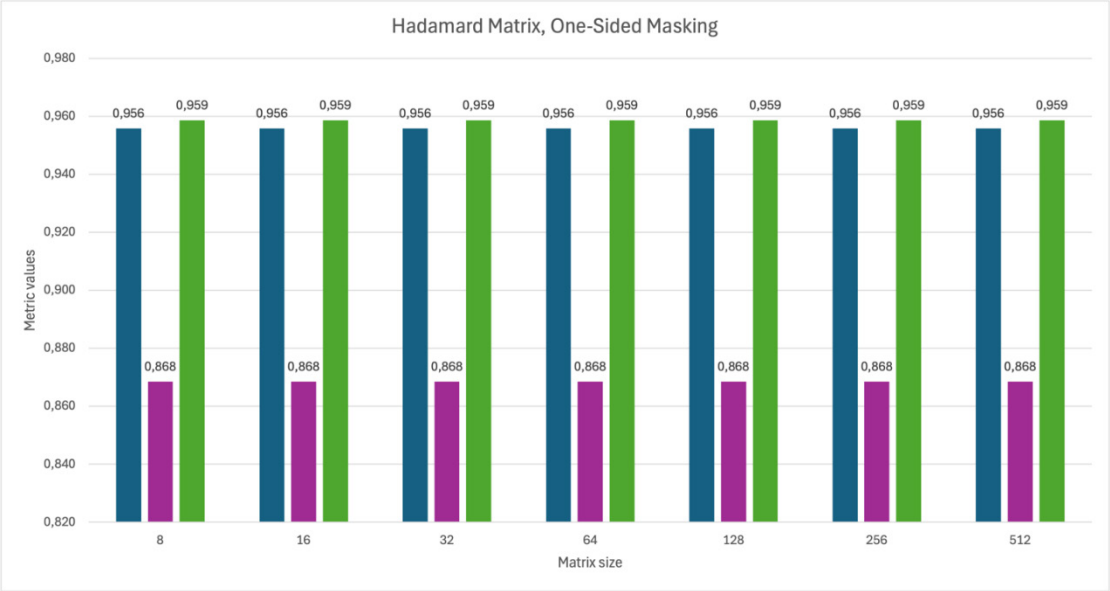


Fig. 6. 1.5.07. SSIM, correlation coefficient, MS-SSIM

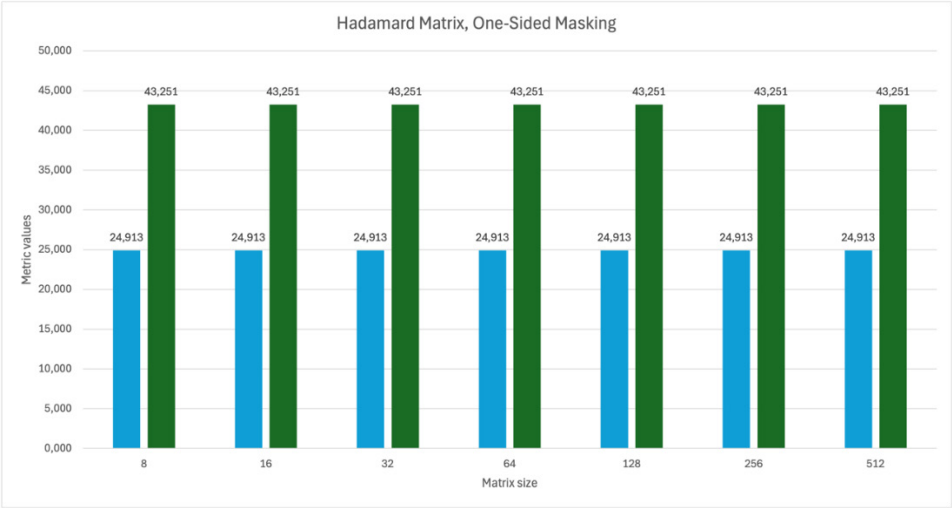


Fig. 7. 1.5.07. PSNR, BRISQUE

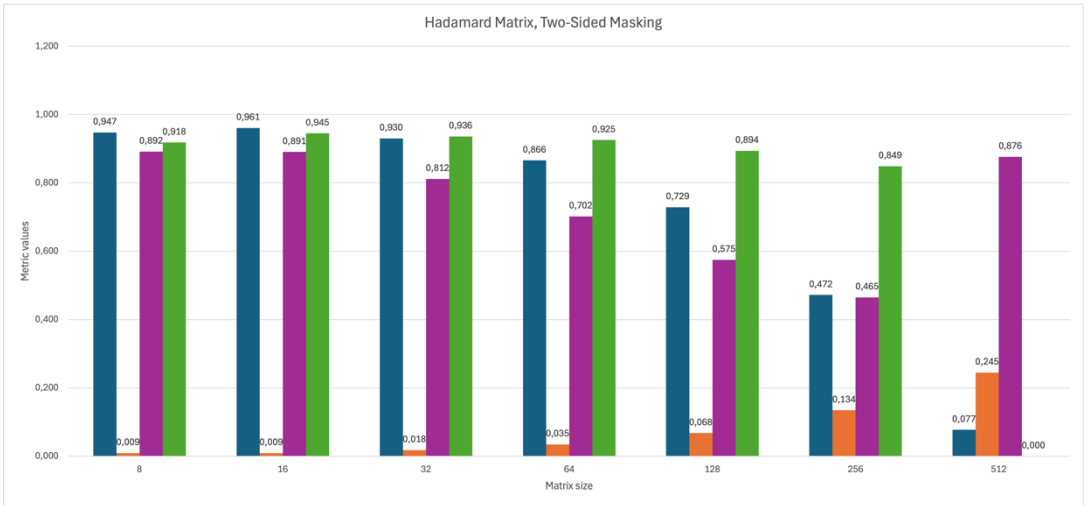


Fig. 8. Lena. SSIM, MSE, correlation coefficient, MS-SSIM

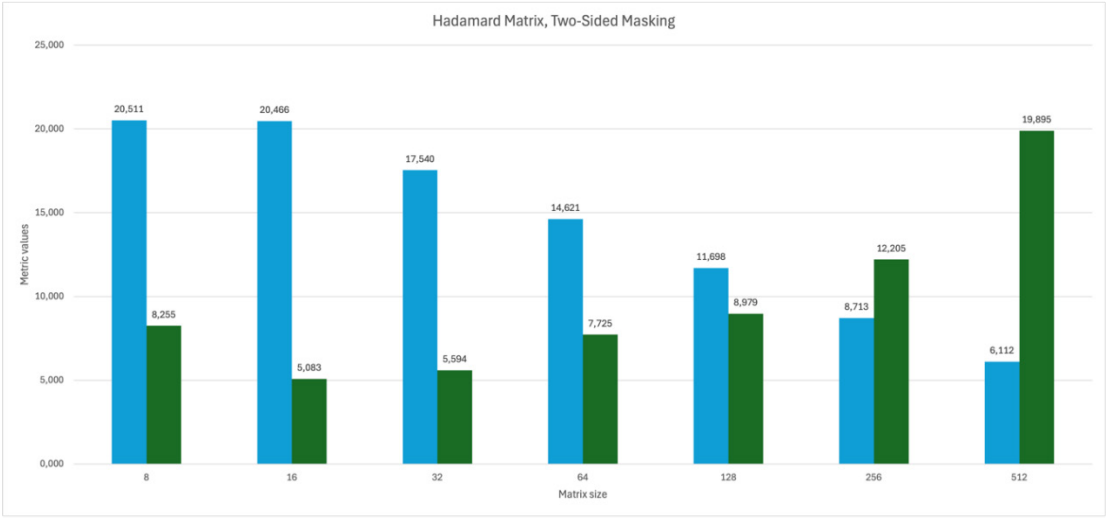


Fig. 9. Lena. PSNR, BRISQUE

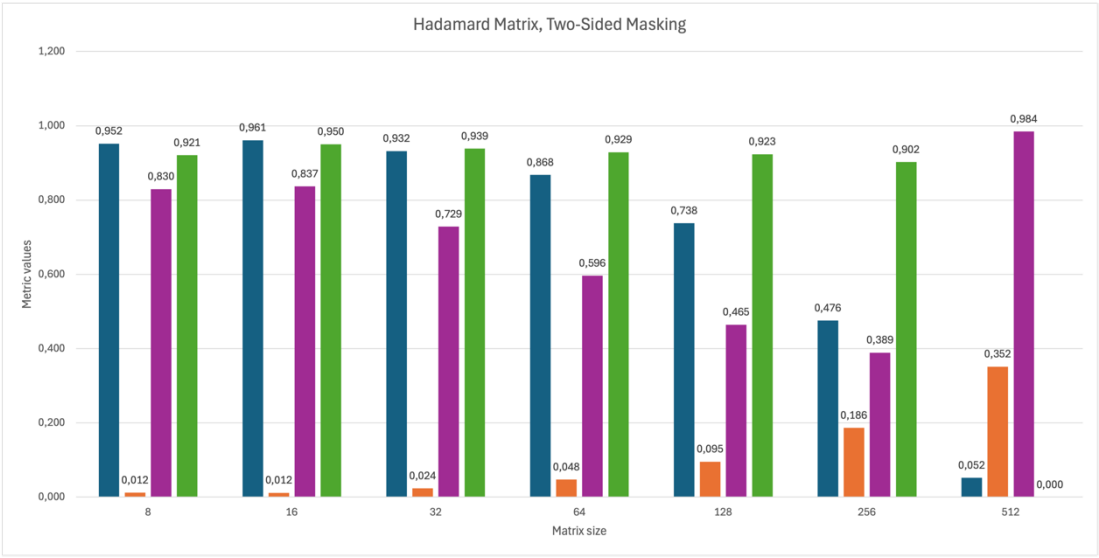


Fig. 10. San-deigo. SSIM, MSE, correlation coefficient, MS-SSIM

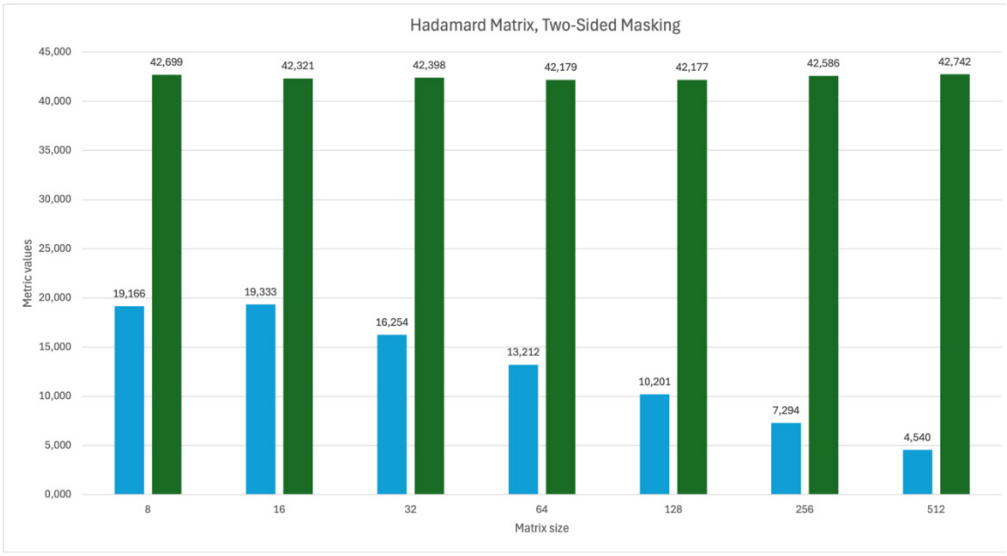


Fig. 11. San-diego. PSNR, BRISQUE

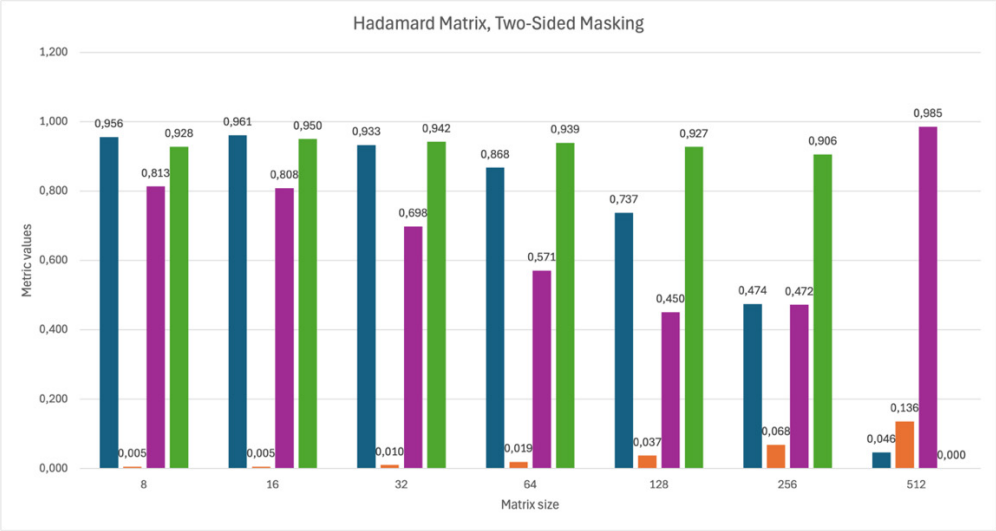


Fig. 12. 1.5.07. SSIM, MSE, correlation coefficient, MS-SSIM

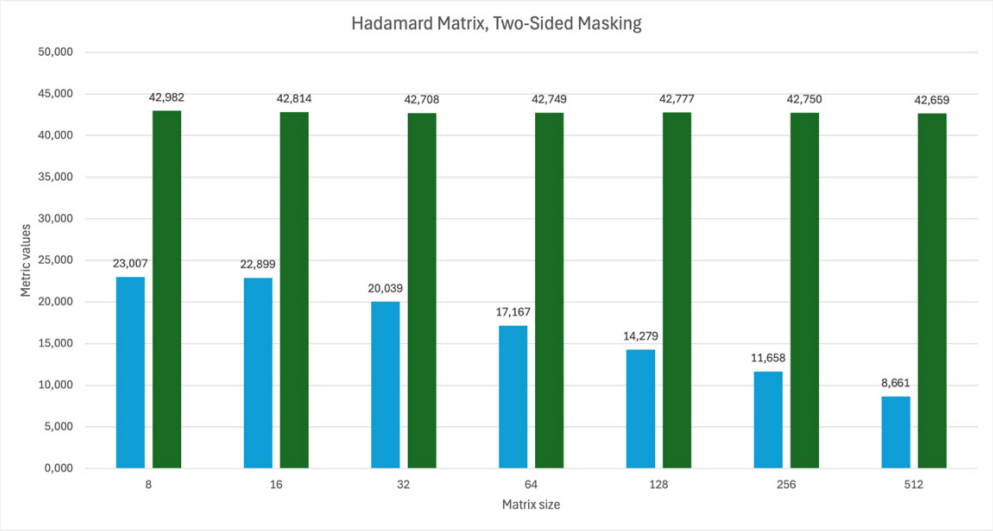


Fig. 13. 1.5.07. PSNR, BRISQUE

Conclusion

Based on the results of the experiment, one-sided masking performs better. The metric results can be classified into the following ranges: SSIM evaluates the quality of the restored image as excellent, MSE – excellent, the correlation coefficient – excellent, with some cases showing good quality, MS-SSIM – excellent, PSNR – average, and BRISQUE – excellent, although in some cases, this metric falls into the range of medium image quality. For two-sided masking, the metric values depend on the matrix order, and for larger sizes, they show results corresponding to poor image quality. Considering that two-sided masking is computationally more expensive, it makes sense to prefer one-sided masking, which shows good and stable results.

Currently, the number of quasi-orthogonal matrices has increased significantly, so it is advisable to further study the noise immunity of images using matrices of other structures and families.

References

1. Research and development of intelligent distributed sensing network for surface monitoring from shipborne UAVs / V. A. Nenashev, K. YU. Ryzhov, S. A. Nenashev [et al.] // Proceedings of the Krylov State Research Centre. – 2023. – № S1. – C. 45-56. – DOI 10.24937/2542-2324-2023-1-S-I-45-56.

2. Raskopina, A. S. Use of Deep Learning in Diagnosing Pneumonia from X-ray Images / A. S. Raskopina, V. V. Bozhenko, T. M. Tatarnikova // Journal of Higher Education Institutions. Instrument Engineering. – 2024. – T. 67, № 4. – C. 315-320. – DOI 10.17586/0021-3454-2024-67-4-315-320.

3. Kucher A. I. Study of the noise stability of the methods for estimating the internal structures functional state of biological objects at electroimpedance tomography (on the example of human lungs) / A. I. Kucher, G. K. Aleksanyan, I. D. Shcherbakov // Modern high-tech technologies. – 2019. – № 2. – С. 100-104.

4. Stepanyan I. V. Biomathematical system of the nucleic acids description / I. V. Stepanyan // Computer Research and Modeling. – 2020. – Т. 12, № 2. – С. 417-434. – DOI 10.20537/2076-7633-2020-12-2-417-434.

5. Gosteve I. M. Application of image processing methods in computational biology / I. M. Gosteve // Vestnik of the Russian University of Friendship of Peoples. Series: Mathematics, Informatics, Physics – 2012. – № 1. – С. 44-49.

6. Mironovsky L. A. Strip-Method of image and signal transformation: Monograph / St. Petersburg: Polytechnical, СПб., 2006. 163 с.

7. Sergeev A. M. Methods of image transformation and signal coding in distributed system channels using special quasi-orthogonal matrices: specialty 05.12.13 "Telecommunication systems, networks and devices": thesis for the degree of candidate of technical sciences / Sergeev Alexander Mikhailovich. – St. Petersburg, 2020. – 153 с.

8. Khvoshch S. T. Hadamard matrices in space communications/ S. T. Khvoshch // Engineering Bulletin of the Don. – 2024. – № 1(109). – С. 210-222.

9. Balonin N. A. Special matrices: pseudo-inverse, orthogonal, Hadamard and Cretan / N. A. Balonin, M. B. Sergeev. – St. Petersburg: "Polytechnic" Publishing House, 2019. – 196 с. – ISBN 978-5-7325-1155-0. – DOI 10.25960/7325-1155-0.

DEVELOPMENT OF AN INTELLIGENT ALARM SYSTEM FOR HOME SECURITY

Interlice Emanuele Maria, Nunzio Vincenzo Di Bella
Computer Engineering and Networks Laboratory – Kore University of Enna – Italy
Email: {emanuelemaria.interlice, nunziovincenzo.dibella}@unikorestudent.it

Abstract. *The proposed system, a home security alarm system, is designed to ensure the safety and protection of living spaces by detecting potential intrusions or anomalous situations. The system integrates motion and noise sensors, whose data is transmitted via a Zigbee wireless network to a central controller. The latter, through a wired Ethernet network, sends the processed signals to an actuator that triggers the sound alarm in case of critical situations.*
The design is based on a combination of wireless and wired technologies to ensure efficiency, reliability, and low energy consumption, which is particularly important for battery-powered sensors. The main system nodes include the sensors, the central controller, and the actuator, each playing a specific role in event detection, processing, and response. This integrated approach allows for the simulation of a realistic and scalable system suitable for various home application scenarios.

Introduction

A smart alarm system is an innovative home automation solution designed to improve the security of residential spaces. By combining advanced sensors, wireless and wired communication networks, and intelligent control algorithms, these systems can quickly detect and respond to potential intrusions or unusual situations. They can be used not only in homes but also in offices and public spaces, offering flexible and customizable security solutions.

The proposed system includes battery-powered motion and noise sensors, a Zigbee wireless network for data transmission, and an Ethernet network that connects the central controller to the alarm actuator. This setup ensures energy efficiency, fast response times, and high reliability, making the system effective in enhancing security and reducing reaction times in case of a threat.

Related works

Smart home security systems often face challenges with device interoperability, as many rely on proprietary protocols. To overcome this, several studies have explored the use of wireless sensor networks (WSNs) integrated with Ethernet gateways to enable seamless communication and data sharing across existing infrastructures [1]. This approach facilitates Internet connectivity, allowing for remote monitoring and better scalability in smart alarm systems [2]. A common challenge in these systems is the power consumption of wireless sensors. Research suggests solutions such as high-capacity batteries to extend sensor lifetime or selective power-down of sensor circuits when inactive to enhance energy efficiency [3]. Furthermore, the use of cloud computing to process data in real time has proven beneficial in improving the responsiveness of smart home systems [5]. Commercial solutions like Ring Alarm and Ajax Systems demonstrate the practical implementation of these technologies, offering customizable, scalable, and energy-efficient smart alarm solutions [6].

The proposed approach

According to these considerations, we decided to realize the network architecture on the standards IEEE 802.3 (Ethernet) and IEEE 802.15.4 (ZigBee).

Network 1 (wireless) includes noise and motion sensors that detect environmental events and transmit data to the gateway, which acts as a bridge between the two networks.

The gateway manages communication between the wireless sensors and the devices in Network 2 (wired), which includes an activity controller. This controller processes the received data and, if necessary, triggers the alarm actuator to signal potential danger.

The graphical interface provides intuitive, real-time system monitoring. At the center, two circular indicators display the motion value and noise level (dB), facilitating event analysis. Above them, the alarm system is highlighted with an activation icon, indicating the security status. On either side, the two networks are represented: Network 1 Wireless, which visualizes the sensor structure, and Network 2 Ethernet, which shows the configuration of wired devices. This interface enables clear and immediate control of environmental conditions, enhancing the effectiveness of the security system.

Scenario

The present study investigates a simulation scenario implemented in MATLAB/Simulink and the TrueTime toolbox.

The simulation duration is set to 150 seconds, during which noise and motion signals are generated using a dedicated MATLAB function.

The choice of employing a MATLAB function was dictated by the necessity of achieving a realistic representation of these two parameters while ensuring a low probability of triggering false alarms.

Specifically, the noise generation module produces values within the range of 0 to 90 dB, with a 0.3% probability that the noise level exceeds 40 dB, a threshold distinguishing standard ambient noise from potentially suspicious or high-intensity disturbances.

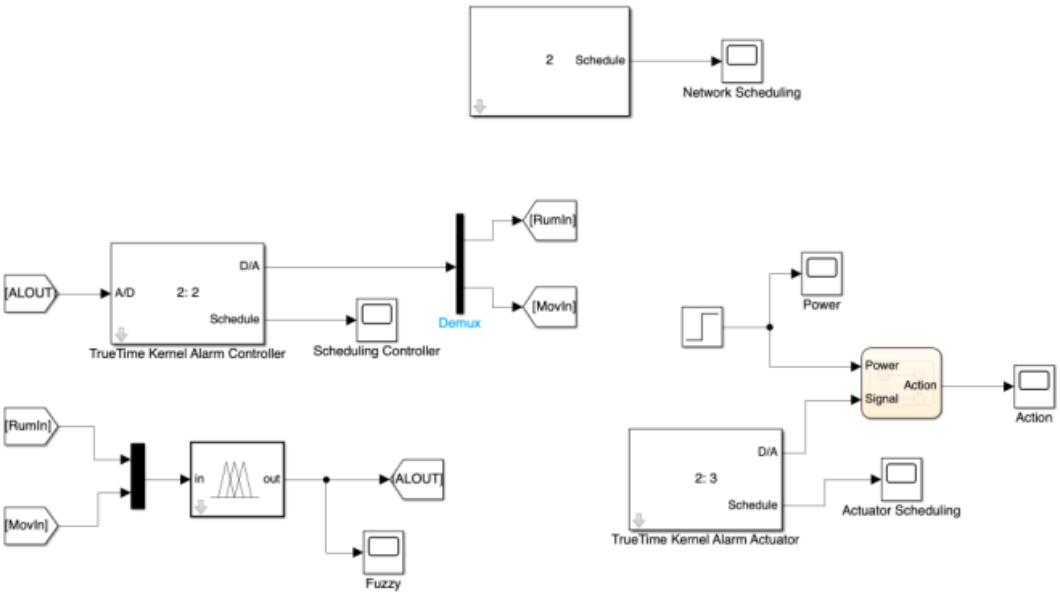
Similarly, the motion generation module yields values in the 0 to 1 range, with a 0.07% probability that motion values exceed 0.5, a threshold indicating significant movement within the monitored environment.

The system comprises two independent TrueTime Kernel-based sensors, each tasked with acquiring analog signals related to the monitored parameters (noise and motion). The acquired data is subsequently transmitted to a centralized controller via a gateway, which serves as an intermediary node between the wireless and wired networks.

The gateway does not feature direct inputs but is equipped with dedicated ports for scheduling management and power consumption monitoring. Both the sensors and the gateway incorporate a TrueTime Battery module, enabling the simulation of energy consumption dynamics.

To facilitate performance assessment, a monitoring system is employed to track the number of transmitted packets via a dedicated scope, allowing the evaluation of sensor activity over time. Additional monitoring tools include:

- Input, output, and scheduling scopes,
- Go-To and From blocks, facilitating communication within the TrueTime Wireless Network.



The wired Ethernet network integrates a TrueTime Network block, which governs inter-component communication, a controller, responsible for sensor data analysis and alarm decision-making, and an actuator, tasked with executing alarm activation based on the controller's commands.

The controller, designated as Node 2 within the Ethernet network, operates through a TrueTime Kernel integrated with a fuzzy logic controller. Its primary function is to process sensor data received from the gateway, classify the incoming signals based on predefined fuzzy logic rules, and generate an appropriate control output, which is subsequently relayed to the actuator.

The fuzzy logic system categorizes sensor input data as follows:

- Noise: Five classification levels (None, Low, Medium, Suspicious, High)
- Motion: Three classification levels (Low, Medium, High)

In contrast to conventional threshold-based methodologies, the fuzzy logic approach facilitates a more adaptive and robust interpretation of environmental variations.

Based on the analyzed data, the alarm classification variable assumes one of three possible states:

- No Action: No anomalous condition detected.
- Monitoring: Indication of potentially suspicious activity.
- Alarm-Intervention: Detection of critical noise and motion levels, requiring immediate response.

By leveraging fuzzy logic inference, the system enhances adaptability to real-world conditions, effectively minimizing false alarms while improving the reliability of environmental monitoring.

The actuator, identified as Node 3 in the Ethernet network, is responsible for executing alarm activation based on the control signals received from the fuzzy logic-based controller.

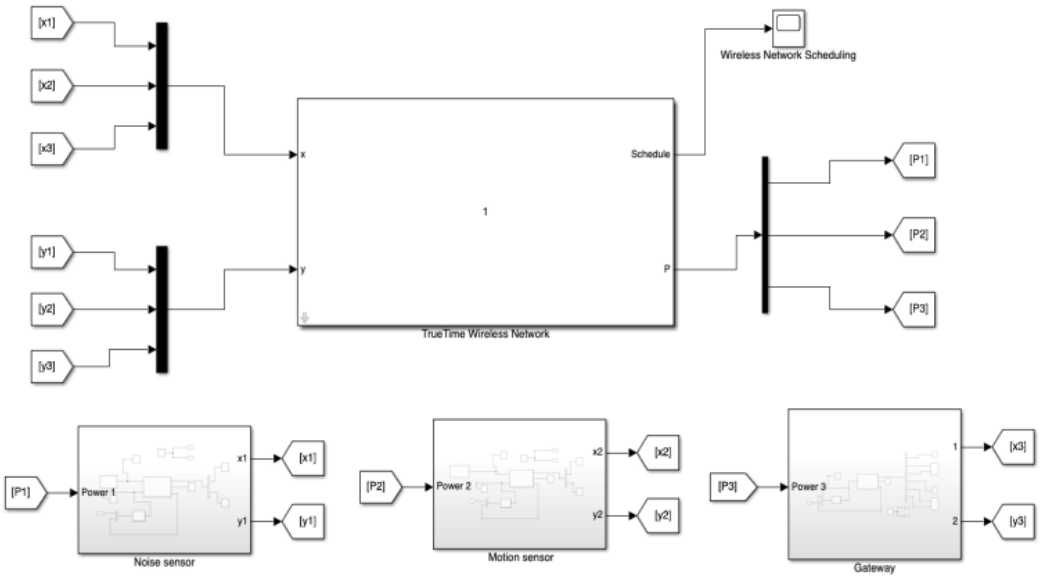
The actuation mechanism is implemented through a TrueTime Kernel, configured without direct input interfaces, which transmits control signals to a state chart implementing Model-Based Software Design (MBSD) logic. The state chart governs the siren light activation, while an additional step function is employed to simulate the activation and deactivation of the alarm system.

State Chart: Dynamically manages siren lighting behavior in response to received signals.

To ensure optimal performance and system reliability, the framework integrates various monitoring subsystems, including:

- Energy consumption analysis: Evaluating the battery status of sensors and the gateway.
- Scheduling management: Assessing the real-time scheduling of processes executed by TrueTime Kernels.
- Packet loss and transmission latency monitoring: Validating the reliability of network communications.

This structured approach ensures an efficient, adaptive, and resilient alarm system, capable of dynamically responding to environmental conditions while optimizing resource utilization.



Performance evaluation

The simulations were conducted to evaluate the network's performance and its ability to fulfill its designated tasks.

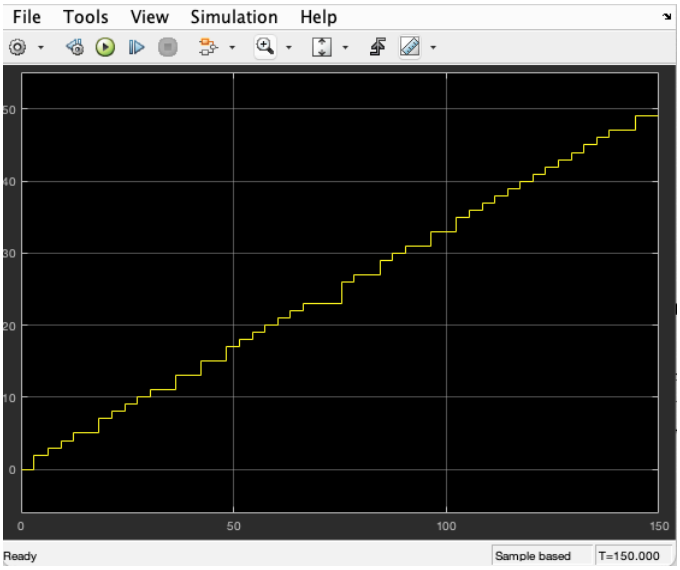
To assess network efficiency, we measured packet loss rate during transmission and propagation delay.

The motion and noise sensors periodically transmit data to the gateway, including a unique, sequentially increasing identifier. Under ideal conditions, all data packets should be received in order. However, the presence of interference or network congestion may result in packet loss.

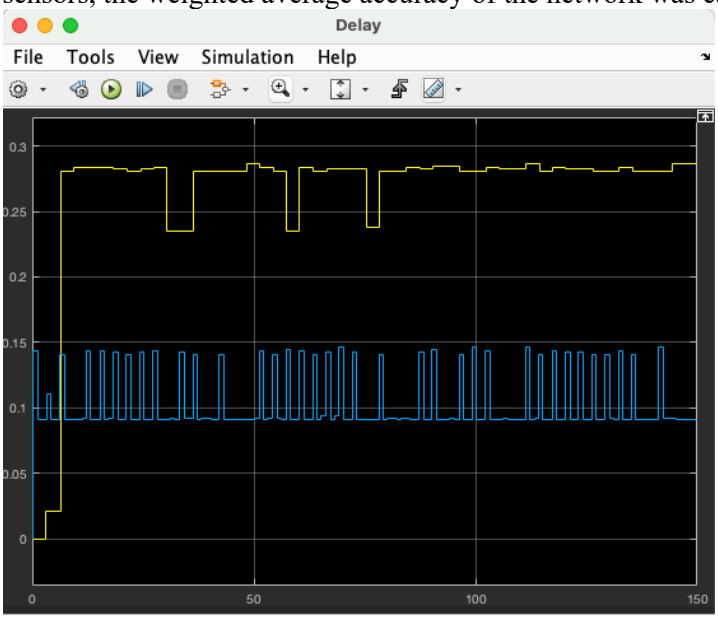
To detect lost packets, the system compares the received identifier with the expected value. If the received identifier is greater than the expected one, it indicates that certain packets have been lost during transmission. The exact number of missing packets is determined by computing the difference between the expected and received identifiers.

The system continuously updates and records packet loss statistics to monitor transmission quality.

During the simulations, a 20% probability of packet loss was introduced. As a result, the motion sensor achieved an effective transmission rate of 90.60%, with 135 packets received out of 149 transmitted. Conversely, the noise sensor successfully delivered 38 packets out of 49, corresponding to an accuracy of 77.55%.



Considering both sensors, the weighted average accuracy of the network was calculated to be 87.94%.



Regarding propagation delay, which represents the time elapsed between the transmission and reception of a packet, the ttCurrentTime primitive was utilized to log the exact transmission and reception timestamps. The delay was computed using the following formula:

$$time_delay = ttCurrentTime - sp.time$$

The collected data was compared against performance evaluation metrics, demonstrating an overall system accuracy of 87%. Based on this result, the network falls within the "Good Performance" category, confirming the effectiveness of the adopted configuration.

Conclusions

The project developed in MATLAB/Simulink with TrueTime demonstrated the effectiveness of an integrated smart alarm system, combining real-time detection, analysis, and actuation. By leveraging sensors, wireless networks, and fuzzy control algorithms, the system accurately identifies anomalies, processes sensor data, and triggers appropriate responses, such as alarm activation. The implementation of advanced packet loss management and delay calculation techniques ensured reliable communication between sensors and

actuators, enhancing system efficiency and responsiveness—critical for security applications. This work highlights how MATLAB/Simulink and TrueTime facilitate the development of robust, scalable, and reliable security management solutions.

References

[1] Al Ameen, M., Liu, J., & Kwak, K. (2012). "Security and privacy considerations for Wireless Sensor Networks in smart home environments."

[2] Alrawais, A., Alhothaily, A., Hu, C., & Cheng, X. (2017). "Fog Computing for the Internet of Things: Security and Privacy Issues."

[3] Liu, Y., & Yu, J. (2011). "Energy-efficient design of wireless sensor networks: A survey."

[4] Alaba, F. A., Othman, M., Hashem, I. A. T., & Alotaibi, F. (2017). "Internet of Things security: A survey."

[5] Ring Alarm e Ajax Systems.

DEVELOPMENT OF A LIGHT MONITORING SYSTEM BASED ON A WI-FI MICROCONTROLLER

Kirill Ivanov
Saint Petersburg State University of Aerospace Instrumentation
e-mail: kirillivanov200230@gmail.com

Abstract. *The article describes the development of a system for real-time light level monitoring using the ESP32 microcontroller and the APDS-9960 sensor. The system provides remote access to light data through a web interface and includes an FTP server for storing web interface data. The methods for implementing the Wi-Fi server, interaction with the sensor, and processing the received data are discussed.*
Keywords: *microcontroller, ESP32, light monitoring, Wi-Fi, APDS-9960, FTP server, HTTP server, web interface.*

Introduction

Modern information and measurement systems in instrument engineering and radio electronics provide high accuracy and efficiency in various fields, including light control in smart buildings, laboratories, and industrial premises. This article presents the development of an integrated system for monitoring light levels using the APDS-9960 sensor, the ESP32 microcontroller, and data transmission via Wi-Fi. The system provides the user with a web interface for remote control and can be integrated into more complex management systems.

System Description

1. Software Used

The system was programmed using the PlatformIO environment with the Arduino core, which ensured ease of development and broad compatibility with various libraries [1]. Specifically, the following were used:

– SparkFun_APDS9960 for working with the light sensor;
– FTPServer for implementing the FTP server, through which data exchange occurs;
– WebServer for creating a web server that provides interaction with the user via an internet browser;
– SPIFFS – a file system for storing web pages and configurations.

2. Components Used

The system is based on the following key components:

– ESP32 – a Wi-Fi-enabled microcontroller that performs data collection and processing from the APDS-9960 sensor and manages the web interface for remote interaction;
– APDS-9960 – a light sensor that provides high-precision information about external light levels via the I2C interface.

3. System Structure

The system includes several key components:

– ESP32 microcontroller – a Wi-Fi-enabled platform used for collecting data from the light sensor via the I2C interface and providing a web interface for user interaction;
– Data reading from the sensor – the ESP32 microcontroller reads data about external light levels from the APDS-9960 sensor via the I2C interface;
– Data processing – the received light level data is processed and transmitted to the web interface;
– Web server – a web server running on the ESP32, providing the user with the ability to view the current light level in real-time, as well as turn the light sensor on and off;
– FTP server based on SPIFFS – used for storing the web interface code, allowing for quick addition of functionality.

System Implementation

1. Web Server and User Interaction

To provide remote monitoring of light levels, a web server was created on the ESP32, allowing the user to interact with the system through a browser [2].

It is worth noting that the ESP32 operates within a local network. The microcontroller connects to the router via Wi-Fi, and if another device with a browser is connected to the same router, the web interface can be accessed by entering the IP address assigned to the ESP32 by the router.

HTTP request processing occurs in an infinite loop using a method from the WebServer library, as shown in Fig. 1.

```
server.handleClient(); // Handling HTTP requests
```

Fig. 1 – HTTP request processing

2. FTP Server for Data Storage

An FTP server implemented using the FTPServer library [3] is used to store the web page code on the device. The server allows the developer to quickly upload and download server interface files. FTP request processing also occurs in the server's infinite loop.

```
ftpSrv.handleFTP();     // Handling FTP requests
```

Fig. 2 – FTP request processing

3. Reading Data from the Sensor

To obtain data on the current light level, the APDS-9960 sensor connected via the I2C interface is used. The SparkFun_APDS9960 library is used to work with the sensor, providing the readAmbientLight() function, which reads the light level and stores it in a variable, as shown in Fig. 2. This function runs in an infinite loop, just like HTTP and FTP request processing.

```
apds.readAmbientLight(ambient_light); // Reading current illumination
```

Fig. 2. Reading data from the light sensor

When an HTTP request is received to a specific URL, for example, /light, the server calls the handle Light Date() function, which returns data on the current light level to the user, Fig. 3.

```
server.on("/light", HTTP_GET,handleLightData); // Route for illuminance data
```

Fig. 3. Calling the function for sending illumination data

Fig. 4 shows the function of sending illumination data itself, it is performed when a request is received from the client side.

```
void handleLightData() {
// Send current light data in JSON format
  String response = "{\"ambient_light\": " + String(ambient_light) + "}";
  server.send(200, "application/json", response);
}
```

Fig. 4. A function for processing a request for a weaving light level

4. User interaction interface

A web interface is available for the user, which allows you to control the status of the sensor and receive data on the current light level. This interface includes buttons for turning the sensor on and off, as well as dynamically updating light level data. The interface has the form shown in Fig. 5. The Ambient Light line displays the current illumination.

The web interface, implemented using HTML, CSS, and JavaScript, provides the ability to display the current sensor status and light level, as well as control the sensor status (on/off) [4].

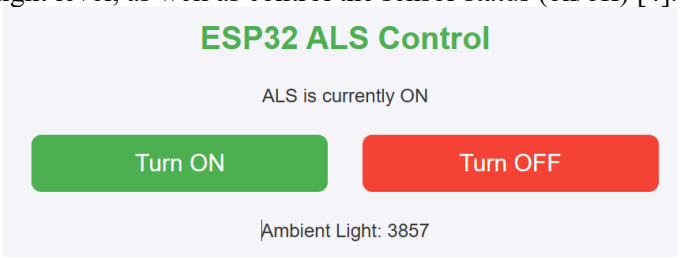


Fig. 5. User interface

For dynamic data updates on the client side, the fetch method is used to periodically request data from the server at /light, Fig. 6.

```
<script>
  // Function to get light data
  function fetchLightData() {
    fetch('/light')
      .then(response => response.json())
      .then(data => {
        // Update the light level on the page
        document.getElementById('ambient-light').innerText = data.ambient_light;
      })
      .catch(error => {
        console.log('Error fetching light data:', error);
      });
  }

  // Update light data every second
  setInterval(fetchLightData, 500);
</script>
```

Fig. 6. Updating data in the web interface

Conclusion

The presented light monitoring system based on the MSP 32 microcontroller and the APDS-9960 sensor offers the user a simple and effective way to monitor the light level with the possibility of remote access via a web interface and an FTP server. This solution is ideal for applications in smart buildings, offices, and production facilities where constant monitoring and control of light levels are required.

This system opens up new possibilities for integration into more complex automation systems, providing flexibility and real-time availability.

References:

1. Espressif 32 – PlatformIO latest documentation [Electronic resource]: https://docs.platformio.org/en/latest/platforms/espressif32.html

2. Веб-сервер ESP32 (ESP8266) в среде Arduino IDE [Electronic resource]: https://arduino-tex.ru/news/15/urok-1-veb-server-esp32-esp8266-v-srede-arduino-ide.html

3. FTP server on esp8266 and esp32 [Electronic resource]: https://www.hackster.io/xreef/ftp-server-on-esp8266-and-esp32-b03bd8

4. Using JavaScript for dynamic web content [Electronic resource]: https://appmaster.io/ru/blog/javascript-dinamicheskii-veb-kontent

AUTOMATIC SPACE MAPPING USING LIDAR

Mikhail Kalinichenko
Saint Petersburg State University of Aerospace Instrumentation, Ivangorod Branch, Saint Petersburg, Russia, E-mail: jayli04@mail.ru

Abstract. *The paper considers the problem of automatic mapping of space using a wheeled robot moving along grid cells. The article describes the process of forming this grid using LiDAR installed on a mobile platform. The article provides a mathematical model for determining impassable grid cells based on the positioning of the mobile platform and LiDAR data. The article ends with a roadmap for the development of the topic.*
Keywords: *automatic mapping, automation, LiDAR, wheeled robot.*

3D scanning is used in various tasks. This may be the task of obtaining a virtual copy, archeology, speleology, searching for victims in rooms under rubble, scanning rooms and others [1][2][3]. Most often, remotely controlled and autonomous robotic systems are used in such tasks. For autonomous systems, it is necessary to automate the process of orientation in space, drawing up a plan of the area and route. This will allow not only to map the space but also to return the robot to the starting point.

Let's consider a robotic wheeled system based on Arduino Uno. It has low computing power compared to more expensive analogs. For spatial orientation, we can use LiDAR 360°, but we will consider its analog based on a regular LiDAR module and a 360° servo motor. Instead of a point cloud from LiDAR 360°, we will get a set of angles and distances. Fig. 1 shows the concept of building such a system. On the left is the control board, servo motor and LiDAR, on the right are Motor Shield and 4 electric motors.

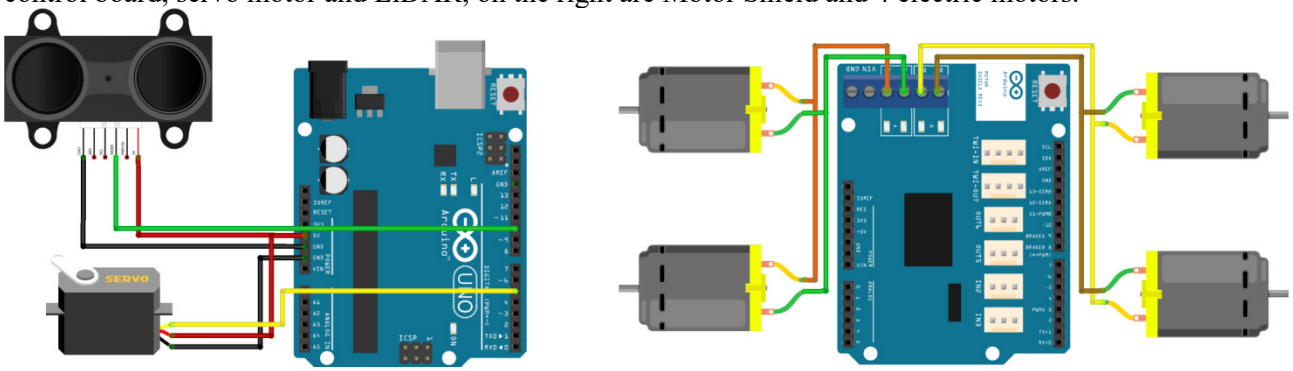


Fig.1. Concept of a mobile platform equipped with LiDAR

When using radar, it is possible to determine not only the distance to the reflective surface, but also its classification [4]. When using a camera and the Open CV library, it is possible to find out other properties of objects around. LiDAR only returns the distance to the obstacle. For low-cost devices, the data set is small enough for them to process.

The problem of small memory and low computing power requires a mathematical approach to determining the robot's position in space. Depending on the complexity of the problem being solved, the ability to precisely control the motors, and other factors, one can rely on a complex model of precise positioning [5] or a simpler option focused on the robot's location and its movement along the cells in the grid [6].

Let's consider the second case, as less resource-intensive. The space is divided into squares, one of which is occupied by the robotic platform. In the controller, this is a matrix, where each cell is passable by default. We will assume that the robot occupies exactly one cell on the grid (a cell in the matrix). Using LiDAR, the distance to objects is estimated with a certain angular step. An obstacle encountered in the matrix is marked as an impassable cell. An example is shown in Fig. 2.

The distance from the robot to the obstacle and the angle at which the LiDAR was deployed are known (Fig. 3). It is possible to determine which cell of the map should be marked as impassable. To do this, we will consider the distance as the hypotenuse of a right-angled triangle, and the displacements along the *Ox* and *Oy* axes as the legs of a right-angled triangle.

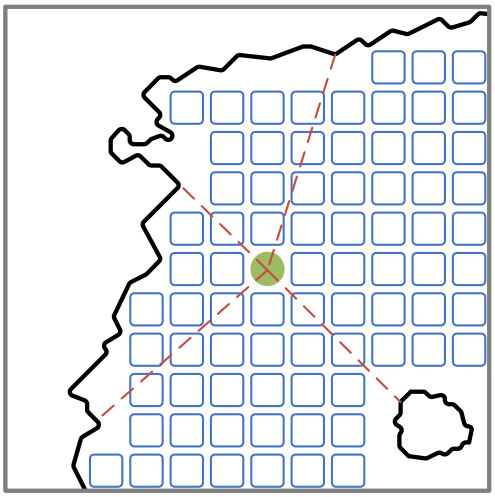


Fig.2. Robot with LiDAR searches for obstacles

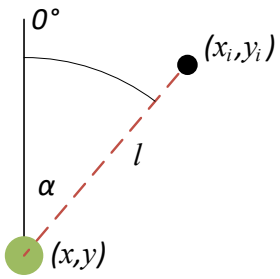


Fig.3. Distance- and angle-based coordinates

LiDAR is placed exactly in the center of the cell:

$$\begin{cases} x_i = l\sin\alpha_i \\ y_i = l\cos\alpha_i \end{cases}, \tag{1}$$

where l is the distance to the object, α is the angle, i is the conventional operation number. The values of α will be limited to the interval 0°-90°. For objects behind the platform, we will subtract this value from 90° (mirror reflection). The real direction will be used for displacement in the positive or negative direction along the *Ox* and *Oy* axes (to the right for α from 0° to 180° and forward for α from 270° to 90°).

We get the offset from (1). We divide x_i and y_i by the cell size (s), reduce this value by half a cell, and round the results up:

$$\begin{cases} x_i = \pm\left\lceil \dfrac{(l\sin\alpha_i)}{s} - 0.5 \right\rceil \\ y_i = \pm\left\lceil \dfrac{(l\cos\alpha_i)}{s} - 0.5 \right\rceil \end{cases}. \tag{2}$$

Thus, we calculated the coordinates of the impassable cell relative to the one in which the robotic platform is located.

If the LiDAR on the platform is not installed in the center (Fig. 4), it is necessary to shift it again in (2):

$$\begin{cases} x_i = \pm\left\lceil \dfrac{(l\sin\alpha_i) + \delta x}{s} - 0.5 \right\rceil \\ y_i = \pm\left\lceil \dfrac{(l\cos\alpha_i) + \delta y}{s} - 0.5 \right\rceil \end{cases}. \tag{3}$$

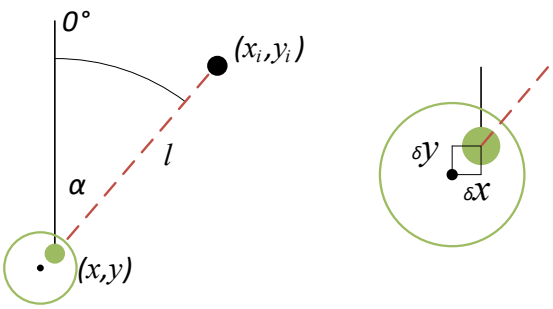


Fig.4. Offsetting the anchor point from the center

To minimize further calculations, you can remember the values of the sine and cosine for the angle – use memoization. If the controller's performance is not enough for calculations, we set fixed angular steps in advance, then the sines and cosines of these values can be stored as constants, rather than calculated each time. Offsets can also be calculated and stored as constants.

Model (3) is not ideal. It can be improved in various ways. For example, reduce the cell size by half (the platform size will be not *s*, but *2s*). This will increase the detail of the map (Fig. 5). Then (3) can be written as follows:

$$\begin{cases} x_i = \pm\left\lceil \dfrac{(l\sin\alpha_i)+\delta x}{s} - 1 \right\rceil \\ y_i = \pm\left\lceil \dfrac{(l\cos\alpha_i)+\delta y}{s} - 1 \right\rceil \end{cases}.$$

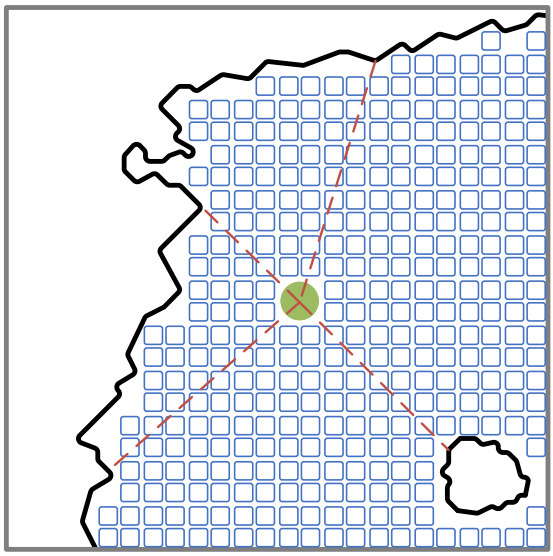


Fig.5. Robot with LiDAR searches for obstacles smaller than itself

Such a mobile platform can be a wheeled robot, a tracked robot, a drone [7]. The task of automating mapping will be based not only on a previously written mathematical model, but also on minimizing calculations and minimizing memory consumption for storing the map and various data.

The model of automatic spatial mapping using LiDAR described in this article allows not only to obtain a map of the space approximated by squares, but also to plot routes using various algorithms. The topic is promising for further research. Roadmap: the next stage is modeling for debugging the algorithm, assembly and a full-scale experiment, transition from the movement model [6] to the movement model [5]. This will allow for more accurate positioning of the autonomous robot and obtaining more accurate maps.

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

3. Dolgov, E. “LiDAR-based autonomous 3D scanning of rooms”. Bulletin of the UNESCO Chair “Distance education in engineering” of the SUAI. I. 9. 2024. Pp. 58-61.

4. Sorokin A.A., Shepeta A.P. “Classification of earth surface types by small-sized radars”. III international scientific and practical conference Innovation in Information Technology: from Concept to Implementation. 2024. Pp. 12-17.

5. 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.

6. 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.

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.

RESEARCH OF THE AUDIO DATA NOISE-IMMUNITY MASKED BY HADAMARD MATRICES

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

Abstract. *The problem of assessing the noise immunity of audio data masked by Hadamard matrices is considered. In this paper, we study the method of audio data masking, which ensures the confidentiality of audio data with a short period of relevance. However, the loss and distortion during transmission have not been assessed before. The audio signal quality was assessed using well-known audio data restoration quality metrics, such as RMS, SNR, PSNR, LSD, NCM, THD and PESQ. Based on the experimental results, we can conclude that bilateral masking provides higher noise immunity compared to one-sided masking, especially when using large matrix sizes. The obtained results open up new directions for studying the influence of different types and sizes of matrices on the effectiveness of masking.*
Keywords: *matrix masking, quasi-orthogonal matrices, Hadamard matrices, noise immunity.*

Introduction

Multimedia data, like any other type of information transmitted over computer networks, consists of files with a finite size. These files are transmitted across the network as a sequence of packets, the size of which is governed by current network standards and architectural decisions. Ideally, each packet should be delivered to the recipient without any errors, ensuring the integrity and correctness of the transmitted data. However, in practice, the data transmission process is often accompanied by errors caused by interference, equipment malfunctions, or network congestion. As a result, some packets may arrive with corrupted bits, while others may be completely lost during transmission.

In error-correcting coding theory, packets lost in the network are classified as erasures. To mathematically describe the processes occurring in data transmission channels with such errors, the erasure channel model is used [1].

The erasure channel model is particularly relevant for analyzing the transmission of multimedia data, such as audio, video, and images, since their quality directly depends on the integrity and timely delivery of packets.

In the field of robust image transmission over erasure channels, the use of methods based on orthogonal and quasi-orthogonal structured matrices is known [2-5]. Examples include Hadamard, Belevitch, and Mersenne matrices, etc.

In the field of robust audio data transmission over erasure channels, matrix methods for erasure protection are also known. For example, using structured Cauchy and Vandermonde matrices [6]. However, in this case, problems may arise with finding inverse matrices during decoding.

In work [7], a method of protective coding (masking) of audio data with a short period of relevance was proposed, which is a development of the ideas laid out in works [2-3]. The essence of the method is the preprocessing of audio data – dividing it into packets, forming a matrix from the packets, and then multiplying it by a quasi-orthogonal structured matrix [5], such as a Hadamard matrix. It is known [2-4] that for images, this method performs well in transmitting data over an erasure channel; however, similar research has not been conducted for audio data.

Therefore, the aim of this work is to evaluate the noise robustness of the matrix masking method for audio data during transmission over erasure channels using objective audio quality metrics.

Metrics overview

The following audio signal quality metrics were used in this experiment:

1. Root Mean Square **(RMS)**: This is a measure showing the average amplitude of the signal. A smaller RMS value indicates a smaller error, meaning the difference between the original and processed signal is less, and thus the processed signal is closer to the original.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A[i])^2},$$

where: $A[i]$— amplitude of the audio signal at a moment in time I, N— total number of samples.

1. Normalized Cross-Magnitude (NCM): A metric used to evaluate the similarity between two signals or time series. The NCM value ranges from 0 to 1, where 0 indicates no correlation and 1 indicates complete correlation.

$$NCM = \left(\frac{\sum_{i=1}^{N}(x_i - x)(y_i - y)}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y)^2}}\right)$$

where: x_i – values of the original signal, y_i – values of the processed signal, x – mean value of the original signal, y – mean value of the processed signal.

2. Perceptual Evaluation of Speech Quality (PESQ) is a widely used standard for the objective assessment of speech quality [8]. It mimics the psychoacoustic processes of human hearing to compare the original and distorted signals. At the output, PESQ typically provides MOS-LQO and Raw MOS scores.

3. MOS-LQO (Mean Opinion Score – Listening Quality Objective): This is an objective metric for assessing sound quality, designed to predict human subjective sound quality ratings. Scores are given on a scale from 1 to 5, where 1 is "bad quality" and 5 is "excellent quality." The higher the value, the better the perceived quality of the signal.

$$MOS - LQO = f(PESQ_{score})$$

where f – function that converts the PESQ value to the MOS scale.

4. Raw MOS (Raw Mean Opinion Score): This is a subjective quality assessment metric obtained algorithmically, using a model that predicts subjective quality perception. Scores are given on a scale from 1 to 5, where 1 is "bad quality" and 5 is "excellent quality." The higher the value, the better the perceived quality of the signal.

$$RawMOS = \frac{1}{N}\sum_{i=1}^{N} R[i]$$

where: $R[i]$— rating given by the i-th listener, N— total number of ratings.

5. Signal-to-Noise Ratio **(SNR)**: This is the ratio of the average signal power to the average noise power. The higher the SNR, the better the signal quality.

$$SNR = 10 \bullet \log 10\left(\frac{RMS_{signal}^2}{RMS_{noise}^2}\right)$$

where: RMS_{signal}— root mean square value of the signal, RMS_{noise} – root mean square value of the noise.

6. Peak Signal-to-Noise Ratio **(PSNR)**: This is a signal quality measure that compares the maximum possible power of a signal to the power of noise, expressed in decibels (dB). The higher the PSNR, the better the signal quality.

$$PSNR = 10 \bullet \log 10\left(\frac{MAX^2}{MSE}\right)$$

where: MAX – maximum amplitude value of the signal, MSE – mean squared error.

7. Log-Spectral Distortion (LSD): This metric measures the discrepancy between logarithmic spectra [9]. The lower the LSD, the better the quality.

$$LSD = \frac{1}{N-1}\sum_{i=2}^{N}\left|20\log 10\left|x_{original}(k)\right| - 20\log 10\left|x_{processed}(k)\right|\right|$$

where: N – total number of frequency components (FFT size), $X_{original}$ (k) – k-th spectral component of the original signal, $X_{processed}$ (k) – k-th spectral component of the processed signal.

8. Total Harmonic Distortion (THD): Measures harmonic distortions in the signal compared to the reference signal, i.e., the original signal fed into the system under test [10]. The lower the THD, the cleaner the sound.

$$THD = \frac{\sqrt{A_2^2 + A_3^2 + A_4^2 + ... + A_n^2}}{A_1} \times 100\%$$

where: A_1 – amplitude of the fundamental frequency, A_2, A_3, …, A_n – amplitudes of the harmonic components.

Description of the Experiment

A computer experiment was conducted to evaluate the transformation parameters. A Hadamard matrix was used as the key matrix. Two types of masking were also implemented.

In one-sided masking, the transformation on the transmitting side is performed by multiplying the original message matrix X_n of size $n \times n$ by the key matrix M_n of the same size, as follows:

$$Y_n = X_n \bullet M_n, \quad (1)$$

where Y_n – the digitally transmitted secure audio signal.

The inverse transformation on the receiving side is performed according to the expression:

$$X_n = Y_n \bullet M_n^{-1}. \quad (2)$$

In two-sided matrix masking, the transmitting side multiplies the matrix of the original message X_n of size n × n by the key matrix of the same size on the left and the transposed key matrix M_n^T on the right:

$$Y_n = X_n \bullet M_n \bullet M_n^T, \quad (3)$$

The inverse transformation on the receiving side is performed using the expression:

$$X_n = Y_n \bullet M_n^{-1} \bullet (M_n^T)^{-1}. \quad (4)$$

For the experiment, several audio files from different classes were selected, namely: File 1 and File 2. "File 1.wav" is a file containing human speech, while "File 2.wav" is a file containing animal sounds. Audio files with a brief description of their content are presented in Table 1.

Table 1

Description of audio files

File 1.wav	«Внимание! Говорит и показывает Москва»
File 2.wav	Sounds of a squirrel

1) The input to the transformation program was a sound signal from files (see Table 1).

2) The sound signal was loaded and transformed into a one-dimensional array to simplify further operations.

3) If necessary, the number of audio file samples was adjusted to fit the size of the key matrix [7].

4) Two types of masking (unidirectional and bidirectional) and the Hadamard matrix were used for masking the sound signal. The sound signal was transformed into a matrix with data block distribution, allowing operations with the masking matrix.

5) The masking matrix was applied to the input sound signal, resulting in the formation of a masked sound file. During the masking process, a simulated loss of a certain number of rows (10 rows) in the masked signal was also imitated.

6) Metrics were calculated for the masked sound file: root mean square (RMS), signal-to-noise ratio (SNR), peak signal-to-noise ratio (PSNR), Log Spectral Distance (LSD), normalized cross-correlation coefficient (NCM), and total harmonic distortion (THD) in percentage.

7) After demasking the sound signal, the same metrics were recalculated for the demasked signal, allowing for an assessment of the quality of the restored signal.

8) For evaluating the quality of the demasked signal, PESQ (Perceptual Evaluation of Speech Quality) methods were used.

The experiments were conducted using special software implemented in MATLAB.

Simulation results

The simulation results are presented in Figures 1-8. Figures 1-4 show the results for the audio file "File 1.wav" with one-sided and two-sided masking, while Figures 5-8 show the results for "File 2.wav" with one-sided and two-sided masking, respectively. On all figures, the following color coding is used: red represents the Root Mean Square (RMS) metric, dark orange represents the Normalized Cross-Correlation Coefficient (NCM), light orange represents the Mean Opinion Score Listening Quality Objective (MOS-LQO), yellow represents the subjective Mean Opinion Score from listeners (Raw MOS), blue represents the Signal-to-Noise Ratio (SNR), light blue represents the Peak Signal-to-Noise Ratio (PSNR), purple represents the Log Spectral Distance (LSD), and green represents the Total Harmonic Distortion (THD) in percentage.

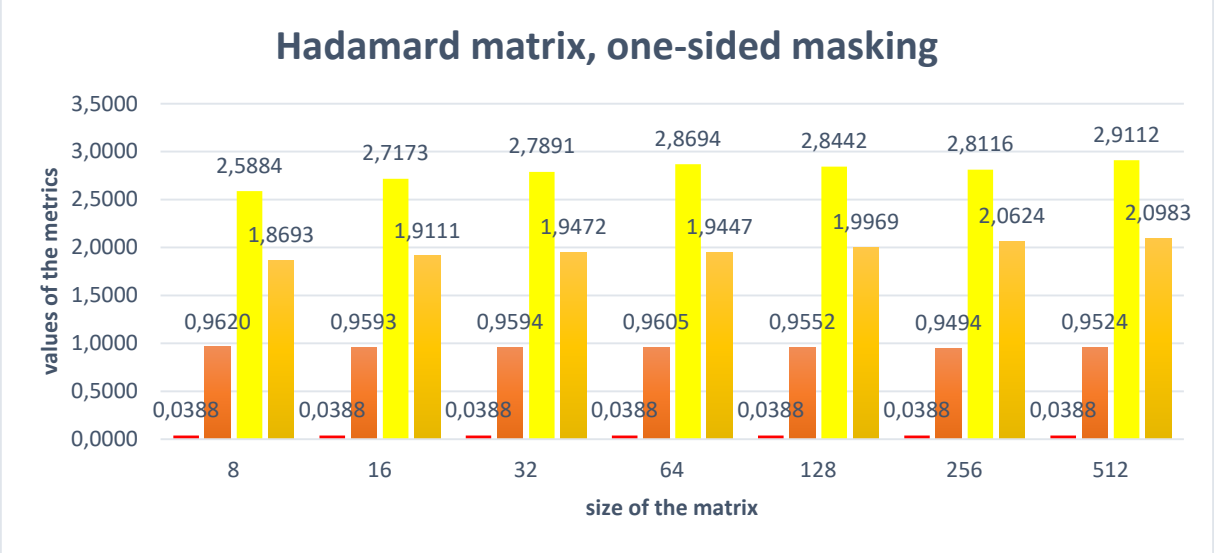


Fig. 1. File 1.wav. Values of the metrics RMS, NCM, Raw MOS, MOS-LQO. One-sided masking

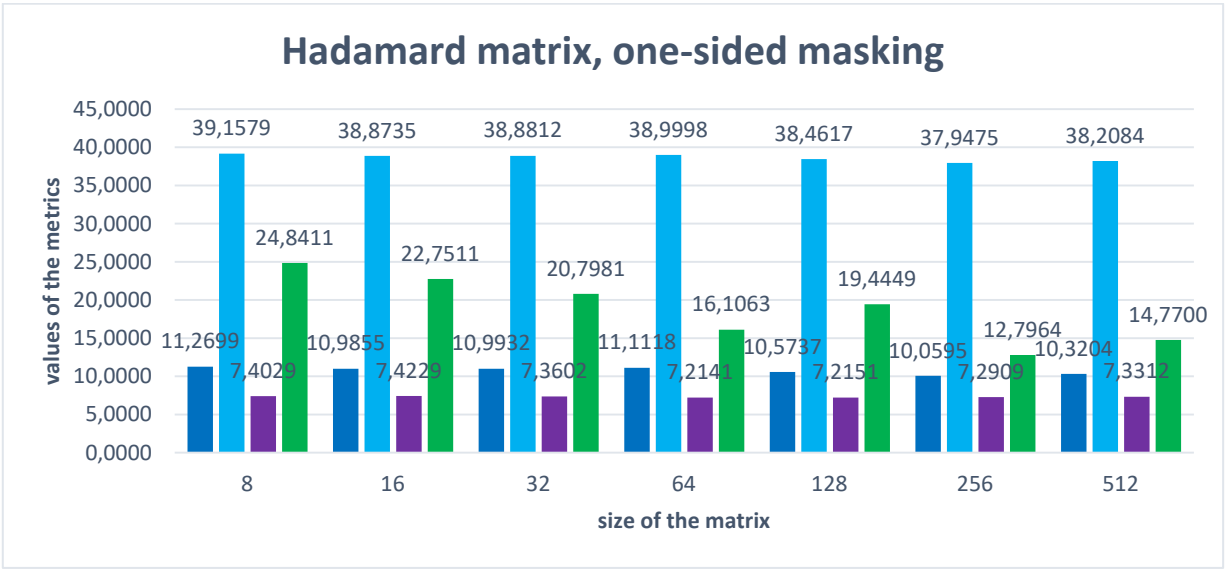


Fig. 2. File 1.wav. Values of the metrics SNR, PSNR, LSD, THD. One-sided masking

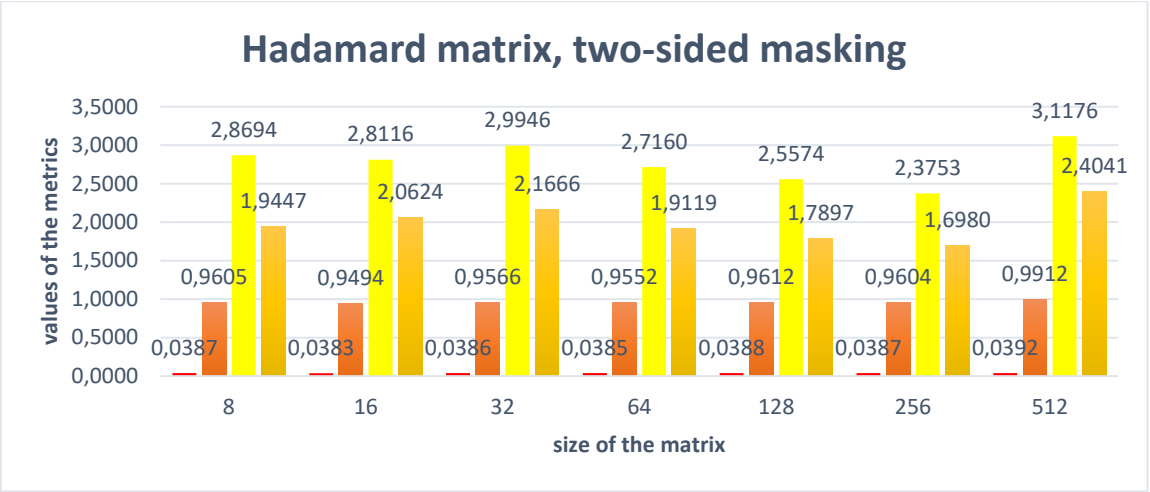


Fig. 3. File 1.wav. Values of the metrics RMS, NCM, Raw MOS, MOS-LQO. Two-sided masking

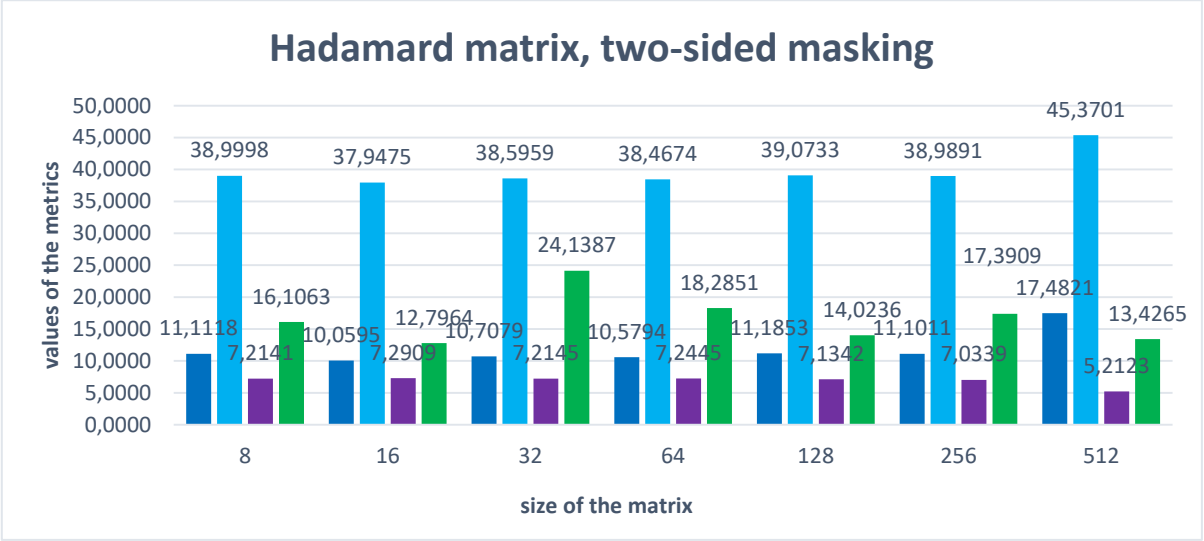


Fig. 4. File 1.wav. Values of the metrics SNR, PSNR, LSD, THD. Two-sided masking

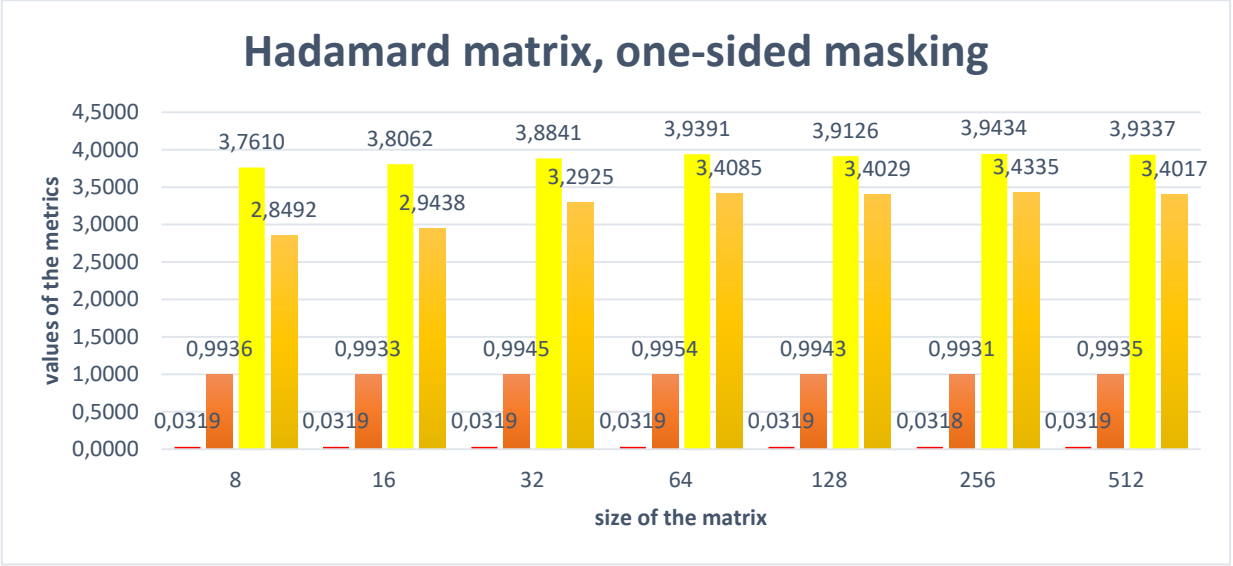


Fig. 5. File 2.wav. Values of the metrics RMS, NCM, Raw MOS, MOS-LQO. One-sided masking

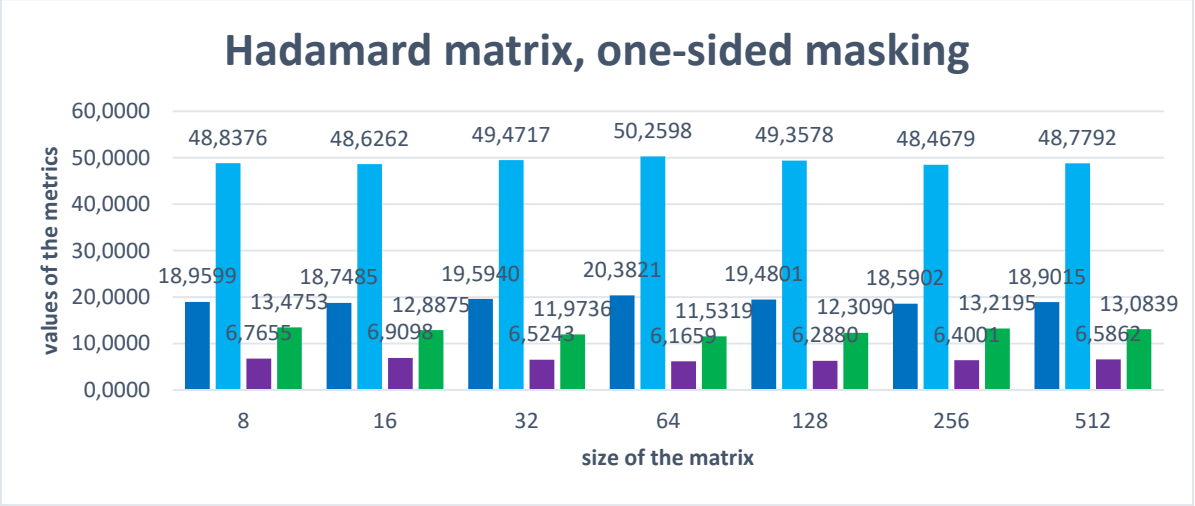


Fig. 6. File 2.wav. Values of the metrics SNR, PSNR, LSD, THD. One-sided masking.

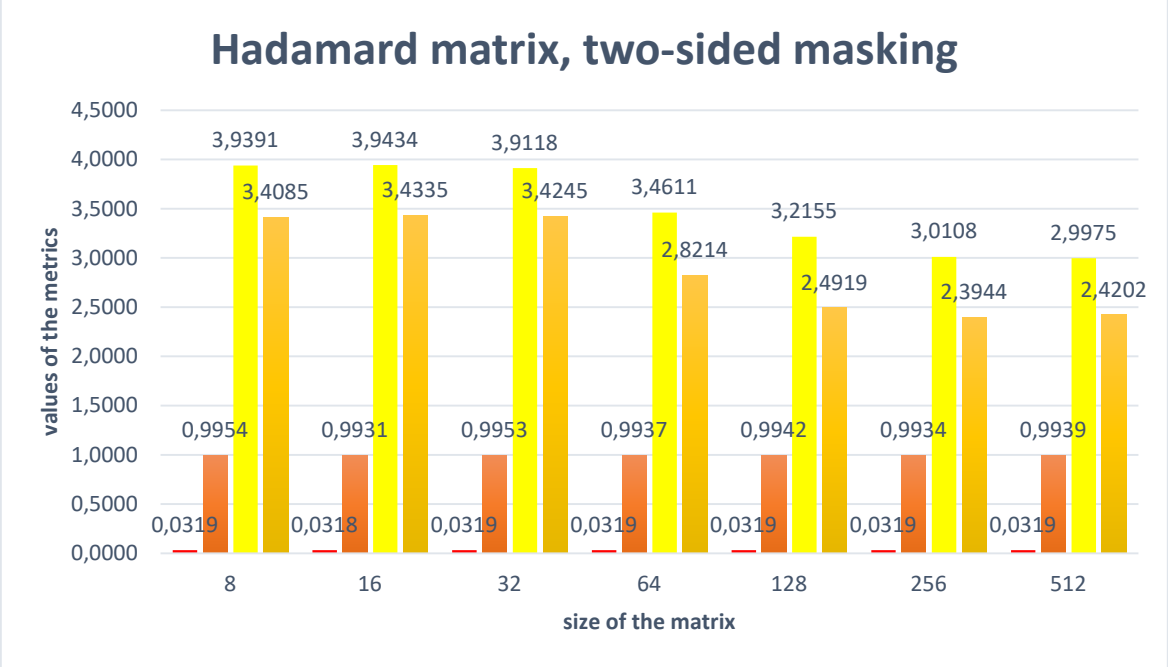


Fig. 7. File 2.wav. Values of the metrics RMS, NCM, Raw MOS, MOS-LQO. Two-sided masking.

The RMS values for both masking types in both files remain within a similar range, indicating a stable root mean square value of the signal. This suggests that masking primarily affects the phase or frequency characteristics, not the overall signal energy. The THD values are also similar in both files for one-sided and two-sided masking. In the first file, the THD percentage ranges from 10.6% to 24.8%, while in the second file, the range is from 11.53% to 13.52%.

Let's first analyze the graphs obtained from processing "File 1.wav."

Two-sided masking achieves significantly higher SNR and PSNR values with a matrix size of 512. In other cases, the difference is not as large, but two-sided masking still leads.

One-sided masking consistently yields higher LSD values than two-sided masking, indicating greater spectral distortion. Only when the matrix size is 64 is the LSD value slightly better (by 0.0304) for one-sided masking.

Two-sided masking with a matrix size of 512 yields an NCM approaching 1, which represents near-perfect correlation, indicating minimal distortion. In other cases, the values are close.

Raw MOS and MOS-LQO scores: Two-sided masking with a 512 matrix yields the best values, corresponding to better quality.

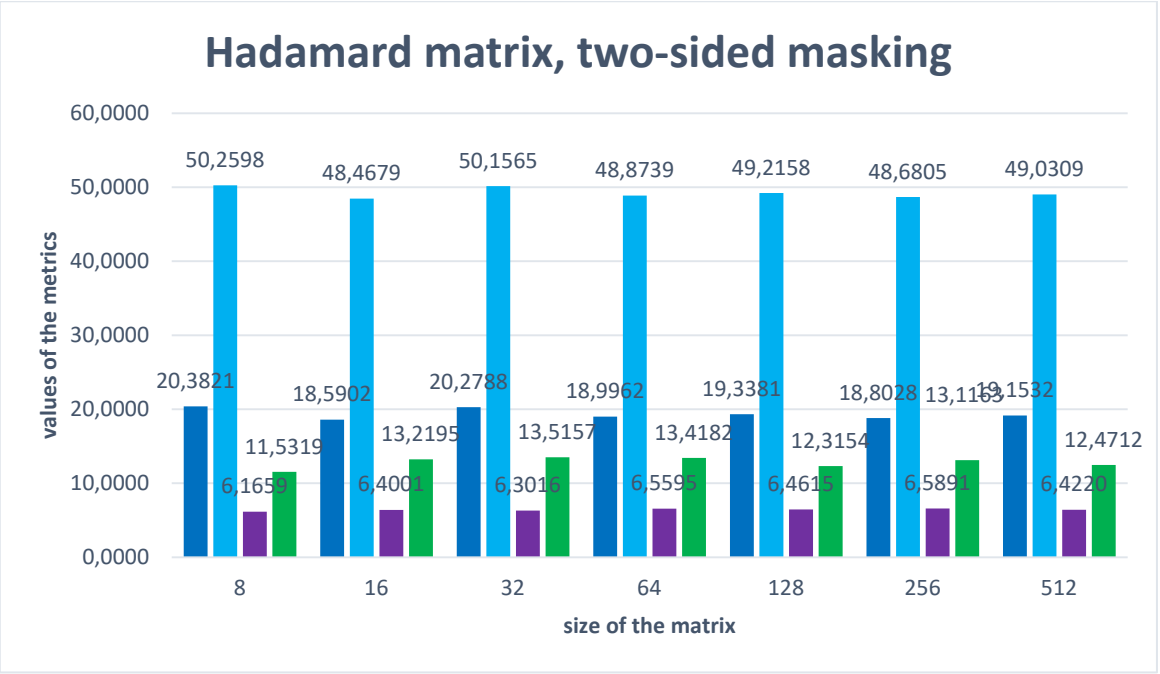


Fig. 8. File 2.wav. Values of the metrics SNR, PSNR, LSD, THD. Two-sided masking.

All these factors indicate that for two-sided masking, a larger matrix size allows for more effective consideration of long-term dependencies in the signal and more accurate modeling of the masking effect. For one-sided masking, no significant dependence on the matrix size is observed.

Let's also analyze the graphs obtained from processing "File 2.wav."

Overall, the difference between one-sided and two-sided masking is much smaller than in File 1. The SNR, PSNR, and NCM values for one-sided and two-sided masking are almost identical, except for LSD, Raw MOS, and MOS-LQO, where there is a minor difference. LSD values for one-sided masking show worse results in 4 out of 7 cases, with masking matrix sizes of 8, 16, 32, and 512. The Raw MOS and MOS-LQO values for one-sided masking show better results in 4 cases, namely with larger masking matrix sizes of 64, 128, 256, and 512.

The dependence of the metrics on the matrix size is not strongly expressed when processing "File 2.wav."

The audio files after distortion were also listened to. When listening to the first audio file, human speech was clearly audible, but interference was present, especially with one-sided masking. When listening to the second audio file, it was more difficult to assess the noise immunity by ear, as it contained squirrel sounds that mixed with the interference.

Conclusion

Therefore, two-sided masking is better used when high accuracy is important and long-term dependencies in the signal need to be considered. It allows for more effective modeling of complex masking effects but requires more computational resources. This is particularly useful for signals with complex structure and long-term dependencies. For "File 1.wav" in this dataset, it clearly performs better, especially with a matrix size of 512.

One-sided masking is better used when low computational complexity and fast processing are required. For "File 2.wav" in this dataset, it shows comparable, and sometimes even better, results than two-sided masking.

Different classes of audio files may have different statistical properties, which affects the effectiveness of masking.

A small matrix size does not take into account long-term dependencies in the signal; it may be insufficient for effective masking of complex sounds. A large matrix size considers long-term dependencies but requires more computational resources.

In conclusion, the choice of the optimal masking method and matrix size depends on the specific task and characteristics of the processed audio signal.

Currently, the number of quasi-orthogonal matrices has significantly increased, and further research on this issue is advisable, using matrices from other families and structures.

Acknowledgment

The author expresses gratitude to the senior lecturer of the Department of Computer Systems and Networks Grigoriev E. K., for his scientific guidance in preparing this article.

References

1. Shinkarenko, K. V. Noise-immune coding of multimedia data in computer networks / K. V. Shinkarenko, A. M. Korikov // Bulletin of the Tomsk Polytechnic University. – 2008. – Vol. 313, No. 5. – Pp. 37-41.
2. Mironovsky, L. A. Strip-method of image and signal transformation: Monograph / St. Petersburg: Politekhnika, 2006. 163 p.
3. Methods of image transformation and signal coding in distributed system channels based on special quasi-orthogonal matrices: Specialty 05.12.13 "Systems, networks, and devices of telecommunications": Dissertation for the degree of Candidate of Technical Sciences / Sergeev Alexander Mikhailovich. – St. Petersburg, 2020. – 153 p.
4. Khvoshch, S. T. Hadamard matrices in space communication / S. T. Khvoshch // Engineering Bulletin of the Don. – 2024. – No. 1(109). – Pp. 210-222.
5. Balonin, N. A. Special matrices: Pseudoinverse, orthogonal, Hadamard, and Cretan / N. A. Balonin, M. B. Sergeev. – St. Petersburg: Politekhnika Publishing House, 2019. – 196 p. – ISBN 978-5-7325-1155-0. – DOI 10.25960/7325-1155-0.
6. Aydarkin, E. E. Experimental study of matrix methods for erasure protection in digital data transmission channels / E. E. Aydarkin, V. M. Deundyak, E. A. Pozdnyakova // News of Higher Educational Institutions. North-Caucasian Region. Technical Sciences. – 2017. – No. 3(195). – Pp. 97-104. – DOI 10.17213/0321-2653-2017-3-97-104.
7. Grigoriev, E. K. Quality assessment of matrix masking for digital audio data / E. K. Grigoriev, A. M. Sergeev // Proceedings of Educational Institutions of Communication. – 2023. – Vol. 9, No. 3. – Pp. 6–13. – DOI 10.31854/1813-324X-2023-9-3-6-13.
8. ITU-T Recommendation P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs // International Telecommunication Union (ITU-T). – URL: https://www.itu.int/rec/T-REC-P.862.
9. Wu, Y., Wang, D. A two-stage algorithm for one-microphone reverberant speech enhancement // Interspeech 2008. – 2008. – Pp. 2674-2677. – DOI: 10.21437/Interspeech.2008-697. – URL: https://www.isca-archive.org/interspeech_2008/wu08b_interspeech.pdf.
10. Total Harmonic Distortion (THD) // MathWorks Help Center. – URL: https://ch.mathworks.com/help/sps/powersys/ref/totalharmonicdistortion.html.

ANALYSIS OF THE QUALITY OF MATRIX MASKING OF VISUAL INFORMATION WITH CONSIDERATION OF HUMAN PERCEPTION

Tatiana Kazakevich
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E – mail: kazakevitchtatjana@yandex.ru

Abstract. *The paper examines an approach to assessing the quality of matrix masking of visual information using quasi-orthogonal matrices. The method is based on the analysis of the two-dimensional correlation coefficient and subjective data from observers. The results enable a numerical evaluation of masking effectiveness for matrices of different structures and confirm the possibility of achieving sufficient quality at a correlation threshold not exceeding 0.45.*
Keywords: *matrix masking, quasi-orthogonal matrices, masking quality, confidentiality assurance, visual experiment.*

Introduction

Modern methods for protecting digital images and video information are becoming increasingly relevant amid the growing threats of cyberattacks and data leaks [1]. An actively developing approach to safeguarding visual information is matrix masking [2], which involves the use of low-level structured quasi-orthogonal matrices, such as Hadamard, Mersenne, and others [3]. This concept relies on the application of linear algebraic operations to transform visual information into a form that conceals its content while allowing data recovery with the corresponding masking matrix.

There are two schemes of matrix masking using these matrices: one-sided and two-sided. In the first case, the original image, its fragment, or a video frame is multiplied by a key matrix of the same size on the right. In the second case, it is multiplied by the key matrix on the left and its transposed version on the right. The use of quasi-orthogonal matrices $\mathbf{M}_n$ for masking, for which $(M_n)^{-1} = M_n^T$, simplifies the inverse transformation: for one-sided masking, it is enough to multiply the image pixel matrix by the transposed matrix M_n^T on the right, and for two-sided masking – by on M_n^T the left and on M_n the right.

An open question in this area is to find an objective metric for assessing the quality of masked images that would provide a correlation between the computed estimates and those obtained from live observers. Classical metrics such as PSNR (Peak Signal to Noise Ratio) and SSIM (Structural Similarity Index), do not always allow adequately assessing the effectiveness of masking. For example, even with high PSNR values, clear contours of objects can be preserved, which makes partial image restoration possible [4]. In this regard, alternative methods of quality assessment are actively being studied.

Literature review

Modern studies of matrix masking of visual information using structured quasi-orthogonal matrices (e.g. Hadamard, Mersenne, Walsh) demonstrate its effectiveness as a privacy method [7, 8, 9]. However, the key problem remains the lack of a universal metric that would objectively evaluate the quality of masking and correlate with human perception. Despite significant progress in developing masking algorithms, current approaches to evaluating their effectiveness rely on metrics originally designed for other purposes, such as compression or data transmission, limiting their applicability to this task.

The most commonly used metrics for image quality assessment are full-reference metrics such as MSE (Mean Squared Error), PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index) – are also used in the context of matrix masking [10, 11]. These methods compare the original image with the masked one, providing quantitative indicators of distortion. However, their effectiveness in assessing masking is questionable. For example, as noted in [10], when using small matrices, the MSE and PSNR values can be high, indicating significant distortion, but the contours of objects in the image remain distinguishable. This allows for the potential recovery of the content of visual information even without a key, which undermines the purpose of masking. SSIM, which focuses on structural similarity, also has shortcomings. Although it takes into account brightness, contrast, and texture, its results do not always reflect the degree of information hiding. In [12], it is

emphasized that SSIM can give high scores for images with preserved structural elements despite distortion, which makes it unreliable for assessing the quality of masking. An example from [10] shows that with two-sided masking by a 16×16 Hadamard matrix, the PSNR value was -18.41 dB and the MSE was 69.35, but visual analysis revealed remaining contours, which confirms the discrepancy between the numerical indicators and the actual masking effectiveness.

These limitations are justified by the fact that classical metrics are not adapted to the specifics of human perception in the context of masking. The human visual system is sensitive to contours, symmetry, and semantic elements that can remain recognizable even under significant distortions [13]. Thus, high numerical distortion rates do not guarantee sufficient content concealment, which makes the search for a specialized metric a priority task.

In an attempt to overcome the limitations of classical metrics, researchers turn to alternative methods. For example, in [11], an approach based on the analysis of the spectral proximity of masked data to white noise is proposed. This method has proven effective for audio data, but its applicability to visual information requires additional research, since the spectral characteristics of images differ significantly from audio signals. Similarly, hybrid models such as VMAF (Video Multi-Method Assessment Fusion), which integrate texture characteristics and noise, demonstrate improved results in video quality assessment [14]. However, in [15], it is noted that VMAF does not always outperform SSIM and can be artificially inflated without taking into account the specifics of masking, which limits its use in this area.

No-reference metrics, such as NIQE (Natural Image Quality Evaluator) or BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator), could theoretically be suitable for evaluating masked images without access to the original, but their application in masking tasks remains underexplored [5]. These methods focus on identifying artifacts (noise, blur) rather than analyzing the degree of structure concealment, which makes them less relevant for the current task.

Materials and methods

In [16], a method was developed based on the analysis of pixel correlation and the two-dimensional correlation coefficient between the original and masked images. This approach allowed numerical confirmation that two-sided masking using symmetric structure matrices (Mersenne matrices structured according to Walsh [9]) provides higher masking quality compared to one-sided masking or the use of cyclic structure matrices (Mersenne matrices based on the M-sequences [2]). In addition, it was found that a two-dimensional correlation coefficient value of no higher than 0.4 guarantees sufficient destruction of the image structure, making its contents visually unrecognizable.

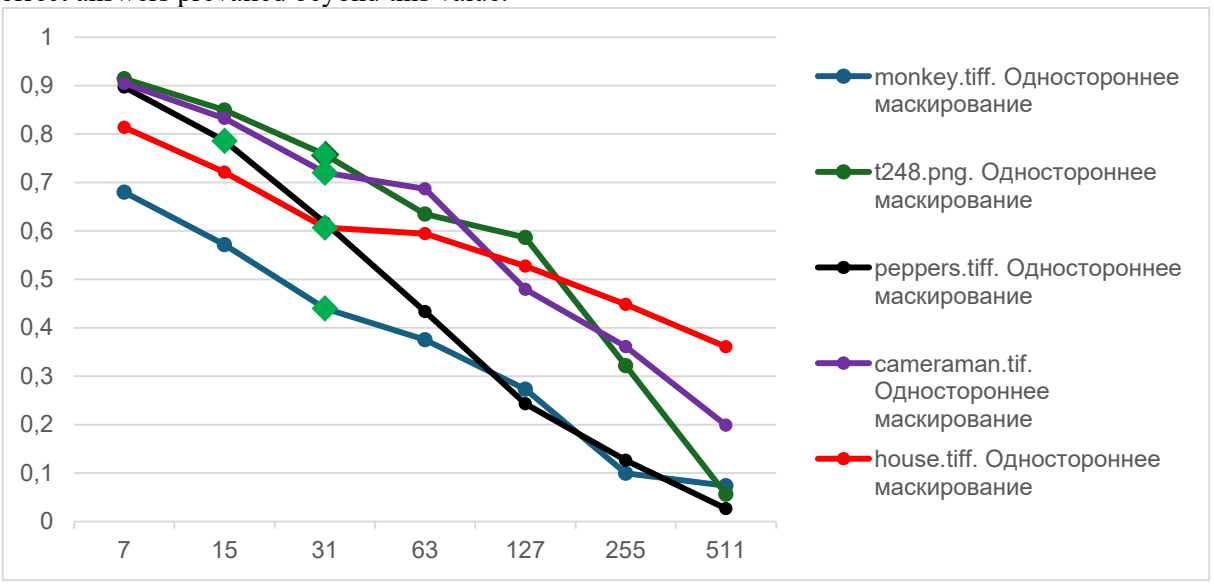
The experiment in this study aims to develop the proposed approach with an emphasis on subjective assessment of masking quality and refinement of correlation thresholds. The main objective was to expand the empirical base by increasing the number of participants and the image bank. Unlike the previous study [16], where images were shown in descending order of the masking matrix size, a random order of display was used here, which minimized the effect of sequence on observers' perception. To analyze subjective responses, the Sentence BERT neural network model was used, which ensured accurate classification of responses based on their semantic similarity to the correct options.

The experiment included the following stages:

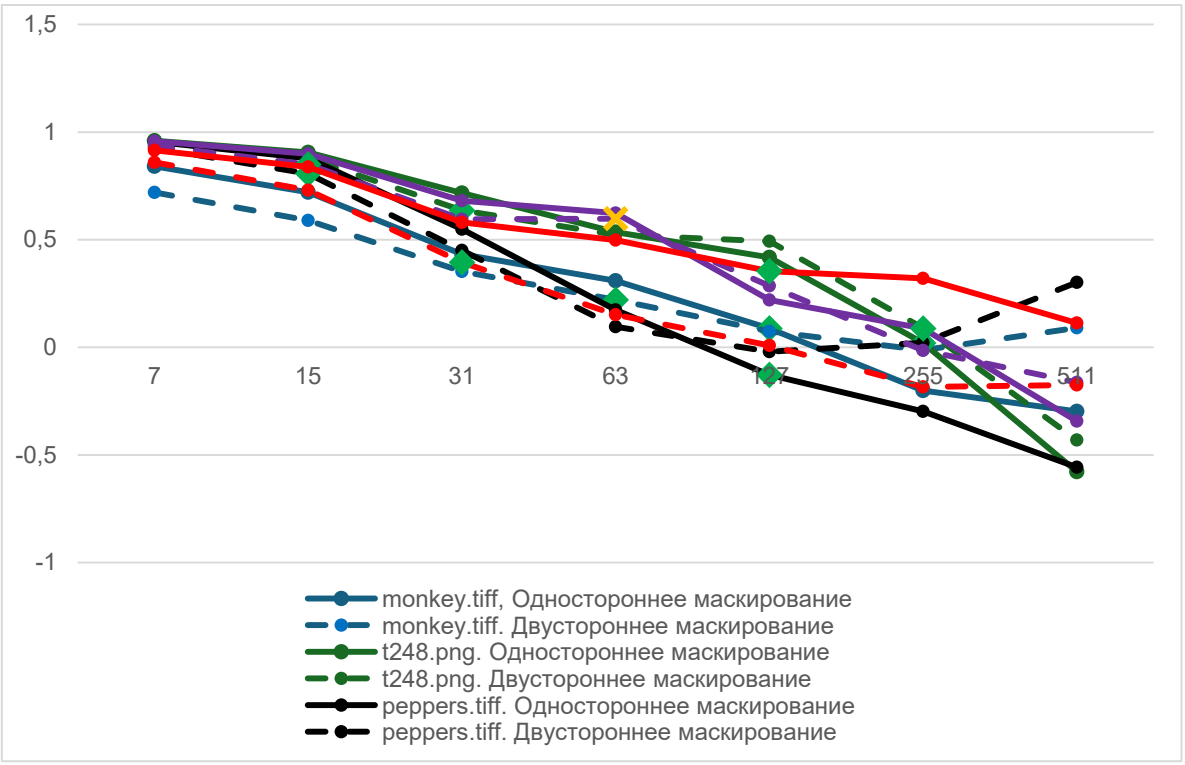
1. Image database preparation: Three classes of images were used – aerial photographs, textures and images with distinct objects – to more accurately assess the impact of the level of detail of the image content on the masking quality.
2. Image masking: both one-sided and two-sided matrix masking [20] were used using Mersenne matrices of symmetric and cyclic structure. The sizes of the masking matrices varied from 7×7 to 511×511 pixels. This step was implemented in MATLAB, which was also used to compute correlation characteristics.
3. Data collection: Participants were randomly shown images and asked to guess their original content. Responses were recorded and processed using automated methods. Over 30 independent respondents with no prior exposure to the original images took part in the experiment. Subjects represented a variety of age, ethnicity, and gender groups to ensure a representative sample.
4. Results analysis: The two-dimensional correlation coefficient values were calculated for each image, and subjective responses were classified as “correct” or “incorrect” based on the similarity index (threshold ≥ 0.5). The obtained data were compared with the results of visual analysis and correlation metrics to determine the threshold values that ensure sufficient masking quality.
5. Plotting: plots of the two-dimensional correlation coefficient versus the size of the masking matrix were created.

Results

Figures 1–3 show graphs demonstrating the results of experimental data processing, illustrating how the two-dimensional correlation coefficient changes depending on the size of the masking matrix. Three types of images were considered in the study: aerial photographs, textures, and salient objects. The resulting graphs also contain special markers that allow for better interpretation of the results. Green diamonds indicate the range of the correlation coefficient, starting from which the number of correct answers of the subjects begins to prevail over the erroneous ones. Yellow crosses indicate those values at which correct answers were also more common than incorrect ones, but were not dominant in the entire sample, since incorrect answers prevailed beyond this value.

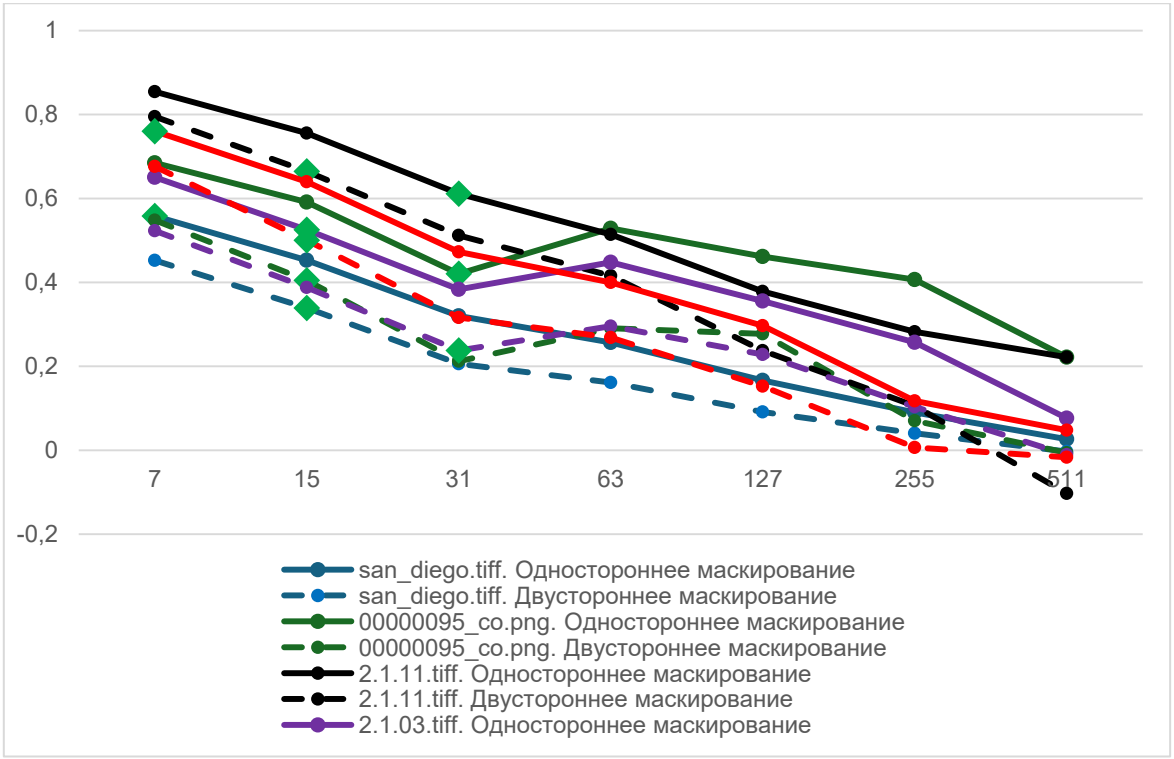


a)

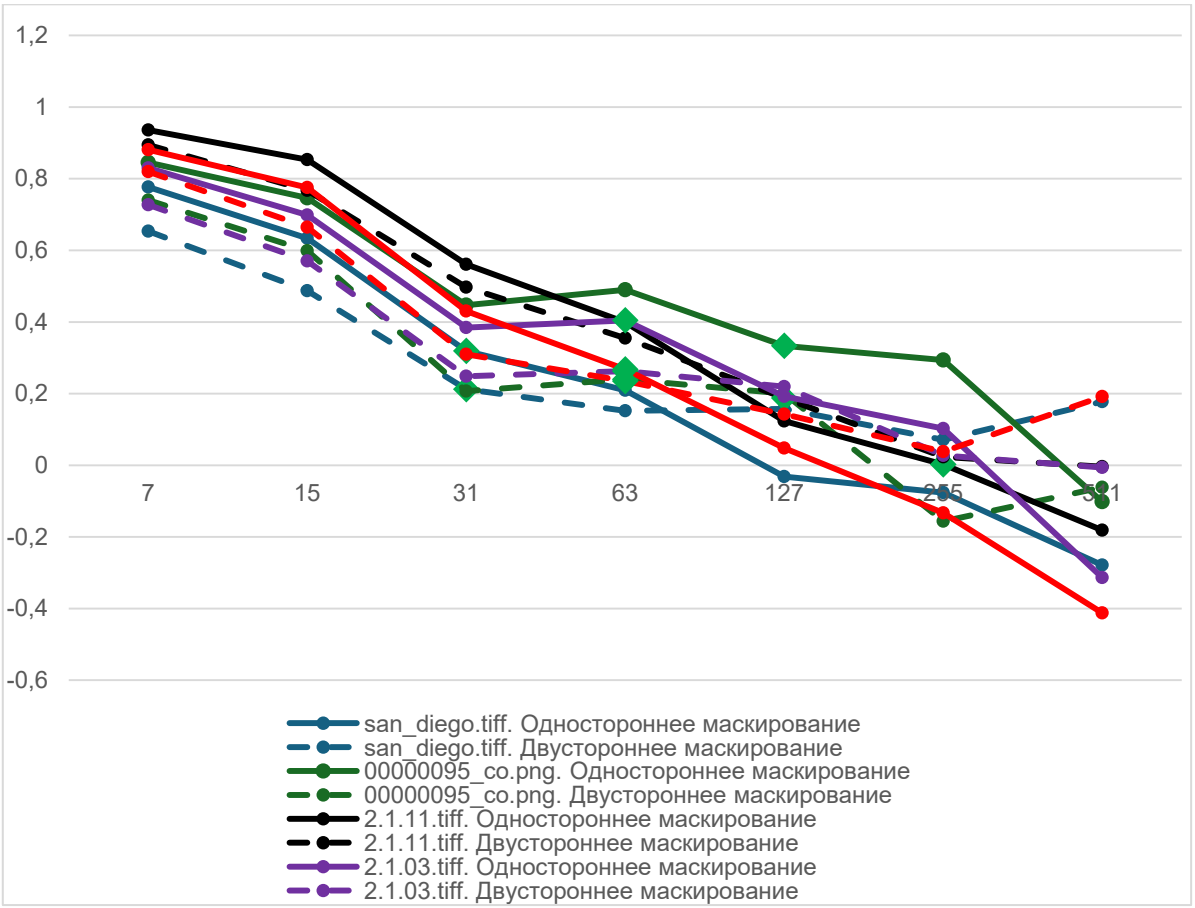


b)

Fig. 1. Graph of the dependence of the two-dimensional correlation coefficient on the size of the matrix of a symmetric structure (a), a cyclic structure (b) for images of clearly visible objects

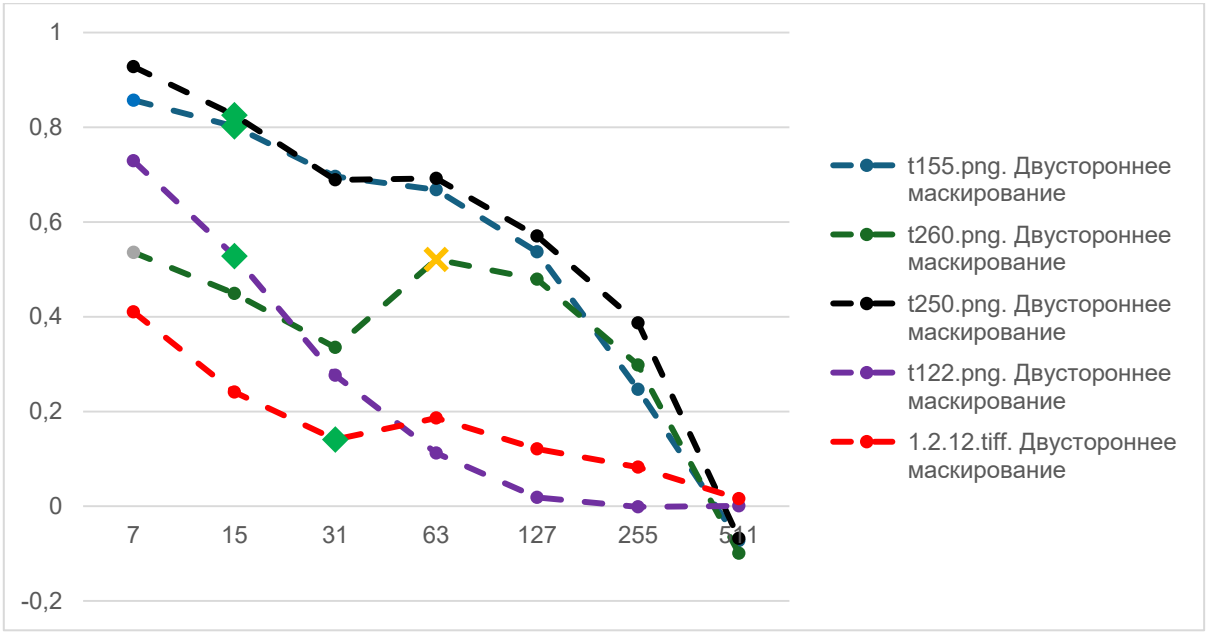


a)

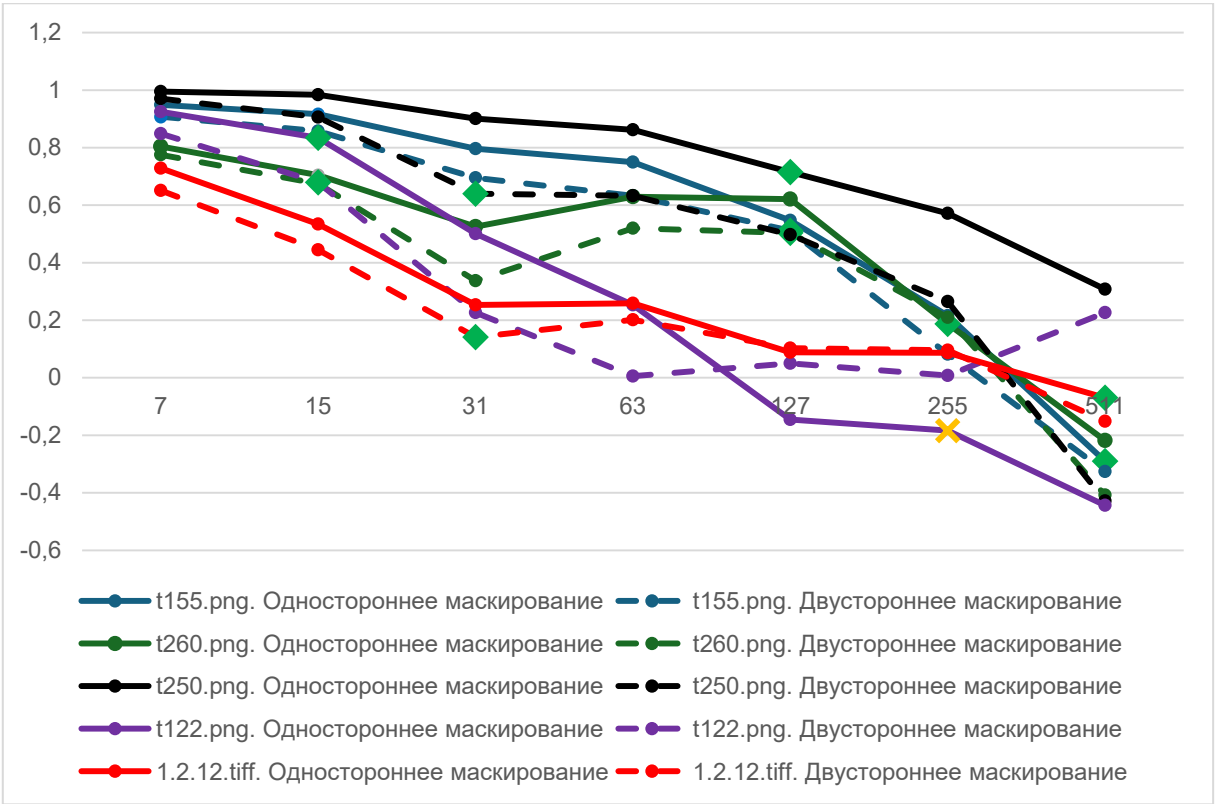


b)

Fig. 2. Graph of the dependence of the two-dimensional correlation coefficient on the size of the matrix of a symmetric structure (a), a cyclic structure (b) for aerial photographs



a)



b)

Fig. 3. Graph of the dependence of the two-dimensional correlation coefficient on the matrix size of a symmetric structure (a), a cyclic structure (b) for texture images

Analysis of the survey results showed that respondents frequently identified the images correctly. Specifically, the average two-dimensional correlation coefficient at which identification became possible was relatively low. For aerial photographs, the identification threshold was 0.49; for textures, it was 0.37; and for images with prominent objects, it was 0.47. On average, to achieve acceptable masking quality, the two-dimensional correlation coefficient should not exceed approximately 0.45 when it is not possible, for

any reason, to ensure that the masking matrix size matches the original image size (the best possible masking quality).

Conclusion

This study compared various methods of matrix masking of images using quasi-orthogonal matrices, such as Mersenne matrices structured according to Walsh and Mersenne matrices based on M-sequences. Particular attention was paid to evaluating masking effectiveness and analyzing existing metrics used for the quantitative assessment of noisy image quality.

The proposed method of subjective masking assessment based on the analysis of observers' responses using the Sentence BERT neural network model allowed us to obtain objective data on the visual discernibility of images. It was found that when using two-sided masking with symmetrical structure matrices, the structure of masked images was disrupted better than when using cyclic structure matrices or one-sided masking, which ensured lower recognition of the original content during the visual experiment.

Analysis of experimental data showed that there is a threshold value of the two-dimensional correlation coefficient (not higher than 0.4) at which sufficient masking quality is achieved.

Acknowledgements

The author expresses gratitude to the senior lecturer of the Department of Computing Systems and Networks Grigoriev E.K. for scientific guidance in the preparation of this article.

References

1. https://www.infowatch.ru/analytics/analitika

2. 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.

3. Balonin N. A., Sergeev M. B. Matrices of local maximum of determinant // Information and control systems. 2014. No. 1(68). P. 2-15.

4. Erosh I. L., Sergeev A. M., Filatov G. P. On the protection of digital images during transmission over communication channels // Information and control systems. 2007. Vol. 5. No. 30. P. 20-22.

5. Dost, Shahi & Saud, Faryal & Shabbir, Maham & Khan, Muhammad Gufran & Shahid, Muhammad & Lövström, Benny. (2022). Reduced reference image and video quality assessments: review of methods. EURASIP Journal on Image and Video Processing. 2022. 10.1186/s13640-021-00578-y.

6. Sadou, Besma & Atidel, Lahoulou & Bouden, Toufik. (2018). PFF-RVM: A new no reference image quality measure. Mathematics and Computers in Simulation . 167. 10.1016/j.matcom.2018.11.005.

7. Digital masking with Mersenne matrices and its special images / Yu. N. Balonin, A. A. Vostrikov, E. A. Kapranova [et al.] // Fundamental research. – 2017. – No. 4-1. – P. 13-18.

8. On the selection of matrices for image masking and unmasking procedures / A. A. Vostrikov, O. V. Mishura, A. M. Sergeev, S. A. Chernyshev // Fundamental research. – 2015. – No. 2-24. – P. 5335-5339.

9. Sergeev, A. M. Walsh-structured two-level and modular two-level quasi-orthogonal matrices for image masking // Bulletin of higher educational institutions. Instrument engineering. – 2023. – Vol. 66, No. 5. – P. 399-408. – DOI 10.17586/0021-3454-2023-66-5-399-408.

10. Chernyshev S. A. Development and research of the method of matrix masking of video information in globally distributed systems: diss . ... Cand. of Technical Sciences. St. Petersburg, 2018. 120 p.

11. Grigoriev E.K., Sergeev A.M. Assessment of the quality of matrix masking of digital audio data. Transactions of educational institutions of communication. 2023; 9 (3): 6-13. https://doi.org/10.31854/1813-324X-2023-9-3-6-13

12. Starovoytov V. V. The SSIM index is not a metric and poorly assesses the similarity of images // Informatics. 2019. No. 2. P. 1–17

13. N.N. Krasilnikov. Digital processing of 2D and 3D images. – SPb.: «BHV-Petersburg», 2011. – 608 p.

14. Blog, Netflix Technology (2016-06-06). "Toward A Practical Perceptual Video Quality Metric". Netflix TechBlog .

15. Bampis, Christos G.; Bovik, Alan C. (2017-03-02). "Learning to Predict Streaming Video QoE : Distortions, Rebuffering and Memory." arXiv:1703.00633

16. Kazakevich, T. About one approach to assessing the quality of masking visual information. Bulletin of the UNESCO Chair "Distance education in engineering" of the SUAI: Collection of the papers. – SPb .: SUAI, 2024. – P. 254

MAIN NON-GAUSSIAN PROBABILITY DENSITIES USED IN MODELING TECHNICAL SYSTEMS

Vadim Khudyakov
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia,
E-mail: vadym.khudyakov@yandex.ru

Abstract. *This paper presents two-dimensional probability densities of stochastic processes used in mathematical modeling of automation systems. These models account for both the probability distribution laws of the simulated processes and their correlation-spectral characteristics. Such algorithms are widely applied in designing complex automation systems throughout their lifecycle.*
Keywords: *modeling algorithm, two-dimensional probability densities, non-Gaussian field, correlation-spectral characteristics, non-Gaussian processes.*

Traditional descriptions of stochastic processes, including non-Gaussian ones, rely on multivariate probability densities [1]. However, determining multivariate densities for real-world processes is practically infeasible. Moreover, validating the adequacy of hypothetical multivariate densities against real signal densities is fundamentally impossible due to the excessively large sample sizes required under identical experimental conditions. Therefore, practical approaches are limited to ensuring the adequacy of marginal distribution laws and corresponding correlation-spectral characteristics between models and real signals.

Under these constraints, experimental validation of mathematical models becomes achievable, as experiments yield reliable estimates of both one-dimensional probability densities and correlation functions or signal spectra. The challenges in obtaining multivariate non-Gaussian densities and the sufficiency of two-dimensional densities for experimental validation (model verification) justify restricting the analysis to two-dimensional densities [2]. Hence, this work focuses solely on two-dimensional distributions applied in non-Gaussian process and field modeling [3]:

1. Nakagami Distribution (m-Distribution)

This distribution describes the joint probability density of radio signal envelopes:

$$w(U_1,U_2,m,\sigma_1,\sigma_2,r_2)=(1-r_2)^m\sum_{n=0}^{\infty}\frac{(m)_n}{n!}r_2^n w(U_1,m,\sigma_1)\,w(U_2,m,\sigma_2), \tag{1}$$

where

$$w(U_i,m\ ,\sigma_i)=\left\{\frac{2\sigma_i^{-(m+n)}\left(m+n^{m+n}\right)}{\Gamma(m+n)(1-r_2)^{m+n}}\right\}\cdot U_i^{2(m+n)-1}\cdot\exp\left(-\frac{(m+n)U_i^2}{\sigma_i(1-r_2)}\right),\ i=1,2; \tag{2}$$
$$(m)_n=m(m+1),\ldots,(m+n-1);\quad (m)_0=1;$$

Here r_2 – is the cross-correlation coefficient between squared envelopes U_1 Ц U_2, m and σ_i – are parameters.

If U_1 Ц U_2 are uncorrelated with $m_1\neq m_2$:

$$w(U_1,U_2,m_1,m_2,\sigma_1,\sigma_2)=\frac{4}{\Gamma(m_1)\Gamma(m_2)}\left(\frac{m_1}{\sigma_1}\right)^{m_1}\left(\frac{m_2}{\sigma_2}\right)^{m_2}\times$$
$$\times U_1^{2m_1-1}U_2^{2m_2-1}\exp\left\{-\frac{m_1}{\sigma_1}U_1^2-\frac{m_2}{\sigma_2}U_2^2\right\}. \tag{3}$$

2. Bivariate Rayleigh-Rice Distribution (Generalized Rayleigh Distribution) [5].

This distribution describes the two-dimensional density of the sum of a stationary Gaussian process $n(t) \sim N\left(0, \sigma_{\cdot}^2\right)$ and a deterministic signal $s(t) = a(t)\cos(\omega_0 t)$. The resulting non-stationary process has a time-dependent density:

$$w(\xi_1, \xi_2, \tau, t) = \frac{\xi_1 \xi_2}{\sigma^4\left(1 - R^2(\tau)\right)} \exp\left\{-\frac{\xi_1^2 + \xi_2^2 + a_1^2 + a_2^2 - 2a_1 a_2 R(\tau)}{2\sigma^2\left(1 - R^2(\tau)\right)}\right\} \times$$

$$\times \sum_{m=0}^{\infty} \varepsilon_m I_m\left[\frac{R(\tau)\xi_1\xi_2}{\sigma^2\left(1 - R^2(\tau)\right)}\right] I_m\left[\frac{a_1 - R(\tau)\xi_1\xi_2 a_2}{\sigma^2\left(1 - R^2(\tau)\right)}\right] I_m\left[\frac{a_2 - R(\tau)\xi_1\xi_2 a_1}{\sigma^2\left(1 - R^2(\tau)\right)} r_2\right], \tag{4}$$

where $\xi_1 > 0,\ \xi_2 > 0,\ \varepsilon_0 = 1,\ \varepsilon_m = 2,\ m > 0,\ R(\tau)$ – is the normalized correlation function.

For a harmonic signal ω_0 with amplitude U_0, то $a_1 = a_2 = U_0$ and the probability distribution density of a stationary random process $\xi(t)$ It will take the form:

$$w(\xi_1, \xi_2, \tau) = \frac{\xi_1 \xi_2}{\sigma^4\left(1 - R^2(\tau)\right)} \exp\left\{-\frac{\xi_1^2 + \xi_2^2}{2\sigma^2\left(1 - R^2(\tau)\right)}\right\} \exp\left\{-\frac{U_0^2}{\sigma^2\left(1 + R\ (\tau)\right)}\right\} \times$$

$$\times \sum_{m=0}^{\infty} \varepsilon_m I_m\left[\frac{R\ (\tau)\xi_1\xi_2}{\sigma^2\left(1 - R^2(\tau)\right)}\right] I_m\left[\frac{U_0\xi_1}{\sigma^2\left(1 + R\ (\tau)\right)}\right] I_m\left[\frac{U_0\xi_2}{\sigma^2\left(1 + R\ (\tau)\right)}\right],\ \xi_1 > 0,\ \xi_2 > 0. \tag{5}$$

As $\tau \to \infty,\ R(\tau) \to 0,$ (5) simplifies to:

$$w(\xi_1, \xi_2) = \frac{\xi_1}{\sigma^2} \exp\left\{-\frac{\xi_1^2 + a_1^2}{2\sigma^2}\right\} I_0\left(\frac{\xi_1 a_1}{\sigma^2}\right) I_0\left(\frac{\xi_2 a_2}{\sigma^2}\right) \frac{\xi_2}{\sigma^2} \exp\left\{-\frac{\xi_2^2 + a_2^2}{2\sigma^2}\right\}. \tag{6}$$

3. *Laplace Probability Density.*

$$w(\xi_1, \xi_2, \tau_2) = w(\xi_1, | \xi_2, \tau_2) w(\xi_2), \tag{7}$$

$w(\xi_1 \mid \xi_2, \tau)$ – is the conditional density, and $w(\xi_2)$ is the one-dimensional density:

$$w(\xi_1 \mid \xi_2, \tau) = \frac{\sqrt{2}}{\sqrt{\sigma_\xi^2\left(1 - R^2(\tau)\right)}} \exp\left\{\frac{\sqrt{2}\left|\xi_1 - R(\tau)\xi_2\right|}{\sqrt{\sigma_\xi^2\left(1 - R^2(\tau)\right)}}\right\};$$

$$w(\xi_2) = \frac{\sqrt{2}}{\sigma^2} \exp\left\{-\frac{\sqrt{2}(\xi_1(}{\sigma^2}\right\}. \tag{8}$$

4. *Log-Normal Distribution.*

This distribution, or rather its multidimensional generalization, is used to describe the joint probability density distribution of the envelopes of input signals (information and interfering) technical complexes [3]:

$$w\left(U_1, U_2, \sigma_1^2, \sigma_2^2, \mu_1, \mu_2, r_{12}\right) = \frac{U_1^{-1} U_2^{-1}}{2\pi\sigma_1\sigma_2\sqrt{1 - r_{12}^2}} \exp\left\{-\frac{1}{2\left(1 - r_{12}^2\right)} \times\right.$$

$$\left.\times\left[\frac{\left(\ln U_1 - \mu_1\right)^2}{\sigma_1^2} + \frac{\left(\ln U_2 - \mu_2\right)^2}{\sigma_2^2} - \frac{2r_{12}}{\sigma_1\sigma_2}\left(\ln U_1 - \mu_1\right)\left(\ln U_2 - \mu_2\right)\right]\right\}, \tag{9}$$

where $\sigma_1^2, \sigma_2^2, \mu_1, \mu_2$ – are variances and means of the log-envelopes:

$$\sigma_i^2 = \left|\left(\ln U_i - \mu_i\right)^2\right|;\ \mu_i = \left|\ln U_i\right|; \tag{10}$$

r_{12} – is the cross-correlation coefficient between log-envelopes.

Mean envelope values:

$$m_{U_i}\left(\mu_i, \sigma_i\right) = \exp\left(\frac{\sigma_i^2}{2} + \mu_i\right),\ i = 1, 2. \tag{11}$$

Variance:

$$\sigma_{U_i}^2\left(\mu_i, \sigma_i\right) = \exp\left(2\mu_i + \sigma_i^2\right)\left[\exp\left(\sigma_i^2\right) - 1\right], \tag{12}$$

where

$$\begin{cases} \mu_i\left(m_{U_i}, \sigma_{U_i}^2\right) = \ln\left(\dfrac{m_{U_i}}{\sqrt{1 + \sigma_{U_i}^2 m_{U_i}^{-2}}}\right); \\ \sigma_i^2\left(m_{U_i}, \sigma_{U_i}^2\right) = \ln\left(1 + \sigma_{U_i}^2 m_{U_i}^{-2}\right). \end{cases} \tag{13}$$

The function of mutual correlation between envelopes at coincident time points:

$$\begin{aligned} B_{U_1,U_2}^{(\tau)} &= \int_0^\infty \int_0^\infty \left(U_1 - m_{U_1}\right)\left(U_2 - m_{U_2}\right) w\left(U_1 U_2\right) dU_1 dU_2 = \\ &= \exp\left(\mu_1 + \frac{\sigma_1^2}{2}\right)\exp\left(\mu_2 + \frac{\sigma_2^2}{2}\right)\left\{\exp\left(r_{12}\sigma_1\sigma_2\right) - 1\right\}. \end{aligned} \tag{14}$$

Correlation coefficient:

$$r_{12}\left(m_{U_1}, \sigma_{U_1}^2, r_{U_1U_2}\right) = \frac{B_{U_1U_2}}{\sigma_{U_1}\sigma_{U_2}} = \frac{\ln\left\{1 + \dfrac{r_{U_1U_2}\sigma_{U_1}\sigma_{U_2}}{m_{U_1}m_{U_2}}\right\}}{\sqrt{\ln\left(1 + \dfrac{\sigma_{U_1}^2}{m_{U_1}^2}\right)\ln\left(1 + \dfrac{\sigma_{U_1}^2}{m_{U_1}^2}\right)}}. \tag{15}$$

One-dimensional log-normal density:

$$w\left(U, \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2 U}}\exp\left\{-\frac{\left(\ln U - \mu\right)^2}{2\sigma^2}\right\}. \tag{16}$$

The coefficient of asymmetry:

$$k_a = \sqrt{\exp\left(\sigma^2\right) - 1}\left(\exp\left(\sigma^2\right) + 2\right). \tag{17}$$

Kurtosis coefficient:

$$k_э = 3 + \left(\exp\left(\sigma^2\right) - 1\right)\left(\exp\left(3\sigma^2\right) + 3\exp\left(2\sigma^2\right) + 6\exp\left(\sigma^2\right) + 6\right). \tag{18}$$

5. *Bivariate Density of a Stationary Narrowband Gaussian Process Envelope [6,7]:*

$$w\left(\xi_1, \xi_2, \tau\right) = \frac{\xi_1\xi_2}{\sigma^4\left(1 - R^2\left(\tau\right)\right)}\exp\left\{-\frac{\xi_1^2 + \xi_2^2}{2\sigma^2\left(1 - R^2\left(\tau\right)\right)}\right\} I_0\left\{\frac{R\left(\tau\right)\xi_1\xi_2}{\sigma^2\left(1 - R^2\left(\tau\right)\right)}\right\}. \tag{19}$$

where $\tau \to 0;\ R(\tau) \to 0;\ w(\xi_1, \xi_2, \infty) = \frac{\xi_1}{\sigma^2} \exp\left(-\frac{\xi_1^2}{2\sigma^2}\right) \frac{\xi_2}{\sigma^2} \exp\left(-\frac{\xi_2^2}{2\sigma^2}\right)$.

The density of the transition probability distribution has become widespread in applied problems:

$$w(\xi_2, \tau \mid \xi_1) = \frac{w(\xi_1, \xi_2, \tau)}{w(\xi_1)} = \frac{\xi_2}{\sigma^2\left(1 - R^2(\tau)\right)} \exp\left\{-\frac{R^2(\tau)\left(\xi_1^2 + \xi_2^2\right)}{2\sigma^2\left(1 - R^2(\tau)\right)}\right\} \times$$
$$\times I_0\left\{\frac{R(\tau)\xi_1\xi_2}{\sigma^2\left(1 - R^2(\tau)\right)}\right\},\ \ \xi_1 > 0,\ \xi_2 > 0; R(\tau) = \exp\left(-\beta(\tau()\right). \quad (20)$$

Correlation function of the envelope of a narrow-band stationary Gaussian random process:

$$B_U(\tau) = \frac{\pi}{2}\sigma^2\left\{1 + \frac{1}{4}R^2(\tau) + \sum_{n=2}^{\infty} \frac{\left[(2n-3)!!\right]^2}{2^{2n}(n!)^2} R^{2n}(\tau)\right\} =$$
$$= \sigma^2\left\{2E\left[R(\tau)\right] + \left[1 - R^2(\tau)\right]K\left[R(\tau)\right]\right\}, \quad (21)$$

where K and E – are complete elliptic integrals of the first and second kinds.

Here are the basic and simplest two-dimensional probability distribution densities used for the synthesis and analysis of technical systems. From this consideration, the complexity of the task of synthesizing scalar random processes with similar distributions of instantaneous values or parameters immediately becomes obvious. The problem of synthesis of vector (multidimensional) processes and fields is even more difficult. Therefore, when choosing a model, its simplicity is also important, especially since the simplicity of the model allows synthesizing the most efficient modeling algorithms in terms of speed.

The adequacy requirements of a non-Gaussian two-dimensional model for real signals and its simplicity are mutually contradictory, and the compromise in solving this problem strongly depends on the specific subject area.

References

1. Shalygin A.S., Palagin Y.I. Applied methods of statistical modeling- L.: Mashinostroenie. Leningr. Department, 1986. 320 p.

2. Blaunstein N.S., Sergeev M.B., Shepeta A.P. Applied aspects of electrodynamics. St. Petersburg: Agraf+, 2016. 272 p.

3. 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.

4. Shelukhin O.I., Belyakov I.V. Non-Gaussian processes. -St. Petersburg: Polytechnic, 1992. – 312 p., ill.

5. Podoplekin Y.F., Shepeta D.A., Ivanov I.N. Modeling of an anisotropic stochastic spatial field distributed according to the Weibull law. Marine radio electronics. 2024. No. 3 (89). pp. 48-53.

6. Levin B.R. Theoretical foundations of statistical radio engineering. – 3rd ed., revised Moscow: Radio and Communications, 1989. 486 p., ill.

7. Isakov V.I., Shepeta D.A. Modeling of location signals reflected from the edge of the earth-sea // Information and control systems. 2017. No. 5 (90). pp. 89-94.

OPTIMIZATION OF DIFFRACTION GRATING PARAMETERS FOR MULTI-CRITERIA SPECTRAL ANALYSIS TASKS

Veniamin Kitaev
post-graduate student
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
E-mail: kitaev.veniamin@mail.ru

Abstract. *This paper proposes a method for optimizing the parameters of a diffraction grating to improve the efficiency of spectral analysis. An algorithm and its software implementation in MATLAB have been developed to increase the intensity of light in higher diffraction orders. The results demonstrate a significant improvement in the grating's performance for operation in the second and fourth orders.*
Keywords: *diffraction grating topology, higher orders, optimization, simulation, optics, transmission function.*

Introduction

Optical diffraction gratings are one of the key components in modern optical systems used in spectroscopy, laser technologies, telecommunications, and other fields. Their primary function is to spatially separate light into spectral components, enabling high-precision analysis of the spectral composition of radiation [1-4]. The formation of multiple diffraction orders by the grating allows for the creation of several spectral analysis channels. However, the energy of the optical radiation is distributed among the diffraction orders, which reduces the sensitivity of the instrument. The efficiency of a diffraction grating is determined by its topology, which includes parameters such as the grating period (dependent on the width of transparent and opaque stripes), the depth and shape of the lines, the coating material, and the geometric profile of the surface.

The relevance of this work is driven by the increasing demands for the accuracy and efficiency of optical systems [5], as well as the need to develop new approaches to the design of diffraction gratings for modern applications such as quantum optics, biosensing, and integrated photonics. This paper proposes a comprehensive approach to optimizing the topology of a diffraction grating, utilizing a mathematical model implemented in MATLAB. This model describes the determination of the grating period based on the required light intensity in the desired diffraction order. The proposed method represents one of the ways to optimize the performance of a diffraction grating.

1. Topology of a High-Order Diffraction Grating

The topology of a high-order diffraction grating consists of alternating transparent and opaque stripes of varying widths. A key feature of such a grating is that each line is not equal to the others [6].

The transmission function of such a diffraction grating can be written as:

$$T(\xi)=\begin{vmatrix}1,\xi\in[0,a]\\0,\xi\in[a,a+b]\\1,\xi\in[a+b,2a+b]\\0,\xi\in[2a+b,2(a+b)]\\1,\xi\in[2(a+b),3a+2b]\\0,\xi\in[3a+2b,3a+2b+c]\end{vmatrix}, \quad (1)$$

where «*a*» is the width of the transparent element of the grating, and «*b*» and «*c*» are the widths of the opaque lines, with $a \neq b \neq c$.

According to the previously proposed model, the radiation is modulated in space according to the transmission function of the diffraction grating *T(ξ)*, which, for convenience, can be represented as a complex Fourier series:

$$T(\xi)=\sum_{-\infty}^{\infty}C_n\exp(in\frac{2\pi}{T_g}\xi), \quad (2)$$

where the expansion coefficients C_n determine the intensity of the diffracted light in the n-th order, and T_g is the period of the diffraction grating.

The coefficients C_n are calculated using the known formula:

$$C_n = \frac{1}{2T_g} \int_{-\frac{T_g}{2}}^{\frac{T_g}{2}} T(\xi) e^{-in\frac{2\pi}{T_g}\xi} d\xi. \quad (3)$$

Despite the increase in light intensity in the second and fourth diffraction orders, this topology requires further optimization.

2. Optimization Model for a High-Order Diffraction Grating

The optimization problem requires establishing a criterion for optimizing the parameters of the diffraction grating. For this purpose, the criterion can be defined as the search for the maximum value in the desired diffraction orders:

$$P = \begin{vmatrix} \max C_n(n), \\ \min C_n(n-1). \\ \min C_n(n+1) \end{vmatrix} \quad (4)$$

The values of the coefficients C_n depend on the transmission function of the diffraction grating $T(\xi)$, which, in turn, can be expressed through the parameters of the grating lines «*a*», «*b*», and «*c*», with the period $Tg = a+b+a+b+a+c$, as illustrated in Fig. 1.

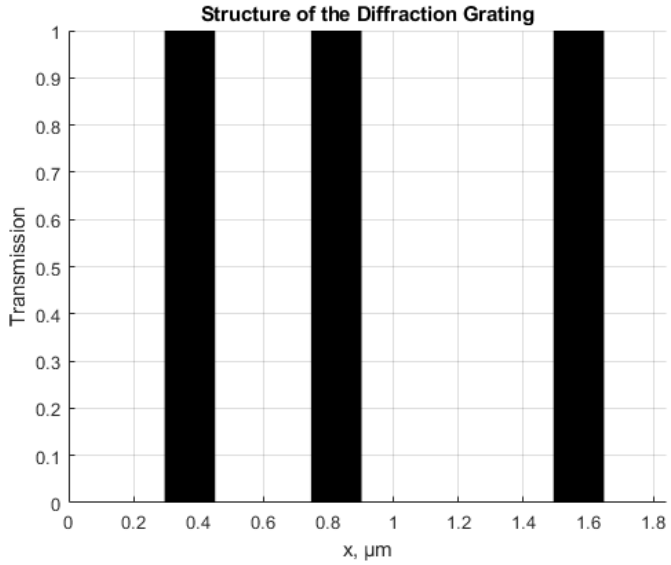


Fig. 1. Optimization Parameters of the Grating Transmission Function

The algorithm for finding the optimal solution for the corresponding diffraction order involves iterating through possible parameter values and finding the maximum value of the coefficient C_n for the desired diffraction order. The algorithm for finding the optimal transmission function for the corresponding diffraction order is expressed as a sequence of steps. The parameters of the diffraction grating transmission function are calculated in micrometers.

Main Steps of the Algorithm:

1. Initialization:
— Set the target expansion coefficients (*target_coeffs*).
— Set the initial grating parameters (*initial_params*).
— Define the model parameters (*sampling frequency fs*, *x-axis*, *range n*).
2. Optimization of Grating Parameters:
— Use the *fminsearch* function to minimize the objective function (objective_function).

```
clc;
clear all;

% Целевые коэффициенты разложения (пример)
target_coeffs = [0, 1, 0, 1, 0, 0];  % G0, G1, G2, G3, G4, G5

% Параметры решетки (начальные предположения)
initial_params = [0.28, 0.18, 0.67];  % a, b, c (в микрометрах)

% Параметры модели
fs = 10^3;          % Интервал дискретизации
x = 0:1/fs:1;       % Ось X (временная ось)
n = 0:5;            % Диапазон расчета коэффициентов разложения (включая нулевой порядок)

% Оптимизация
options = optimset('Display', 'iter', 'MaxIter', 1000);
optimized_params = fminsearch(@(params) objective_function(params, target_coeffs, x, n), initial_params, options);

% Вывод результатов
a = optimized_params(1);  % Ширина прозрачной полосы
b = optimized_params(2);  % Ширина первой непрозрачной полосы
c = optimized_params(3);  % Ширина второй непрозрачной полосы

fprintf('Оптимизированные параметры решетки:\n');
fprintf('a = %.4f мкм, b = %.4f мкм, c = %.4f мкм\n', a, b, c);

% Вычисление коэффициентов для оптимизированной решетки
optimized_coeffs = calculate_coeffs([a, b, c], x, n);
fprintf('Вычисленные коэффициенты разложения:\n');
disp(optimized_coeffs);

% Визуализация функции пропускания
figure(1); clf;
T_optimized = create_grating([a, b, c], x);
plot(x, T_optimized, 'LineWidth', 2);
title('Функция пропускания решетки');
xlabel('x, мкм');
ylabel('T(x)');
grid on;

% Визуализация структуры решетки
figure(2); clf;
hold on;

% Позиции непрозрачных элементов
opaque_positions = [a, 2*a + b, 4*a + 2*b];

% Закрашиваем непрозрачные области
for pos = opaque_positions
    fill([pos, pos + b, pos + b, pos], [0, 0, 1, 1], 'k', 'EdgeColor', 'none');
end

% Настройка графика
xlim([0, 3*a + 2*b + c]);
ylim([0, 1]);
xlabel('x, мкм');
ylabel('Пропускание');
title('Схема дифракционной решетки');
grid on;

% Визуализация коэффициентов разложения
figure(3); clf;
plot(n, target_coeffs, 'o-', 'LineWidth', 2, 'DisplayName', 'Целевые коэффициенты');
hold on;
plot(n, optimized_coeffs, 's--', 'LineWidth', 2, 'DisplayName', 'Вычисленные коэффициенты');
title('Сравнение коэффициентов разложения');
xlabel('Порядок дифракции');
ylabel('Интенсивность');
legend;
grid on;

% Проверка оптимизированных значений с помощью ряда Фурье
fprintf('Проверка оптимизированных значений с помощью ряда Фурье:\n');
for k = n
    C = 1/(3*a + 2*b + c) * T_optimized .* exp(-1j * (2*pi*k/(3*a + 2*b + c)) * x);
    G_k = abs(trapz(C, x));
    fprintf('G%d (ряд Фурье) = %.4f\n', k, G_k);
end

% Функция для создания решетки
function T = create_grating(params, x)
    a = params(1);  % Ширина прозрачной полосы
    b = params(2);  % Ширина первой непрозрачной полосы
    c = params(3);  % Ширина второй непрозрачной полосы

    % Позиции непрозрачных элементов
    opaque_positions = [a, 2*a + b, 4*a + 2*b];

    % Инициализация функции пропускания
    T = ones(size(x));  % По умолчанию все прозрачно

    % Задаем непрозрачные элементы
    for pos = opaque_positions
        T((x >= pos) & (x <= pos + b)) = 0;
    end
end

% Функция для вычисления коэффициентов разложения
function coeffs = calculate_coeffs(params, x, n)
    T = create_grating(params, x);
    a = params(1);
    b = params(2);
    c = params(3);
    Tg = 3*a + 2*b + c;  % Период решетки

    % Вычисление коэффициентов разложения
    coeffs = zeros(1, length(n));
    for k = n
        C = 1/Tg * T .* exp(-1j * (2*pi*k/Tg) * x);
        coeffs(k+1) = abs(trapz(C, x));  % k+1, так как индексация в MATLAB начинается с 1
    end
end

% Функция ошибки (среднеквадратичная ошибка)
function error = objective_function(params, target_coeffs, x, n)
    a = params(1);  % Ширина прозрачной полосы
    b = params(2);  % Ширина первой непрозрачной полосы
    c = params(3);  % Ширина второй непрозрачной полосы

    % Проверка на положительность параметров
    if a <= 0 || b <= 0 || c <= 0
        error = inf;  % Возвращаем большую ошибку, если параметры не положительны
        return;
    end

    % Вычисление коэффициентов разложения
    coeffs = calculate_coeffs([a, b, c], x, n);

    % Вычисление ошибки (MSE)
    error = sum((coeffs - target_coeffs).^2);
end
```

Fig. 2. Program Code for Optimizing the Diffraction Grating Transmission Function

— The objective function calculates the mean squared error between the target and current expansion coefficients.

3. Output of Results:

— Obtain the optimized grating parameters (*a*, *b*, *c*).

— Calculate the expansion coefficients for the optimized grating (*optimized_coeffs*).

4. Visualization:

— Plot the grating transmission function.

— Visualize the grating structure.

— Compare the target and calculated expansion coefficients.

5. Verification of Optimized Values:

— Use the Fourier series to verify the expansion coefficients.

A program implementing the above algorithm was written in MATLAB. Text comments are provided to simplify understanding of the program's operation. During the simulation, the *rectpuls* (*a,b*) function was used, which generates a rectangular pulse of duration *a* and coordinate shift *b*.

The program code is shown in Fig. 2. This program is designed to find the optimal topology for operation in the first and third diffraction orders. However, by changing the parameters, it can be used to find topologies for operation in other desired diffraction orders, such as the first and fourth or second and fourth.

As a result of the program's execution, the transmission function optimized for operation in the first and third diffraction orders was obtained, along with the values of the expansion coefficients of this transmission function, which determine the intensity of the diffracted light in the corresponding diffraction orders. The transmission function (a) and the calculation of the expansion coefficients (b) are shown in Fig. 3.

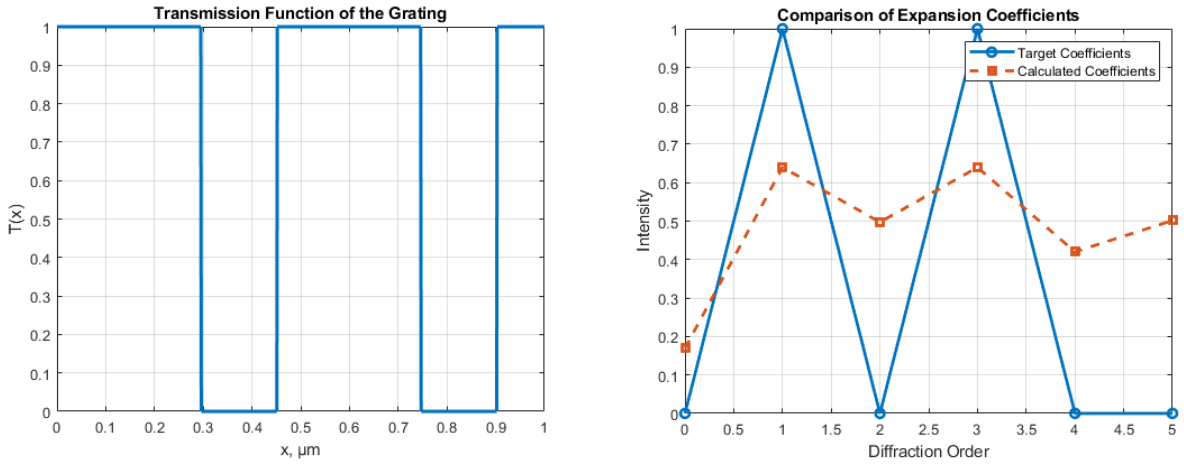


Fig. 3. Program Results: Optimized Transmission Function (a) and Expansion Coefficients (b) for Operation in the Second Order

The results of the calculations of the expansion coefficients of the optimized diffraction grating transmission function are compared with the coefficients for a standard diffraction grating, where the widths of the transparent and opaque lines are equal. The analysis showed that the developed optimization model significantly increased the intensity of the diffracted light in the second diffraction order. At the same time, the intensity of the light in the zero, first, and third orders was significantly reduced, indicating effective redistribution of energy in favor of the target order. This confirms that the proposed model allows for control over the energy distribution between orders, minimizing losses in non-target directions [6].

Conclusion

This work investigates the possibility of improving the spectral resolution of spectral instruments based on diffraction gratings by operating in higher diffraction orders. It is shown that the use of higher orders increases the angular dispersion and, consequently, improves the resolution of the instrument. However, to implement this approach, it is necessary to ensure high diffraction efficiency in the target orders, which requires optimization of the grating parameters.

Based on the proposed model, an algorithm for optimizing the topology of a diffraction grating was developed. The algorithm was implemented in MATLAB, allowing for numerical simulation and

optimization of the grating parameters. As a result, a configuration was found in which the diffraction efficiency in the second order was almost equal to that in the first order for a standard grating. At the same time, the intensity of the light in the neighboring orders (zero, first, and third) was significantly reduced, confirming the effectiveness of the proposed approach.

The obtained results demonstrate that optimizing the topology of a diffraction grating significantly improves its characteristics for operation in higher orders. This opens up new possibilities for creating spectral instruments with increased resolution and sensitivity. In addition, the developed algorithm and software implementation can be used to design gratings with specified properties, making them a universal tool for solving a wide range of problems in optics and spectroscopy.

In future research, it is planned to study the application of the proposed approach to gratings with more complex topologies, as well as to adapt the algorithm for operation in other spectral ranges.

References

1. Belyakov, Yu. M. Spectral Instruments: Textbook / Yu.M. Belyakov, N.K. Pavlycheva; Kazan: Kazan State Technical University Publishing House, 2007. 203 p. (in Russ.).

2. Tarasov, K. I. Spectral Instruments / K. I. Tarasov; 2nd ed. Leningrad: Mashinostroenie, 1977. 367 p. (in Russ.).

3. Malyshev, V. I. Introduction to Experimental Spectroscopy / V. I. Malyshev; Moscow: Nauka, 1979. 480 p. (in Russ.).

4. Palmer, C. A. Diffraction Grating Handbook / C. A. Palmer, E. G. Loewen. N. Y.: Newport Corporation, 2005. 271 p.

5. Moharam M. G., T. K. Gaylord. Rigorous Coupled-Wave Analysis of Planar-Grating Diffraction / Moharam M. G., T. K. Gaylord // J. Opt. Soc. Am. 71, 811-818 (1981).

6. Kazakov, V. I. Measurement of Optical Spectra by a Grating Spectral Instrument in Higher Diffraction Orders / V. I. Kazakov, O. D. Moskalets // Sensors and Systems. 2018. №12. P. 22-27. (in Russ.).

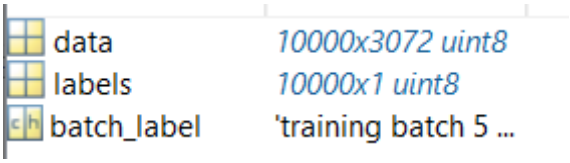
IMPLEMENTATION OF A CONVOLUTIONAL NEURAL NETWORK ALGORITHM IN MATLAB

Boris Kleshnin
Saint Petersburg State University of Aerospace Instrumentation
e-mail: boria456@gmail.com

Abstract. *Convolutional neural networks (CNNs) are widely used in computer vision tasks such as object recognition, image classification, and anomaly detection. This article discusses the process of implementing CNN in MATLAB. In the process, data preprocessing, network architecture construction, and model training will be used.*
Keywords: *convolutional neural network, CIFAR-10, ReLU, CNN, dataset*

In today's world, automated image analysis systems cannot be dispensed with, especially in areas such as medicine, security, and drone technology. Convolutional neural networks (CNNs) have become a key tool in computer vision, as they make it possible to efficiently identify image features and classify them with high accuracy. MATLAB provides powerful tools for building and training such networks, facilitating the process of model development even for users who do not have deep knowledge in the field of machine learning[1].

To train the model, a pre-created CIFAR-10 dataset is used, which contains 60,000 images divided into 10 classes[2]. The choice of this particular dataset was influenced by the convenience of working with it in the MATLAB environment. This is primarily due to the fact that the data is presented in the .mat format. This format contains pre-prepared images and labels for them. For convenience, the images are presented in the form of a vector, which can be reassembled.



Picture 1. Data stored in one of the parts of the training dataset

For further work, you need to use the arrayDatastore command, which creates a special MATLAB object designed to work with large amounts of data, such as images or numeric data. Next, the specification "'IterationDimension', 4" is prescribed, which indicates that iterations will occur along the 4th dimension (i.e., along the axis of the number of images). As a result, a 4-D uint8 matrix is constructed[3]. Using this algorithm allows you to efficiently load images in batches, instead of loading the entire array into memory at once (which may be inefficient with large data). A similar process must be carried out with labels. Next, the label data and image data are loaded into a single array, with which the neural network will interact[4].

Next, you need to configure the architecture of the neural network. To do this, it is necessary:

- Enter the input layer. In a pre-selected dataset, all images are structured and adjusted to a single standard [32x32x3]. A neural network needs to work with this standard. If the image does not match the standard, an error will occur.
- Create a 2D convolutional layer using the convolution2dLayer(3, 32, 'Padding', 'same') command. This is where the frames that will "slide" across the image are selected, comparing sets of pixels. The size of the 3x3.32 sliding window means the number of filters (cores). Next, the Padding is filled in and adjusted, which sets a limit, meaning that the input size will be equal to the output 'same'. These settings are set by default and it is undesirable to change them.reluLayer() задает слой активации ReLU. Эта технология применяется для выхода каждого фрейма. Она преобразует входные значения по следующему правилу:

$$f(x) = \max(0, x)$$

Thus, if the value of x>0, it remains unchanged, otherwise it is replaced by 0. Such a transformation is necessary because it allows the neural network to model complex dependencies in the data, which speeds up learning. The ReLU function does not cause problems with a fading gradient and provides easy calculations – it is fast, as it simply replaces negative values with zeros[5].

- Next, you need to create another convolutional layer. On the second layer, the model combines simple features from the first layer to detect more complex features, such as the shapes of objects or their parts.
- Then you need to create a Fully Connected Layer (or Dense Layer) – this is a layer where each neuron is connected to all the neurons of the previous layer. In this case, the layer has 10 neurons, which corresponds to the number of classes in the CIFAR-10 dataset.
- The softmax layer is an activation function that converts the output values of a fully connected layer into probabilities. Softmax normalizes the output values so that their sum is equal to 1. This allows us to interpret the outputs of the model as probabilities of belonging to each class.
- classificationLayer() is a layer that calculates the loss function and returns the predicted class. It compares the predicted probabilities (from softmax) with the true labels and calculates the error (for example, cross-entropy).

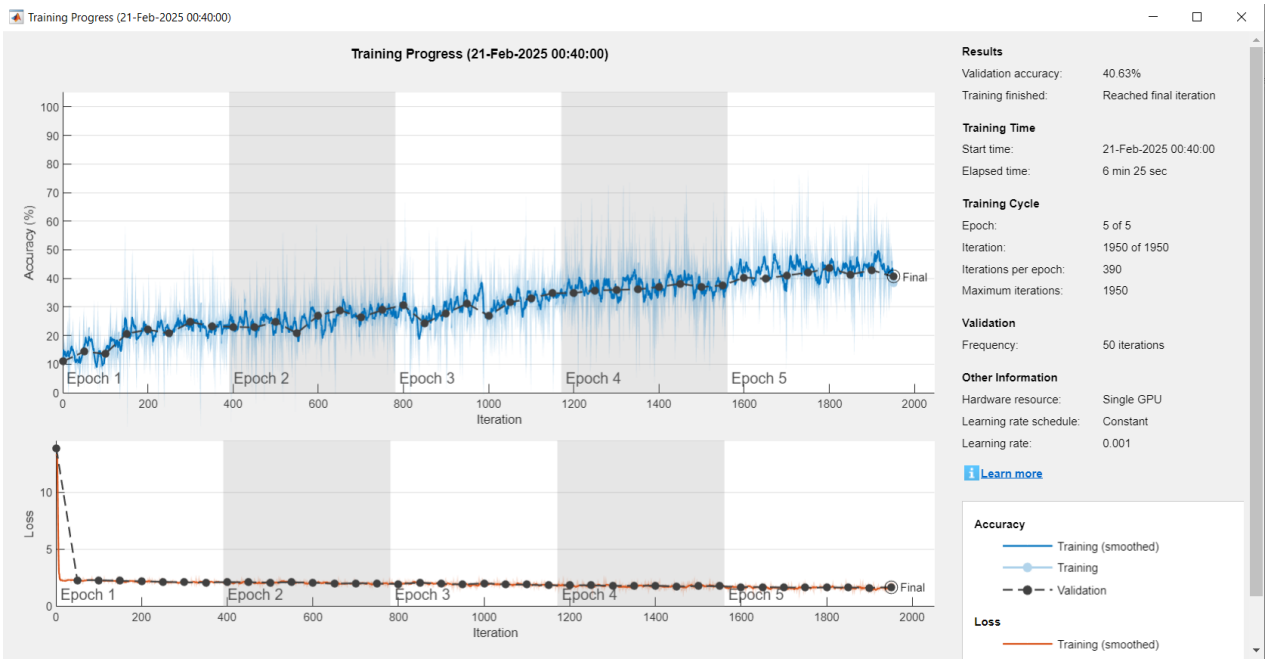
After configuring the network architecture, you need to enter training parameters such as the number of epochs, learning rate, etc.

Code Listing:

```
% Загрузка обучающих данных
XTrain = [];
YTrain = [];
for i = 1:5
    filePath = fullfile(dataFolder, ['data_batch_' num2str(i) '.mat']);
    data = load(filePath);
    XTrain = [XTrain; data.data]; % Объединение данных
    YTrain = [YTrain; data.labels]; % Объединение меток
end
% Загрузка тестовых данных
testFilePath = fullfile(dataFolder, 'test_batch.mat');
testData = load(testFilePath);
XTest = testData.data;
YTest = testData.labels;
% Преобразование данных в правильный формат (32x32x3)
XTrain = reshape(XTrain', 32, 32, 3, []); % Обучающие данные [32x32x3x50000]
XTest = reshape(XTest', 32, 32, 3, []); % Тестовые данные [32x32x3x10000]
% Преобразование меток в categorical
YTrain = categorical(YTrain); % Обучающие метки
YTest = categorical(YTest); % Тестовые метки
% Создание arrayDatastore для изображений
adsTrain = arrayDatastore(XTrain, 'IterationDimension', 4); % Обучающие данные
adsTest = arrayDatastore(XTest, 'IterationDimension', 4); % Тестовые данные
% Создание arrayDatastore для меток
labelsTrain = arrayDatastore(YTrain);
labelsTest = arrayDatastore(YTest);
% Объединение данных и меток
dsTrain = combine(adsTrain, labelsTrain);
dsTest = combine(adsTest, labelsTest);
% Создание архитектуры сверточной нейронной сети
layers = [
    imageInputLayer([32 32 3]) % Входной слой (32x32x3)
    convolution2dLayer(3, 32, 'Padding', 'same') % Сверточный слой
    reluLayer() % Слой активации ReLU
    maxPooling2dLayer(2, 'Stride', 2) % Слой подвыборки
    convolution2dLayer(3, 64, 'Padding', 'same') % Еще один сверточный слой
    reluLayer() % Слой активации ReLU
    fullyConnectedLayer(10) % Полносвязный слой (10 классов)
    softmaxLayer() % Слой softmax
    classificationLayer() % Выходной слой
];
% Настройка параметров обучения
```

```
options = trainingOptions('sgdm', ...
    'MaxEpochs', 5, ... % Количество эпох
    'InitialLearnRate', 0.001, ... % Скорость обучения
    'Shuffle', 'every-epoch', ... % Перемешивание данных каждую эпоху
    'ValidationData', dsTest, ... % Данные для валидации
    'Verbose', false, ... % Отключение вывода в командное окно
    'Plots', 'training-progress'); % Визуализация процесса обучения
% Обучение сети
net = trainNetwork(dsTrain, layers, options);
```

As a result of the program, the training process window opens, in which you can see the graphs (Fig. 2). The first graph shows the accuracy, the second shows the loss depending on the iterations. You can also understand that the neural network analyzes several epochs. Each epoch means that the model has analyzed the entire dataset 1 time. The more epochs, the greater the accuracy and less loss[6].



Picture 2. The interface of the convolutional neural network

In the right menu, you can see the data on the model's training results. Based on these results, it can be concluded that the training accuracy of the model is only 40.63%. This is a pretty bad indicator. The reason for this was the small number of epochs. For optimal model training, it is worth using 8-10 epochs. However, this is a very time-consuming process in terms of computing power.

References

1. Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. – MIT Press, 2016. – 800 p.
2. Haykin, S. Neural Networks and Learning Machines. – 3rd ed. – Pearson, 2008. – 906 p.
3. Russell, S., Norvig, P. Artificial Intelligence: A Modern Approach. – 4th ed. – Pearson, 2020. – 1136 p.
4. Gorban, A. N., Dunin-Barkovsky, V. L., Kirdin, A. N. Neural Networks: Fundamentals of Theory. – Moscow: Fizmatlit, 2010. – 496 p.
5. Osowski, S. Neural Networks for Information Processing. – Moscow: Finance and Statistics, 2004. – 344 p.
6. Rutkowska, D., Pilinski, M., Rutkowski, L. Neural Networks, Genetic Algorithms, and Fuzzy Systems. – Moscow: Goryachaya Liniya – Telekom, 2006. – 452 p.

DEVELOPMENT OF A RELIABILITY MANAGEMENT SYSTEM FOR THE MACHINE-BUILDING INDUSTRY BASED ON MODERNIZATION OF THE TECHNICAL READINESS COEFFICIENT CALCULATION

Timofey Komarov
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
tim1kom@yandex.ru

Abstract. *The article is dedicated to the development of a reliability management system (RMS) in the machine engineering industry, based on the modernization of the calculation of the technical readiness coefficient. The reliability of equipment is a critical factor for improving production efficiency, safety, and economic sustainability of enterprises. In the context of growing competition and demands for product quality, the implementation of RMS becomes essential for minimizing downtime, reducing repair costs, and increasing productivity. The article discusses key aspects of reliability management, such as failure prevention, lowering maintenance costs, and enhancing operational safety. Methods for assessing reliability are presented, including probabilistic models and the calculation of the readiness coefficient, as well as an analysis of factors affecting equipment reliability levels. Special attention is given to simplifying calculations and reducing the impact of human factors on the accuracy of computations.*
Keywords: *Equipment reliability, reliability management system, technical readiness coefficient, production efficiency, maintenance costs, mean time between failures (MTBF), probability of failure-free operation, reliability factors, machine engineering industry.*

Reliability of equipment is a key factor determining the efficiency of production, safety, and economic stability of enterprises. In the context of growing competition and increasing demands for product quality, the development of a reliability management system (RMS) becomes a necessary tool for minimizing downtime, reducing repair costs, and improving productivity.

The reliability management system aims to:

- Prevent failures.
- Reduce maintenance costs.
- Improve operational safety.
- Extend the service life of equipment. [1]

According to research, the implementation of an RMS can reduce repair costs by 20-30% and increase the service life of equipment by 15-25%. [1] The following factors are considered when managing RMS:

1. Technical factor: design and production features of the equipment;
2. Human factor: quality of maintenance and repair performed by personnel;
3. Environmental factor: the surrounding environment affecting equipment operation;
4. Material factor: quality of replaceable spare parts and consumables used in maintenance and repair;
5. Methodological factor: production processes regulating maintenance and repair, directly or indirectly affecting the level of equipment reliability and the costs of maintaining it. [2]

To assess reliability, it is common to analyze probability—the likelihood that equipment will operate without failures for an extended period. The probability of failure-free operation P(t)$P(t)$ can be calculated using formula (1):

$$P(t) = e^{-\lambda t}, \quad (1)$$

Another significant indicator in reliability analysis is the Mean Time Between Failures (MTBF) – the average time between failures. This indicator can be calculated using formula (2):

$$P(t) = e^{-\lambda t} = e^{-0,001*1000} \approx 0,368$$

Another significant indicator in reliability analysis is the Mean Time Between Failures (MTBF)—the average time between failures. This indicator can be calculated using formula (2): [3]

$$MTBF = \frac{1}{\lambda} \quad (2)$$

For example, with a failure rate of λ=0.001λ=0.001 failures/hour, the Mean Time Between Failures (MTBF) will be:

$$MTBF = \frac{1}{0,001} \approx 1000 \text{ hours.}$$

One of the most important indicators for assessing reliability is the availability coefficient *Ky*. It shows the proportion of time the equipment is in working condition and can be calculated using formula (3):

$$K_y = \frac{MTBF}{MTBF + MTTR}, \quad (3)$$

Where MTTR is the mean time to repair (average time to restore after a failure). For example, if MTTR=10$MTTR$=10 hours and MTBF=1000$MTBF$=1000 hours, the availability coefficient will be calculated as follows:

$$K_y = \frac{1000}{1000 + 10} \approx 0,99 .$$

However, all elements of the currently used formula are calculated values, which complicates the process of operational reliability analysis and increases the likelihood of calculation errors due to the human factor.

In the production process, the most obvious and measurable parameters are equipment downtime and operating time. Based on these, it is possible to calculate the technical readiness coefficient KtKt using formula (4):

$$K_t = \left(1 - \frac{T_{np}}{T_p}\right) \times 100\%, \quad (4)$$

where Tpr is the total downtime spent on maintenance and repairs over a certain period; Tr is the calendar operating time over the same period.

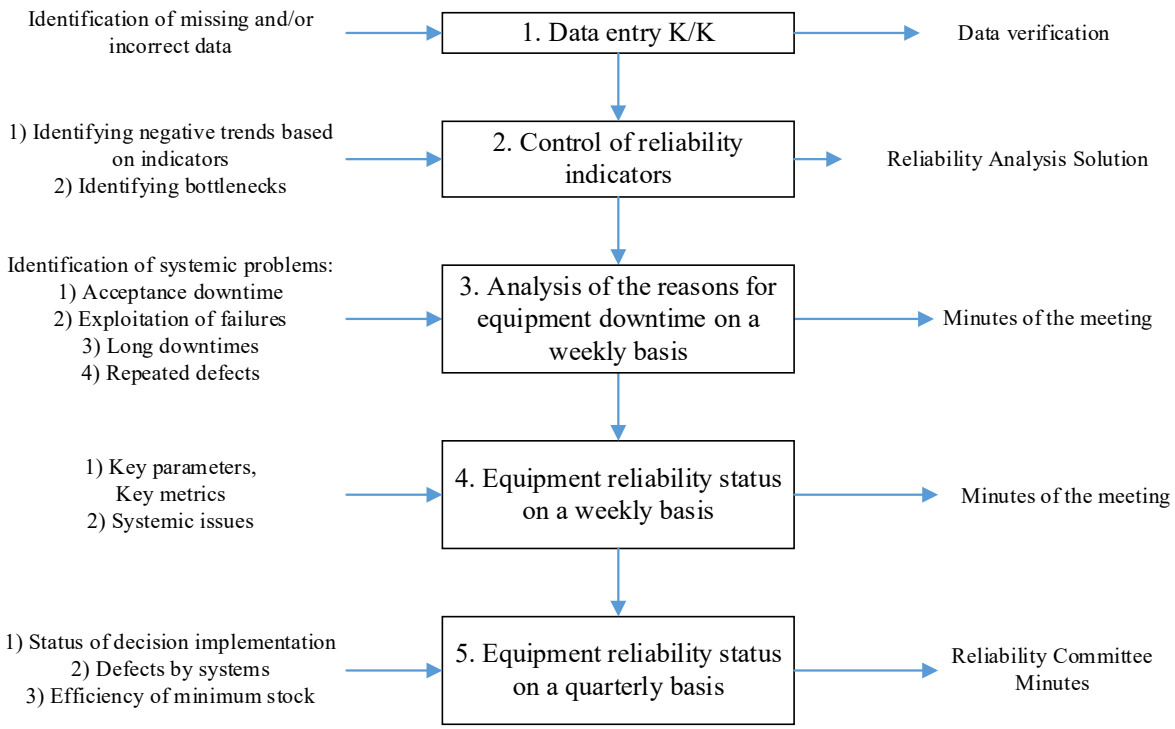


Fig. 1. Equipment Reliability Monitoring System

For example, with a total downtime of Tpr=0.56 hours and operating time of Tr=3.8 hours, the technical readiness coefficient will be:

$$K_t = \left(1 - \frac{0,56}{3,8}\right) \times 100\% = 85\%.$$

The equipment reliability monitoring system is a multi-level system operating on the principle: monitoring – problem identification – action – effectiveness – monitoring.

The general scheme of the equipment reliability monitoring system is shown in Fig. 1.

A detailed algorithm for conducting reliability monitoring is shown in Fig. 2.

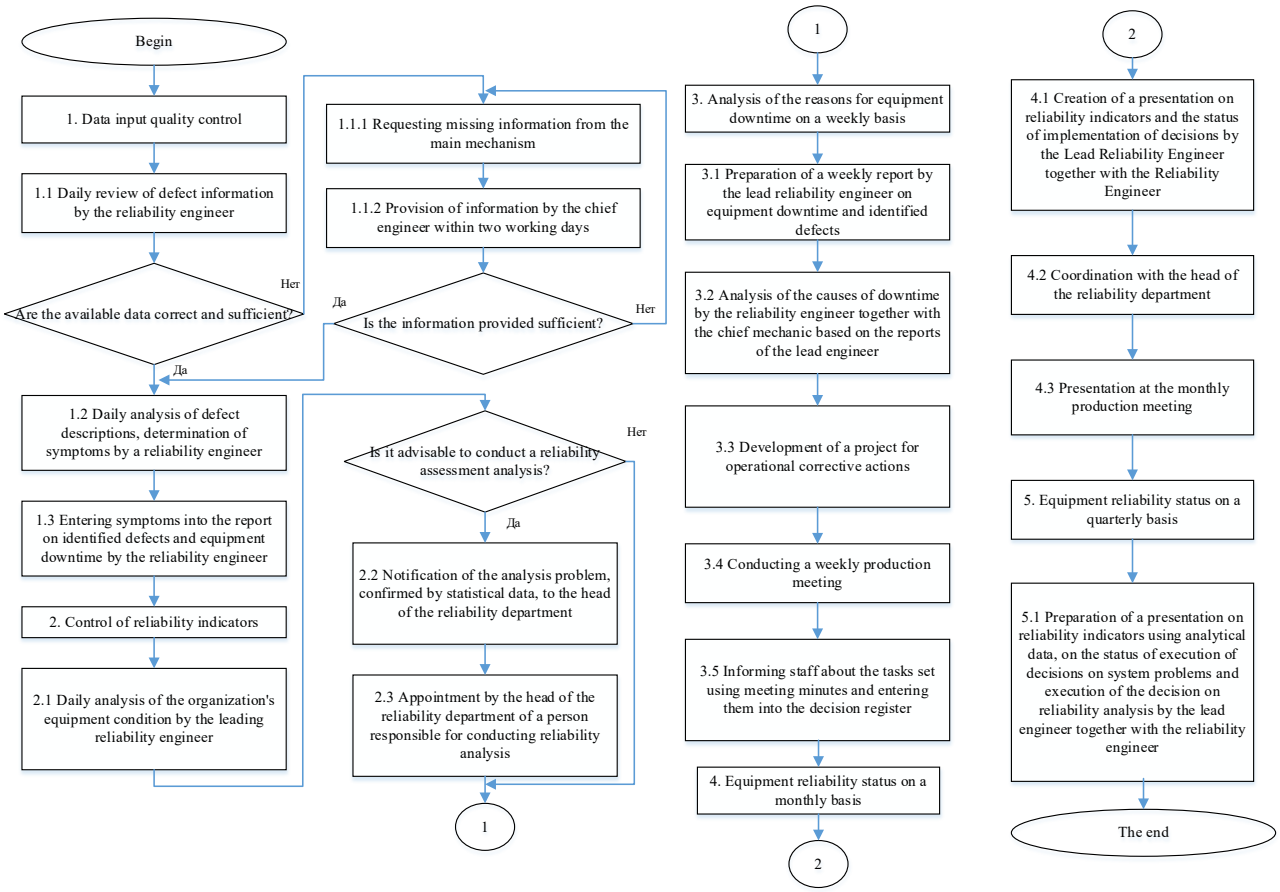


Fig. 2. Algorithm for Conducting Reliability Monitoring

The development of a reliability management system (RMS) is a strategically important process that not only improves equipment efficiency but also significantly impacts the economic performance of an enterprise. The implementation of an RMS minimizes downtime, reduces repair and maintenance costs, and enhances safety and product quality. In the context of increasing demands for product quality and safety, the development and implementation of an RMS become a necessary condition for the successful operation of any industrial enterprise.

References

1. Klubnikina K.V. Mathematical Methods for Assessing the Reliability of Heat Supply. – Tyumen: Tyumen Industrial University, 2021. – Pp. 205–212.

2. GOST R 27.003-2010 «Reliability in Engineering. Terms and Definitions»

3. Gryaznov M.V., Mikhailov D.V. The Role of Ensuring the Reliability of Road Transport in the Implementation of the Russian Transport Strategy until 2030. – Stavropol: Science in Modern Society, 2014. Pp. 133–136.

DEVELOPMENT OF A METHODOLOGY FOR ASSESSING THE NEED FOR IMPLEMENTING CORRECTIVE OR PREVENTIVE ACTIONS

Victoria Komarova
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
vikap1999@mail.ru

Abstract. *The article discusses the necessity of developing a methodology for assessing the feasibility of implementing corrective and preventive actions in modern business systems. In the context of increasing competition and demands for product quality, identifying and eliminating nonconformities have become critically important for maintaining a company's reputation. The author emphasizes that the implementation of such measures requires significant resources, making it essential to conduct a thorough assessment of their necessity. The article describes methods for analyzing the causes of nonconformities, such as the "5 Whys" method and the logical fault tree analysis, and provides examples of their application. Factors contributing to nonconformities are examined, and measures for their prevention and elimination are proposed.*
Keywords: *Corrective actions, preventive actions, feasibility assessment, nonconformities, quality management system, "5 Whys" method, logical fault tree analysis, cause analysis, business processes.*

In the modern business environment, where competition and demands for product and service quality are constantly increasing, ensuring compliance with established standards and regulations has become critically important. However, even in the most efficient management systems, discrepancies can arise, negatively affecting processes, product quality, customer satisfaction, and ultimately, the company's reputation. To address such discrepancies, corrective actions are developed, aimed at eliminating the root causes of problems and preventing their recurrence.

However, implementing corrective actions requires significant resources: time, financial, and human. Therefore, before deciding to implement such measures, a thorough assessment of their feasibility is necessary. This not only optimizes resource utilization but also focuses on the most significant issues that have the greatest impact on business processes.

This article discusses an approach to assessing the need for corrective actions, including the development of evaluation criteria, analysis of the impact of discrepancies, and making informed decisions. [1].

Upon receiving a signal about a detected discrepancy (potential or actual) and after reviewing the initial information, a quality management department specialist (or another authorized department) conducts a root cause analysis to determine the underlying cause of the discrepancy.

Typically, the "5 Whys" method is used, which involves sequentially asking the question "Why" about the problem and the answers received (usually at least five times). The question is asked until the root cause is identified or until it is impossible to answer the question. [2]

Example:

Failure: "Sudden engine shutdown"
Why? Timing belt failure;
Why? The timing belt wore out;
Why? No information on the need for replacement;
Why? No scheduled replacement system in place;
Why? Lack of monitoring of the current condition.

Conclusion: The root cause of the sudden engine shutdown is the lack of monitoring of the engine components' current condition

For more complex cases, the logical tree of failure causes method can be used to determine the root cause. [3] This method involves constructing a cause-and-effect diagram that displays failure events, types of failures, hypotheses about failure causes, primary physical causes of failure, human factor influence, and systemic causes of failure. The scheme of the logical tree of failure causes method is shown in Fig. 1.

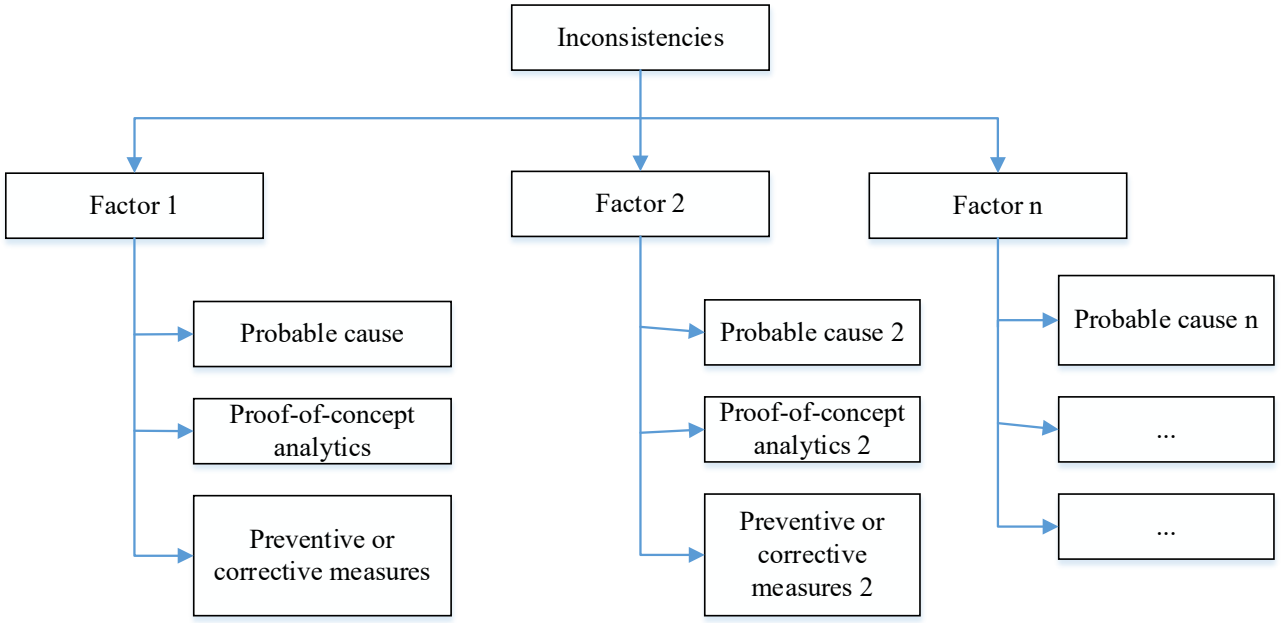


Fig. 1. General Scheme of the Logical Tree of Failure Causes Method

The demonstration of the method is shown in Fig. 2.

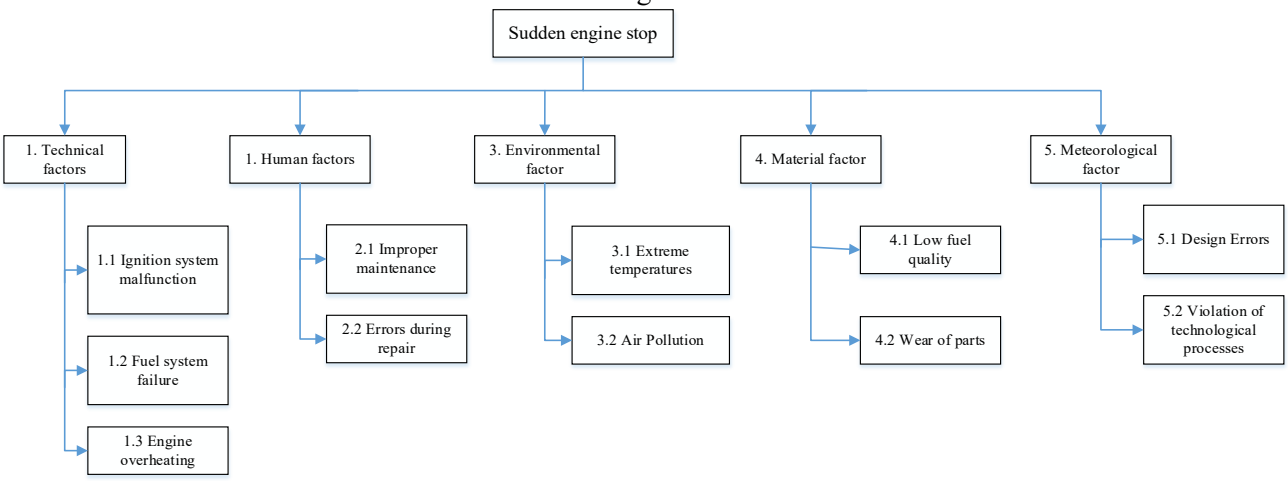


Fig. 2. Logical Tree Method for the Failure Cause «Sudden Engine Shutdown»

The identification of factors leading to discrepancies (i.e., their analysis, cause, preventive and corrective actions) is presented in Table 1.

Table 1

Identification of Factors Leading to Discrepancies

Factor Name	Cause	Analysis	Preventive Actions	Corrective Actions
1. Technical Factors				
1.1 Ignition System Failure	Failure of spark plugs or ignition coil	20% of engine failures are related to ignition system issues	1) Regular diagnostics of the ignition system (checking spark plugs, coils, wires) 2) Periodic replacement of the fuel filter and checking the fuel pump	1) Replacement of faulty ignition system components 2) Cleaning or replacement of the fuel filter, repair of the fuel pump
1.2 Fuel System Failure	Clogged fuel filter or faulty fuel pump	15% of engine failures are due to fuel supply issues		
1.3 Engine	Cooling system	10% of failures are		

Factor Name	Cause	Analysis	Preventive Actions	Corrective Actions
Overheating	failure (coolant leak, thermostat failure)	related to overheating	3) Monitoring engine temperature and timely maintenance of the cooling system	3) Elimination of coolant leaks and replacement of the thermostat
2. Human Factor				
2.1 Improper Maintenance	Untimely oil or filter replacement	25% of engine failures are due to maintenance schedule violations	1) Training personnel on maintenance and repair rules 2) Implementing checklists for maintenance control 3) Using only certified parts and materials	1) Conducting quality audits of maintenance and repairs 2) Correcting errors made during previous repairs 3) Introducing a system of accountability for maintenance violations
2.2 Repair Errors	Incorrect installation of parts or use of low-quality parts	5% of failures are due to repair errors		
3. Environmental Factor				
3.1 Extreme Temperatures	Engine operation in severe cold or heat	8% of failures are related to operation in unfavorable climatic conditions	1) Installing additional filters to protect against dust and sand 2) Using antifreeze and oils adapted to extreme temperatures 3) Regular cleaning of the air filter and intake system	1) Replacing components damaged by contamination 2) Adapting the cooling system for operation in extreme conditions 3) Eliminating corrosion or wear caused by the environment
3.2 Air Contamination	Dust or sand entering the air filter	3% of failures are caused by intake system contamination		
4. Material Factor				
4.1 Low-Quality Fuel	Use of fuel with impurities or low octane rating	12% of failures are related to the use of low-quality fuel	1) Using only high-quality fuel with the recommended octane rating 2) Regular oil and filter changes to prevent component wear 3) Quality control of materials and components at the procurement stage	1) Cleaning the fuel system of impurities and replacing low-quality fuel 2) Replacing worn components (pistons, rings, bearings) 3) Eliminating the consequences of using low-quality materials
4.2 Component Wear	Natural wear of engine components (pistons, rings, bearings)	18% of failures are due to wear		
5. Methodology Factor				
5.1 Design Errors	Insufficient component strength or incorrect cooling system design	7% of failures are related to design errors	1) Conducting a thorough analysis of engine design at the development stage 2) Implementing quality standards at all production stages 3) Regular audits of technological processes	1) Revising the engine design to eliminate identified shortcomings. 2) Correcting errors made during production (e.g., replacing defective parts) 3) Implementing new technologies to improve product quality
5.2 Violation of Technological Processes	Non-compliance with standards during engine production	5% of failures are caused by production defects		

Once the failures and their causes are identified, and corrective and preventive actions are developed, it is important to determine the feasibility of their implementation. To assess the need for implementing the developed measures, a methodology was created that allows for a decision to be made through a survey on whether to adopt or reject the measures.

The essence of the methodology is to assess, by answering questions, the need for implementing corrective or preventive actions. Each question is assigned a score from 1 to 5. After the assessment, all scores are summed up. The questionnaire is presented in Table 2

Table 2

Questionnaire for Assessing the Need for Implementing Developed Measures

Question Group	Question	1	2	3	4	5
1. Impact Scale						
1.1 How much does this cause affect other units or types of equipment?						
1.2 How recurring is the problem being addressed?						
2. Feasibility of Elimination						
2.1 Can the cause be eliminated through operational measures?						
2.2 Is production/operation stoppage required to eliminate the cause?						
3. Resources and Cost						
3.1 How much personnel is needed to implement the measures?						
3.2 How much time is needed to implement the measures?						
3.3 How much material is needed to implement the measures?						
3.4 What budget is needed to implement the measures?						
3.5 Assess the economic efficiency of the measures (implementation cost / problem cost over time)*						
3.6 Are changes to systemic processes required?						
4. Effectiveness						
4.1 How effectively will this measure eliminate or minimize the impact of the cause?						
4.2 Are there alternative, more effective measures?						
TOTAL:						

Note: Measures addressing safety issues related to threats to life and health are not evaluated in terms of economic efficiency.

Typically, corrective actions may include:

- Revising the interval for technical maintenance and repair of equipment;
- Revising the method of performing technical maintenance and repair of equipment;
- Revising work process charts;
- Developing new work process charts;
- Changing the list of non-reducible and circulating stock;
- Performing modifications on equipment;
- Conducting audits;
- Developing instructions, procedures;
- Training personnel;
- Changing the list and quantity of emergency and current stock. [4]

To properly assess the need for implementing the developed measures, scoring scales were created, presented in Tables 3-6.

Table 3

Scoring Scale for the Question Group «Impact Scale»

Question	Score	Criterion	Value
1.1 How much does this cause affect other units or types of equipment?	5	No impact	The cause has no effect on other units or types of equipment. The system's operation remains unchanged, and there are no consequences.
	4	Minor impact	The cause has minimal impact on units or types of equipment. The impact is weakly expressed and does not lead to significant changes in the system's operation.
	3	Moderate impact	The cause has limited impact on individual units or types of equipment. The impact is noticeable but not critical and does not lead to serious consequences for the system as a whole.
	2	Significant impact	The cause affects several units or types of equipment, causing noticeable disruptions or deterioration in their operation. May lead to partial process stoppage or reduced efficiency.
	1	Critical impact	The cause has a strong negative impact on most units or types of equipment. Leads to mass failures, disruptions, or serious consequences for the system's operation as a whole. Requires immediate elimination.

Question	Score	Criterion	Value
1.2 How recurring is the problem being addressed?	5	Single occurrence	The problem occurred once and has not recurred. The probability of recurrence is extremely low. No additional actions are required unless it occurs again.
	4	Rarely recurring	The problem occurs in isolated cases and is not systematic. May be caused by random factors or external conditions. No urgent measures are required.
	3	Periodically recurring	The problem occurs from time to time but has no clear pattern. May be related to certain conditions or factors. Requires attention but is not critical.
	2	Frequently recurring	The problem occurs often but not always. Observed in a significant portion of cases, indicating a stable trend. Detailed study and action are necessary.
	1	Constantly recurring	The problem occurs regularly and systematically. Observed in most cases or every time the system/equipment is used. Requires urgent resolution to prevent further recurrences.

Table 4

Scoring Scale for the Question Group «Feasibility of Elimination»

Question	Score	Criterion	Value
2.1 Can the cause be eliminated through operational measures?	5	Cannot be eliminated through operational measures	The cause cannot be eliminated operationally. Requires deep analysis, long-term planning, and significant resources. Operational measures are ineffective or impossible.
	4	Difficult to eliminate through operational measures	The cause requires significant effort, resources, or time to eliminate. Operational measures may be insufficient, and long-term solutions need to be considered.
	3	Can be eliminated through operational measures but with moderate effort	The cause can be eliminated operationally but requires more time, resources, or coordination among participants. A solution is possible but requires planning.
	2	Can be eliminated through operational measures with little effort	The cause can be eliminated operationally but requires some additional actions or resources. A solution is possible in the short term.
	1	Easily eliminated through operational measures	The cause can be eliminated quickly and without significant resource expenditure. Requires minimal effort and does not need long-term planning or complex solutions.
2.2 Is production/operation stoppage required to eliminate the cause?	5	Full production stoppage	To eliminate the cause, full production or equipment operation stoppage is required. Work cannot continue until the problem is resolved.
	4	Partial production/operation stoppage	To eliminate the cause, stoppage of individual sections, lines, or units is required. Production or operation can continue in a limited mode.
	3	Temporary process suspension	To eliminate the cause, a short-term stoppage or suspension of processes is required. Production or operation can be quickly resumed after the problem is resolved.
	2	Stoppage not required but delays possible	To eliminate the cause, production or operation stoppage is not required, but delays or reduced efficiency are possible. The problem can be resolved during operation.
	1	Stoppage not required	To eliminate the cause, production or operation stoppage is not required. The problem can be resolved without any work interruptions.

Table 5

Scoring Scale for the Question Group "Resources and Cost"

Question	Score	Criterion	Value
3.1 How much personnel is needed to implement the measures?	5	Very high resource level	Implementing the measures requires significant resources: financial, human, time, and material. Additional capacity or external support is needed.
	4	High resource level	Implementing the measures requires a large amount of resources but within the organization's capabilities. Additional funds, personnel, or equipment need to be allocated.
	3	Moderate resource level	Implementing the measures requires a moderate amount of resources. Existing resources need to be reallocated or partially supplemented.
	2	Low resource level	Implementing the measures requires minimal resources. The solution can be implemented by the current personnel within existing budgets.
	1	Very low resource level	Implementing the measures requires almost no additional resources. The solution can be executed without significant costs or changes to current processes.
3.2 How much time is needed to implement the measures?	5	Very high time requirement	Implementing the measures requires a significant amount of time. Long-term planning and coordination are necessary.
	4	High time requirement	Implementing the measures requires a considerable amount of time but within the organization's capacity. Additional time allocation is needed.
	3	Moderate time requirement	Implementing the measures requires a moderate amount of time. Some time reallocation or partial supplementation is necessary.
	2	Low time requirement	Implementing the measures requires minimal time. The solution can be implemented within the current schedule.
	1	Very low time requirement	Implementing the measures requires almost no additional time. The solution can be executed quickly without disrupting current processes.
3.3 How much material is needed to implement the measures?	5	Very high material requirement	Implementing the measures requires a significant amount of material. Additional procurement or external support is needed.
	4	High material requirement	Implementing the measures requires a large amount of material but within the organization's capacity. Additional material allocation is needed.
	3	Moderate material requirement	Implementing the measures requires a moderate amount of material. Some material reallocation or partial supplementation is necessary.
	2	Low material requirement	Implementing the measures requires minimal material. The solution can be implemented with existing resources.
	1	Very low material requirement	Implementing the measures requires almost no additional material. The solution can be executed with minimal resource expenditure.
3.4 What budget is needed to implement the measures?	5	Very high budget requirement	Implementing the measures requires a significant budget. Additional funding or external support is needed.
	4	High budget requirement	Implementing the measures requires a large budget but within the organization's capacity. Additional funding allocation is needed.
	3	Moderate budget requirement	Implementing the measures requires a moderate budget. Some budget reallocation or partial supplementation is necessary.
	2	Low budget requirement	Implementing the measures requires minimal

Question	Score	Criterion	Value
			budget. The solution can be implemented within the current budget.
	1	Very low budget requirement	Implementing the measures requires almost no additional budget. The solution can be executed with minimal financial expenditure.
3.5 Assess the economic efficiency of the measures (implementation cost / problem cost over time)*	5	Very low economic efficiency	The cost of implementing the measures significantly exceeds the potential costs of addressing the problem's consequences. The solution is economically unjustified and requires reconsideration.
	4	Low economic efficiency	The cost of implementing the measures is higher than the potential costs of addressing the problem's consequences. The solution may be justified only by non-economic factors (e.g., safety or reputation).
	3	Moderate economic efficiency	The cost of implementing the measures is comparable to the potential costs of addressing the problem's consequences. The solution is economically neutral but may prevent additional costs in the future.
	2	High economic efficiency	The cost of implementing the measures is lower than the potential costs of addressing the problem's consequences. The solution is economically justified and beneficial.
	1	Very high economic efficiency	The cost of implementing the measures is significantly lower than the potential costs of addressing the problem's consequences in the long term. The solution brings substantial economic benefits.
3.6 Are changes to systemic processes required?	5	Complete systemic process overhaul	To eliminate the cause, a radical change in systemic processes is required. Key work stages need to be revised and restructured, requiring significant time and resource expenditure.
	4	Significant systemic process changes	To eliminate the cause, substantial changes to systemic processes are required. Individual work stages need to be revised, and new approaches or procedures need to be implemented.
	3	Moderate systemic process changes	To eliminate the cause, partial changes to systemic processes are required. Individual elements or stages need to be adjusted without a complete overhaul.
	2	Minor systemic process changes	To eliminate the cause, minimal changes to systemic processes are required. Small adjustments or clarifications to existing procedures are needed.
	1	No systemic process changes required	To eliminate the cause, no changes to systemic processes are required. The problem can be resolved within existing procedures and approaches.

Table 6

Scoring Scale for the Question Group "Effectiveness"

Question	Score	Criterion	Value
4.1 How effectively will this measure eliminate or minimize the impact of the cause?	5	No effect	The measure does not eliminate or minimize the impact of the cause. The problem remains unchanged, and the solution is ineffective.
	4	Minor elimination of the cause	The measure has a minimal impact on the cause. The solution is minimally effective and does not resolve the problem in the long term.
	3	Moderate elimination of	The measure partially reduces the impact of the

Question	Score	Criterion	Value
		the cause	cause, but the problem remains relevant. The solution requires refinement or supplementation with other measures to increase effectiveness.
	2	Significant elimination of the cause	The measure significantly reduces the impact of the cause but does not eliminate it completely. The solution is effective but may require additional actions in the future.
	1	Complete elimination of the cause	The measure completely eliminates the cause, removing its impact on the system or process. The solution is maximally effective and long-term.
4.2 Are there alternative, more effective measures?	5	Significantly more effective alternatives exist	There are alternative measures that are much more effective than the proposed solution. They allow for complete elimination of the cause with fewer resources and time.
	4	More effective alternatives exist	There are alternative measures that are more effective than the proposed solution. They allow for better elimination of the cause or minimization of its impact with less effort.
	3	Comparable effectiveness alternatives exist	There are alternative measures that are similarly effective to the proposed solution. They can be considered as additional or replacement options.
	2	Less effective alternatives exist	There are alternative measures, but they are less effective than the proposed solution. Their implementation is not advisable as they do not fully resolve the problem.
	1	No alternative measures exist	There are no alternative measures, or they cannot be implemented. The proposed solution is the only possible option.

For example, for the failure "Sudden engine shutdown," a detailed analysis identified the most likely cause as untimely oil or filter replacement (factor – improper maintenance). To address the causes of the discrepancy, the following corrective actions were proposed:

1. Replacement of faulty ignition system components;
2. Cleaning or replacement of the fuel filter, repair of the fuel pump;
3. Elimination of coolant leaks and replacement of the thermostat.

The assessment of the need for implementing the developed measures is presented in Table 7.

Table 7

Assessment of the Need for Implementing Developed Measures for the Failure "Sudden Engine Shutdown"

Question Group	Question	1	2	3	4	5
1. Impact Scale						
1.1 How much does this cause affect other units or types of equipment?			+			
1.2 How recurring is the problem being addressed?			+			
2. Feasibility of Elimination						
2.1 Can the cause be eliminated through operational measures?	+					
2.2 Is production/operation stoppage required to eliminate the cause?			+			
3. Resources and Cost						
3.1 How much personnel is needed to implement the measures?		+				
3.2 How much time is needed to implement the measures?		+				
3.3 How much material is needed to implement the measures?		+				
3.4 What budget is needed to implement the measures?		+				
3.5 Assess the economic efficiency of the measures (implementation cost / problem cost over time)*	+					
3.6 Are changes to systemic processes required?		+				
4. Effectiveness						
4.1 How effectively will this measure eliminate or minimize the impact of the cause?			+			
4.2 Are there alternative, more effective measures?			+			
TOTAL:	27					

To interpret the obtained data, the scoring scale presented in Table 8 can be used.

Table 8

Scale for Interpreting the Need for Implementing Developed Measures

Rank	Score	Criterion	Value
1	12-17	Critical need for implementation	The developed measures are absolutely necessary. The problem has catastrophic consequences, affects key processes, safety, or reputation. Without implementation, significant financial losses, production stoppage, or other serious risks are possible.
2	18-23	Very high need for implementation	The developed measures are extremely important. The problem has a strong negative impact on processes, product quality, or customer satisfaction. Implementation will significantly reduce risks and improve the situation.
3	24-29	High need for implementation	The developed measures are important. The problem has a noticeable impact on processes or results but is not critical. Implementation will prevent worsening of the situation and increase efficiency.
4	30-35	Moderate need for implementation	The developed measures are of medium importance. The problem has a limited impact, but its resolution can improve certain aspects of work or prevent minor risks.
5	36-41	Low need for implementation	The corrective measures are of low importance. The problem has a minimal impact, and its resolution will not lead to significant improvements. Implementation can be considered in the long term.
6	42-47	Very low need for implementation	The corrective measures are insignificant. The problem has almost no impact on processes, and its resolution will not bring noticeable improvements. Implementation is not a priority.
7	48-53	Extremely low need for implementation	The developed measures are practically unnecessary. The problem has no impact, and its resolution has no practical value. Implementation is not advisable.
8	54-60	No need for implementation	The developed measures are not required. The problem is absent or has no impact on processes. Implementation makes no sense.

Thus, the need for implementing the developed measures for the failure "Sudden engine shutdown" has a rank of "High need for implementation," meaning the developed measures are important, the problem has a noticeable impact on processes or results but is not critical, and implementation will prevent worsening of the situation and increase efficiency.

Assessing the need for implementing corrective actions is an important step in quality management, allowing companies to make informed decisions and allocate resources effectively. Using a structured methodology, including scoring based on key criteria, helps prioritize and focus on the most significant issues. [5]

It is important to remember that corrective actions should not only address symptoms but also eliminate the root causes of discrepancies. This requires deep problem analysis, involvement of all stakeholders, and continuous monitoring of results.

Implementing corrective actions is not a one-time event but part of a continuous improvement process. Companies that pay attention to assessing and eliminating discrepancies can not only improve the quality of their products and services but also strengthen their competitiveness in the market.

References

[1] Mezentseva T.A. Improving the Effectiveness of the "Corrective and Preventive Actions" Process within the Quality Management System of SIBUR-Khimprom JSC. – Tyumen: Tyumen Industrial University, 2017. – pp. 332-335.

[2] Petrovsky E.A., Kazantseva A.V. Improving the Efficiency of Corrective and Preventive Actions Using a Process Quality Management Model. – Moscow: Innovations and Investments, 2012. – pp. 41-43.

[3] Mamazhonov A.A., Kuldashev D.U. Corrective and Preventive Management Measures. – Moscow: UNIVERSUM: Technical Sciences, 2022. – pp. 31-33.

[4] Kuzhuget T.V. Development of Corrective and Preventive Measures to Improve the Effectiveness of Quality Management Processes at TETZ JSC. – Tomsk: STT LLC, 2024. – pp. 445-449.

[5] Makarova L.V., Tarasov R.V., Baukova N.S. Assessing the Effectiveness of Corrective and Preventive Actions in the Production of Capacitive Equipment. – Penza: Regional Architecture and Construction, 2020. – pp. 46-55.

STUDY OF THE AUDIO SIGNAL ENTROPY CALCULATION ALGORITHM FOR SYNTHESIZED VOICE RECOGNITION

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

Abstract. *The study proposes a method for identifying synthesized speech based on the calculation of the entropy of an audio signal. For this purpose, the ordinate axis is divided into n+k intervals. Of these, n intervals cover the range from the minimum to the maximum average value, while k intervals are arranged such that the first interval includes the global minimum and the last interval includes the global maximum. The entropy is then calculated using Shannon's formula. The results showed that the entropy of synthesized speech is significantly higher than that of natural speech. Additionally, synthesized speech is more evenly distributed across the intervals, and the nature of this distribution remains consistent across different recordings. In contrast, the probability distribution for natural speech is more chaotic and contains more empty intervals, particularly in the extreme regions.*
Keywords: *synthetic speech detection, spoofing, Shannon's entropy, binary classification.*

Introduction

Currently, the generation of synthetic speech using artificial intelligence (AI) has reached a high level, primarily due to the rapid evolution of deep learning models and the availability of large datasets for training such models. However, as the quality of synthetic speech improves, the number of cyberattacks using synthetic voices is also increasing, imposing special requirements on systems where ensuring the confidentiality of users' personal data is critical. To address this issue, several approaches to synthetic speech detection have been proposed, which can be broadly categorized into those that rely on machine learning models and those that use only the mathematical properties of the audio signal. One method for detecting synthetic speech using machine learning involves classifying an audio signal based on the properties of the windowed Fourier transform [1], through which training set features such as constant-Q transform coefficients, Mel-frequency cepstral coefficients (MFCCs), and linear frequency cepstral coefficients (LFCCs) are computed. These features are then used to train convolutional neural network (CNN) models. The main drawbacks of this approach include the computational complexity of feature extraction, the dependence of cepstral coefficients on the recording's volume level, and their potential distortion by random noise. However, this method demonstrates relatively high reliability in detecting synthetic speech. Another known approach to synthetic speech detection is based on bispectral analysis, as proposed in [2]. This method requires significant computational resources for transforming spectral power into a bispectrum, followed by its normalization and averaging. Among newer methods, an approach based on training a Transformer model using segmented Mel-spectrogram features should be noted [3]. Mel-spectrograms, obtained using the Hanning window method, are divided into equal parts and converted into a 768-dimensional vector space, which serves as the training dataset. It is evident that such transformations impose a considerable computational burden on computing systems. A promising approach involves speaker identification through voice analysis using entropy calculation of the audio signal [4]. The key advantages of this method include the absence of computationally expensive preprocessing and the uniqueness of entropy values for different speakers. However, a notable drawback is the uneven distribution of elementary message intervals. Each of the aforementioned approaches offers certain advantages but does not fully solve the problem, serving only as an auxiliary tool in combating cyberattacks that utilize synthetic speech. The absence of an effective universal solution highlights the relevance of exploring new methods for synthetic speech detection. The objective of this study is to propose a modified method for detecting synthetic speech based on the entropy calculation of an audio signal.

A modified algorithm for Calculating the Entropy of an Audio Signal

The developed algorithm for calculating the entropy of an audio signal is a modified version of the algorithm proposed in [4]. Initially, the average extrema are calculated, and the signal undergoes preprocessing (Fig. 1). The ordinate axis of the signal is divided into $n+k$ intervals such that n n intervals are within the segment $[min_{average};max_{average}]$, while k intervals are positioned so that the 0th and $n+k-1$ intervals

encompass the global minimum and global maximum, respectively. Unlike the original entropy calculation method, the modified algorithm also divides the abscissa axis of the signal into *m* equal segments, computing entropy separately for each segment. The final entropy value is then calculated using the following formula:

$$H(X) = \frac{1}{M}\sum_{i=1}^{M} H(x_i).$$

The proposed method is based on a fundamental concept of information theory – **entropy**, a measure that represents the average level of uncertainty in a given random variable. To calculate entropy, Shannon's formula is used:

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i,$$

where p_i – is the probability of the occurrence of the i-th elementary message.

Shannon's entropy function has several important properties, such as non-negativity and an upper bound that depends on the number *n* of elementary messages. In the general case, the following inequality holds for this value:

$$H(X) = -E[\log_2 p_i] = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} \le \log_2 \sum_{i=1}^{n} 1$$

In the case where all elementary events are equally probable, the entropy of the signal will be equal to the base-2 logarithm of the number of messages *n*. However, before proceeding with entropy calculation, it is necessary to first remove parts of the signal that do not contain useful information. For signal preprocessing, the average extrema method is used, which allows for the removal of signal fragments where the average minimum and maximum values are below a certain threshold. The formal expression of this rule is as follows:

$$S_{out} = \{S_p \subseteq S : |\min_{среднее}(S_p)| + \max_{среднее}(S_p) > 2\varepsilon\}, \varepsilon > 0$$

where S is the set of all signal values, S_p is a subset of signal values within a certain interval, and $min_{average}$ и $max_{average}$ are the average minimum and maximum of the signal within this interval. The processing result is shown in Fig. 1.

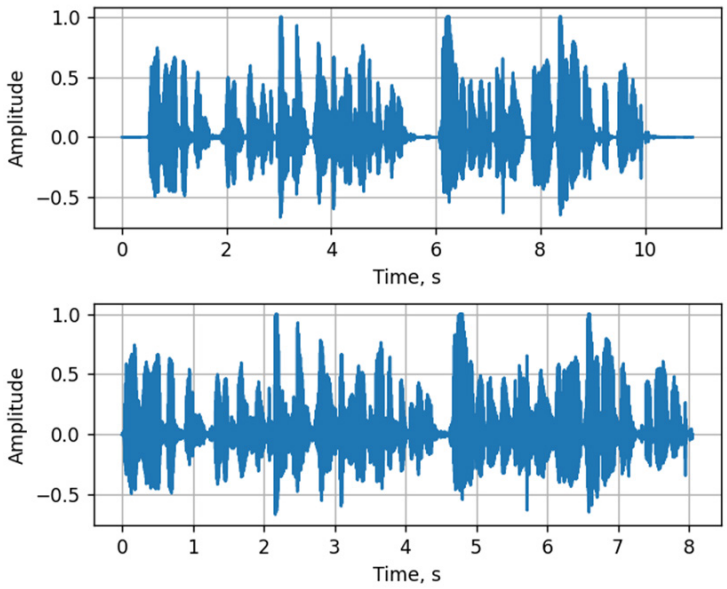


Fig. 1. The input signal (above) and the preprocessing result (below)

After preprocessing, Shannon entropy for the discrete signal is calculated using the following algorithm:

1. The envelope of the minima and maxima of the function is constructed, and the average values of the minimum and maximum of the signal are calculated.

2. The segment is divided into n equal parts, each of which represents an elementary message.

3. The set of elementary messages is supplemented with the necessary number of segments k on the Y-axis, so that the values of the global minimum and maximum are included in the extreme segments. Thus, the total number of elementary messages becomes $n+k$.

4. The values of the discrete signal are sorted, and the membership of each value to one of the segments on the Y-axis is determined.

5. The entropy is calculated using the formula.

$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

The natural numbers n and k, which define the number of partitions, are chosen arbitrarily. According to the definition of entropy, the larger the number of partitions, the greater the value of the function $H(X)$. However, as n and k increase, the rate of growth of $H(X)$ slows significantly when $n+k>100000$. In an experimental comparison of the entropy of four pairs of 10-second audio signals (synthesized and real voices) pronouncing the same phrase, the graph of entropy growth over time shows significant differences, as shown in Fig. 2. The phrases for analysis are presented in Table 1.

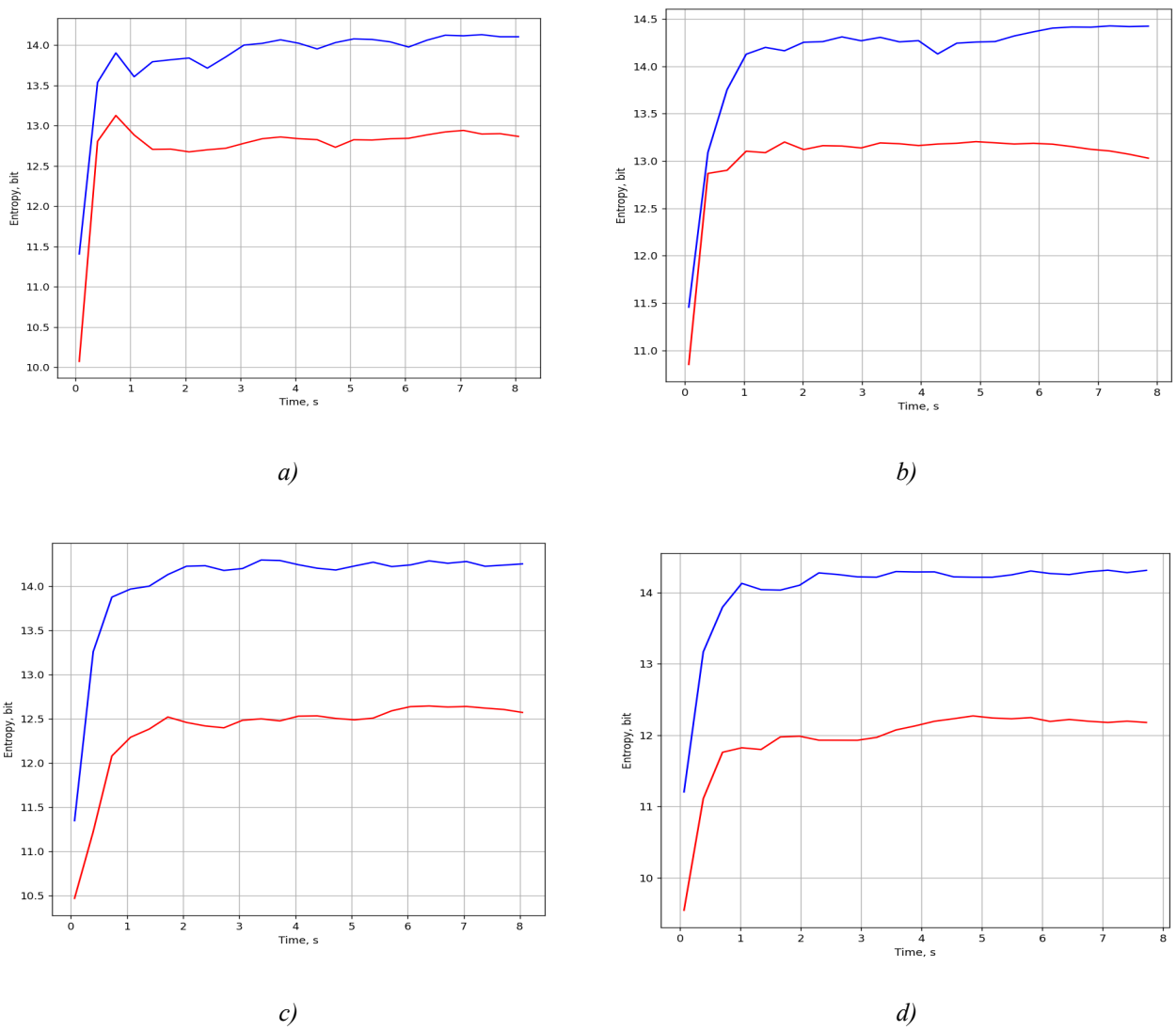
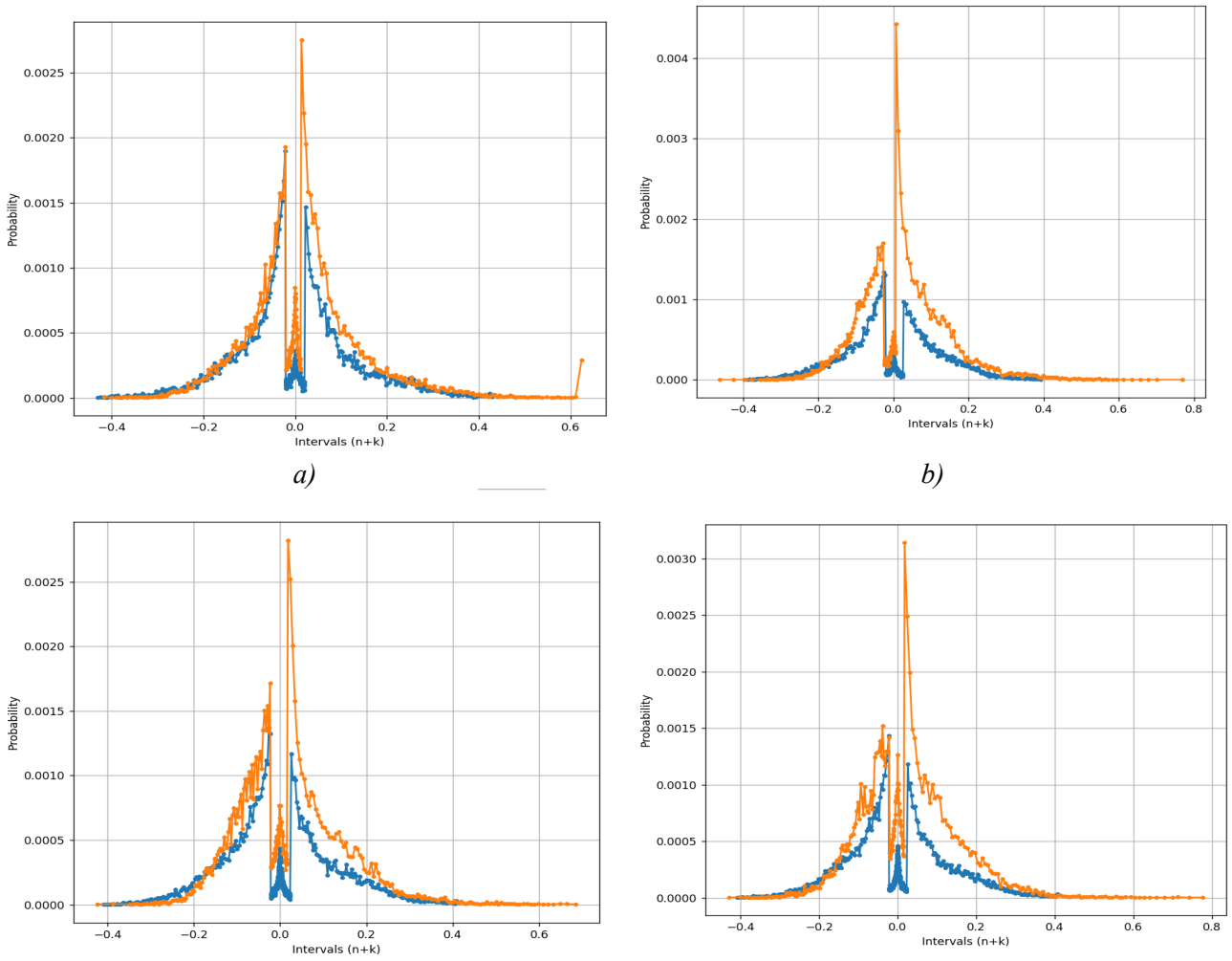


Fig. 2. Comparison of the entropy of real and synthesized signals for 5-second recordings

Table 1

10-second phrases spoken by AI and a real person

Phrases №	Text
1	They're very different. Physics answers the question: what is the nature of the universe? Geology answers the question: you know, what'd I just trip over?
2	Well, we had dinner, played some games, and then I spent the night. Oh, and you'll be happy to know that I now have a much better understanding of 'friends with benefits'.
3	Oh, and one last thing. If you find yourself working with a male scientist who's as smart as me, as tall as me and has hair like Thor, well, then I want you to step away from the situation
4	Apparently you can't hack into a government supercomputer and then try to buy uranium without the Department of Homeland Security tattling to your mother.



a) *b)*

c) *d)*

Fig. 3. The dependence of the probability of a signal value falling into an interval on the midpoint of each interval

During the experiment, it was found that the entropy of synthesized speech is always significantly higher than the entropy of a real speech recording, as shown in Fig. 2, regardless of the volume of the audio recording. Also, if the probability distribution is graphically represented, as shown in Fig. 3, for the signal values falling into the intervals obtained by dividing the ordinate axis into $n+k$ parts, the synthesized voice exhibits a characteristic “freeze” of probability around zero, which does not occur in the real signal.

Conclusion

In this work, a modification of the signal entropy calculation algorithm was proposed. This algorithm was tested for its potential use as a means to counter cyber-attacks involving voice impersonation. The main advantages of the proposed solution are the absence of multi-level preprocessing, independence from volume, and the simple mathematical calculations underlying it. The entropy values for different segments of the signal, obtained through processing, can be used either as an auxiliary tool for analyzing the input signal or as features for training AI models that solve tasks related to clustering the properties of real and synthetic speech. Empirically, significant features inherent only to synthetic speech were identified, which, upon further study, could form the basis for more advanced and universal systems to counter cyber-attacks involving voice impersonation.

Acknowledgment

The author expresses gratitude to the senior lecturer of the Department of Computer Systems and Networks Grigoriev E. K., for his scientific guidance in preparing this article.

References

1. J. Yang и R. K. Das, «Long-term high frequency features for synthetic speech detection», Digit. Signal Process., т. 97, с. 102622, Feb. 2020, doi: 10.1016/j.dsp.2019.102622.

2. E. AlBadawy, S. Lyi, и H. Farid, «Detecting AI-Synthesized Speech Using Bispectral Analysis», CVPR Workshop, сс. 104–109, 2019.

3. A. K. S. Yadav, Z. Xiang, K. Bhagtani, P. Bestagini, S. Tubaro, и E. J. Delp, «Compression Robust Synthetic Speech Detection Using Patched Spectrogram Transformer», 21 February 2024 г., *arXiv*: arXiv:2402.14205. Available on: http://arxiv.org/abs/2402.14205

4. V. A. Suvorova, A. P. Shkaraputa. Development and application possibilities of the algorithm for calculating the entropy of a sound wave, Bulletin of Perm University, vol. 1(32), 2016. (in Russian)

5. Nenasheva Yu. A. Prosodic factor in speech synthesis. Cognitive studies of language. 2024. – № 2-2(58). – pp. 630-634. (in Russian)

6. Murtazin R. A., Kuznetsov A. Yu., Fedorov E. A. Synthetic voice detection algorithm based on cepstral coefficients and convolutional neural network. Scientific and technical journal of information technologies, mechanics and optics. 2021. – vol. 21, № 4. – pp. 545-552. – DOI 10.17586/2226-1494-2021-21-4-545-552 (in Russian)

LIBRARY SYSTEM IMPLEMENTED IN NODE.JS AND MYSQL

Marina Kreyzo
Saint Petersburg State University of Aerospace Instrumentation
e-mail: kreyzo010101@gmail.com

Abstract. *This article presents the development and analysis of a multi-user library system built on the Node.js platform using the MySQL DBMS. It examines the architectural features, key implementation aspects, and testing results. The proposed solution provides access to the book catalog while differentiating user rights (reader/administrator). By leveraging modern web technologies and a relational database, the system offers a convenient interface and reliable data storage.*
Keywords: *Node.js, MySQL, library, database management, user roles, web application*

Introduction

Modern libraries are increasingly shifting toward electronic methods of managing collections, enabling users to quickly locate needed books and obtain information on availability. By utilizing Node.js alongside MySQL, one can develop a high-performance, easily scalable solution. The goal of this work is to create a prototype web application for library management, featuring book inventory, search and reservation capabilities, as well as administrative functions for adding and editing records.

System Architecture

2.1 Node.js Server

- Processes user requests via a REST API (or GraphQL).
- Manages access to the MySQL database using libraries (mysql2, Sequelize, etc.).
- Implements authentication and role-based access control (reader, administrator).

2.2 MySQL Database

- users table: stores login credentials (username, password, role).
- books table: contains information on books (title, author, year, status).
- borrow_records table: tracks borrowing history (who took which book and when).
- Indexes on key fields simplify fast searching and data retrieval.

2.3 Software

- Node.js (JavaScript or TypeScript) and the Express.js framework for request handling.
- MySQL as a relational DBMS for data storage and organization.
- Additional packages (e.g., JWT) for authentication and resource protection.

2.4 User Interface

- A web page or SPA (React, Vue.js, Angular) for interaction with the server.
- Readers: search for books, make reservations, view history.
- Administrators: add, edit, and delete book and user records.

Methods of Implementation and Testing

3.1 Hardware

- A server (VPS or local machine) running Node.js and MySQL.
- At least 1 GB of RAM and 10 GB of disk space are recommended for basic usage.

3.2 Software Logic

- Initializing the Node.js server, connecting to MySQL (via mysql2 or the Sequelize ORM).
- Handling routes: registration, authentication, book searches, and CRUD operations for administrators.
- Using sessions or JWT tokens for security and access control to the API.

```
/*
  Example code snippet (simplified)
  for managing books in Node.js
  Version: 1.0
*/

const express = require('express');
const mysql = require('mysql2/promise');
```

```
const app = express();

// Database connection
const db = await mysql.createConnection({
  host: 'localhost',
  user: 'db_user',
  password: 'db_password',
  database: 'library_db'
});

// Fetch list of books
app.get('/books', async (req, res) => {
  try {
    const [rows] = await db.execute('SELECT * FROM books');
    res.json(rows);
  } catch (err) {
    res.status(500).json({ error: 'Database error' });
  }
});

// Add a new book (admin)
app.post('/books', async (req, res) => {
  try {
    const { title, author } = req.body;
    await db.execute(
      'INSERT INTO books (title, author, status) VALUES (?, ?, ?)',
      [title, author, 'available']
    );
    res.json({ message: 'Book added successfully' });
  } catch (err) {
    res.status(500).json({ error: 'Database error' });
  }
});

// Other routes for CRUD operations...

app.listen(3000, () => {
  console.log('Library system running on port 3000');
});
```

3.3 Experimental Setup

Below are examples of the application's two main screens.

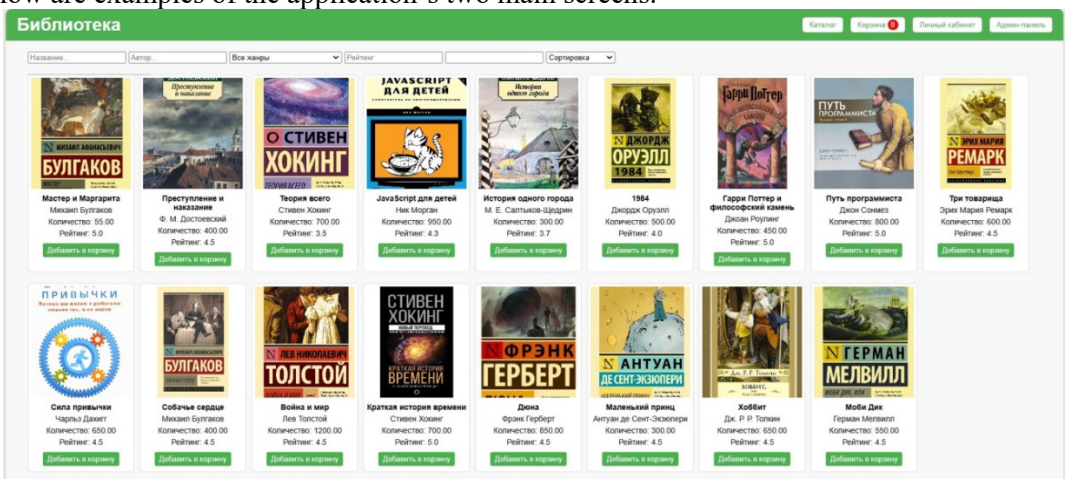


Fig. 1. Main Page

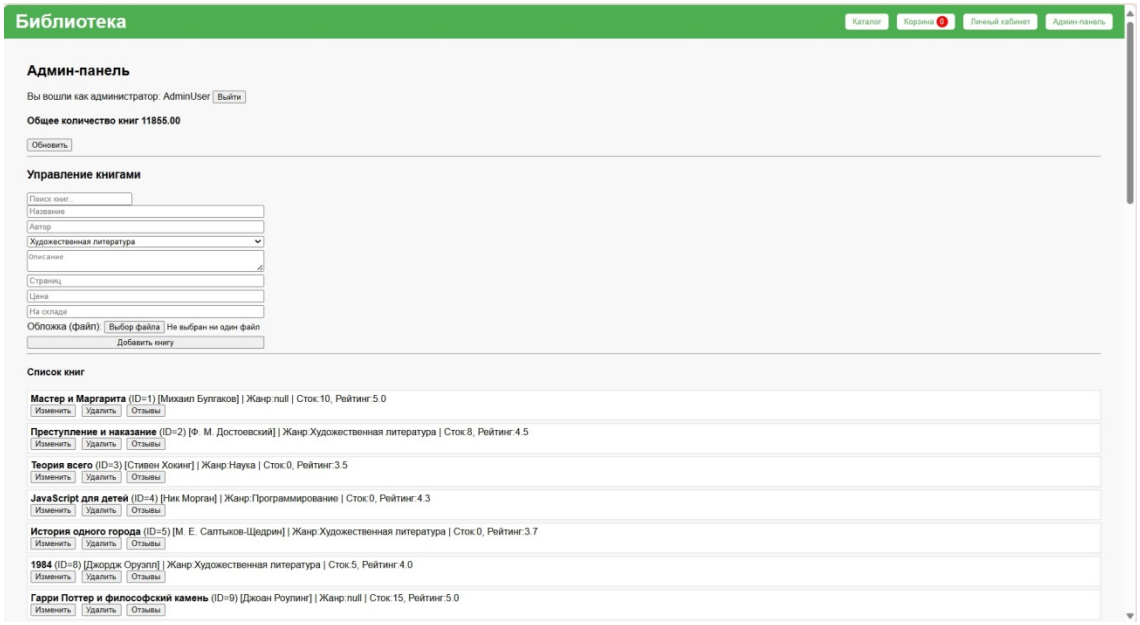


Fig. 2. Admin Panel

To verify performance and functionality, a test configuration was deployed on a VPS with 1 GB of RAM. Testing was conducted using JMeter and k6 with 100–200 concurrent users. The system handled requests stably with an average latency of about 200 ms. Properly configured MySQL transactions prevented conflicts when multiple users attempted to reserve the same book concurrently.

Conclusion

The library system built on Node.js and MySQL has demonstrated ease of use, flexible expandability, and reliable data storage. Test results indicated resilience under moderate loads and correct role-based access for both readers and administrators.

Key Advantages:

- Scalability: As the user base grows, the system can be expanded with new modules or split into microservices.
- Ease of Development: Node.js has a wide ecosystem of packages that simplify the addition of new features.
- Relational Structure: MySQL supports data consistency and transaction handling.

Prospects for Future Development:

- Adding functionality for uploading and managing e-books (ePub, PDF).
- Implementing a recommendation system based on borrowing history.
- Integrating with external services (email notifications, push notifications about return deadlines).

References

1. Node.js Documentation [Electronic resource]: https://nodejs.org/en/docs (accessed: 02/18/2025).
2. Express.js Official [Electronic resource]: https://expressjs.com (accessed: 02/18/2025).
3. MySQL Developer Zone [Electronic resource]: https://dev.mysql.com/doc (accessed: 02/18/2025).
4. Sequelize ORM [Electronic resource]: https://sequelize.org (accessed: 02/18/2025).

DEVELOPMENT AND ANALYSIS OF MATHEMATICAL MODEL OF ULTRASONIC RANGE FINDER

Vladislav Kulikov
Saint Petersburg State University of Aerospace Instrumentation
e-mail: vladikkulikov83@gmail.com

Abstract. *The paper deals with the process of developing the layout of the ultrasonic sensor based on the Arduino Nano platform. The main attention is paid to the mathematical model of distance determination based on echolocation. Theoretical justifications, calculation algorithms, as well as experimental results of the sensor testing are given.*
Keywords: *ultrasonic sensor, echolocation, mathematical model, Arduino Nano, HC-SR04, MATLAB.*

Modern ultrasonic sensors are widely used in navigation systems, industrial automation and robotics. This paper presents a mockup of ultrasonic sensor developed on the basis of HC-SR04 and Arduino Nano microcontroller. The main task is to investigate the mathematical model of distance calculation and to identify the measurement errors.

The HC-SR04 ultrasonic sensor works on the principle of echolocation, emitting a sound pulse with a frequency of 40 kHz and recording the time of its return after reflection from the object [1]. The distance to the object is calculated by the formula:

$$L = \frac{v \cdot t}{2},$$

where v is the speed of sound (343 m/s at 20°C), t is the signal travel time. Division by 2 takes into account the round trip.

To improve accuracy, a data averaging method based on multiple measurements is applied [2]:

$$L_{cp} = \frac{1}{N} \cdot \sum_{i=1}^{N} L_i$$

where N is the number of measurements set programmatically.

In addition, errors due to changes in air temperature can be taken into account. Adjustment of the sound velocity is made according to the formula:

$$c_T = 331{,}1 + 0{,}606 \cdot T,$$

where T is the ambient temperature (°C).

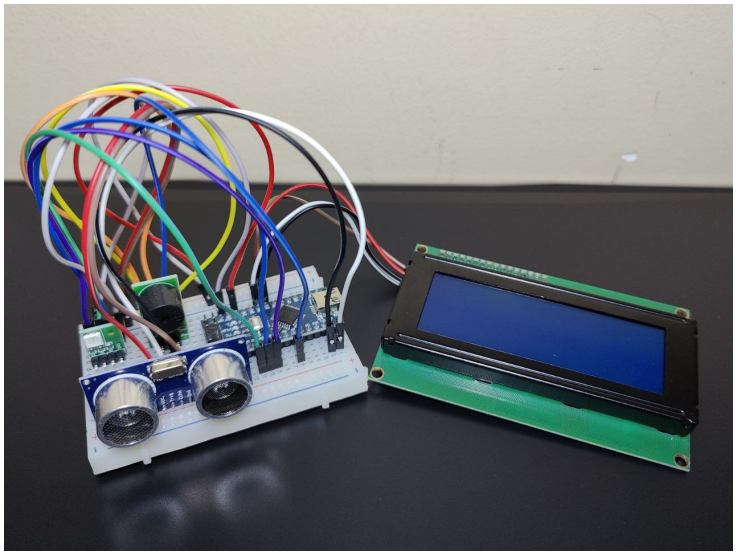


Fig. 1. Layout of the complete device

Measurements and mathematical analysis were performed on a device layout consisting of the following components:

- Arduino Nano microcontroller responsible for data processing and system control [3].
- HC-SR04 ultrasonic module used for distance measurement.
- An RGB LED provides a visual indication of the distance range.
- Buzzer to activate audible signals when objects are close by.
- LCD display for real-time display of measured distance and system status.

All components are connected using a breadboard and connecting wires, which simplifies prototyping and testing. The software is implemented in C++ using the Ultrasonic.h library. The LEDs and buzzer are controlled based on measured distance, providing instant feedback to the user.

Experimental testing was carried out under controlled conditions with changing distances from 10 to 100 cm with a step of 10 cm. The average error was less than 2%, which confirms the correctness of the applied mathematical model. Fig. 2 presents a bar chart showing the measurement error of the HC-SR04 ultrasonic sensor depending on the true distance (in centimeters).

You can use MATLAB to analyze the measurement data by applying visualization of the results. The code below loads the data, performs processing, and plots the data:

```
% MATLAB code for ultrasonic sensor data processing

clc; clear; close all;

% Example data from HC-SR04 sensor
measured_distances = [10.2, 19.8, 29.5, 39.7, 49.3, 59.8, 69.1, 79.5, 89.9, 99.2];
true_distances =  [10, 20, 30, 40, 50, 60, 70, 80, 90, 100];
% Error calculation
error = abs(measured_distances – true_distances) ./ true_distances * 100;

% Graphing
figure;
bar(true_distances, error);
xlabel('True distance (cm)');
ylabel('Uncertainty (%)');
title('Measurement error of the HC-SR04 ultrasonic transducer');
grid on;
```

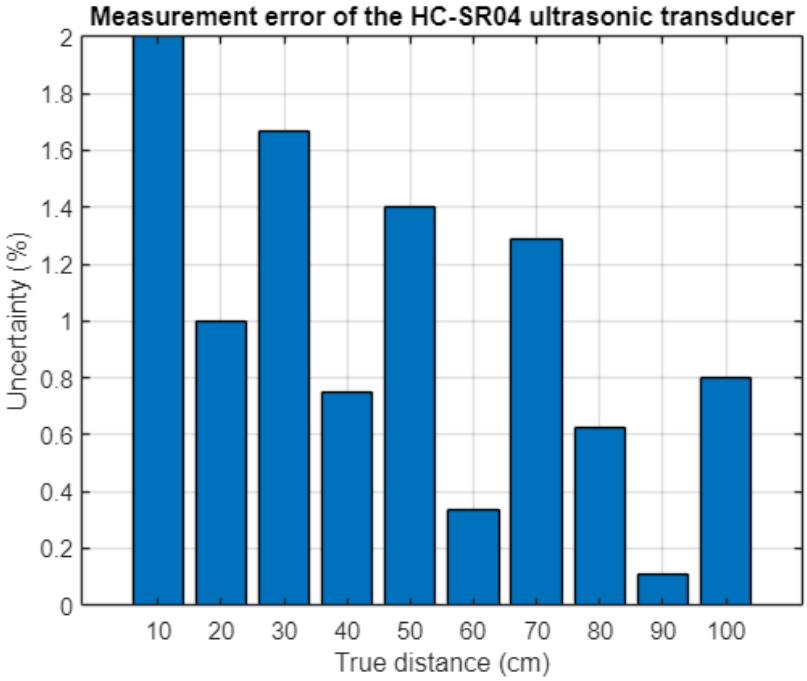


Fig. 2. Graph of measurement errors

At small distances (10 cm, 20 cm) the error is maximum (about 1.8%). At medium distances (30-70 cm) the error varies, but has local peaks. At large distances (80-100 cm) the error decreases, but remains noticeable.

Assumptions about the causes of errors:

- The error at short distances can be explained by reflection of ultrasonic waves from nearby objects and interference of signals.
- Small variations in the data at long distances may be due to the angle of reflection of the sound signal from the surface of the object.
- Also, one should not forget about the hardware limitations of the HC-SR04 module related to the accuracy of the signal return time measurement.

The developed model of the ultrasonic sensor demonstrated sufficient accuracy and stability of measurements. Data analysis in MATLAB allowed to additionally visualize the error distribution and analyze the obtained results. In the future it is possible to integrate Bluetooth and Wi-Fi modules for remote data monitoring and optimize filtering algorithms to improve the accuracy of measurements.

References

1. Zhmud V. A., Kondratyev N. O., Kuznetsov K. A. et al (2017). Ultrasonic distance measurement sensor HC-SR04. Automation and Software Engineering.
2. Pestrikov V. M. (2023). Mathematical Methods in Engineering. SPb.: VShTE SPbGUPTD.
3. Arduino Nano Documentation [Electronic resource]: https://docs.arduino.cc/.

DEVELOPMENT OF A REAL-TIME LIGHT MONITORING ALGORITHM

Yuri Kuzmenko
Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia
E-mail: spider22boy@mail.ru

Abstract. *The article presents the development of an algorithm for monitoring illumination in real time, aimed at improving the efficiency of managing light resources in various environments. The main goal of the study was to create a system that can quickly collect and process data on the level of illumination, providing accurate and continuous tracking of the light mode.*
As part of the work, an algorithm for monitoring illumination was implemented, and the operation of this algorithm was tested. Based on the obtained data, a heat map of illumination and a cumulative distribution function are constructed.

Introduction

Illumination monitoring is an integral part of modern management of illuminated spaces, having a significant impact on various aspects of society's life. Historically, the first attempts to measure and control illumination date back to the beginning of the industrial revolution, when it became necessary to optimize working conditions in factories and enterprises. With the development of light measurement technologies such as photometers and luxmeters, the practice of light monitoring has become more accurate and accessible, which has expanded its application from industrial areas to domestic and urban spaces.[1]

There are many factors that make light monitoring necessary. First, proper lighting increases productivity and improves the quality of life, providing a comfortable environment for work, leisure and learning. Second, effective light management plays a key role in saving energy and reducing the environmental footprint, allowing you to optimize electricity consumption and reduce greenhouse gas emissions. In addition, light monitoring is important to ensure safety on roads and in public places, preventing accidents and helping to create a favorable living environment.

Advanced technologies for light monitoring are rapidly developing, offering innovative solutions for real-time data collection and analysis. Among them, a special place is occupied by the Internet of Things (IoT) and sensor networks, which allow you to create scalable and flexible monitoring systems integrated with other intelligent infrastructures. Artificial intelligence and machine learning provide opportunities for processing large amounts of data, allowing you to predict changes in light conditions and automatically adjust light sources depending on your needs. In addition, the development of wireless technologies and improved sensor accuracy open up new horizons for implementing efficient and reliable light monitoring algorithms, contributing to the creation of smart cities and sustainable public spaces.

Thus, the development of a real-time light monitoring algorithm is an urgent and promising task that combines historical experience, modern requirements and advanced technological solutions to create a comfortable, safe and environmentally sustainable environment.

Development of an algorithm for monitoring light conditions for intelligent lighting control

Modern lighting control systems play a key role in providing a comfortable environment for users, as well as in optimizing energy consumption. Real-time light monitoring allows you to adapt the light level according to current needs and external conditions, which leads to increased energy efficiency, reduced energy costs and a favorable atmosphere. In addition, this approach helps extend the life of lighting devices and provides flexibility in managing lighting in various operating scenarios.[2]

Input data:

1. Current ambient light level: measured in lux using specialized sensors.
2. Desired light level: set by the user according to the preferences and requirements for a particular space.
3. Historical illumination data: previously collected and uploaded to the system information about light levels for previous periods, taking into account the time of day and weather conditions.

Output data:

1. Current light condition:
- Over-illumination.
- Insufficient illumination.

- Optimal light level.

2. Recommended actions:

- Increase the brightness of the lighting.
- Reduce the brightness of the lighting.
- Keep the current level unchanged.
- Detect deviations from the normal light level.

Key performance indicators of the algorithm:

- Speed of response to changes in ambient light.
- Precise adjustment to achieve the desired light level.
- Power consumption of the lighting control system.

Stages of the algorithm:

1. Initializing and configuring threshold values:

- Determination of the permissible range for the user-set light level. For example, for a given level of 500 lux, the acceptable range can be from 480 to 520 lux.

2. Data collection and processing:

- Measurement of current illumination:sensors constantly record the level of light in the room or on the street.
- Historical data availability analysis: checks the availability of relevant data for the current time and conditions, which allows you to take into account seasonal and weather changes when making decisions.

3. Compare the current light level with the desired one:

- If the measured value is within the set range, the system classifies the condition as "optimal".
- Exceeding the upper limit of the range indicates "over-illumination".
- A drop below the lower limit indicates "insufficient illumination".

4. Analyze historical data and identify corrective actions:

- Optimal state: The system activates the mode of maintaining the current light level unchanged.
- Over-illumination: it is recommended to reduce the brightness of the lighting to achieve an optimal level.
- Insufficient illumination: it is suggested to increase the brightness of the lighting to improve the conditions.

5. Output results and update the database:

- Recording the current state: records the results of light assessment and suggested actions.
- Update of historical data: current light indicators and relevant conditions are saved for further use and analysis.

Example of a real-time light monitoring algorithm

To implement the example, we will set the initial conditions using synthetic data, the data models the light level for a 24-hour period with an interval of 10 minutes.

Synthetic data displays the nature of illumination to simulate a typical day-night cycle:

Overnight (0:00 – 2:50): 0-10 Lc

Dawn (3:00 – 5:50): 10-100 Lc

Day (6:00 – 17:50): 100-500 Lc

Twilight (18:00 – 20:50): 10-100 Lc

Overnight (21:00 – 23:50): 0-10 Lc

Let's assume a target illumination level of 250 Lux. This is the level at which, ideally, the room should be illuminated for optimal user comfort.

A deviation of ±20 lux from the specified light level is considered acceptable. That is, the optimal illumination level in this example is considered to be in the range of 230-270 lux [3].

Code operation diagram for performing simulation of the described algorithm

For each data point (timestamp), the current light level is compared with the desired light level.

If the current light level is within the threshold range (230-270 lux), the state is set to "optimal" and the suggested action is set to"maintain".

If the current illumination level exceeds the threshold range (more than 270 lux), then the state is set as "excessive illumination", and the suggested action is set to "reduce brightness".

If the current light level is below the threshold value (less than 230 lux), then the state is set as "insufficient light", and the suggested action is "increase brightness".

These states and actions provide real-time feedback on lighting conditions and determine possible adjustments to the lighting system to achieve the desired light level.

The pseudo-code for modeling the algorithm is presented below.

```
function monitor_illumination(desired_illumination, historical_data):
    THRESHOLD_RANGE = 20  # Example value
    lower_bound = desired_illumination – THRESHOLD_RANGE
    upper_bound = desired_illumination + THRESHOLD_RANGE
        current_ambient_light = measure_light_sensor()
        if lower_bound <= current_ambient_light <= upper_bound:
        state = "optimal"
        action = "maintain"
    elif current_ambient_light > upper_bound:
        state = "over-illuminated"
        action = "decrease brightness"
    else:
        state = "under-illuminated"
        action = "increase brightness"
        if historical_data:
        predicted_change = analyze_historical_data(historical_data)
        if predicted_change == "increase":
            action = "prepare to increase brightness"
        elif predicted_change == "decrease":
            action = "prepare to decrease brightness"
        update_database(current_ambient_light)
        return state, action
```

Results of modeling the algorithm operation

Fig. 1 shows the light levels over a 24-hour period. The green dotted line represents the desired illumination level of 250 lux. Fluctuations in the graph represent changes in the light level, reflecting typical transitions from night to dawn, daytime, dusk, and again night.

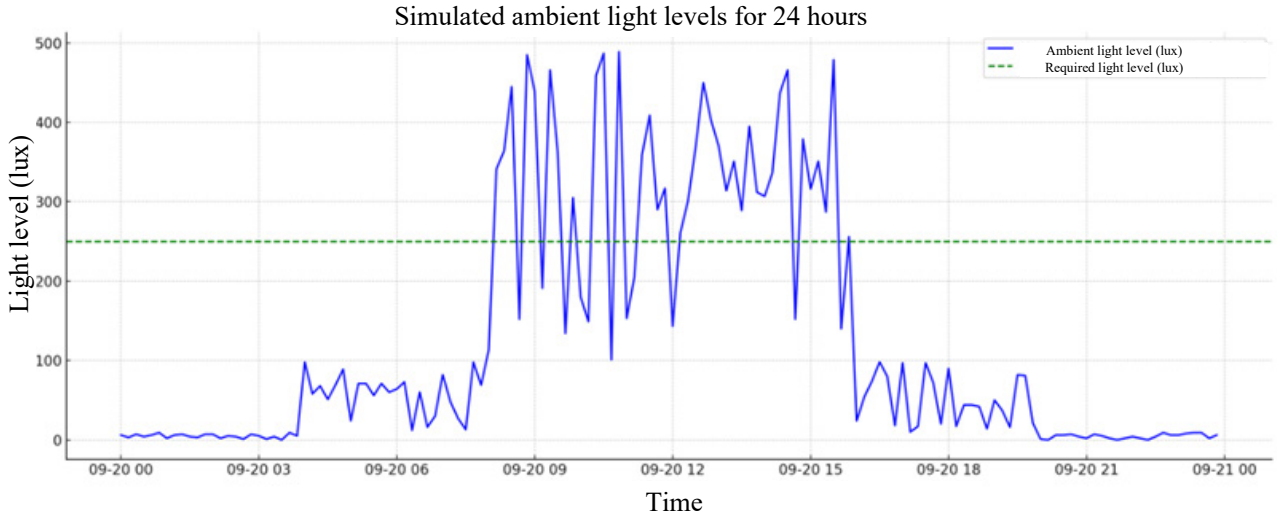


Fig. 1. Graph of the light level per day

This can serve as a guide for understanding the example and the decisions made during its development.

Fig. 2 shows a representation of the average light level for each hour of the day. Darker colors indicate a lower light level, while brighter colors indicate a higher one.

The Cumulative Distribution Function of Light Levels (CDF) shows the proportion of data points whose value is less than or equal to each light level on the x-axis.

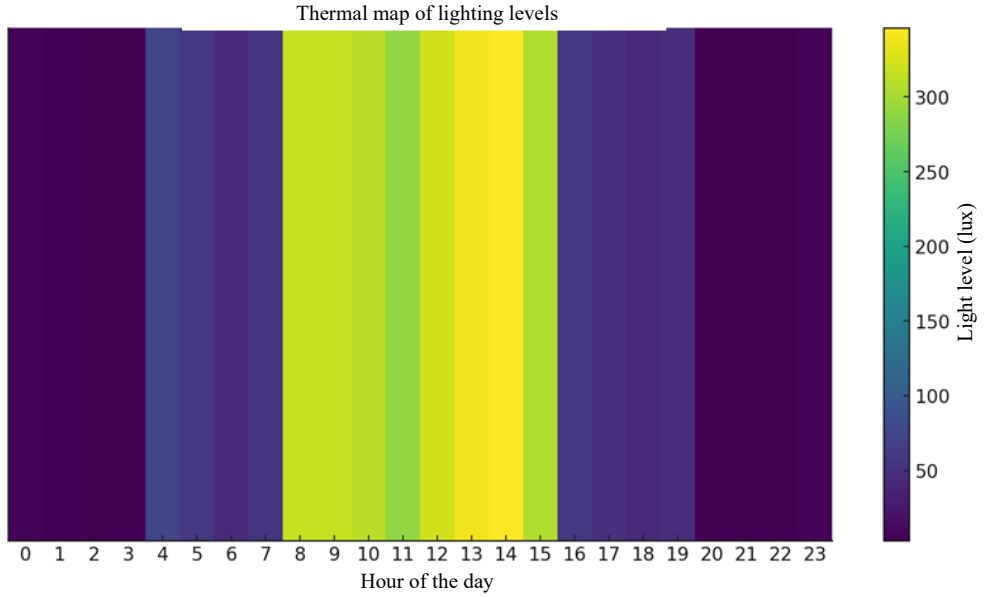


Fig. 2. Heat map of lighting levels

A heat map of light levels provides a visual representation of how the light level fluctuates during the day. This is particularly useful for lighting engineers and designers looking to optimize lighting systems for different environmental conditions. Recognizing these lighting patterns can help you develop more adaptive and energy-efficient systems [4].

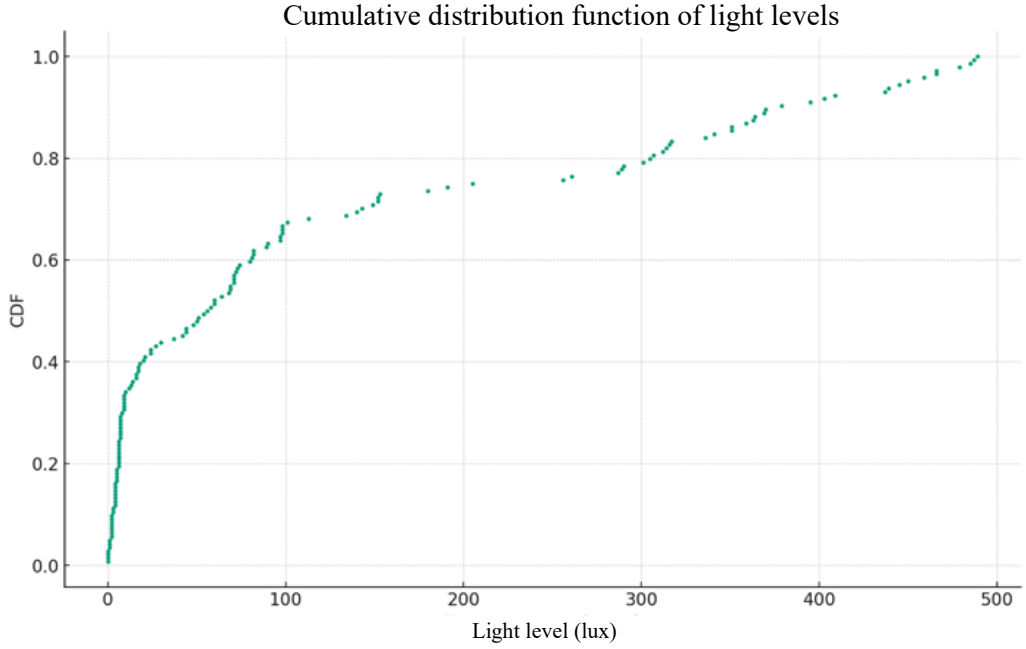


Fig. 3. Plot of the cumulative distribution function

Fig. 3 shows a graph of the cumulative distribution function (CDF), which provides an understanding of how light levels are distributed in the data set. This allows you to estimate the probability that the illumination will be within a certain range. Such information plays an important role in setting thresholds for intelligent lighting control systems, ensuring that lighting adjustments are both subtle and efficient [5].

The advantages of this algorithm include the following:

- Adaptability: theability of the system to automatically adapt to changes in the environment.
- Energy efficiency: optimization oflighting levels helps to reduce electricity consumption.
- Increased comfort: oproviding a stable light level creates a comfortable environment for users.
- Equipment durability: yincreased use of lighting devices extends their service life.

- Flexible management: theability to integrate with other smart home or office management systems for more comprehensive control. Thus, the proposed algorithm for real-time illumination monitoring is an effective solution for creating intelligent light control systems that help both save resources and improve the quality of life of users.

This work and the modeling study carried out in it emphasize the importance of combining the algorithmic functionality of modern digital technologies with responsiveness to dynamic changes in the environment. As we move deeper into the realm of smart cities and smart habitats, research like this will undoubtedly play a key role in shaping a sustainable, comfortable and energy-efficient future.

Conclusion

In conclusion, it should be noted that the developed algorithm for real-time illumination monitoring demonstrates high accuracy and stability in various lighting conditions. Thanks to the use of modern data processing methods and optimization of computing processes, the system is able to quickly respond to changes in the light level, providing reliable control and management of lighting in real time.

Thus, the presented work lays a solid foundation for further research and development in the field of intelligent lighting control, opening up prospects for creating more efficient and adaptive systems for monitoring illumination in real time.

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 illuminationindicators, 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. 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 devices], Moscow, 2021. https://fs.guap.ru/dissov;://fs.guap.ru/dissov;

2. Wang Z., Tan Y.K., Illumination control of LED systems based on neural network model and energy optimization algorithm // Energy and Buildings 62. 2016. pp. 514–521.

3. Kuzmenko V. P., Soleny S. V. Predictive service model for quality management of LED lighting networks // Bulletin of Magnitogorsk State Technical University named after G. I. Nosov. 2023. Vol. 21. No. 3. pp. 155-169.

4. Segovia-Muñoz D. & Serrano Guerrero, Xavier & Barragán-Escandón, Antonio. Predictive maintenance in LED street lighting controlled with telemanagement system to improve current fault detection procedures using software tools. Renewable Energy and Power Quality Journal, 2022, vol. 20, pp. 379-386. 10.24084/repqj20.318.

5. Gong Chen & Xu, Haiping & Yuan, Zengquan & Liang, Jinhua. The Accelerated Life Test Investigation and Lifetime Prediction Method for LED Driver, 2022, pp. 355-359. 10.1109/ICPEA56363.2022. 10052233.

IOT IMPLEMENTATION IN RAILWAY ROLLING STOCK DIAGNOSTICS: ANALYSIS OF MEASUREMENT ACCURACY AND DATA CORRECTION METHODS

Vyacheslav Leontyev
Saint Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia
E-mail: alexvake365@gmail.com

Abstract. *Implementation of the Internet of Things (IoT) in the diagnostics of railway rolling stock allows to significantly increase the accuracy and efficiency of their technical condition monitoring. The performance of units and assemblies is assessed based on the analysis of parameters transmitted by sensors in real time. Correctness of the obtained data depends on operating conditions and characteristics of measuring devices, such as signal stability, interference level, temperature fluctuations and vibration loads.*
An IoT diagnostics system must comply with established regulations governing measurement tolerances and environmental parameters such as temperature, humidity, and electromagnetic radiation. Violation of these conditions can lead to distorted analysis results and incorrect maintenance decisions. Despite the high reliability of modern sensors and data processing algorithms, it is impossible to completely eliminate the probability of errors caused by external factors. However, the use of intelligent filtering methods and self-correction algorithms allows minimizing the influence of interference and ensuring the reliability of diagnostic data.

Deviations of electrical control and measuring instruments

Deviations of railway rolling stock operating parameters from the normative values are diagnostic errors. Detection, analysis and minimization of such errors are the key tasks of technical condition monitoring systems. The introduction of Internet of Things (IoT) technologies makes it possible to record and process data on the condition of rolling stock units and assemblies in real time, which significantly increases the accuracy of diagnostics.

Diagnostic errors can be classified by various criteria depending on the sources of their occurrence. The main division is carried out by computational and manifest deviations. Computational errors include absolute, relative and reduced deviations of equipment operation parameters. Using these indicators, the reliability of monitoring systems operation is analyzed and the need for maintenance is determined. By the nature of manifestation errors are divided into random, systematic and gross. For example, systematic deviations may occur due to incorrect calibration of sensors or changes in operating conditions. Gross errors are most often due to individual sensor failures or data transmission failures [1].

All monitoring devices used in the railway industry can be globally divided into analog and digital devices. Analog sensors record equipment operation parameters by measuring physical quantities, while digital devices transmit data in a discrete format. The methodology of data processing for these types of systems is different. For example, with analog sensors, interpolation and parallax errors associated with visual readings must be considered. In digital devices the key is quantization error due to data discretization [2]:

$$\Delta X_K = \frac{X_{\max} - X_{\min}}{2^n}$$

where: $X_{\max}$, $X_{\min}$ – the range of the measured value; n – ADC (analog-to-digital converter) digit capacity.

In addition, the influence of external factors on sensor performance and measurement accuracy plays an important role. Temperature fluctuations, vibration, electromagnetic interference and voltage spikes can distort the transmitted data, leading to false alarms or missing critical faults. For example, the parallax error associated with visual fixation of analog instruments is calculated using the formula:

$$\Delta X_p = \frac{y \cdot z}{l}$$

where: y – the distance between the scale and the pointer; z – shifting the operator's gaze; l – distance from the observer to the scale.

The IoT diagnostics systems for railroad trains must comply with the established regulations governing the permissible measurement errors and operating conditions. An important parameter is the

reproducibility of diagnostics, which determines the accuracy of repeated measurements under the same conditions. The combination of all factors, such as sensor accuracy, reliability of data transmission and analysis algorithms, determines the quality of the monitoring system [3].

One of the important parameters of diagnostic performance is the probability of fault detection, which is calculated according to the exponential law of reliability:

$$P_d = 1 - e^{-\lambda \cdot t}$$

where: λ – failure rate (number of failures per unit time); t – system uptime.

This indicator allows you to evaluate the effectiveness of predictive analytics in IoT systems and optimize the maintenance schedule.

The probability of diagnostic errors should be highlighted separately. Data analysis can be considered in terms of the probability of fault detection and the risk of false alarms. These indicators have a direct impact on the reliability of train operation and transportation safety. Intelligent analysis algorithms minimize diagnostic errors and ensure high accuracy of equipment condition monitoring [4].

Statistical analysis of errors in IoT-diagnostics systems of railway trains

Measurement errors recorded by IoT sensors in railroad rolling stock monitoring systems can be viewed from a statistical point of view, as they obey certain distribution laws. IoT devices are designed to display real physical quantities with high accuracy, but various factors such as sensor wear, electromagnetic interference and temperature fluctuations lead to the appearance of measurement errors.

It can be concluded that the calculated value of sensor error tends to a certain mathematical expectation. However, in the process of operation, the accuracy of sensors decreases, which leads to an increase in the scatter of measurements. Despite this, the readings of the device continue to tend to the real value of the measured quantity, but with large deviations. This effect confirms the hypothesis about statistical behavior of errors and their dependence on operating conditions.

The most noticeable among all deviations are gross errors, which arise for various reasons and have a probabilistic nature. They represent statistical outliers that significantly distort the diagnostic results. For example, sudden voltage spikes in the onboard network can cause abnormal sensor readings, leading to false alarms or, conversely, hiding a potential fault.

Analysis of error distribution

When considering measurement errors from a statistical point of view, we can use methods to estimate their distribution. In IoT diagnostic systems, the most common errors are those that obey the normal (Gaussian) distribution law. This can be verified using statistical criteria such as:

- Shapiro-Wilk criterion,
- Kolmogorov-Smirnov criterion,
- Lilliefors criterion.

These criteria allow us to determine if the error distribution conforms to the normal law. In case of conformity, standard methods of sensor calibration and error correction can be used [3].

If the errors do not conform to a normal distribution and the data contain a significant number of outliers, then more sophisticated filtering methods such as:

- median filtering,
- adaptive smoothing algorithms,
- machine learning methods for anomaly prediction.

When analyzing IoT sensor data, a confidence interval indicator can be used, which is calculated using the formula:

$$X_{mid} \pm t_a \cdot \frac{\sigma}{\sqrt{n}}$$

where: $X_{\text{ср}}$ – average value of measurements, t_a – critical value of the t-statistic, σ – standard deviation, n – sample size.

This indicator allows you to assess how accurate the measurements are and within what limits you can trust the data received from the sensors [5].

It is critical to detect gross errors in IoT-based train diagnostics systems, as they can lead to incorrect decisions about the condition of the rolling stock. For example, if the system mistakenly detects overheating

of the axlebox due to a sensor failure, this can lead to false train stoppage and unnecessary maintenance costs.

Emissions can be detected using the three-sigma method:

$$X \notin [X_{mid} - 3 \cdot \sigma, X_{mid} + 3 \cdot \sigma]$$

If the measured value falls outside these limits, it is considered an outlier and must be further checked.

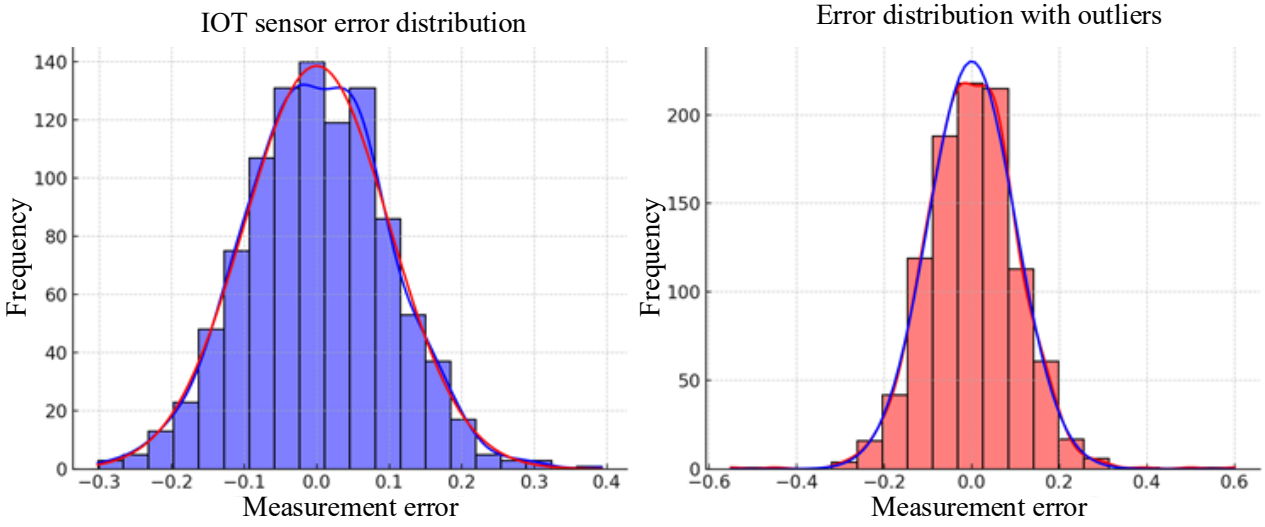
Additionally, the interquartile range (IQR) method can be used:

$$IQR = Q3 - Q1$$

$$X \notin [Q1 - 1{,}5 \cdot IQR, Q3 + 1{,}5 \cdot IQR]$$

where: Q1 – first quartile (25% of the data), Q3 – third quartile (75% of the data).

These methods can automatically filter out anomalous values and improve the quality of the data fed into the monitoring system.



a) *b)*

Fig. 1. Histograms representing IoT sensor measurement error distributions. a) The histogram shows the distribution of errors obeying a normal (Gaussian) distribution, b) The histogram includes outliers due to electromagnetic interference, sensor wear and tear and other factors.

The study showed that before the introduction of errors, the deviations in the performance of instrumentation and control (I&C) 1 and 2 were within ±1.5 % and the average accuracy values were 0.6 % and 0.665 % respectively. However, three measurements were out of limits after the introduction of deviations, increasing the average accuracy to 0.737 % and 0.833 % [6].

The graphical analysis of the error distribution shows a change in the shape of the curve:

- An increase in the variance of values, indicating a decrease in sensor accuracy.
- Appearance of statistical outliers indicating possible systematic errors.
- A shift in the mathematical expectation, which can lead to incorrect diagnostic conclusions.

To improve the reliability of measurements in IoT monitoring systems, it is necessary to apply error correction methods. Among the most effective ones are:

- Median filtering, which allows to exclude outliers and reduce the influence of random noise.
- Adaptive smoothing algorithms correcting the dynamics of data changes.
- Machine learning methods that analyze time series and predict anomalies in sensor operation.

Application of these methods will minimize the influence of random and systematic errors, as well as increase the accuracy of diagnostics.

Identification of critical errors using statistical criteria

Evaluation of IoT sensor data quality is possible using statistical criteria:

Shapiro-Wilk criterion – checking if the distribution conforms to the normal law.

Three sigma method – identifying values outside the X_cp±3σ limits

Inter-quartile range (IQR) method – detection of abnormal readings.

If the deviations of IoT sensors exceed acceptable limits, the devices need to be calibrated or replaced.

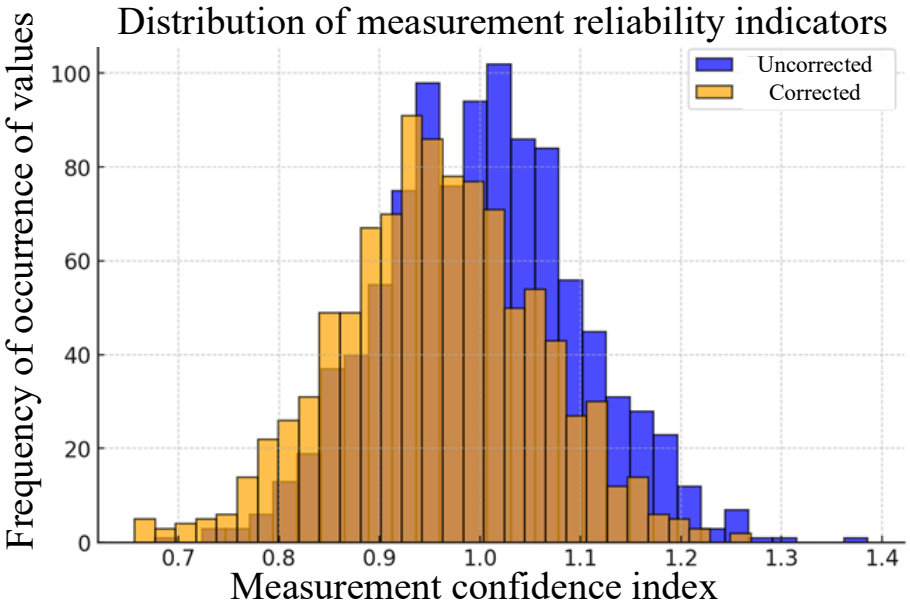


Fig. 2. The histogram shows the distribution of confidence values for two cases: without correction (blue histogram) and with correction (orange histogram)

The histogram illustrates the distribution of measurement confidence indices in two different cases:

- Without correction (blue histogram) – reflects the raw data obtained without applying filtering and error correction methods.
- With correction (orange histogram) – shows how the confidence distribution changed after applying error correction methods such as median filtering and adaptive smoothing.

X-axis ("Measurement Confidence Score") – reflects the numerical value of IoT sensor data confidence. The closer to one, the higher the measurement accuracy [7, 8]. Y-axis ("Frequency of occurrence of values") – shows how often different validity values occur in the studied data.

Uncorrected (blue graph):

- The distribution of plausibility values has a wider spread.
- Outliers are observed – extreme values that may indicate systematic errors or strong interferences in the measurements.
- Data are less concentrated around the mean value, indicating high variability in measurements.

With correction (orange graph):

- After correction, the distribution became more clustered around the center value, indicating improved measurement accuracy.
- The number of outliers decreased and the spread of values decreased.
- On average, the confidence scores have shifted to a narrower range, confirming the effectiveness of the correction methods.

Conclusion

The introduction of IoT in railway rolling stock diagnostics offers significant advantages in monitoring the technical condition of equipment. However, the measurement accuracy of IoT sensors is affected by various external factors such as temperature changes, electromagnetic interference and mechanical vibrations. This can lead to systematic and random errors that significantly affect the reliability of diagnostic data.

To improve the efficiency of IoT monitoring in the railroad industry, it is necessary not only to use high-quality sensors, but also to apply modern data processing methods. This includes median filtering, adaptive smoothing, and machine learning algorithms for anomaly detection and prediction.

The application of these techniques will not only improve the accuracy of measurements, but also ensure timely detection of faults, which will improve the operational safety of railroad trains and reduce their

maintenance costs. Thus, the combination of IoT technologies with intelligent data processing is becoming a key element of reliable diagnostics in the transportation industry.

References

1. GOST R 8.615-2013. “Reliability and requirements for methods of verification of measuring instruments”. – Moscow: Standardinform, 2013.

2. GOST 8.207-76. “Direct measurements with multiple observations. Methods of processing the results of observations”. – Moscow: Izd vo standards, 1976.

3. Shapiro, S.S., Wilk, M.B. “Analysis of distribution of random variables for conformity to the normal law” // Biometrika, 1965.

4. David J. Hand. “Principles of statistical data analysis”. – Moscow: DMK Press, 2018.

5. Gavrilov, I.Yu., Kuznetsov, V.I. “Methods of data processing in the systems of monitoring the technical condition of railway transport” // Vestnik Transportnoy nauki, 2020.

6. IEEE Internet of Things Journal. “Reliability and Accuracy of IoT Sensors for Railway Condition Monitoring”. – IEEE, 2021.

7. Vorobyev, A.A., Smirnov, K.P. “Methods of filtering and adaptive smoothing in the systems of rolling stock diagnostics” // Izvestiya Vuzov. Mashinostroenie, 2019.

8. Smirnov, V.V., Ivanov, D.A. “The use of machine learning algorithms for time series analysis in railway transport monitoring” // Automation and Telemechanics, 2021.

MODELING OF PLANT GROWTH PROCESS

Sofya Lisovenko
Saint Petersburg State University of Aerospace Instrumentation
e-mail: lisovenko123123123@gmail.com

Abstract. *This paper examines the process of plant development from the perspective of mathematical modeling. A mathematical model has been developed to describe the sequential phases of plant growth, including germination, vegetative development, flowering, and fruit formation. For each stage, key parameters influencing growth dynamics are defined, and time dependencies are constructed. Additionally, an implementation of the model based on a microcontroller, a real-time module, and a display is proposed, enabling visualization and control of the growth process.*
Keywords: *seed germination, vegetative growth, Lockhart equation, flowering, fruiting, growth stage, ESP-WROOM-32, mathematical model.*

Introduction

The plant growth process is a complex interaction of biological, chemical, and physical factors accompanying each stage of development. From the moment of seed germination to fruit maturation, plants undergo sequential stages regulated by both internal and external conditions. The application of mathematical modeling allows for a detailed description of growth dynamics and prediction of future changes. In modern conditions, technologies such as microcontrollers and real-time sensors enable automated monitoring and increase observation accuracy.

Main Stages of Plant Growth

Four main stages of plant growth can be identified: germination, vegetative growth, flowering period, and fruit ripening.

Seed Germination – This stage covers the period from seed planting to the emergence of the first sprout, where the primary factor is water absorption and seed swelling.

The mathematical model is given by:

$$W(t) = W_0 + k \cdot t, ,$$

where $W(t)$ is the seed mass at time, W_0 is the initial seed mass, and k is the water absorption coefficient. When the seed mass reaches W_{max}, determined experimentally, the next stage begins. This stage usually lasts 48-72 hours.

Vegetative Growth – This phase includes the development of the root system and shoots. During this stage, active cell division and biomass increase occur, which can be described by the Lockhart equation:

$$\frac{dL}{dt} = \frac{1}{n} \times (P - Y),$$

where L is the plant length, n is the viscosity of the cell wall, P is the turgor pressure, and Y is the threshold pressure value. The transition occurs when the plant reaches a critical length L_{crit}, for example, 20 cm. This stage can last from 7 to 21 days, depending on the plant species.

Flowering and Fruit Formation – This stage involves the emergence of buds and flowers. Growth regulators and hormones such as auxin and gibberellins are activated. The photosynthesis model is used:

$$A = \frac{V_c (C_i - \Gamma^*)}{C_i + K_c (1 + O / K_o)},$$

where A is the photosynthesis rate, V_c is the maximum carboxylation rate, C_i is the CO_2 concentration, and Γ* is the CO_2 compensation point. The phase transition occurs when the accumulated carbohydrate amount reaches the critical value C_{crit}, for example, 5 g per plant. The duration is approximately 14-28 days, depending on conditions.

Fruit Ripening and Seed Maturation – At this stage, the plant redirects resources to fruit and seed development. The biomass accumulation model is described as:

$$B(t) = B_0 + R \cdot t,$$

where B(t) is the fruit mass at time, B_0 is the initial fruit mass, and R is the biomass accumulation rate. The stage ends when the fruit mass reaches the maximum value B_{max}. Typically, this lasts from 20 to 40 days [1, 2].

These stages contribute to creating a detailed plant growth model that allows predicting phase transitions based on real-time and experimental data.

Implementation of the Model

To integrate the obtained data into practical applications, the ESP-WROOM-32 microcontroller, a display, and a real-time clock module are used. The ESP32, developed by Espressif Systems, is a system-on-chip with built-in Wi-Fi and Bluetooth modules. It is based on the Tensilica Xtensa LX6 core, which provides high computational power and advanced peripheral capabilities. Due to its affordability and extensive functionality, ESP32-based boards are widely used among developers and electronics enthusiasts [3].

The DS1302 module ensures accurate timekeeping even when power is lost due to a backup lithium battery [4]. The ESP32 microcontroller retrieves time data from the DS1302 module and determines the current stage of plant growth based on this data. All information, including the growth stage, completion percentage, and progress bar, is displayed on the screen. Time and stage data are stored in non-volatile memory, allowing recovery after a system reboot.

An example of program output is shown in Fig. 1.

Fig. 1. Example of program operation when the plant is in the vegetative growth stage

To enhance understanding of what is displayed, Fig. 2 provides data labels.

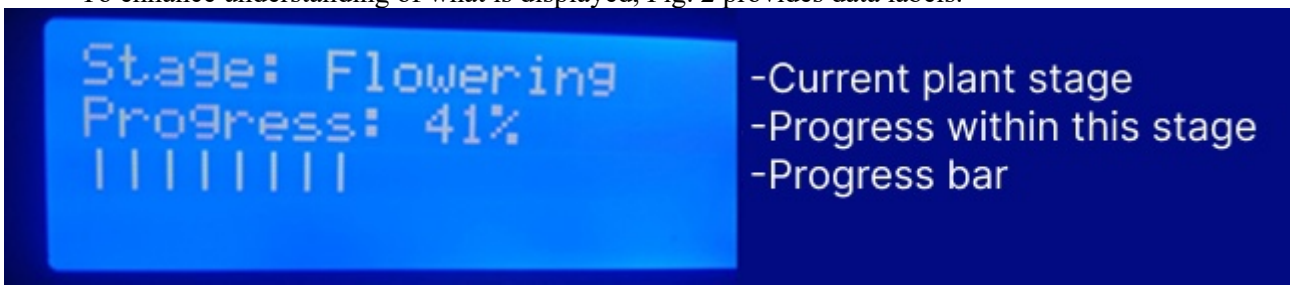


Fig. 2. Description of the data lines on the display

The examined plant growth model mathematically describes development stages and enables forecasting future trends in this process. Its practical implementation allows tracking plant growth dynamics, monitoring formation stages, and controlling growth and maturation rates.

References

1. Plant Growth Modeling. URL: https://link.springer.com/chapter/10.1007/978-3-540-71992-2_89 (дата обращения 23.09.2024)

2. Mathematical principles and models of plant growth mechanics: from cell wall dynamics to tissue morphogenesis. URL: https://academic.oup.com/jxb/article/70/14/3587/5498733?login=false (дата обращения 23.09.2024)

3. Программирование устройств на основе модуля ESP32. URL: https://habr.com/ru/companies/epam_systems/articles/522730 (дата обращения 23.09.2024)

4. Часы реального времени Ардуино DS1302 RTC. URL: https://clck.ru/3DU6Qd (дата обращения 23.09.2024)

REDUCING TRAFFIC JAMS IN HISTORIC URBAN INTERSECTIONS USING FUZZY NETWORKS AND SMART TRAFFIC MANAGEMENT: A PRACTICAL IMPLEMENTATION

Simone Giovanni Matraxia, Paolo Francesco Garrasi
Computer Engineering and Networks Laboratory – Kore University of Enna – Italy
Email: {simonegiovanni.matraxia, paolofrancesco.garrasi}@unikorestudent.it

Abstract. *In the transition toward cities of the future, one of the main objectives will be to ensure efficient and safe urban traffic management for both drivers and pedestrians, with particular attention to historic cities. The latter, characterized by infrastructure designed in eras when modern mobility was not a determining factor, present a unique challenge for traffic optimization.*
This project aims to address these challenges by developing an intelligent system based on fuzzy logic for traffic management at existing intersections. The approach involves designing a local Ethernet network to facilitate communication between semaphores and a central controller, along with a theoretical implementation in MATLAB.
The implementation was applied to the intersection in Caltanissetta, highlighting how advanced technologies can be integrated with pre-existing infrastructures. The results demonstrate the system's potential in providing adaptive and efficient traffic control, representing a step forward in the development of more sustainable and technologically advanced cities.

Introduction

Urban traffic management represents one of the main challenges for the development of modern cities, particularly in historic centers characterized by infrastructures designed in eras when road mobility was not a predominant factor. The exponential increase in the number of vehicles and the inadequacy of existing road networks contribute to phenomena such as congestion and inefficiency, making the adoption of intelligent traffic control systems essential.

In this context, fuzzy logic has emerged as a promising methodology for optimizing urban traffic management, enabling dynamic and adaptive control based on parameters such as vehicle density and queue length. This work is inspired by previous research, including that conducted by Lilin Zang, Lei Jia, and Yonggang Luo (2006) [1], which demonstrates the effectiveness of intelligent approaches for semaphore control. However, compared to these studies, the project presented here places greater emphasis on integration with existing infrastructures, utilizing a local Ethernet network for communication between the controller and the semaphores.

1. Related works

The research conducted by Wu Jianhua, Liu Ning Xing Enhong, and Zhao Xin (2010) [2] highlights the challenges associated with urban traffic management and proposes an approach based on the concept of the "green wave." However, this concept presents significant critical issues, particularly:

- Its dependence on regular and predictable vehicle flows, which are difficult to maintain in high-density urban environments;
- Its low adaptability to unforeseen situations, such as accidents, adverse weather conditions, or extraordinary events;
- Implementation difficulties in complex or non-linear infrastructures, such as historic cities.

Other studies in the literature explore fuzzy logic-based approaches for semaphore management [3], but they often focus on ideal scenarios, neglecting integration with existing infrastructures or adaptation to historic and complex urban contexts. Our system stands out by offering dynamic management based on measurable parameters, such as queue length and density, ensuring more precise and adaptive traffic control. Compared to existing commercial solutions, our approach provides greater scalability and flexibility by integrating a fuzzy controller and a local Ethernet network, ensuring better adaptation to existing infrastructures and efficient management of emergencies and nighttime operation.

2. The proposed approach

The system architecture consists of a local Ethernet network, where a central node, the controller, communicates with the semaphore nodes distributed at the intersection using a wired format. In this study, the intersection of "Via Catania" and "Via Calabria" in the city of Caltanissetta is considered, where three semaphores are installed:

Semaphore 3 (Via Calabria towards East): Considered the priority semaphore[1] due to heavy traffic flow.

Fig. 2.1. Semaphore 3

Semaphore 1 (Via Catania towards East): Allows a continuous turn towards the East.

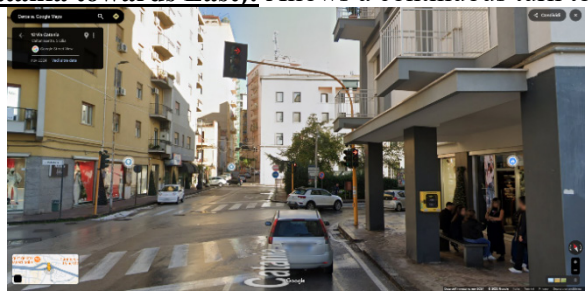


Fig. 2.2. Semaphore 1

Semaphore 2 (Via Catania towards South): Synchronized with Semaphore 1 for managing southbound traffic flows.

Fig. 2.3. Semaphore 2

The area in question represents one of the busiest intersections in the city, as it connects the main avenue and several significant institutions. This necessitates a highly dynamic and adaptive traffic management system capable of responding in real-time to varying traffic flow conditions.

a. Network architecture

Now, we delve into the implemented architecture and the algorithms adopted.

The network enables Ethernet communication between the controller and the semaphores, with a transmission speed of 2.5 Mb/s. The central node (controller) is responsible for coordinating and monitoring traffic. It periodically queries the fuzzy model to calculate the optimal semaphore light timing based on data related to traffic density and queue length.

In a practical implementation, these values would be detected by sensors distributed along the lanes. In the theoretical implementation, however, traffic and density data are artificially generated through simulations[2], as described in the MATLAB/Simulink model.

The wired network approach was chosen to ensure centralized control and fast, secure communication between the controller and the semaphores, minimizing risks associated with wireless communication reliability.

b. Semaphores operation

The semaphores are designed to operate in a basic mode, with a default logic that sets them to flashing yellow in the absence of communication from the controller. This state ensures the safety of vehicles and pedestrians in case of communication failure between the controller and the semaphores. The flashing yellow light signals drivers to proceed with caution, reducing the risk of accidents.

When communication is active, the controller calculates the optimal green light duration based on received traffic data. The semaphores then dynamically adjust their activation times, ensuring an optimal distribution of traffic flow in different directions.

c. Traffic management with the fuzzy model

The core of the control system lies in the use of fuzzy logic. The controller utilizes the fuzzy model to determine the duration of green lights based on measurable parameters such as queue length and traffic density. The fuzzy control algorithm can adapt to real-time traffic conditions, adjusting green light durations as needed.

• Adaptive traffic control: The controller can assign longer green light times to roads with high traffic density, reducing vehicle wait times and optimizing flow distribution.

• Priority bias: Semaphore timing can be influenced by priority bias, allowing certain roads to be favored over others. This is the case with Semaphore 3 (Via Calabria towards East), which has been prioritized to ensure efficient traffic management during peak hours.

d. Special scenarios and emergencies

The architecture also includes special management for specific scenarios, such as nighttime hours or emergency situations. During low-traffic periods (e.g., at night), the semaphores can automatically switch to flashing yellow, reducing the need for manual intervention. In emergencies (e.g., the passage of law enforcement or emergency vehicles), the system can be remotely controlled, allowing temporary suspension of traffic control to facilitate the safe passage of emergency vehicles.

References in this chapter:

[1]: *The road leading to Semaphore 3 crosses the city's main avenue.*

[2]: *The sensor data consists of randomly generated numbers at fixed intervals.*

3. Scenario

The analyzed scenario includes, in addition to the Simulink model, two main MATLAB scripts: *controllore_code.m* and *S_code.m*, which serve as the core of the implemented logic and are essential for the overall system operation. These scripts correspond to periodic tasks created by the controller and the semaphores immediately after initialization. The initialization process is handled by dedicated scripts for each associated kernel: *controllore_init.m*, *S1_init.m, S2_init.m, and S3_init.m.*

The following section will delve into the model, starting with the Simulink file *crossroad.slx*.

a. Simulink model

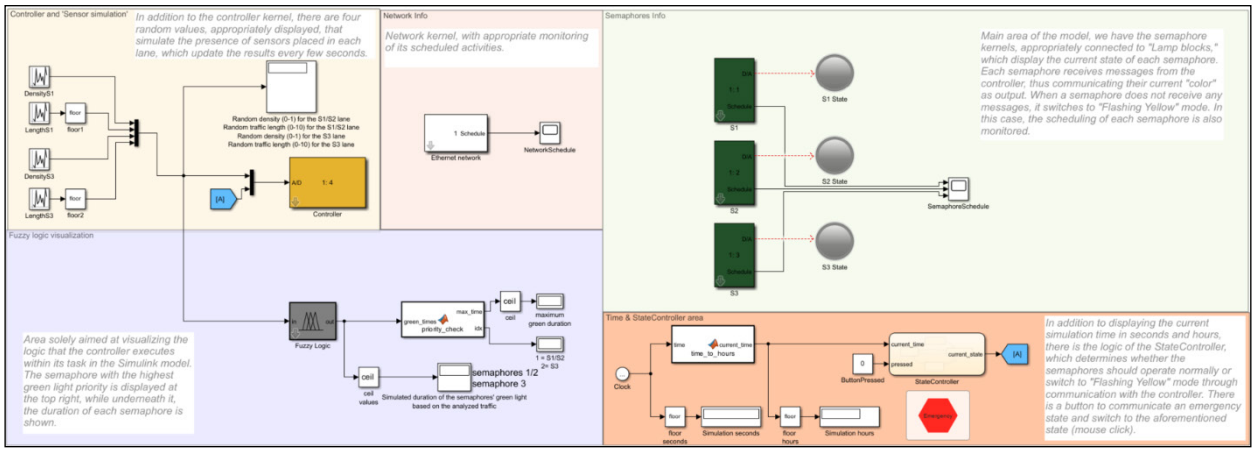


Fig. 3.1. File structure of the scenario

The model is divided into clearly commented sections, reflecting its segmentation into five main subsystems. This structure simplifies model interpretation, especially for external users. Therefore, further subdivision details are unnecessary, allowing the focus to remain on the applied program logic.

The simulation starts at 12:00 PM (43,200th simulation second), as defined in the model properties. This time reference is crucial for initializing the control system, as it marks the point from which traffic management logic is applied and system state transitions occur.

b. State controller

The State Controller manages the system's state logic, implemented through MBSD (Model-Based State Diagram). It regulates the transition between the semaphores' two main states: *Normal* and *Flashing Yellow*. This transition depends on both the simulation time and the detection of an emergency state.

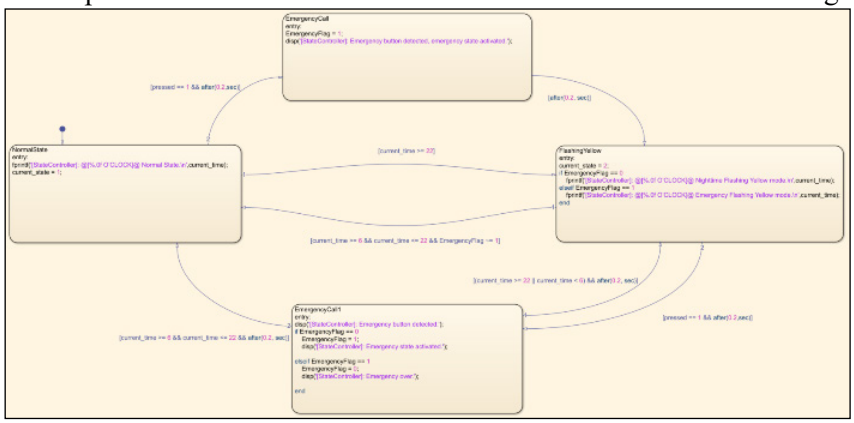


Fig. 3.2. States logic

c. Controller

The controller kernel receives five analog input signals, consisting of two simulated inputs for each lane (in this example, there are two lanes) and the output from the State Controller.

- Normal State: During the normal state, the controller reads real-time analog values from the sensors and feeds them into the fuzzy logic system, which generates two main outputs:
  - The green light duration for lane S1/S2;
  - The green light duration for lane S3.

Semaphore synchronization is managed based on these computed values: the semaphore with the longest duration initiates the cycle by activating the green light and progressively decrementing the remaining time. The last 5 seconds of the green phase are indicated by a transition to yellow. Once the time reaches zero, the semaphore turns red, and the controller assigns the turn to the next semaphore.

- Flashing Yellow State: In this condition, the controller halts all operational activities and waits for an update from the State Controller.

4. Performance evaluation

To evaluate the effectiveness of the proposed system, simulations were conducted in MATLAB/Simulink, analyzing three key metrics: Ethernet network reliability, average waiting times per lane, and the impact of priority assigned to semaphore S3. Data was collected using dedicated scripts (*controllore_code.m*, *S_code.m*) and processed with *performance_metrics.m*. The following section presents the results and their interpretation.

```
>> performance_metrics
--------------------------------------------------------------------------------
Total simulated hours: 10.00
Total messages sent: 107994
Total messages received: 107895
Lost messages: 99
Percentage of lost messages: 0.09%

Total S1/S2 waiting time: 327.43 minutes
Total S3 waiting time (priority lane): 276.32 minutes
Priority impact: S3 has 15.61% less waiting time compared to S1/S2
Average green light time calculated for S1/S2: 63.80 seconds
Average green light time calculated for S3 (priority lane): 75.18 seconds
--------------------------------------------------------------------------------
```

Fig. 4.1. Simulation results

The communication between the central controller and the semaphores was tested using an Ethernet network with a transmission speed of 2.5 Mb/s. Over a 10-hour simulation, 107,994 messages were sent, with 107,895 successfully received, resulting in 99 lost messages.

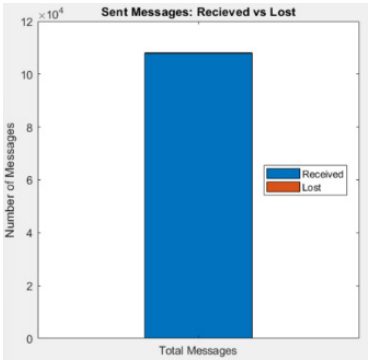


Fig. 4.2. Sent vs. lost packets

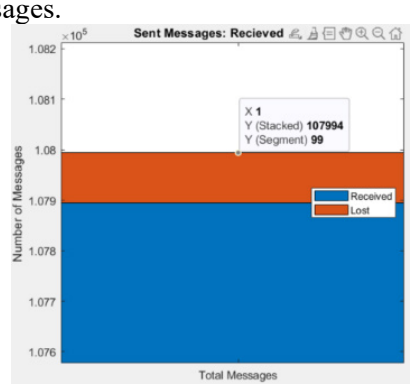


Fig. 4.3. Focus on lost packets

The wired architecture ensures stable and reliable communication. A 0.09% packet loss is negligible, demonstrating that the system can handle high-frequency communication without affecting performance. This is crucial for real-world applications, where network disruptions could compromise semaphore coordination.

The fuzzy controller dynamically calculated green light durations for lanes S1/S2 and S3, adapting to simulated traffic conditions. The results show:

- Average green time for S3 (priority lane): 75.18 seconds
- Average green time for S1/S2: 63.80 seconds

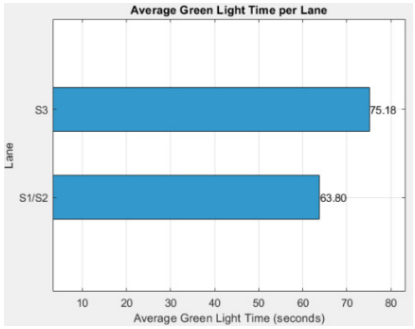


Fig. 4.4. Average green light duration per lane

The fuzzy logic assigns 17.83% more green time to S3 compared to S1/S2. This demonstrates the effectiveness of the priority bias integrated into the model, optimizing traffic flow on the priority road (Via Calabria) and reducing wait times for high-density traffic.

The simulation has shown a significant total wait time reduction for S3, compared to S1/S2:

- Total waiting time for S1/S2: 327.43 minutes
- Total waiting time for S3 (priority lane): 276.32 minutes
- Reduction of waiting time for S3: 15,61%

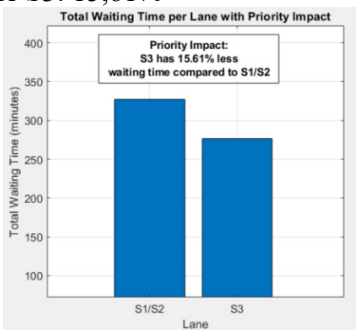


Fig. 4.5. Average green light duration per lane

The priority assigned to S3 substantially decreased the waiting time for the priority lane. This highlights the system's ability to manage traffic asymmetries, a crucial feature in historic cities with nonlinear road infrastructures.

Conclusions

The project demonstrates how a fuzzy logic-based controller can optimize traffic management at intersections, addressing the challenges of historic cities with outdated infrastructures. Unlike traditional methods, this system adapts dynamically by adjusting semaphore timings based on measurable parameters like traffic density and queue length. Using MATLAB and Simulink, the model shows how theoretical simulation and adaptive control can improve traffic flow, enhancing road safety and reducing travel times. In the future, the system could integrate real sensors for more precise management, expand to a network of intersections for urban optimization, and leverage technologies like 5G and IoT. Additionally, adapting to communicate with connected vehicles and analyzing environmental impacts could enhance urban sustainability.

This work sets the stage for smarter, more efficient traffic systems that meet the needs of modern cities and promote a sustainable future.

References

[1] Lilin Zang, Lei Jia and Yonggang Luo, "An Intelligent Control Method for Urban Traffic Signal Based on Fuzzy Neural Network," 2006 6th World Congress on Intelligent Control and Automation, Dalian, China, 2006, pp. 3430-3434, doi: 10.1109/WCICA.2006.1713005. keywords: {Intelligent control;Communication system traffic control;Fuzzy neural networks;Fuzzy control;Traffic control;Control systems;Fuzzy logic;Hybrid intelligent systems;Fuzzy reasoning;Artificial neural networks;Traffic signal control;Fuzzy neural network;optimization;simulation},

[2] Wu Jianhua, Liu Ning, Xing Enhong and Zhao Xin, "Optimal design for traffic light and simulation based on MATLAB," 2010 International Conference on Mechanic Automation and Control Engineering, Wuhan, 2010, pp. 2854-2857, doi: 10.1109/MACE.2010.5536285. keywords: {Traffic control;MATLAB;Cities and towns;Roads;Vehicles;Investments;Lighting control;Constraint optimization;Optimal control;Mechanical engineering;Traffic light;Optimal design;Simulation},

[3] Arena Fabio, Pau Giovanni, Ralescu Anca, Severino Alessandro, You Ilsun, An Innovative Framework for Dynamic Traffic Lights Management Based on the Combined Use of Fuzzy Logic and Several Network Architectures, Journal of Advanced Transportation, 2022, 1383349, 17 pages, 2022.

RECOGNITION OF OBJECTS IN A POINT CLOUD USING DEEP LEARNING

Aleksandr Matveev
Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia
E-mail: am6400@yandex.ru

Abstract. *This article presents a methodology for object recognition in a three-dimensional point cloud obtained from LiDAR systems. The proposed approach is based on deep learning and is combined with point cloud processing algorithms (segmentation, metric calculation, clustering). The text provides a detailed description of data preparation processes, the creation of a digital elevation model, watershed segmentation, as well as the final classification of objects into trees, shrubs, and grass. An analysis of the various processing stages is also presented.*
Keywords: *point cloud, LiDAR, deep learning, segmentation, watershed, clustering, Canopy Height Model, K-Means.*

Introduction

The relevance of object recognition tasks in point clouds is driven by the widespread use of remote sensing and the development of LiDAR technologies. LiDAR scanning provides highly accurate three-dimensional data on terrain, vegetation, and buildings, enabling the automatic analysis of the acquired point clouds. The segmentation and classification of objects—particularly trees and shrubs—plays an important role in the monitoring of forests, agro-landscapes, and urban green spaces [1-5].

Deep learning has proven to be an effective tool for analyzing complex multidimensional data, including images and three-dimensional models. However, working with LiDAR data presents several challenges: large point volumes, variability in cloud density, and noise introduced by measurement equipment. Therefore, proper preprocessing is required (filtering, normalization, interpolation), which, combined with deep learning approaches, makes it possible to reliably solve the task of recognizing and segmenting objects [6-10].

Preliminary data processing

A point cloud in *.laz format was used in this study, containing the coordinates X, Y, Z, and the intensity of the reflected signal. The initial number of points was around 70 thousand. The preliminary data processing consists of the following steps:

1. Cleaning data from outliers

Points lying outside reasonable limits of height and coordinates (for instance, excessively high Z values or obviously erroneous X, Y coordinates) are removed. If duplicates are found (points with identical coordinates), only a single instance is retained.

2. Downsampling

To reduce computational costs and decrease noise in the point cloud, every n-th point is sampled. In practice, this operation offers significant resource savings in subsequent analysis stages. For example, the original data contained 70 thousand points, while after downsampling, there were 14 thousand (Figure 1 (a)).

3. Interpolation and creation of a regular grid

Using the griddata function (from the scipy.interpolate library), the unstructured point cloud is transformed into a uniform three-dimensional grid (Figure 1 (b)). At this stage, it is possible to obtain 10 thousand regular grid nodes, each storing coordinates and reflected signal intensity. This approach simplifies subsequent operations since each grid cell has a fixed position.

4. Ground segmentation

Based on a slope threshold and an analysis of nearby points, a terrain model is formed. Points belonging to the ground are identified and, if necessary, excluded from further analysis, for instance, if the focus is on above-ground vegetation.

5. Height normalization

After determining the terrain, the Z values are converted to above-ground coordinates by subtracting the ground surface height at the corresponding X, Y coordinates. This operation clearly separates the ground and above-ground parts (Figure 2). Violet coloring in the visualization typically indicates lower elevations,

while green indicates higher ones. After cleaning and filtering, about 8,247 points remain, corresponding to objects above ground level.

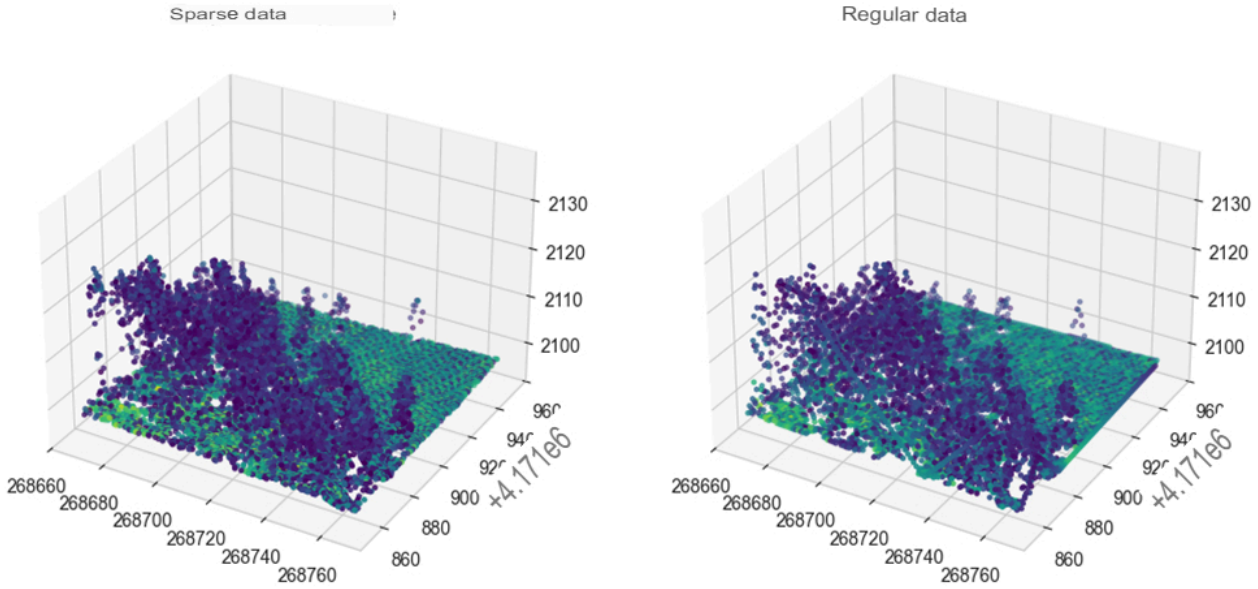


Fig. 1. (a) Downsampled data (left) and (b) Regularly gridded data (right)

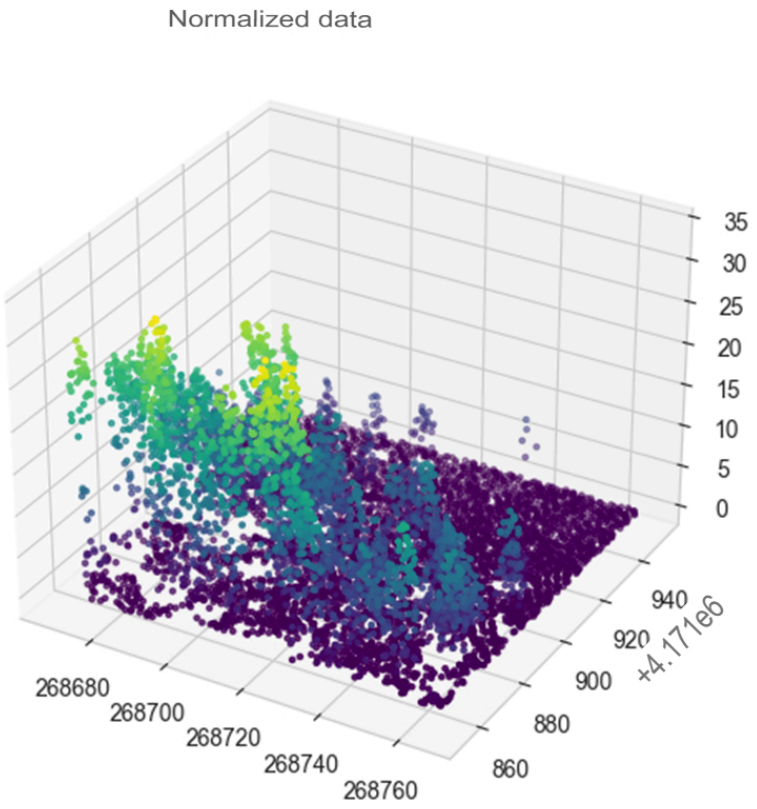


Fig. 2. Normalized data after terrain removal

6. Building the Canopy Height Model (CHM)

On the regular grid, each cell is assigned the maximum (or average) normalized height, forming the Canopy Height Model (CHM) (Figure 3). The CHM reflects the upper boundary of vegetation above the surface and significantly simplifies the subsequent tree apex detection and segmentation procedures.

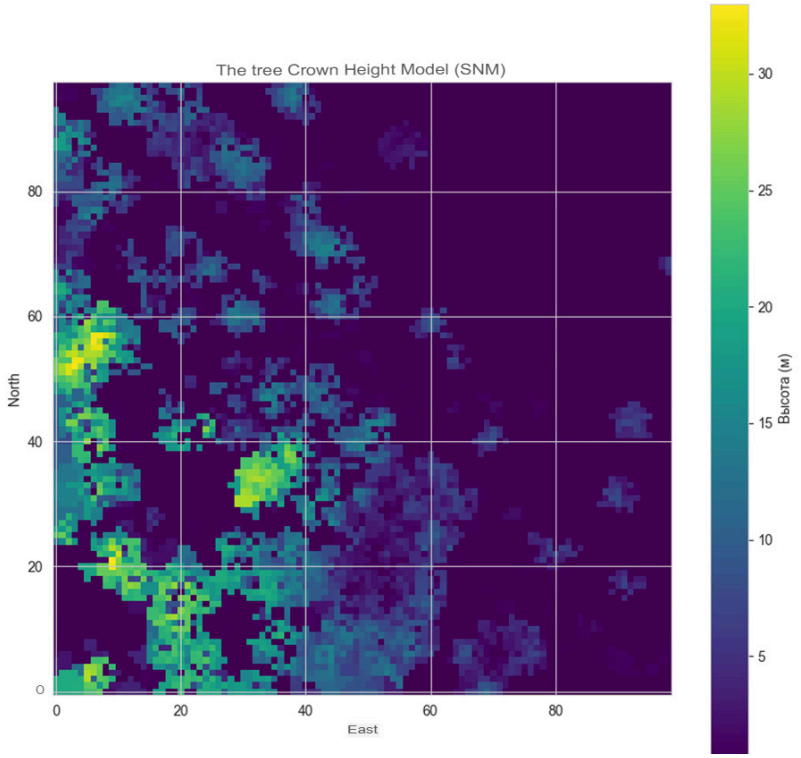


Fig. 3. Tree canopy height model (CHM)

Crown segmentation

1. Tree Top Detection

Figure 4 shows tree tops (Tree Tops) marked with red dots against the CHM background. The algorithm accurately identifies major tree tops even with partial crown overlap. In the implementation, the peak_local_max function from the scikit-image library is used to search for local maxima on the two-dimensional map (CHM).

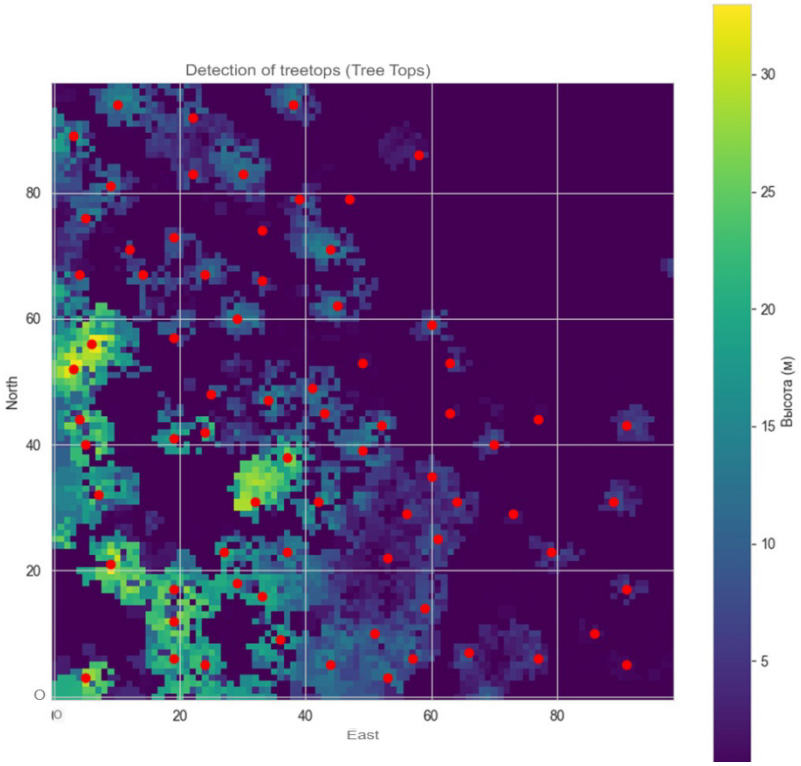


Fig. 4. Detection of tree tops on the CHM

2. Segmentation using the Watershed Method

To divide the CHM into distinct regions, the watershed segmentation method was employed. Starting from local maxima, “filling” proceeds into neighboring areas of lower height. Areas where the fillings from different peaks meet form the boundaries between segments. Figure 5 demonstrates the result of segmentation using the watershed method: each colored region corresponds to a separate object (a tree or group of points).

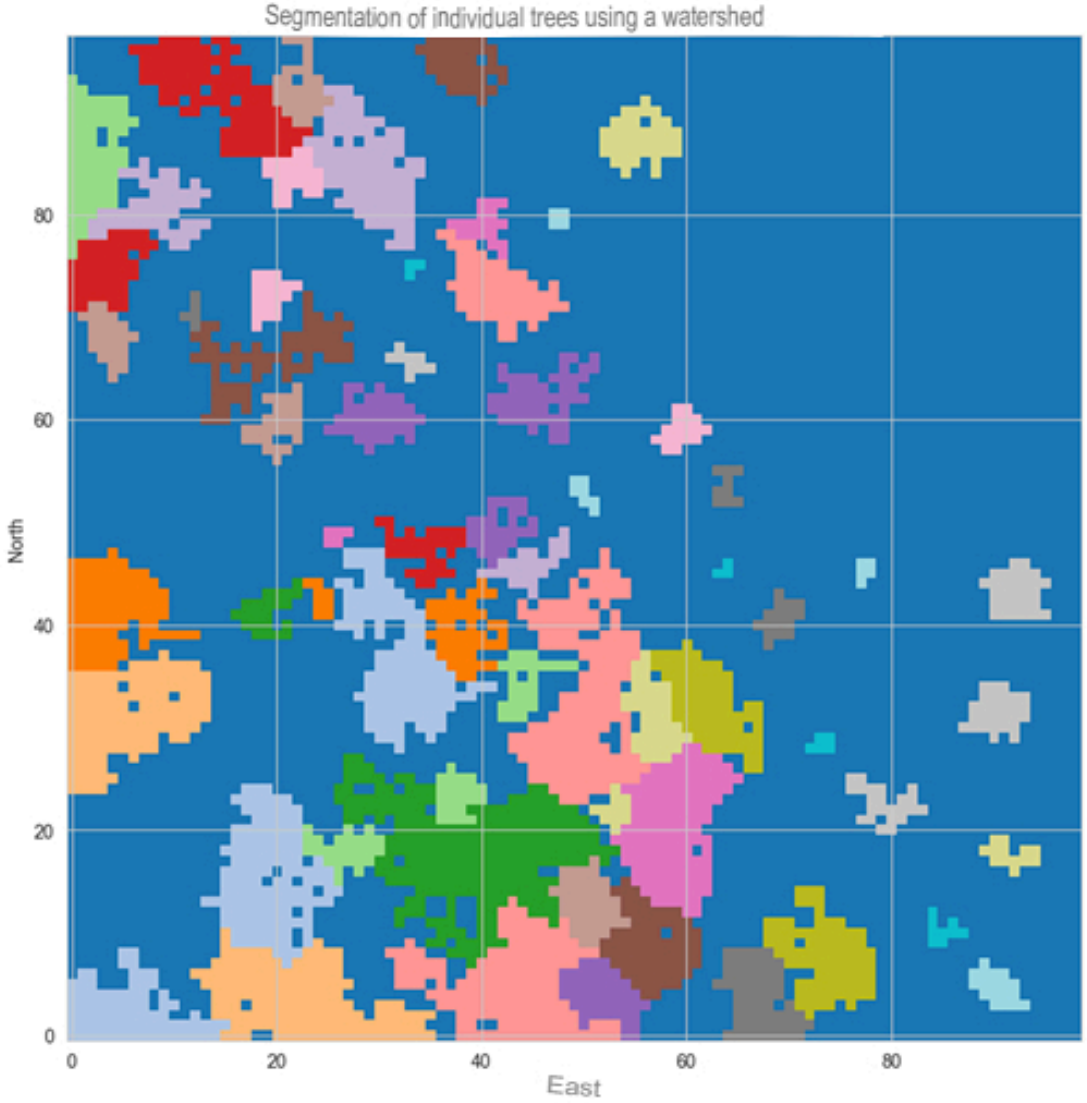


Fig. 5. Segmentation of individual trees using the watershed method

Clustering of objects

After dividing the scene into segments, the task becomes classifying each segment into one of the possible object types: “tree,” “shrub/low tree,” or “grass.” In this work, an approach based on extracting height-based features was applied:

1. Feature Extraction for Segments

For each segment, metrics such as maximum and average height, area (number of pixels), boundary shape, and others are calculated.

2. Scaling and Normalization

To prevent one feature (e.g., height) from dominating, weighting coefficients are introduced. Height is multiplied by 1.25, and crown area by 0.75.

3. K-Means Algorithm (3 clusters)

The centroid with the highest height value is interpreted as “trees,” the lowest as “grass,” and the intermediate one as “shrubs/low trees.” The resulting segments are then displayed on a two-dimensional map (Figure 6). Black areas correspond to empty sections or places not assigned to any segment.

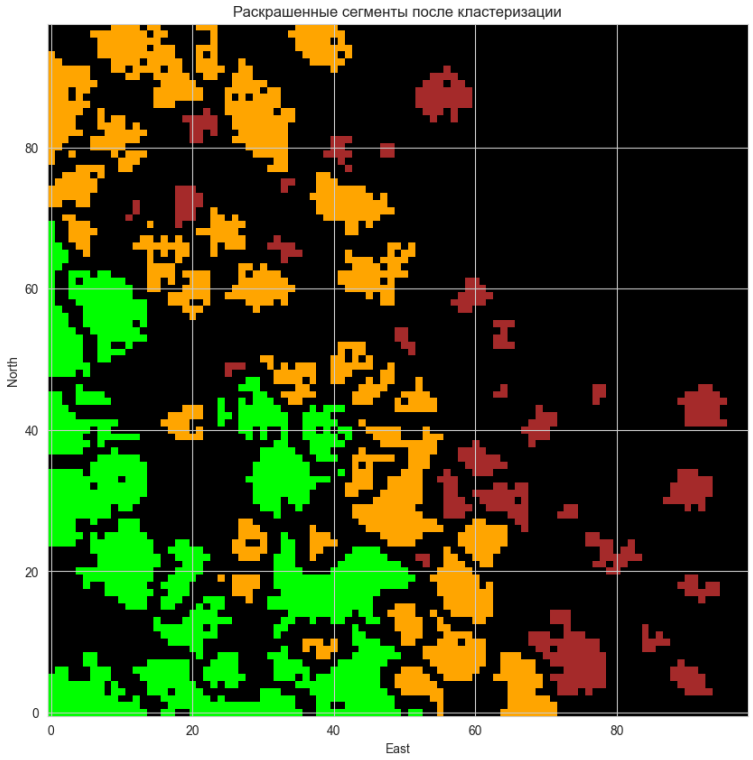


Fig. 6. Clustering of objects (trees, shrubs, and grass)

As a result, 74 objects were correctly identified, of which 16 were trees, 31 shrubs/low trees, and 27 fragments of grass cover.

3D visualization of results

For a clearer representation of the scene's structure, a three-dimensional height map is generated with the clustering results overlaid. Figure 7 shows an example of such a 3D map: the tall "peaks" correspond to the canopies of mature trees (yellow/orange), whereas purple/blue regions indicate low vegetation.

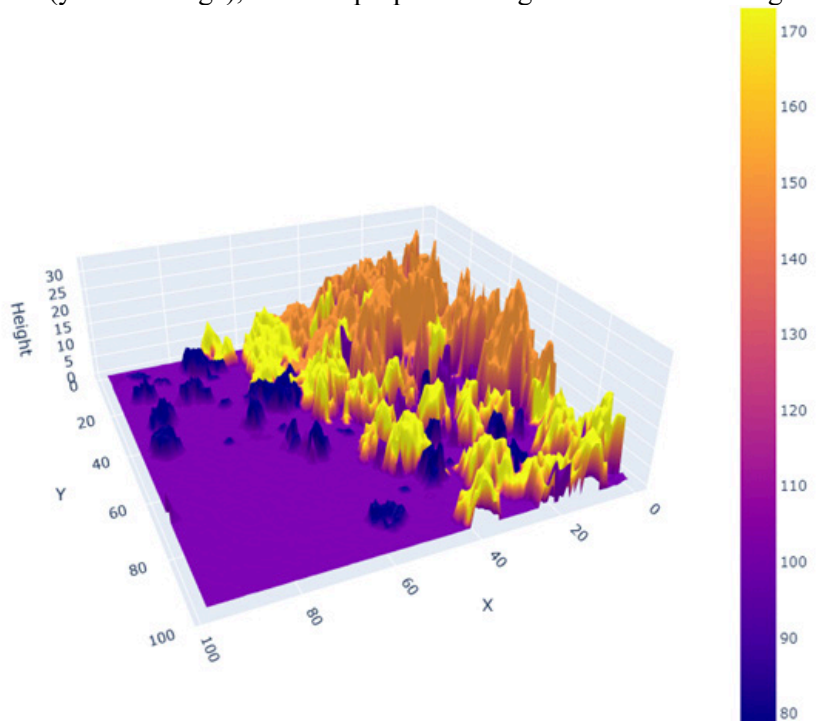


Fig. 7. 3D height map with colored clusters

Conclusion

This article demonstrates a comprehensive process for recognizing tree, shrub, and grass vegetation based on a point cloud. It is shown that combining classic preprocessing methods with deep learning approaches yields high accuracy in automatically recognizing various types of vegetation and simplifies the counting of objects.

Thus, the use of deep learning methods together with classical data processing algorithms ensures effective and noise-resistant segmentation and classification of a point cloud, opening up new possibilities for resource monitoring, territory management, forest management, and related areas.

References

1. Goshin, Ye.V. Segmentation of stereo images with the use of the 3D Hough transform / Ye. V. Goshin, G. E. Loshkareva // CEUR Workshop Proceedings. Vol. 1638 2016. pp. 340–347.

2. Berezhnoy A.A. Point Cloud Processing Using Deep Learning Methods. – Moscow: MSTU Publishing, 2020. 120 p.

3. Razlatsky, S.A. Application of the Hough Transform for Object Detection in a Three-Dimensional Scene / S.A. Razlatsky, P.Yu. Yakimov // Information Technologies and Nanotechnologies (ITNT-2015). -2015.

4. Gribkov I.N., Kuznetsov A.V. Deep Learning Methods for Object Recognition. – Yekaterinburg: UrFU, 2021. 150 p.

5. Gupta, S. Learning rich features from RGB-D images for object detection and segmentation / S. Gupta, R. Girshick, P. Arbelaez, and J. Malik // ECCV, 2014.

6. Efremov V.P. Point Cloud Processing Algorithms. – St. Petersburg: Nauka, 2019. – 200 p.

7. Smirnov, I.P. Foundations of Deep Learning // Electronic Technologies. 2021. Vol. 21, No. 4, pp. 15–22.

8. Fedorov, A.V. Deep Learning Models for 3D Data // Informatics and Education. 2022. Vol. 34, No. 1. pp. 99–105.

9. Frome, A. Recognizing objects in range data using regional point descriptors / A. Frome, D. Huber, and R. Kolluri // ECCV. – 2004. – Vol. 1. – pp. 1–14.

10. Maturana, D. VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition / D. Maturana, S. Scherer // IEEE/RSJ International Conference on Intelligent Robots and Systems. – 2015.

DEVELOPMENT OF A HARDWARE-SOFTWARE SYSTEM USING A SINGLE-BOARD COMPUTER AND NEURAL NETWORK MODELS FOR THE TASK OF EARTH SURFACE OBJECT RECOGNITION

Timofey Pisklenov
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russian Federation
E-mail: tim.kirp@mail.ru

Abstract. *This paper proposes a hardware-software system designed for Earth surface monitoring tasks, based on a single-board computer and electronic components integrated with computer vision techniques. A hardware-software object recognition system based on a single-board computer has been developed. The results of the developed system for recognizing haystacks using the neural network model YOLOv8n are demonstrated.*
Keywords: *computer vision, single-board computers, onboard aviation systems, neural network models, object recognition.*

Introduction

A distinctive feature of the current stage of information-technological development is the active integration of artificial intelligence (AI) systems into onboard aviation complexes aimed at increasing their functionality and safety. Effective real-time data processing solutions require a comprehensive approach that includes optimization of software algorithms and selection of hardware platforms balancing performance and energy efficiency, accuracy, and frame rate. Single-board computers equipped with CPUs featuring multi-threaded computing capabilities enable inference of neural network models, such as convolutional neural networks (CNNs), through efficient computational load distribution. Additional improvement in model accuracy is achieved through data augmentation techniques involving image transformations. This approach allows machine learning algorithms to be adapted to operate under conditions of limited computational resources and diverse observation scenarios.

Hardware-Software Object Recognition System

Developing the hardware configuration for detecting haystacks is a complex process requiring careful selection, integration, and optimization of components to achieve a balance between performance, energy efficiency, and functionality. The hardware-software object recognition system (Fig. 1) consists of: Single-board computer Raspberry Pi 4 Model B (8 GB RAM), Raspberry Camera Module 3 for image capture, Power supply and cooling unit, Wi-Fi wireless communication module, External memory and Software code performing object recognition tasks.

The Raspberry Pi Camera Module V3 with CSI (Camera Serial Interface) [1] is used for capturing frame streams. This camera supports resolutions up to 8 MP at 30 frames per second, ensuring high image quality with low power consumption. It connects directly to the Raspberry Pi 4 Model B (8 GB RAM), simplifying integration and reducing data transmission latency.

The software configuration includes the 64-bit Raspberry Pi OS (formerly Raspbian), based on Debian and optimized for single-board computers. For computer vision tasks, OpenCV and TensorFlow Lite frameworks are utilized. OpenCV provides basic image processing functions such as video stream reading [2], data preprocessing (e.g., normalization, resizing), and result visualization. TensorFlow Lite enables inference execution of the YOLOv8n model on resource-constrained devices with high performance and low power consumption.

The YOLOv8n model runs within a Docker container, ensuring runtime environment isolation and simplifying system deployment. The Docker container includes all necessary dependencies such as Python, TensorFlow Lite, OpenCV, and other libraries, facilitating rapid deployment on new devices. Post-training quantization methods were applied to optimize the YOLOv8n model by reducing its size and accelerating its execution on Raspberry Pi 4 Model B (8 GB RAM). Quantization to 8-bit integers (INT8) reduces model size and accelerates inference while maintaining low energy consumption.

System performance could be significantly improved by connecting external tensor processors (TPs), such as the "Hailo-8" [3]. Tensor processors can substantially accelerate machine learning operations like convolutions and matrix multiplications—core computational steps in YOLOv8n models. The TP supports 8-bit quantization, reducing model size further and speeding up execution on Raspberry Pi 4 Model B (8 GB

RAM). Using a TP could increase processing speed up to 20 frames per second—critical for real-time operation.

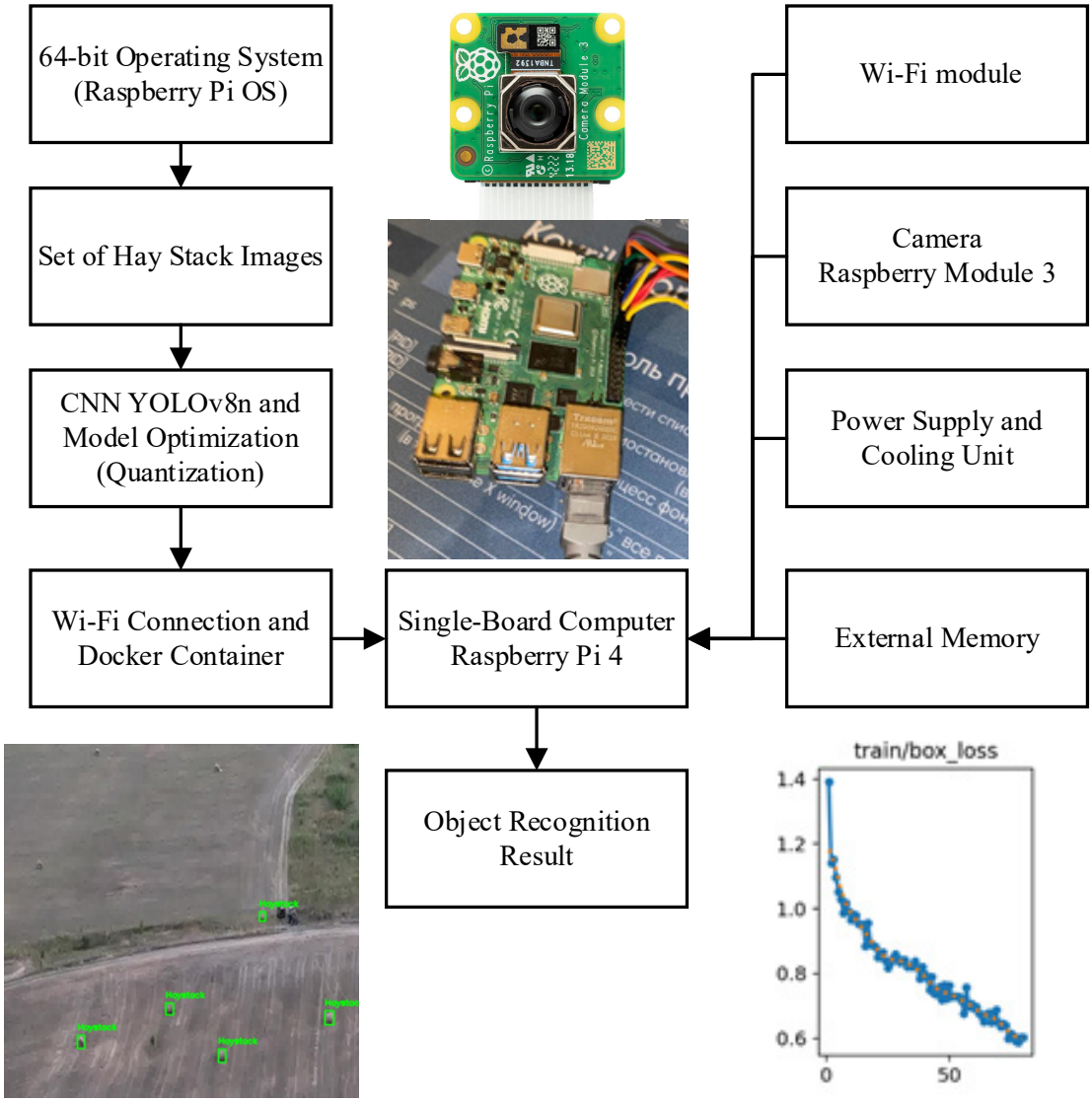


Fig. 1. System integration diagram

Thus, the current configuration based on Raspberry Pi 4 Model B (8 GB RAM) provides sufficient performance for haystack monitoring tasks but has potential for improvement through TP integration. This enhancement would significantly increase data processing speed, reduce power consumption, and ensure real-time operation.

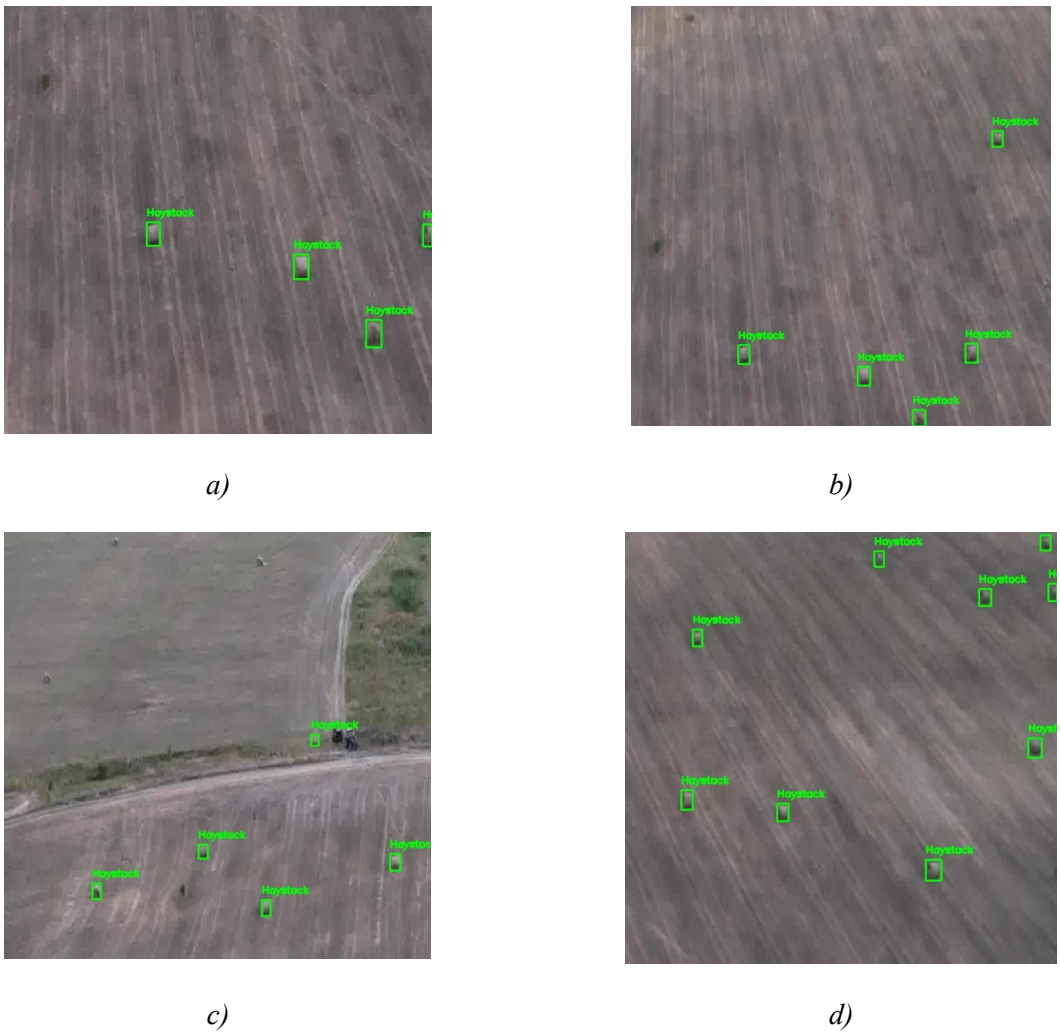
Results from Testing the Developed System

During testing on the developed system, the results (Fig. 2) demonstrate that the proposed haystack recognition system using the YOLOv8n neural network model on the Raspberry Pi 4 Model B platform (8 GB RAM) performs successfully.

The system effectively recognizes haystacks in images captured by the camera, exhibiting high accuracy and robustness to varying lighting conditions and scales. The optimized model ensures stable real-time operation, confirming its suitability for agricultural monitoring tasks.

During the training of the YOLOv8n model for haystack detection, graphs reflecting key metrics and losses at each training stage were obtained. These graphs allow assessing the model's effectiveness and identifying potential issues during training.

Analysis of the graphs indicates that the YOLOv8n model successfully trains and consistently improves its performance throughout the training process. The decreasing losses on both training and validation datasets demonstrate stable learning without overfitting.



a) *b)*

c) *d)*

Fig. 2. Results of the haystack recognition model

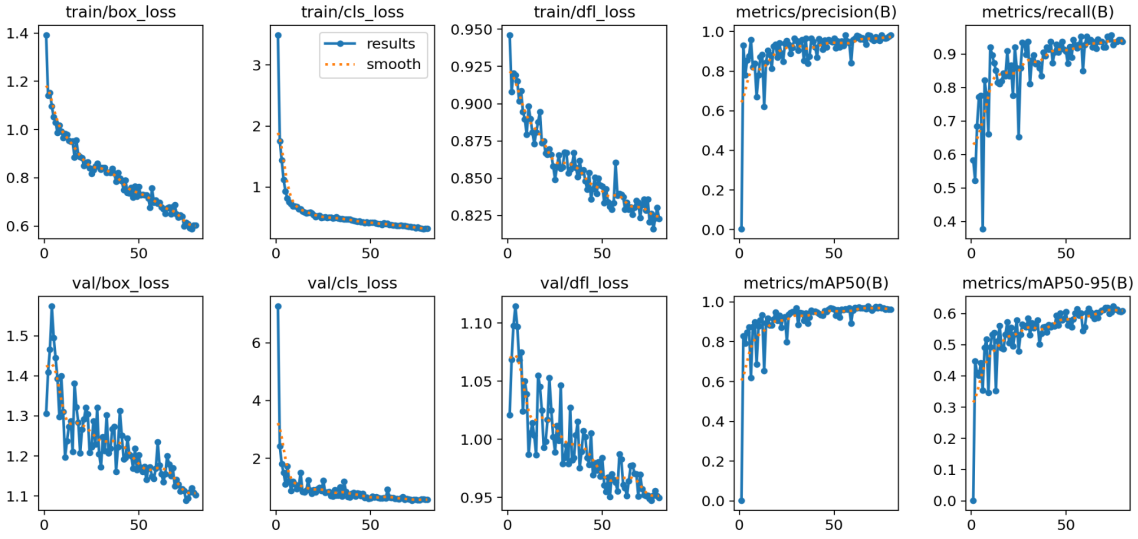


Fig. 3. Evaluation of the developed neural network model

The graphs demonstrate metrics of mAP-50 = 0.875, mAP50-95 = 0.65, with a model size of 8 MB and computational load (FLOPs = 8.7B). These results confirm that the model achieves high accuracy in localization and classification, making it suitable for real-time monitoring tasks.

Conclusions

Thus, this work presents a hardware-software system based on a single-board computer utilizing the YOLOv8 neural network model for object recognition tasks. The effectiveness of the developed system was demonstrated through successful haystack detection using YOLOv8n neural network model implementation.

References

1. Flat-Field and Colour Correction for the Raspberry Pi Camera Module / R. W. Bowman, B. Vodenicharski, J. T. Collins, Ju. Stirling // Journal of Open Hardware. – 2020. – Vol. 4, No. 1. – DOI 10.5334/joh.20.

2. Nenashev, V. A.; Voronov, R. M.; Berezin, A. V. [et al.] Ground Object Recognition in Aviation Systems for Optical Frame Stream Processing // Information-Measuring and Control Systems. – 2024. – Vol. 22, No. 4. – P. 85-90. – DOI 10.18127/j20700814-202404-10.

3. Pavlenko, D. A. Investigation of the Efficiency of the Budget Tensor Processor Google Coral for Solving Practical Image Classification Tasks / D. A. Pavlenko, E. V. Snezko, V. A. Kovalev // Applied Artificial Intelligence: Prospects and Risks: Proceedings of the International Scientific Conference, Saint Petersburg, October 17, 2024. – Saint Petersburg: Saint Petersburg State University of Aerospace Instrumentation, 2024. – P. 82-90.

4. Lebedev, M. S. Implementation of Artificial Neural Networks on FPGA Using Open Tools / M. S. Lebedev, P. N. Beletsky // Proceedings of the Institute for System Programming of the Russian Academy of Sciences. – 2021. – Vol. 33, No. 6. – P. 175-192. – DOI 10.15514/ISPRAS-2021-33(6)-12.

INVESTIGATION OF A METHOD FOR DETECTING UNMANNED AIRCRAFT LANDING SPOTS USING A POINT CLOUD

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

Abstract. *This paper considers a method of processing a point cloud obtained by an airborne laser scanner to determine safe landing zones for an unmanned aircraft. Point clustering and classification algorithms are applied. The result of the study is a point cloud processing algorithm that divides the point cloud into zones labeled with the probability of a successful landing.*
Keywords: *point cloud, lidar, airborne laser scanning, point cloud clustering, point cloud classification.*

Introduction

Automation of processes of search and identification of safe landing zones for unmanned aircrafts (UAVs) is one of the topical tasks in modern unmanned aviation systems. The development of airborne laser scanning (lidar) technologies allows to obtain accurate spatial data in the form of point clouds, which can be used to analyze surface characteristics and detect suitable landing sites.

A point cloud, which is a three-dimensional set of coordinates, allows to reconstruct the terrain in detail and highlight its features. However, processing of such data requires complex analysis methods including clustering and classification of points. Clustering allows to group points based on their proximity and homogeneity, while classification allows to identify geometric surface characteristics such as slope, flatness and obstacles.

Clustering and classification of point cloud

Clustering is a method of dividing a set of points into groups (clusters) based on their spatial position or other attributes. Clustering is used to select areas with homogeneous characteristics in the solution of the landing site detection problem [1].

Density Based Spatial Clustering with Noise (DBSCAN) – clusters points based on their density. Clusters are formed from points that are a small distance apart. In the left part of Fig. 1, three clusters (shown in the right part of Fig. 1) are visually distinguished. As a rule, points that do not belong to these clusters (seven white circles) are considered as noise.

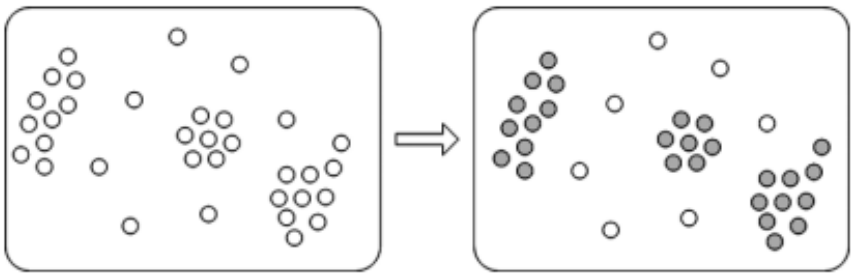


Fig. 1. Clusters as they appear as groups of points

The advantages of this algorithm are its robustness to noise and anomalous data, as well as its ability to select clusters in arbitrary shapes, as can be seen in Fig. 1. At the same time, it has a disadvantage, which is the need to adjust the parameters of the search radius and the minimum number of points to form a cluster.

After clustering is performed for each selected cluster, points are classified in order to assess the surface characteristics. This stage is aimed at determining the criteria that will allow to establish the suitability of the area for landing of UAVs. The method of plane approximation using the least squares method is used for surface estimation. Based on the obtained surface parameters, such characteristics as surface slope and flatness, height anomalies and obstacles are estimated. Surface slope is determined by calculating the angle between the normal and the vertical axis. Areas with a slight slope are considered

potentially suitable for landing. Flatness, on the other hand, is considered as the deviation of a point from the approximation of the plane to assess uniformity. If the deviation is minimal, it is indicative of high surface flatness. Height anomalies and obstructions show detected points that deviate significantly from the surface and may indicate the presence of obstructions such as vegetation, rocks or man-made objects [2-4].

Determination of criteria for the suitability of an area for planting

Based on the results of clustering and classification, the criteria determining the suitability of the area for landing of a UAV are formulated. The surface is considered flat if the deviation of points from the approximation of the plane does not exceed the specified threshold, and the slope is less than 5 degrees relative to the horizontal plane. In addition, the area shall have minimum dimensions corresponding to the dimensions of the unmanned aircraft. If all conditions are met, the area is considered suitable for landing.

The point cloud is divided into two categories of points: terrestrial and non-terrestrial. Non-terrestrial points are automatically marked as dangerous points, as they may pose a threat to landing, the procedure allows you to select and mark dangerous points corresponding to non-terrestrial objects. Visualization of the result gives a visualization of the distribution of hazardous points in the cloud. After all checks are completed, each point receives one of the following labels presented in Table 1.

Table 1

Initial label names

Dangerous	not ground or heavily sloped points.
Unsuitable	terrestrial but with significant obstacles or slope.
Risky	points that are close to, but not directly dangerous or inappropriate.
Suitable	level ground points with no significant obstructions.

Further data processing is carried out with classification of unlabeled points using a block overlap based processing method. This method includes calculation of nearest neighbors for each unlabeled point. If the number of neighbors is less than the set threshold or if there are dangerous points among the neighbors, the point is marked as unsuitable. Otherwise, the indexes of the neighboring points are stored in the cell array, and the attributes of each unlabeled point and its neighbors, described in Table 2, are used for risk assessment.

Table 2

Attributes of risk point classification

Vertical variance	is the dispersion of points along the z-axis. A high dispersion value indicates a significant variation in altitude between points, making the area less suitable for landing.
Relief	is the difference in elevation between the lowest and highest point in the neighborhood. A low terrain value indicates a flatter surface and fewer obstacles in the landing zone.
Slope	is the angle of inclination of the landing plane. A small slope is preferred for landing as it indicates a more stable surface for the vehicle.
Residual	is the roughness of the landing area. A low residual means fewer obstacles and a smoother landing area.

Labeling is performed only for the inner parts of the point clouds. The inner block indices are recalculated relative to the outer labels and updated with the corresponding labels.

The remaining unmarked points that meet the parameter thresholds represent potential safe landing sites. These points are characterized by the fact that there are no dangerous or unsuitable areas around them. These points should be marked as suitable for landing.

The end result will be a data array in which each point is assigned to one of four classes of landing zones, which can be further designated as colors in the image, with the value of each color listed in Table 3 [5-6].

Table 3

Classification of landing zones

Red	Danger points
Brown	Unsuitable points
Yellow	Risky points
Green	Suitable points

The proposed algorithm, combines clustering and point classification methods. The algorithms allow to select zones based on surface characteristics, such as slope and obstacles, and to consider suitability criteria, including zone size and absence of anomalies.

An experiment on detecting landing sites of an unmanned aircraft by point cloud

An experiment was conducted using a point cloud depicting a forest area on a hill. This point cloud has 69909 points and its size is 260 KB, the point cloud is in LAZ format.

Before starting the analysis of laser data, it is necessary to load and visualize them (Fig.2).

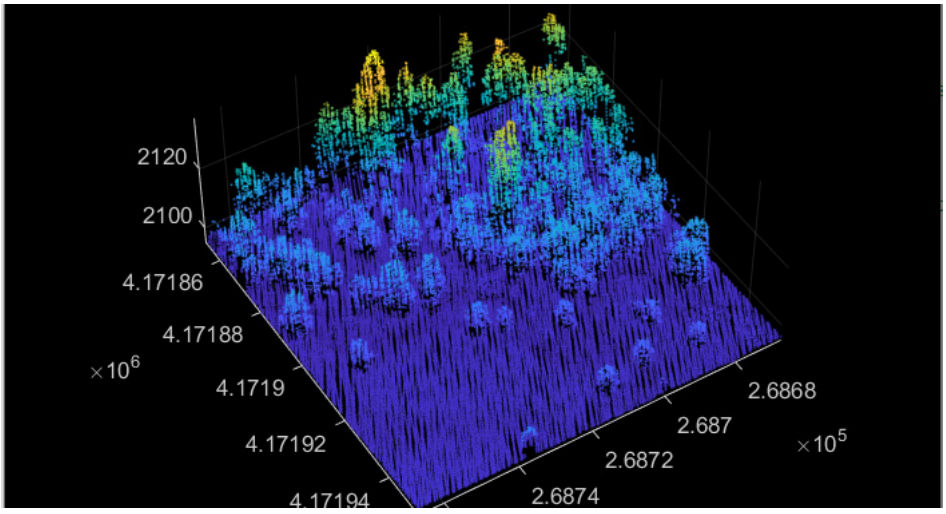


Fig. 2. Input data

This is followed by dividing the point cloud into two categories: terrestrial and non-terrestrial. Non-ground points are automatically marked as hazardous because they may pose a threat to the landing. This procedure highlights and marks the hazardous points in red color as shown in Fig. 3a.

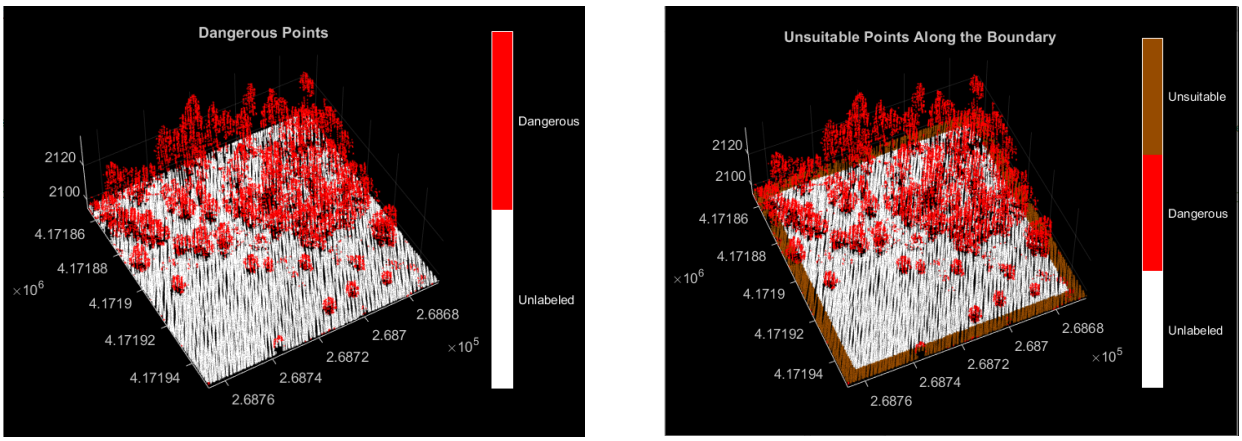
Special attention should be paid to the points located at the boundary of the available data. Since these points do not have complete neighborhoods within the cloud, therefore all unlabeled boundary points should be marked as unusable (Fig. 3b.).

In the process of analyzing the data for it is necessary to classify points as unusable if they are part of trees or other ground structures (Fig. 4a.).

Next, the points are classified using the block overlap-based method, which involves calculating the nearest neighbors for each unlabeled point. If the number of neighbors is less than the set threshold value, or if there are dangerous points among the neighbors, the point is marked as unsuitable. Otherwise, the indexes of the neighboring points are stored in the array of cells, the result is shown in Fig. 4b.

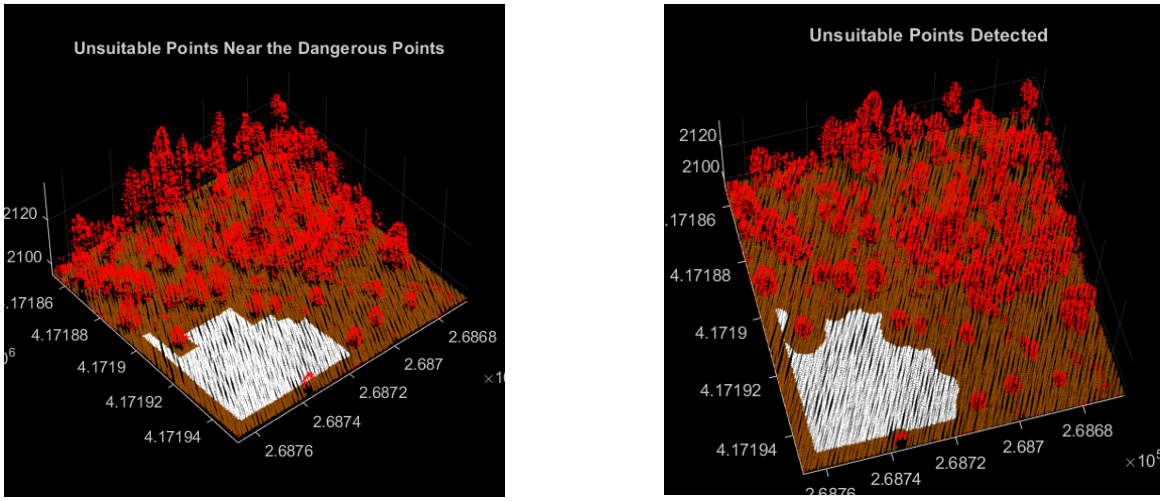
After identification of unsuitable points, unmarked points are analyzed. Each unlabeled point is considered risky if there are unsuitable points in its vicinity. The attributes shown in Table 2 are used for risk assessment. If the value of any attribute of a particular point exceeds the set threshold value, such a point is classified as risky (Fig. 5a.).

The remaining unlabeled points that satisfy the parameter thresholds represent potential locations for safe landings (Fig. 5b).



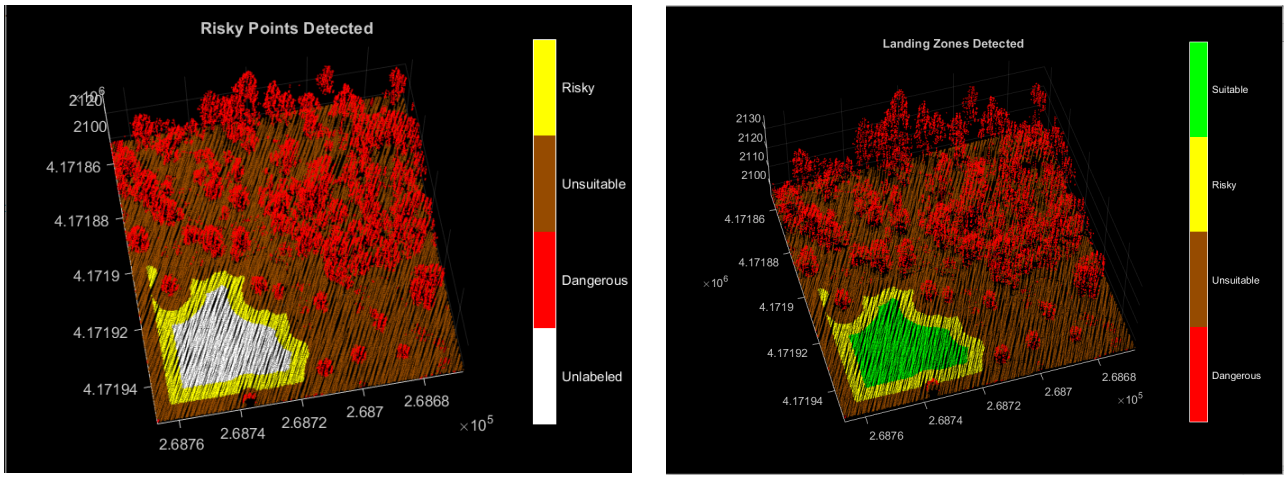
a) Ground and non-ground points *b) Border points*

Fig. 3. Danger points



a) Suitable points of the classification field *b) Suitable points after treatment with the block overlap-based method*

Fig. 4. Identification of suitable points for planting



a) Unspecified points *b) Safe points combined into landing zones of BWCs*

Fig. 5. Classified point cloud with BVS landing zones

Conclusions

The research considered a method of detecting landing spots of an unmanned aircraft based on the analysis of a point cloud obtained by airborne laser scanning. In this work, point clustering and classification algorithms were applied, which made it possible to identify zones with different degrees of suitability for landing. The used algorithms of data processing include consecutive stages of ground and non-ground points selection, surface characteristics estimation and determination of the probability of successful landing of BVS. The methods of cluster analysis and least squares approximation of the plane were also used, which made it possible to classify points depending on their slope, flatness and presence of obstacles.

Experimental studies confirmed the effectiveness of the proposed method. A point cloud depicting a terrain area with forest cover and uneven relief was analyzed. As a result, it was possible to identify areas suitable for landing and visualize them on the basis of color coding. The final model allows not only to identify safe areas, but also to assess the risk of landing in a particular area.

References

1. Afanasyeva V.I. Classification of the observed zones on the generated images of the video frame stream based on segmentation in the onboard systems for monitoring the Earth surface // Collection of reports of the 74th International Student Scientific Conference of GUAP (April 19-23, 2021) – SPb.: GUAP, 2021. – C. 126-133.

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

3. BVS visual positioning system for high-precision autonomous landing / I. E. Sevostyanov, D. V. Devitt, D. V. Trichleb, A. A. Baranova // Problems of mechanical engineering and automation. – 2022. – № 2. – C. 124-130. – DOI 10.52261/02346206_2022_2_124. – EDN HKXUZF.

4. Ilyasov, R. M. Application of aerial photography and airborne lidar imaging for tracking changes in roadway flatness in Arctic conditions / R. M. Ilyasov, V. E. Pushkarev, K. A. Plesovskikh // Scientific Bulletin of Yamalo-Nenets Autonomous Okrug. – 2023. – № 3(120). – C. 6-18. – DOI 10.26110/ARCTIC.2023.120.3.001. – EDN UEVPOM.

5. Gitis L.H. Cluster analysis in classification, optimization and forecasting problems [Text]/ L. H. Gitis. – M. : MGGU, 2001. – 104 c.

6. Marmanis H., Babenko D. Algorithms of intellectual Internet. Advanced methods of data collection, analysis and processing. – Per. from Engl. – SPb.: Symbol-Plus, 2011. – 480 c.

PSYCHOLOGY AND AI: ALTERNATIVES FOR SELF-HELP AND THE FUTURE OF PSYCHOTHERAPY WITH LANGUAGE MODELS

Aleksandra Pogodina
Saint Petersburg State University of Aerospace Instrumentation
E-mail: alexpogodina1@yandex.ru

Abstract. *This article examines the role of artificial intelligence (AI), specifically large language models (LLMs) such as GPT-4, in enhancing psychological support and self-help mechanisms. It elucidates the training processes of these models on extensive datasets to facilitate text comprehension and generation. The significance of precise query formulation for effective interaction with ChatGPT, which can function as a virtual psychologist, is emphasized. An illustrative example of an interaction with ChatGPT in the capacity of a psychologist is presented. While acknowledging the potential of AI to provide emotional support and promote self-reflection, the article also highlights the limitations of AI compared to human psychologists, particularly in terms of empathy and relationship-building. The integration of artificial intelligence into psychological practice is posited as a promising trend that may complement traditional therapeutic methodologies.*

Artificial intelligence (AI) refers to a system capable of imitating human skills. One specific area within AI is large language models (LLMs). LLMs are trained on extensive datasets of textual information and are utilized for text comprehension and generation.

Among the most prominent large language models is GPT-4, which is integrated into the ChatGPT application, providing a dialogic interface for user interaction. The potential applications of this application are numerous, including data analysis, software development assistance, information retrieval, and the generation of various forms of textual content.

It is noteworthy that there exists a variety of large language models, each differing in the types of tasks they are designed to perform. The subsequent discussion will focus on ChatGPT, recognized as the most popular LLM for text generation.

The effectiveness of ChatGPT's output is contingent upon the textual prompt provided. Given that the neural network (the algorithm underlying ChatGPT) generates text based on the probabilistic distribution of words—essentially calculating how frequently words co-occur in texts—and context, it is crucial to formulate prompts as accurately as possible.

Additionally, within the context of using ChatGPT, the concept of a "thread" emerges. This feature allows users to maintain multiple chats on different topics simultaneously, thereby preventing context mixing and facilitating more precise responses [1].

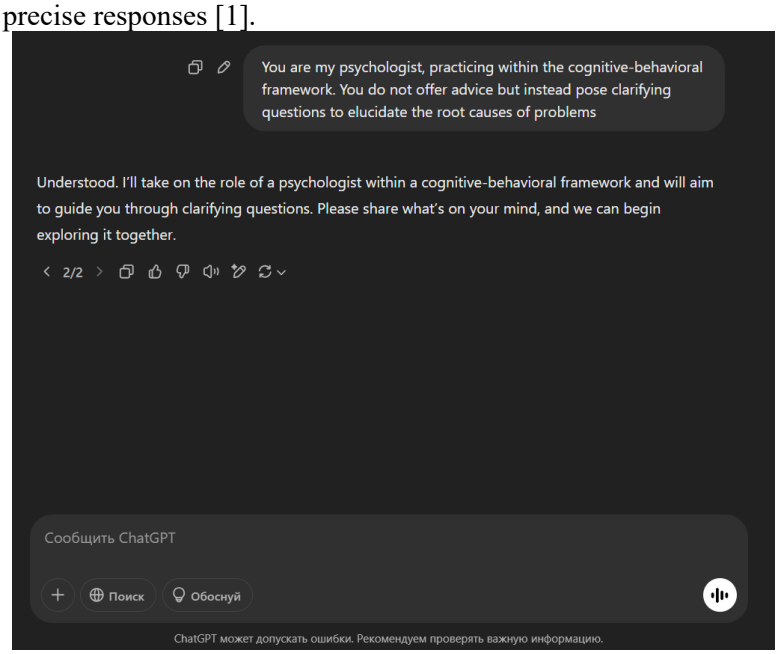


Fig. 1. Example of a request to ChatGPT

One potential application of ChatGPT is in the role of a psychologist. A separate chat can be initiated where a specific prompt is submitted. Here, the limitations are confined only to the user's imagination. The prompt can establish parameters for the desired interaction; for instance, specifying a particular psychotherapeutic approach (such as cognitive-behavioral therapy or Gestalt therapy); instructing that each response will be followed by a probing question to foster deeper self-understanding; or indicating that the dialogue should consider the entire chat history for a more accurate comprehension of the individual's personality. Users may even assign a name to their digital "psychologist." Thus, the quality of dialogue is directly influenced by the instructions provided.

For instance, one could delineate the following parameters for ChatGPT: "You are my psychologist, practicing within the cognitive-behavioral framework. You do not offer advice but instead pose clarifying questions to elucidate the root causes of problems" (Fig. 1).

Assume that the individual seeking assistance is experiencing a negative mood due to the difficulty in making a decision, wherein none of the available options are satisfactory to them (Fig. 2).

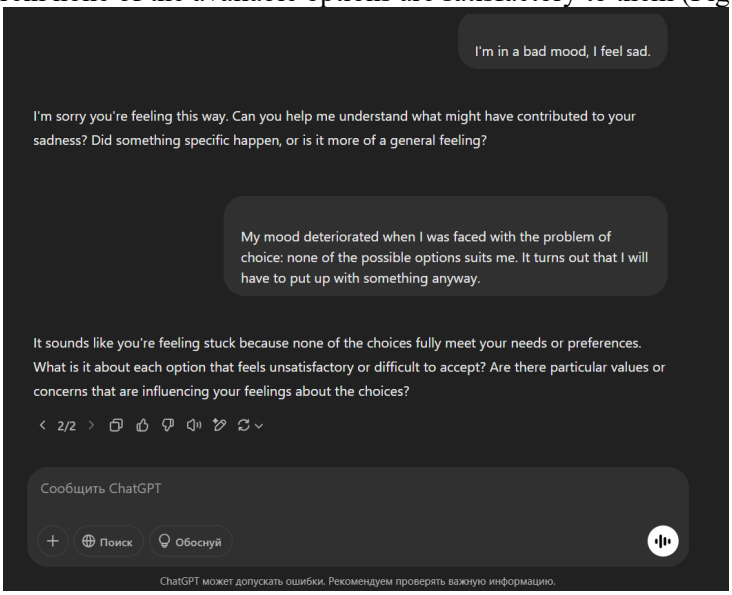


Fig. 2. Example of a conversation with ChatGPT

As evidenced by the aforementioned example, ChatGPT poses clarifying questions as requested by the user. In responding to messages within the chat, individuals may explore their feelings more deeply, identify the root of their issues, and potentially discover solutions. Ultimately, much depends on the individual: their honesty with themselves and their willingness to seek answers to pressing questions.

Thus, ChatGPT, in the role of a psychologist, can serve as one of the tools for self-help or reflection. In one of the articles [1], psychologist Olga Zotova states: «*At a minimum, through conversation with a chatbot, an individual will direct their attention to their complexity, verbalize it, reframe it, read the reformulated version, and examine the situation from various perspectives».* This principle is analogous to written practices, such as when individuals document their automatic negative thoughts or maintain achievement journals. Such exercises activate internal dialogue and reflection. Thus, interaction with a neural network can stimulate one's thinking, enabling the discovery of answers to pressing questions.

However, technology is not yet capable of fully replacing psychotherapists; ChatGPT is not a living person, and the empathy and compassion expressed in chat interactions are artificial and readily perceived as such. As it is well known, humans require human connection. Psychologist Marina Volkova notes [1]: *«Comparing any artificial intelligence—even the most advanced—with psychotherapy is inappropriate: their capabilities differ significantly. In psychotherapy, the formation of a relationship between the client and the helping professional is fundamentally important. This is a universal principle applicable to all methods. It is crucial because most psychological issues are related to relationships or arise within the realm of interpersonal dynamics. The psychological problems we currently face were formed in the past as a means of survival in different circumstances with different people. One of the main postulates of psychotherapy is that difficulties arising in relationships can only be resolved within relationships» [2].*

Numerous technological solutions based on artificial intelligence can serve as tools for self-help and complement therapeutic work with psychologists. These include specialized psychology chatbots, sometimes

grounded in specific types of therapy, such as cognitive-behavioral therapy (CBT), applications for assessing disease probabilities based on symptoms, and options for remote patient support.

Currently, there is active development and integration of artificial intelligence. This field is gaining popularity; however, it represents only the beginning of incorporating such technology into our lives. Therefore, the integration of artificial intelligence into psychology holds significant promise. For example, AI could assist psychologists by aiding in diagnostics. Additionally, virtual environments with integrated artificial intelligence may be utilized for better understanding reactions and training new patterns. This represents just a fraction of the possibilities that could be realized.

Conclusion

At present, artificial intelligence (AI) serves as a valuable tool for self-discovery and self-help. This technology enhances access to psychological assistance, representing a positive trend. Furthermore, the integration of AI into the field of psychology is likely to increase in the future.

References

[1] How I Used ChatGPT as a Psychologist [Electronic resource]. – URL: https://journal.tinkoff.ru/chatgpt-psychologist/ (accessed: January 12, 2025).

[2] The Impact of Artificial Intelligence on the Development of Psychology [Electronic resource]. – URL: https://www.b17.ru/blog/409312/ (accessed: January 12, 2025).

AUTOMATION OF QUALITY MANAGEMENT PROCESSES

Maria Rassykhaeva
Saint Petersburg State University of Aerospace Instrumentation
E-mail: mitschiru@gmail.com

In such conditions, quality management becomes an integral part of success. Traditional approaches are no longer able to cope with high requirements, which makes the automation of quality management processes (QMS) a necessity. The introduction of digital technologies allows you to increase the efficiency, transparency and predictability of processes, minimizing the human factor. Main aspects of automation Fig. 1.

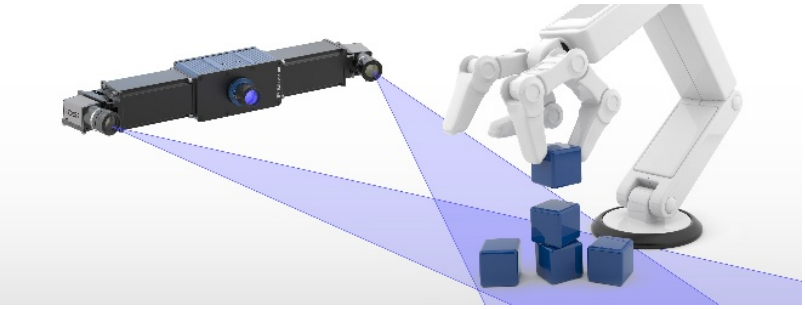


Fig. 1. Video Monitoring Concept

Automation of quality management processes plays a key role in ensuring the efficiency and competitiveness of organizations in today's world. It includes the use of advanced software solutions such as ERP (Enterprise Resource Planning) for enterprise resource planning, MES (Manufacturing Execution Systems) for manufacturing process management, and CRM (Customer Relationship Management) for customer interaction. These systems are integrated into the overall ecosystem of the company, creating an interconnected process for monitoring, analyzing and improving the quality of products or services.

One of the key benefits of automation is the ability to work with big data. Data collection, storage, and processing using artificial intelligence and machine learning algorithms allow not only to identify existing problems, but also to predict possible failures. For example, the use of predictive analytics algorithms helps companies to proactively identify risks associated with production processes and take proactive measures [3–7].

Automation also contributes to the standardization of processes, which makes it easier to comply with international standards such as ISO 9001. With digital technologies, companies can quickly adapt to changing market demands, increasing their flexibility and reducing the cost of adapting new approaches. For example, the implementation of IoT (Internet of Things) allows companies to collect data from connected devices in real time, ensuring quick decision-making and transparency of production processes [1-3].

The role of automation in improving the customer experience is also important. Quality management automation systems can analyze customer feedback, identify patterns, and suggest improvements that directly affect customer satisfaction. Moreover, automation helps to reduce costs, minimize equipment downtime and eliminate the human factor, which is especially important for high-tech industries.

Table. Benefits of automation

Advantage	Description
Cost reduction	Reduced manual labor and product defects result in lower operating costs.
Increased productivity	Digital technologies speed up operations and simplify communication between departments.
Compliance with standards	Automation systems are easily adaptable to ISO 9001 standards, ensuring a high level of quality.
Increase customer satisfaction	Quick response to requests and high quality of products or services strengthen customer confidence.

Automation brings significant benefits to organizations. In addition to the aspects described above, it contributes to deeper data analysis, improved logistics processes and increased accuracy of demand forecasting.

Examples of successful automation.

1. **Manufacturing companies.** The introduction of MES systems allows you to track the execution of orders in real time, which reduces the number of defects by 25%.
2. **Retail.** CRM systems automate customer relationship management, increasing customer loyalty.
3. **Health care.** The use of quality monitoring systems reduces errors and improves the overall patient experience.

Challenges and solutions.

Despite the obvious advantages, automation faces a number of challenges that require an integrated approach and strategic planning:

- **High implementation costs.** The initial investment in automation includes the cost of hardware, software, and staff training. However, proper planning and the use of government subsidies or grants can reduce the financial burden.
- **Resistance to change.** Staff are often wary of new technologies, fearing job loss or increased workload. To minimize this effect, it is important to organize training programs and maintain open communication so that employees understand the benefits of automation.
- **Systems integration.** One of the main challenges is to ensure that new systems are compatible with existing infrastructure. Using hybrid solutions and engaging experienced integrators helps avoid problems.
- **Cybersecurity.** Automation is often associated with the transmission of data over digital networks, which increases the risks of cyberattacks. Companies must invest in data protection, including encryption, regular software updates, and backups.
- **Difficulty measuring results.** Evaluating the effectiveness of automation requires accurate metrics. It is important to establish key performance indicators (KPIs) such as reduced defects, reduced costs, and increased speed of operations in order to objectively measure the success of the implementation.

Development prospects.

In the future, the automation of quality management processes will include increasingly intelligent technologies such as adaptive systems and self-learning algorithms. For example, the use of predictive analytics and IoT will be even more widely integrated, allowing companies to collect and process real-time data with even greater accuracy. In addition, the development of quantum computing can significantly accelerate the process of analyzing big data, leading to deeper process optimization and reduced time costs.** The initial investment can be significant, but it quickly pays for itself through increased efficiency.

- **Resistance to change.** Training staff and explaining the benefits of automation help minimize negative perceptions.
- **Systems integration.** Choosing compatible solutions and managing your project makes it easy to transition to automation.

Conclusion

Automation of quality management processes is an integral part of modern business. It not only helps companies stay competitive, but also contributes to their sustainable development. Cost reduction, increased accuracy and increased customer satisfaction make automation a strategically important focus. The introduction of digital technologies in quality management is not just a trend, but a necessity that allows organizations to meet market requirements and consumer expectations

Bibliography

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. Chabanenko, A.V. Quality Management of Hull Elements of REA. RIA: Magazine: "Standards and Quality". Moscow: 2018. №2. Pp. 90-94.

6. PROGRAM FOR MODELING THE BEHAVIOR OF A POLYMER ROD AT THE STAGE OF HOMOGENIZATION IN AN ADDITIVE UNIT Chabanenko A.V., Nazarevich S.A., Rassykhaeva M.D.Certificate of registration of a computer program RU 2023614958, 09.03.2023. Application No 2023614143 dated 09.03.2023.

7. 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. Pp. 285-289 .

COMPARATIVE CHARACTERISTICS OF LEAST SQUARES AND LEAST MODULI METHODS FOR APPROXIMATING REGRESSION RELATIONSHIPS IN THE PRESENCE OF ANOMALOUS OUTLIERS

Ivan Rybkin
Saint Petersburg State University of Aerospace Instrumentation,
St. Petersburg, Russia,
E-mail: ivan9643721960@gmail.com

Abstract. *The results of the analysis of the characteristics of the estimates of the regression dependence parameters for the least squares and least moduli methods in the presence of rare "large" anomalous outliers in small samples for normal measurement errors and for non-Gaussian errors are presented. Confidence bounds for regression relationships and errors of prediction of future values of the investigated time series are calculated*
Keywords: *linear regression, least squares method, confidence interval*

In the modern world, methods and models for estimating regression dependencies in various fields of science and technology are of great importance. In this case, an important aspect is the analysis of model error distribution, which allows to improve the quality of forecasts and reduce the error of parameter estimates.

In the classical view, when it is assumed that errors are distributed according to the Gaussian law, the best way to obtain estimates of regression relationships is the method of least squares. This method has a simple analytical form, which is why it is the most popular tool for processing regression dependencies.

An alternative method for obtaining estimates of regression dependence parameters is the method of least moduli. This method has a similar idea as least squares least squares, but only if in the case of least squares the minimization of squares of errors is performed, in least moduli the minimum of the function of the sum of absolute deviations is found. According to [1], it is possible to obtain estimates with the help of IRLS.

The use of the method of least moduli for absolutely all cases of error distribution is not optimal. It has been proved that the least moduli is the best way to obtain robust estimates for errors distributed according to the Laplace distribution [2]. But, for example, when using "clogs" of a different nature, the least moduli will also lose to better methods. The "clogs" of a different nature can mean anomalous outliers, the amount and power of which can strongly affect the obtained estimates of regression dependences.

To generate data that have "anomalous" outliers we can use Tukey's model, which is a special case of Huber's model [3]

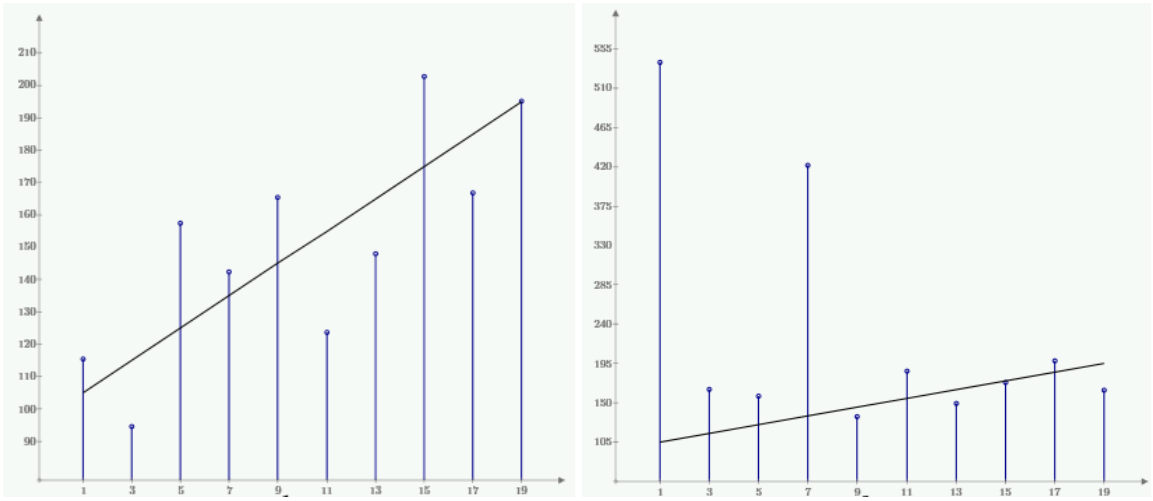
$$(1-\gamma)*\mathrm{N}(0,\sigma_1^2)+\gamma*\mathrm{N}(0,\sigma_2^2)$$

This law allows controlling data generation by two parameters: σ_1,σ_2 – standard deviations of the data, γ – level of clogging (probability of outliers). It is worth noting that it is not so important the standard deviations values themselves as their ratios. The following data parameters were considered:

1. $\frac{\sigma_2}{\sigma_1}=2,\gamma=0.1$ – small outliers in small quantities
2. $\frac{\sigma_2}{\sigma_1}=10,\gamma=0.25$ – "abnormal" outliers in greater numbers

For reference, all data were generated according to a predetermined law: $y=a+bx, a=100, b=5$. Thus, Fig. 1 shows the correlation fields of some of the many data obtained. A correlation field is a dot plot in a rectangular coordinate system in which the scale for one trait (x) is plotted on the abscissa axis and the scale for another trait (y) is plotted on the ordinate axis.

Such an image creates an overall picture of correlation and helps to reveal the presence, direction and closeness of correlation between the features.

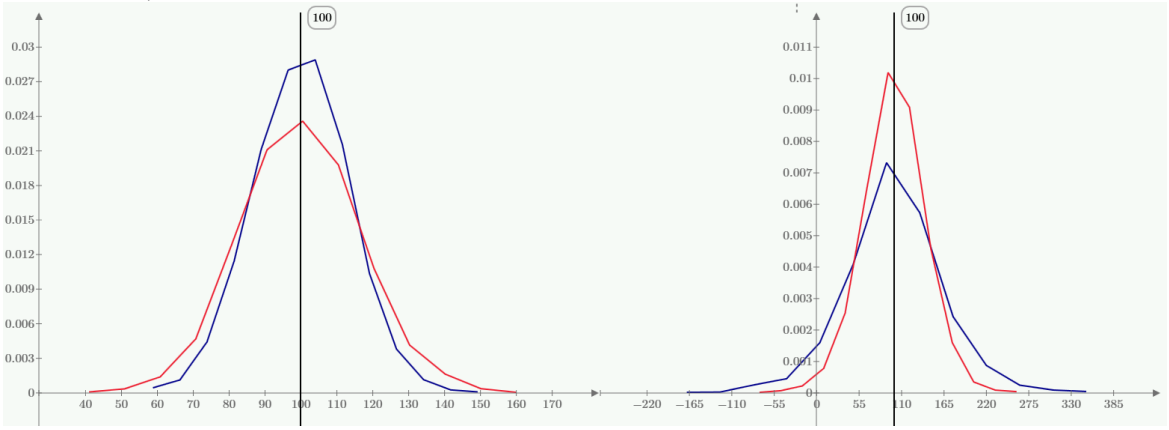


a) *b)*

Fig. 1. Correlation field

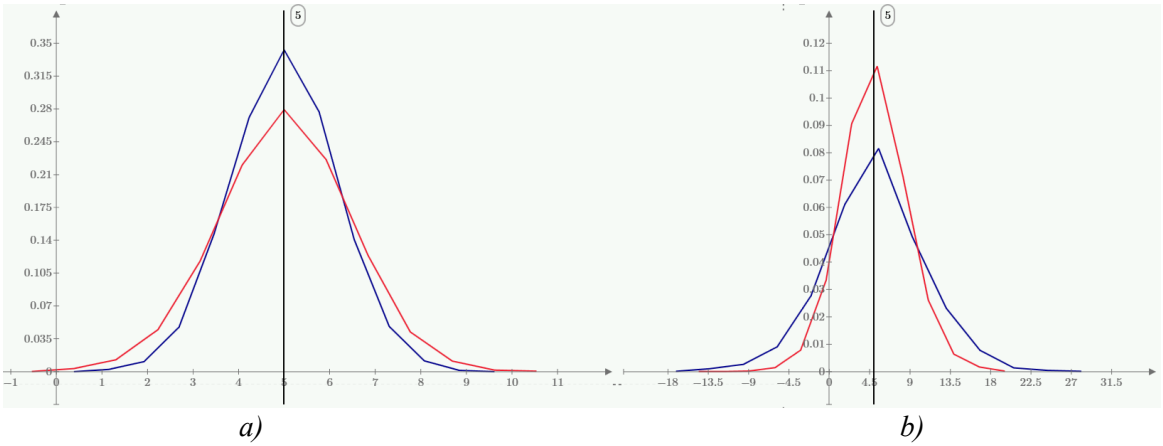
Obtaining an array of estimates of linear pairwise regression coefficients, we can plot their distribution densities. Comparison of these plots allows us to see that on the first set of data (with small outliers), the least squares gives better results. While large outliers are better handled by the least moduli.

Fig. 2 shows the comparison plots for estimating the coefficient a. The right side shows estimates obtained for small emissions, the left side for large emissions. Blue color indicates curves based on least squares estimates, red – on least moduli.



a) *b)*

Fig. 2. Density analysis of parameter estimates at small and large outliers (coefficient a)
Similar plots for the coefficient b are shown in Fig. 3.



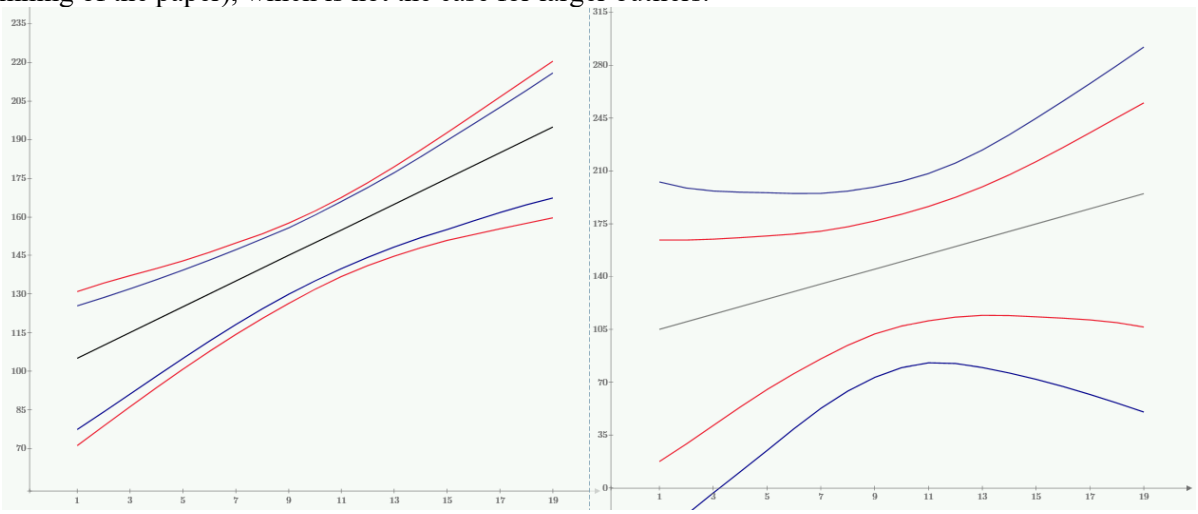
a) *b)*

Fig. 3. Density analysis of parameter estimates at small and large outliers (coefficient b)

Comparison of these plots allows us to see that on the first set of data (with small outliers), the least squares gives better results. While large outliers are better handled by the least moduli.

When obtaining the regression equation and parameter estimates, a very important part of the analysis is the construction of confidence intervals for the regression equation, or conditional mathematical expectation, as well as for the obtained estimates. A confidence interval is an interval that covers an unknown parameter with a given reliability [1]. When it is not easy to understand the nature of the error distribution, it is necessary to use methods other than Student's t-test to obtain parameter estimates. Such methods include the method of constructing the empirical distribution function, and then find the points of its intersection with the straight lines $y = \frac{1-\alpha}{2}$ and $y = \frac{1+\alpha}{2}$, where α is the reliability level.

Fig. 4 shows plots of confidence intervals, with plots at small outliers on the left and plots at large outliers on the right. Blue color shows the plots for least squares, red – for least moduli. It is easy to note that for smaller outliers the least squares gives even better results than the least moduli (which was noted at the beginning of the paper), which is not the case for larger outliers.



a) *b)*

Fig. 4. Confidence intervals

In conclusion, the least squares method, as the simplest method, can indeed cope with data that has a small number of low power outliers, but for data with “anomalous” outliers it is worth using the least moduli method as a more robust method.

References

1. PETER J. HUBER Robust Statistics / Translated from English – M.: Mir, 1984. – 304 p.
2. Estimates of parameters of regression models, URL: https://studwood.net/1560786/ekonomika/otsenki_parametrov_regressionnyh_modeley (in Russian) (Дата обращения 27.02.2025).
3. Zakharova A. Y. Estimation of regression curve parameters using Huber model / A. Y. Zakharova, V. A. Tyurinova, and A. P. Shepeta. Molodoj ucheny`j [Young Scientist], 2022, no. 47.1 (442.1), pp. 52-54. (in Russian). URL: https://moluch.ru/archive/442/96762/ (Дата обращения 27.02.2025).

AUTOMATION OF MEASURING PROCESSES WITH A ROUND-ROBIN METER

Karina Ryndina
Saint Petersburg State University of Aerospace Instrumentation
E-mail: rindina.karina2017@yandex.ru

Abstract. *Automation of measurement processes plays a key role in improving the accuracy and speed of analyzing the geometries of parts. One of the important tools in this field is the round-gauge, which is designed to evaluate the shape parameters of parts, in particular, roundness, cylindricity and concentricity. This article discusses the basic principles of the round-gauge, its application in industry, as well as issues related to measurement automation and data processing.*

Modern production plants place high demands on the accuracy of the geometric parameters of parts. Deviations from the set values can lead to a decrease in assembly quality, increased wear of mechanisms and even equipment failure. In this regard, there is an increasing need for the use of high-precision measuring instruments, among which a special place is occupied by round-gauges.

Round gauges are used to control the shape of the surfaces of parts, providing high measurement accuracy through the use of specialized sensors and data processing algorithms. Automating measurements with a round gauge minimizes human error, improves accuracy and repeatability of results, and optimizes the quality control process.

A round-gauge is a measuring device that registers deviations in the shape of a part from a perfect circle (pic.1).

Picture 1. Round gauge

The main components of a round-measure include:

1. A high-precision displacement sensor that measures microscopic changes in the surface of the part. Modern sensors are capable of detecting deviations within the nanometer accuracy, which makes them indispensable in precision engineering;
2. Precision rotary table that ensures uniform rotation of the part relative to the measuring axis. It is important that the axis of rotation has minimal radial and axial run-outs, as even small deviations can lead to measurement errors;

3. A data management and analysis system that includes specialized software for signal processing and reporting of measurement results. Modern systems are equipped with algorithms for digital filtering and automatic determination of shape parameters.

The principle of operation of a circular meter is based on a comparison of the actual shape of the part with the reference circle. During the measurement process, the sensor detects surface deviations as the part rotates, which allows you to build a profile and calculate shape deviation characteristics such as radial runout, ovality, and concentricity. To improve the accuracy of measurements, methods of multiangular analysis and adaptive interpolation of the data obtained are used.

Automation of measurements using round-gauges is implemented at several levels. Automatic positioning and clamping of the part is carried out using robotic arms and automated parts feeding systems, which significantly reduces the time of preparation for measurement and eliminates the influence of the human factor. Software control of the measurement process is provided by modern round gauges equipped with computer numerical control (CNC) systems, which guarantees precise adjustment of measurement parameters and their reproducibility. Digital signal processing algorithms incorporate filtering, interpolation, and data analysis techniques, minimizing measurement errors and improving the accuracy of results. Integration with industrial information systems such as ERP (Enterprise Resource Planning) and MES (Manufacturing Execution Systems) enables automatic collection, storage and analysis of measurement data, allowing enterprises to manage product quality in a single digital space [1-4].

Round gauges are widely used in various industries due to their high accuracy and reliability in controlling the geometric parameters of parts. In the automotive industry, they are used to inspect the quality of crankshafts, bearings, and other precision components, ensuring the durability and safety of vehicles. In the aerospace industry, round gauges play an important role in measuring the geometry of aircraft engine parts, turbine blades and other critical aircraft structural components, where measurement accuracy is a key factor in reliability. In mechanical engineering, round meters are used to control the quality of gears, cylinders and other parts subject to high loads, which helps to increase the efficiency and durability of mechanisms. In the medical industry, these devices are used to test the quality of implants, surgical instruments, and other products where the accuracy of the geometry directly affects patient safety. Automation of measuring processes based on round meters can significantly reduce the time spent on quality control, increase the accuracy of measurements and reduce the number of rejects, providing a higher level of control and management of product quality.

Round gauges are widely used in various industries due to their high accuracy and reliability in controlling the geometric parameters of parts. In the automotive industry, they are used to inspect the quality of crankshafts, bearings, brake discs and other precision components, which ensures the durability and safety of vehicles. In the aerospace industry, circular meters play an important role in measuring the geometry of aircraft engine parts, turbine blades, rocket engine housings, and other critical aircraft structural components, where measurement accuracy is a key factor in reliability. In mechanical engineering, circular meters are used to control the quality of gears, cylinders, bearings and other parts subject to high loads, which helps to increase the efficiency and durability of mechanisms. In the medical industry, these devices are used to test the quality of implants, surgical instruments, and prostheses, where the accuracy of the geometry directly affects patient safety. Round gauges are also used in machine tool and optical instrument manufacturing, where high-precision shape measurements are critical. Automation of measurement processes based on round-wheel meters can significantly reduce the time spent on quality control, increase the accuracy of measurements, minimize the impact of the human factor and reduce the number of defects, providing a higher level of control and management of product quality.

Round gauges play an important role in ensuring the quality and accuracy of parts used in industry. They are used in the production, assembly and inspection phases, enabling businesses to maintain high standards of geometric accuracy. The use of round-gauges is especially relevant in the conditions of mass production, where it is necessary to quickly identify defects and deviations from the design parameters. Automated round gauges are integrated into production lines, providing non-contact control and instant analysis of the data obtained. This can significantly reduce the likelihood of rejects, improve product reliability and reduce the cost of corrective operations. Round-gauges are also in demand in the production of complex parts with high precision requirements, for example, in microelectronics, optical instrumentation and power engineering.

Thanks to the development of technology, modern round meters are equipped with intelligent data analysis algorithms that allow not only to record shape deviations, but also to predict potential defects, optimizing production processes. The use of artificial intelligence and machine learning in the processing of measurement data makes it possible to identify hidden patterns and increase the predictability of processes,

which is critically important in high-precision production. Integration of data obtained from round-level meters with production control systems contributes to the improvement of process monitoring and the formation of a database for subsequent analysis and optimization of technological operations. The introduction of integrated automated geometry control systems allows enterprises to reduce costs, increase production efficiency and guarantee the high quality of products.

Conclusion

Automating measurement processes with round meters is an important step towards digitalizing production and improving product quality. The introduction of such systems allows minimizing the impact of the human factor, increasing the accuracy and efficiency of control, as well as integrating measurement data into the overall enterprise management system. Further development of automated measurement technology will contribute to improving the efficiency and reliability of production processes.

Bibliography

1. Gostev, V. N., & Kozlov, A. I. (2020). Methods and Instruments of Measurement in Mechanical Engineering. Moscow: MSTU Publishing House.
2. Smith, G. T. (2016). Industrial Metrology: Surfaces and Roundness. Springer.
3. GOST 24642-81. Control of the geometric accuracy of parts. Methods for measuring roundness and cylindricity.
4. ISO 1101:2017. Geometrical product specifications (GPS) – Geometrical tolerancing – Tolerances of form, orientation, location and run-out.

THE INFLUENCE OF THERMAL FLOW TURBULENCE ON THE SPATIAL CHARACTERISTICS OF A LASER BEAM

Yana Ryvkina
Master's Student
Saint Petersburg State University of Aerospace Instrumentation (SUAI),
Saint Petersburg, Russia
yana.ryvkina@bk.ru

Abstract. *Thermal flows significantly affect the spatial characteristics of laser beams, causing changes in their focusing, divergence, and intensity distribution. These effects are due to variations in the refractive index of the medium under heating, leading to distortions in the trajectory and parameters of laser radiation. This study investigates the influence of thermal flows on the spatial characteristics of a laser beam at various distances using computational algorithms for laser beam profile processing.*
Keywords: *thermal convective flow, laser beam, medium turbulence, correlation function, integral-difference function.*

Introduction

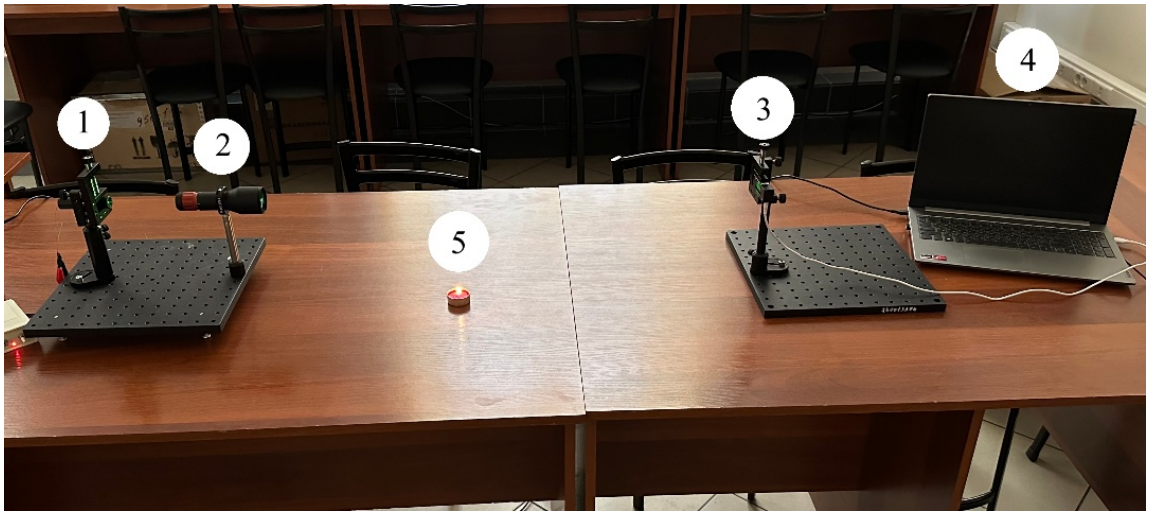
Turbulence of thermal flows has a significant effect on the propagation of laser beams in various media. Temperature fluctuations lead to changes in the refractive index, causing distortions of the phase front and amplitude of radiation. This, in turn, leads to deterioration of the quality of information transmission, decreased pointing accuracy and reduced efficiency of optical communications.

By analyzing the spatial characteristics of laser radiation, it is possible to detect thermal flows. The application of laser-optical technologies for detecting thermal convective flows is a promising area of research [1-3].

1. Laboratory setup

A laboratory setup was developed to study the interaction dynamics of a thermal convective flow (TCF) with a laser beam (Fig. 1). The setup included an LG-D532-15-5 laser module (1) with a wavelength of 532 nm and power of 15 mW, an adjustable beam expander (GBE10-B) from THORLABS (2), and a CCD array SL-TCD-VI (3) from Avesta.

The laser beam passes through the beam expander, ensuring alignment with the receiving area of the CCD array. The distance between the laser and the CCD array can be adjusted.



1 – diode laser; 2 – beam expander; 3 – CCD array; 4 – PC; 5 – TCF source

Fig. 1. Laboratory setup for recording spatial characteristics of optical beams

The PC was equipped with custom-developed software that allowed modification of CCD array operating parameters. The laser beam intensity profile distribution data were saved in txt or CSV files for further processing. The software interface is shown in Fig. 2.

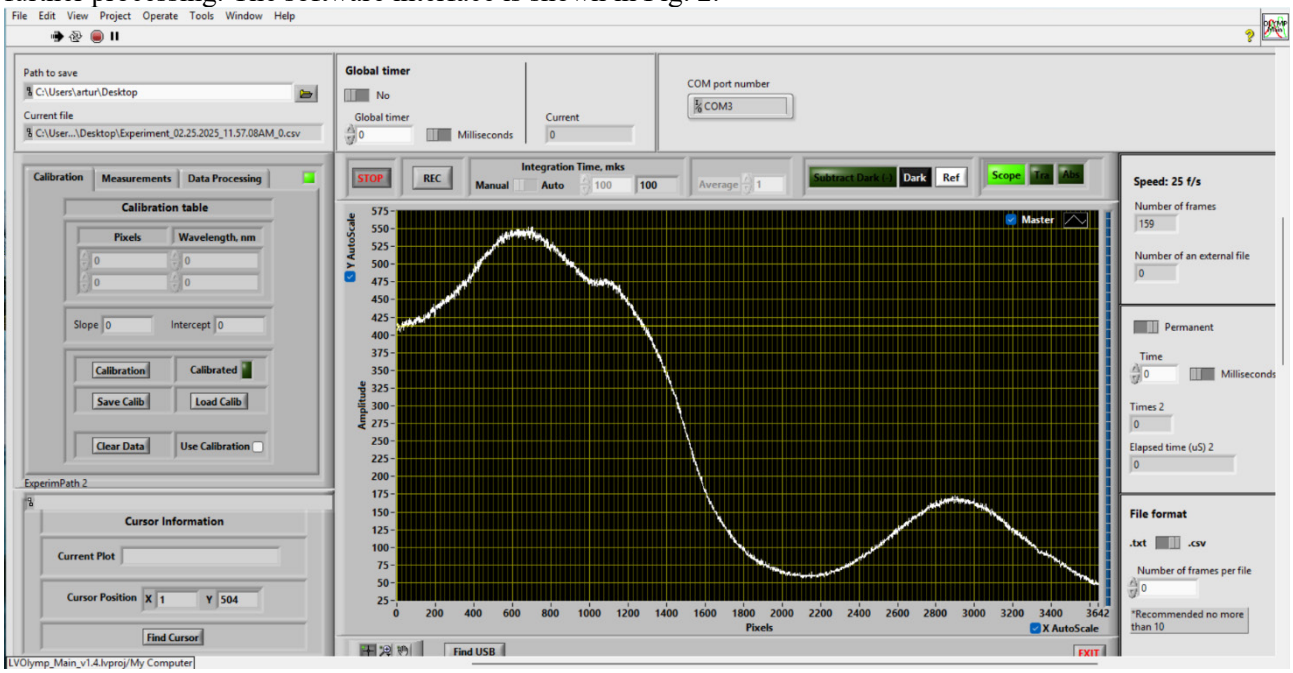


Fig. 2. CCD array control software interface

The software allows adjustment of the following CCD array parameters: the number of active pixels, signal integration time, the number of profile distribution samples for averaging, and the time interval for sample recording.

To study the influence of medium turbulence on the laser beam, the following experiments were conducted:

1. Simulation of the system's normal operating mode. The intensity profile distribution samples were recorded for 30 seconds. The accumulation time for each sample was 10 microseconds.

2. Formation of a TCF. In this series, fluctuations in the intensity profile distribution of the beam were recorded. A candle flame was used as the TCF source. As a result, samples were recorded for 30 seconds.

As a result of the experiment, a RAW file was generated containing data on the laser beam intensity profile distribution. The distance between the laser and the CCD array was varied from 1 to 5 meters in 1-meter increments.

2. Computer processing and interpretation of results

During the experiment, intensity distribution data were obtained at various distances in the absence and presence of a TCF. Beam profiles are presented in Figures 3-7.

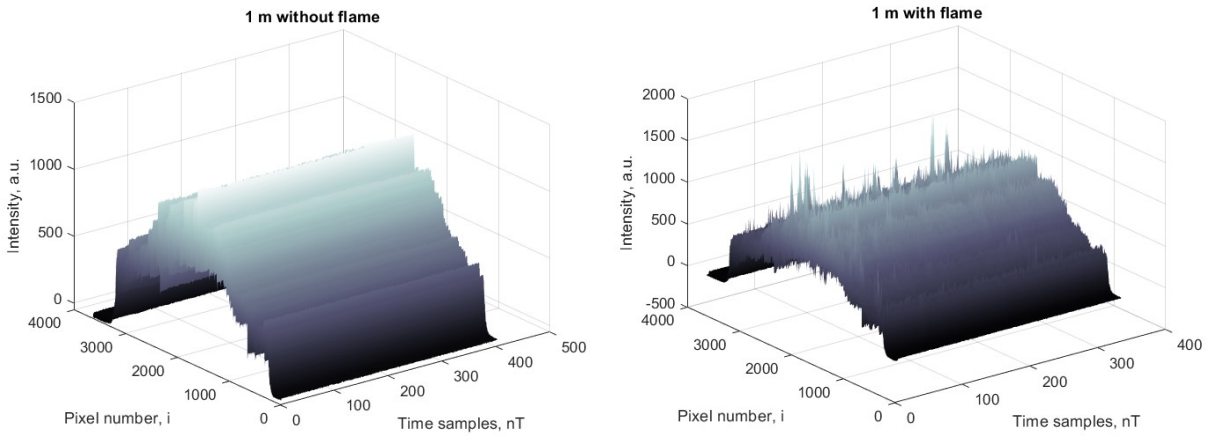


Fig. 3. Dynamics of laser beam intensity profile distribution at different distances l=1 m

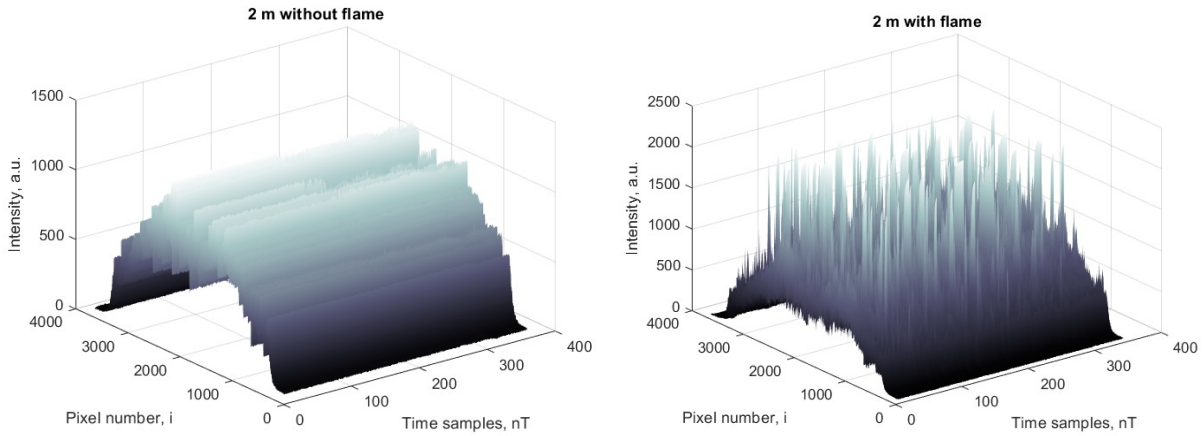


Fig. 4. Dynamics of laser beam intensity profile distribution at different distances l=2 m

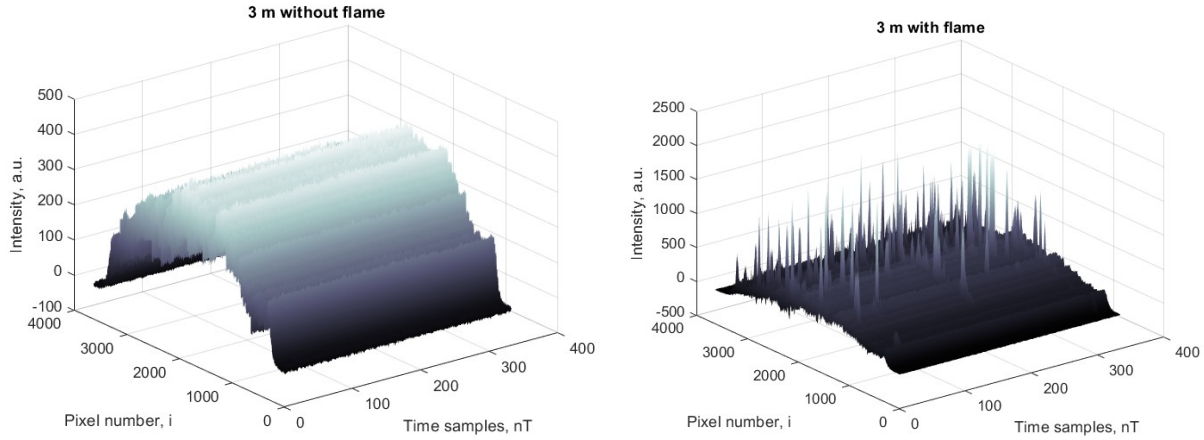


Fig. 5. Dynamics of laser beam intensity profile distribution at different distances l=3 m

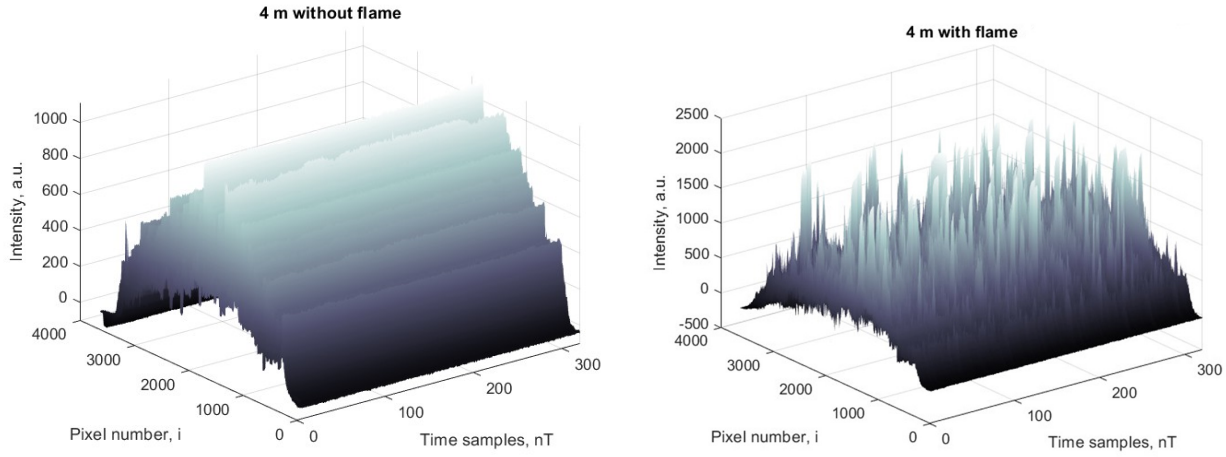


Fig. 6. Dynamics of laser beam intensity profile distribution at different distances l=4 m

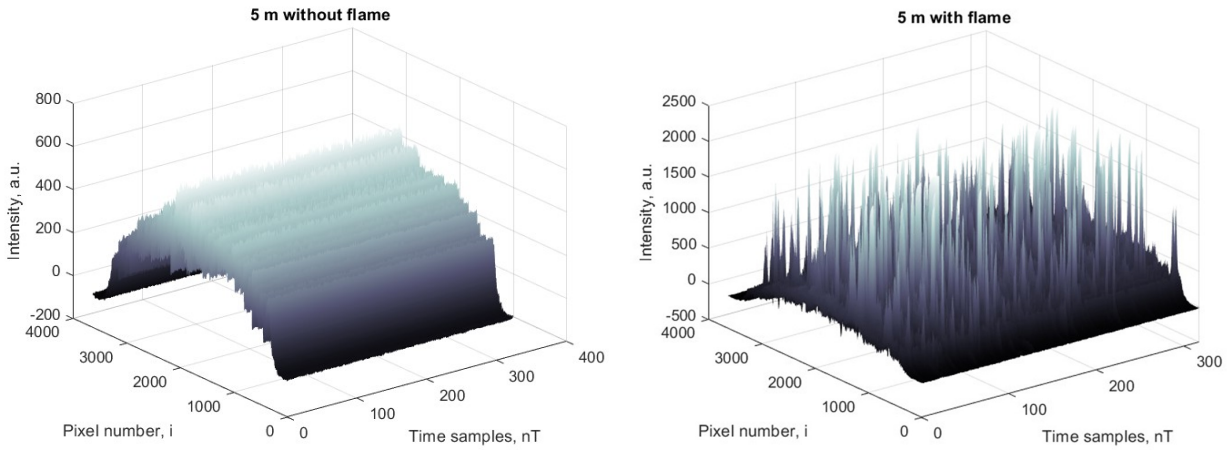


Fig. 7. Dynamics of laser beam intensity profile distribution at different distances l=5 m

For data processing, two approaches were used: correlation and integral-difference methods.

In normal operation, the correlation function remains at its maximum. The presence of a TCF leads to fluctuations, causing a decrease in the maximum value of the correlation function.

Mathematical processing involved forming the dynamic maximum of the cross-correlation function of the discrete signal W(nT), expressed as:

$$W(nT) = \max_l \left[\frac{1}{L} \sum_{i=0}^{L-1} X_i(nT) X_{i-l}([n-1]T) \right], \quad (1)$$

where T is the sample registration interval (1 second), and $X(nT)$, $X([n\text{-}1]T)$ are the beam profile intensity distribution samples recorded in adjacent T intervals.

Additionally, the integral-difference function can be used to process the obtained results, reflecting the pixel-by-pixel difference in the registered intensity distributions of the beam profile. In mathematical form, the proposed integral-difference function for discrete samples of the beam profile intensity distributions can be represented as:

$$R(nT) = \sum_{i=1}^{L} |\Delta X_i(nT)| = \sum_{i=1}^{L} |X_i(nT) - X_i([n-1]T)| . \quad (2)$$

A MATLAB program was implemented to process RAW data using both methods. A program code fragment is shown in Fig. 8.

```
97  for i=1:15
98   CorrX=xcorr(X(i+1,:), X(i,:));
99   CorrY=xcorr(Y(i+1,:), Y(i,:));
100 MAX_X(i)=max(CorrX);
101 MAX_Y(i)=max(CorrY);
102 RaznX(i)=sum(abs(X(i+1,:)-X(i,:)));
103 RaznY(i)=sum(abs(Y(i+1,:)-Y(i,:)));
104 end;
105
106
107 Plot_Max_X=MAX_X/max(MAX_X);
108 Plot_Max_Y=MAX_Y/max(MAX_Y);
109
110 Rez1=horzcat(Plot_Max_X, Plot_Max_Y);
111 Rez2=horzcat(RaznX, RaznY);
112
113 figure(nfig); nfig = nfig + 1; clf;
114 hold on;
115 plot(Rez1,'r'), grid on;
116 title ('Dynamics of the function W(nT) change')
117 xlabel ('nT, s');
118 ylabel ('W(nT)');
119 hold off;
```

Fig. 8. Fragment of the program code for laser beam intensity profile processing

The results of processing using the correlation and integral-difference methods are presented in Figures 9-13 (a) and (b), respectively.

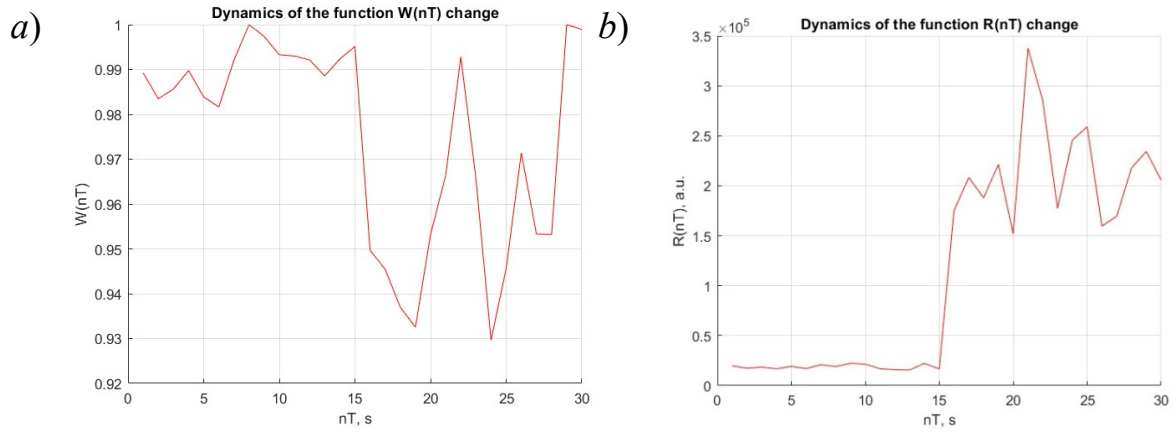


Fig. 9. Laser beam profile distribution processing results using correlation (a) and integral-difference (b) methods at different distances l=1 m

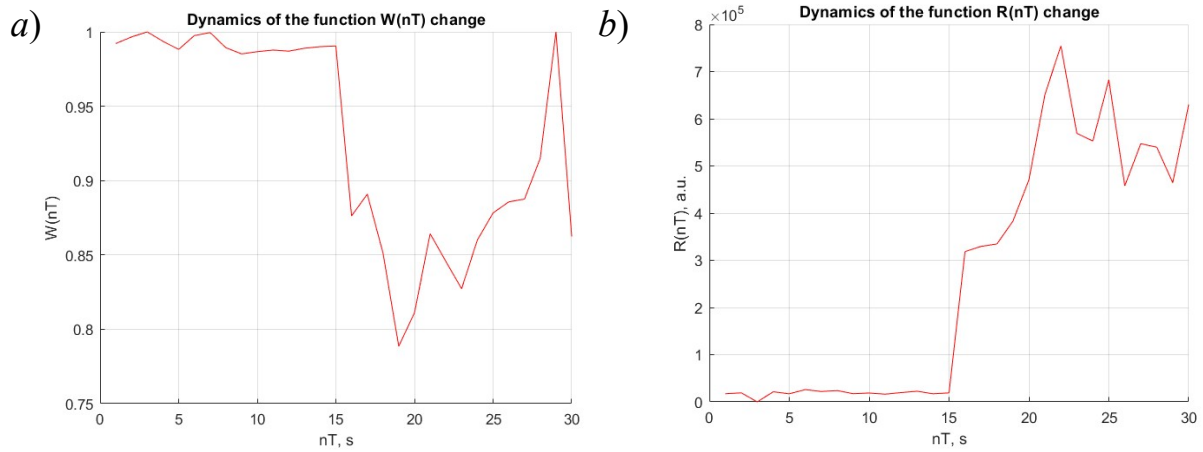


Fig. 10. Laser beam profile distribution processing results using correlation (a) and integral-difference (b) methods at different distances l=2 m

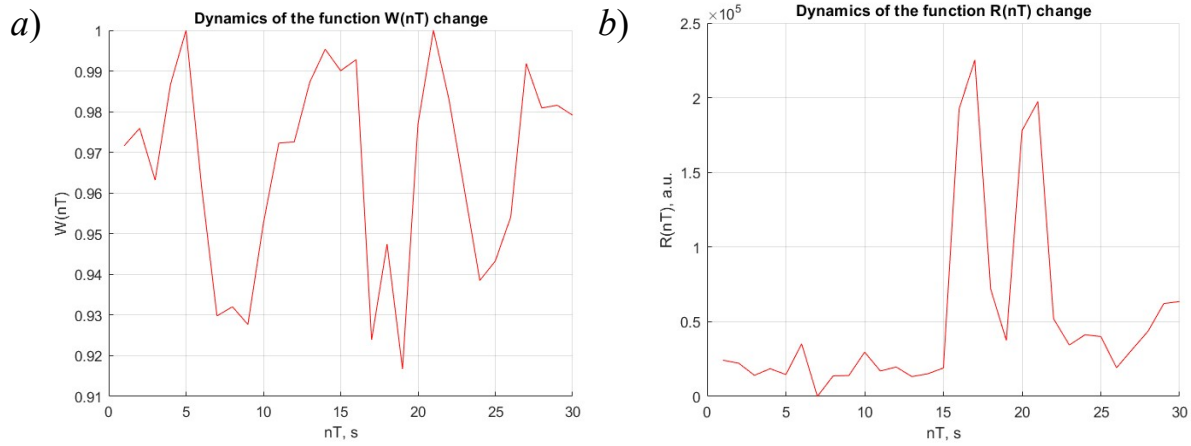


Fig. 11. Laser beam profile distribution processing results using correlation (a) and integral-difference (b) methods at different distances l=3 m

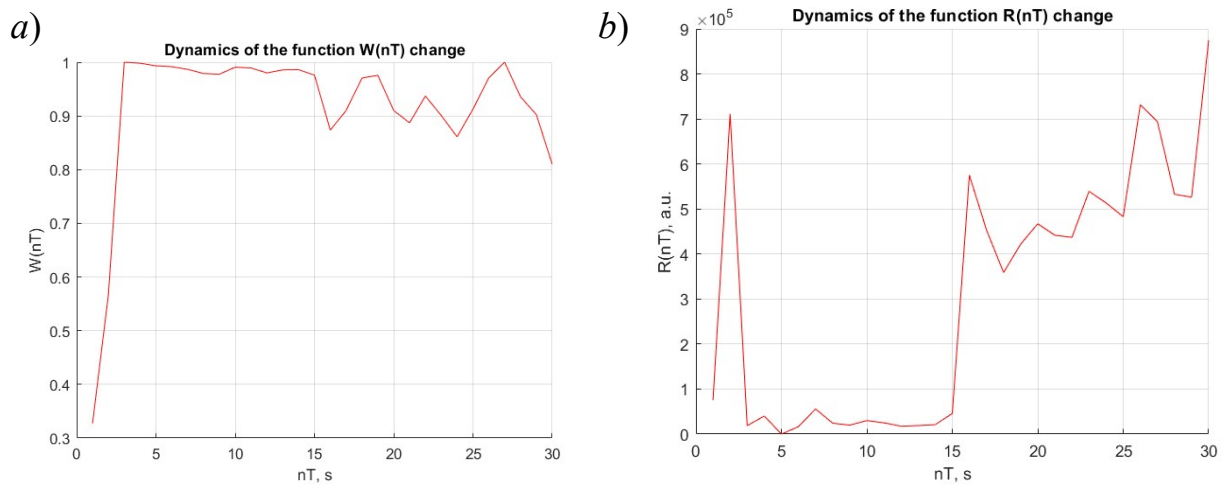


Fig. 12. Laser beam profile distribution processing results using correlation (a) and integral-difference (b) methods at different distances l=4 m

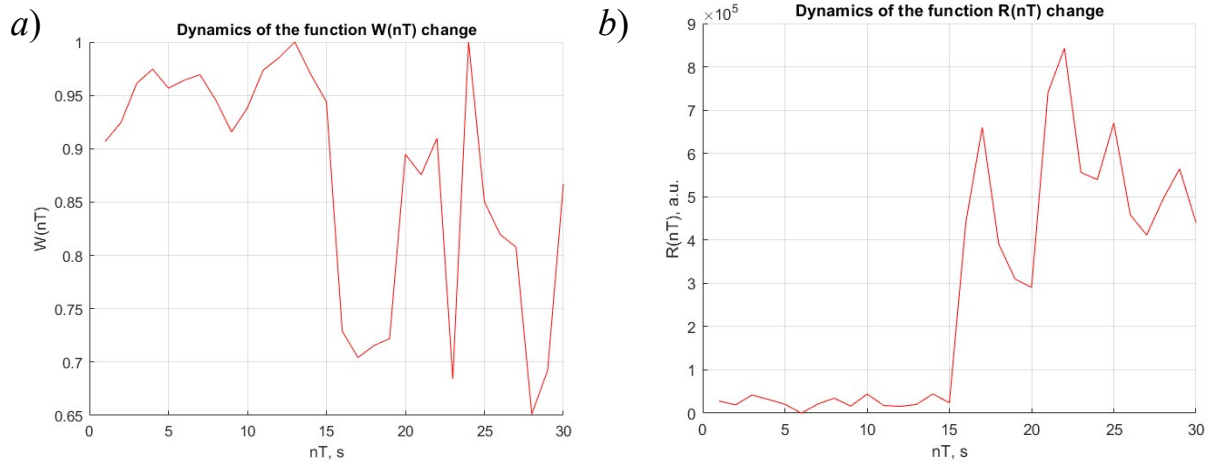


Fig. 13. Laser beam profile distribution processing results using correlation (a) and integral-difference (b) methods at different distances l=5 m

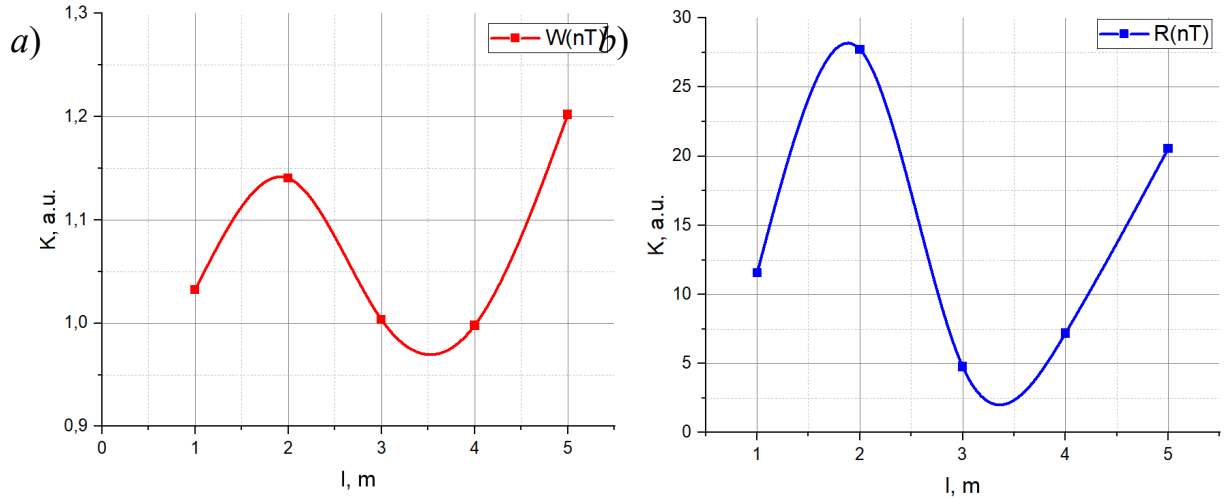


Fig. 14. Evaluation of processing efficiency using correlation (a) and integral-difference (b) methods

To compare the efficiency of both mathematical models, an evaluation criterion was introduced:

$$K_{ef_{corr}} = \frac{\bar{W}_{cTKП}}{\bar{W}_{безTKП}} \tag{4}$$

and

$$K_{ef_{diff}} = \frac{\bar{R}_{безTKП}}{\bar{R}_{cTKП}}. \tag{5}$$

Fig. 14 presents the calculation results of the proposed criteria (4) and (5) for experiments with different interaction lengths of the thermal convective flow with the laser beam.

The graphs clearly demonstrate that both methods effectively detect the presence of a TCF. When a TCF appears between the beam expander and the CCD array (starting at 15 seconds), the target function value changes sharply (decreases in the case of the correlation method and increases in the integral-difference method). The integral-difference method increases the target function value by at least 2.5 times, making it more preferable.

Conclusion

The results obtained from experiments and computer modeling show that the presence of a thermal convective flow affects the distribution of the laser beam intensity profile. The proposed processing methods, correlation and integral-difference, have demonstrated their effectiveness in registering fluctuations in the beam profile intensity distribution. The most effective thermal flow detection occurs at a distance of 1-2 meters; beyond this range, despite an increase in beam profile fluctuations, the effectiveness of the proposed methods decreases. This necessitates the development of more advanced signal processing models using matrix methods, as well as statistical methods applied in optical information processing devices [5].

A promising research direction in this field is the application of machine learning methods [6-7] for creating databases of laser beam profile fluctuations under the influence of various types of convective flows.

References

1. Kazakov, V.I. "Computer Processing of Laser Beam Profile Dynamics Changing During Interaction with a Thermal Convective Flow," Lecture Notes in Networks and Systems, 2023. Vol. 597. P. 972–979.

2. W. M. Isterling, B. B. Dally, Z. T. Alwahabi, M. Dubovinsky, D. Wright, "Beam displacement as a function of temperature and turbulence length scale at two different laser radiation wavelengths," Appl. Opt. 51, 55–63, 2012.

3. Krechov A.A. et al. "Application of a Heat Flow Control Sensor for Early Detection of an Explosive Gas-Air Mixture Under Low-Temperature Conditions," Siberian Fire-Rescue Bulletin, 2019, No. 2, pp. 33-38.

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. V.I. Kazakov, D.O. Moskaletz, O.D. Moskaletz, V.B. Romashova. "Acousto-Optic Modulator as an Element of Signal Processing Systems in the Radio and Optical Range," Proc. SPIE, Vol. 10395, *Optics and Photonics for Information Processing XI*, 1039515, 2017.

6. Xavier K. L. B. L., Nanayakkara V. K. "Development of an Early Fire Detection Technique Using a Passive Infrared Sensor and Deep Neural Networks," *Fire Technology*, 2022.

7. Saeed F. et al. "Convolutional Neural Network-Based Early Fire Detection," *Multimedia Tools and Applications*, 2020.

PROPOSAL FOR AN INTELLIGENT IRRAGATION SYSTEM

Davide Sambito, Erica Schembri
Computer Engineering – Kore University of Enna – Italy
Email: {davide.sambito, erica.schembri}@unikorestudent.it

Abstract. *The use of water in daily life and green space management is highly relevant in the context of the water crisis affecting our country. This document proposes an intelligent irrigation system that utilizes environmental sensors and ICT tools to help end users reduce water waste and improve irrigation efficiency by leveraging process automation. This study presents how an automated irrigation system operates based on data received from a soil moisture sensor and a wind sensor. The main objective is irrigation management aimed at avoiding water waste while ensuring maximum coverage of the irrigated area. The system will dynamically regulate the water flow according to environmental conditions, deciding whether to reduce or increase irrigation when optimal conditions are not met. Simulations and experimental results from the project analysis will be presented.*

Introduction

Water resources are a fundamental asset for all living beings. The proposed irrigation system is designed for domestic use to reduce waste through controls that adapt to environmental variables. The system will be connected to the network to enable remote irrigation management via a web application. The combination of advanced algorithms and soft computing methods, which involve the use of fuzzy functions, optimizes the decision-making process, leading to improved system performance and efficiency. Unlike conventional irrigation systems that rely on predefined and limited conditions, the proposed solution offers a higher level of automation and adaptability, ensuring a more targeted and responsible use of water resources.

Related works

Some simple automatic irrigation systems only control the time and duration of irrigation, which can lead to water waste because they do not take soil conditions and climate details into account. Our system, on the other hand, assesses the actual needs of the soil and weather conditions using sensors, optimizing water usage.

The proposed approach

The smart irrigation system we have designed includes both a wireless and a wired network to collect and manage the data needed for system control regulation. The wireless network consists of a soil moisture sensor and a wind sensor that monitor environmental conditions and provide data to the system. The wired network includes the controller and the regulator, which make decisions based on predefined rules and data from the sensors to adjust irrigation according to the soil's needs. The objectives of our system focus on optimizing water usage and minimizing waste. The system utilizes sensor data to ensure that irrigation is carried out only when necessary and with the right amount of water. The system's management is based on a modular architecture integrating various control technologies, with fuzzy logic determining when and how to activate irrigation. This ensures that resources are used efficiently.

Scenario

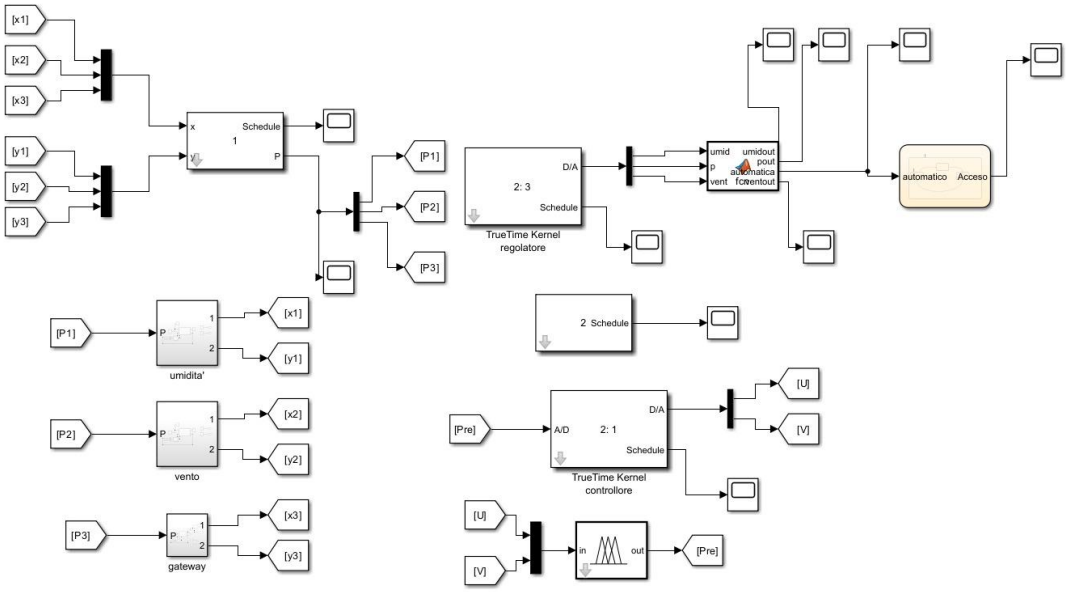
The system setup scenario we developed in Matlab concerns the real-world application of our system. The humidity block represents the soil moisture sensor. To simulate moisture readings, random numbers are generated. This sensor is powered by a battery. The wind block simulates the operation of the wind sensor and has the same characteristics as the humidity block.

The Gateway is responsible for collecting data from the sensors, evaluating performance metrics (packet loss and response time). It is also battery-powered.

The information from the sensors, once received by the controller, will be the input for the Fuzzy function. This function classifies humidity and wind values as low, medium, or high (from 0 to 33.3 as low, from 11.77 as medium, and from 58 to 100 as high), applies rules to determine whether to irrigate, and generates an output value to control irrigation and water pressure.

Another block is the regulator, consisting of a TrueTime Kernel with three outputs: humidity, pressure, and wind. At the output, demultiplexing is performed, and the obtained values are fed into the Matlab Function.

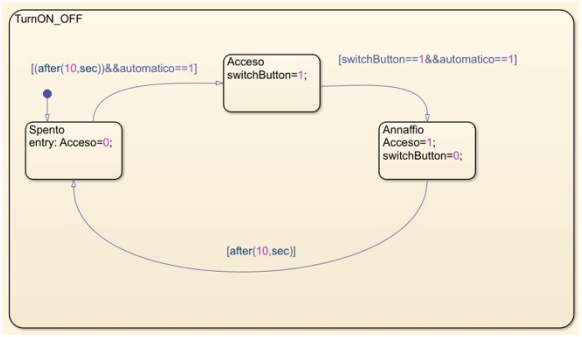
In the Matlab Function, input control is performed to determine the automatic variable to be sent to the Chart block. This variable is determined according to the following algorithm:



If humidity is below 30, irrigation is activated (automatica = 1).

- If wind is greater than 80 and pressure is above 50, the output pressure increases by 20.
- If wind is below 20 and pressure is above 50, the output pressure decreases by 20.

If wind is greater than 90 and pressure is below 10, irrigation is deactivated (automatica = 0).



The MBSD is responsible for controlling the activation and deactivation of the sprinkler based on the value of the variable received from the Matlab Function. There are three system states:

- Spento: Initial state. After 10s, if automatic = 1, it transitions to the On state.
- Acceso: When the system enters this state, the Switch Button is set to 1.
- Annaffio: This state is entered when the Switch Button is 1, and after 10s, it returns to the Off state.

Performance evaluation

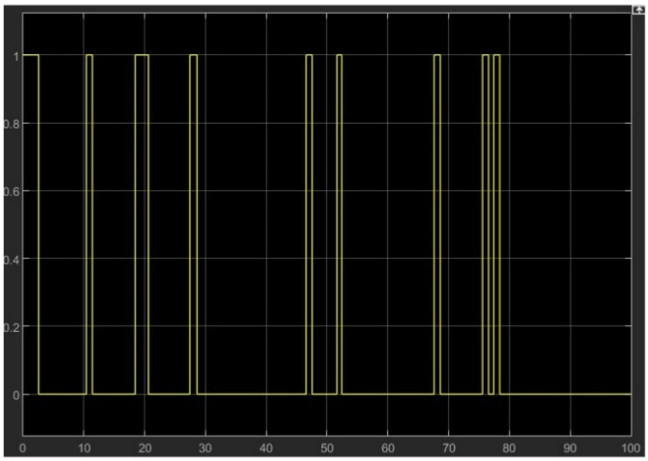
The performance in computer networks is important to ensure the efficiency and reliability of the system. In this project, we focused on performance related to wireless networks as they are more prone to problems compared to wired networks.

A first evaluation is carried out through the count of lost packets. It is important to understand if the data has been forwarded to highlight system problems.

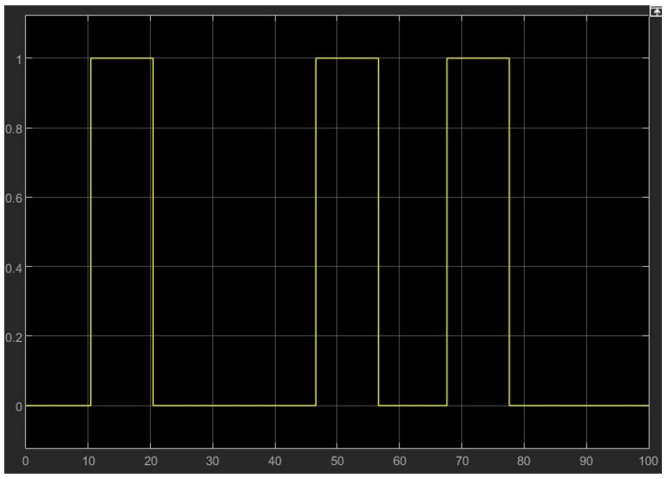
Another performance metric analyzed is the Response Time, which is the time elapsed between the sensor sending the data and the controller starting to use this data.

Both performances have been implemented on the Gateway.

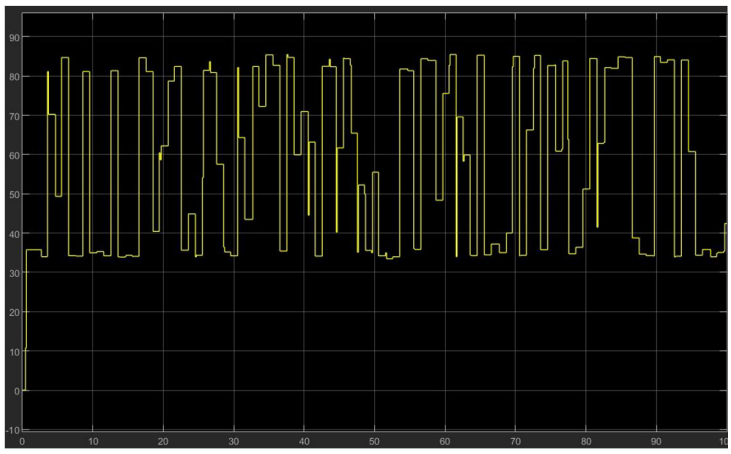
From the 100-second simulation of the irrigation system, we report the following results:



In this graph, the output of the Matlab Function of the variable “automatica” is shown. When the graph shows the value 1, the situation has occurred: low soil humidity and normal wind. Instead, when the value 0 is displayed, the conditions for irrigating the soil have not occurred (sufficiently humid soil or too strong wind with low pressure).



In this graph, the trend of the output variable from the MBSD called “Acceso” is shown. The initial state is off, in fact, a constant 0 is read for the first 10 seconds, then the system switches to the on state because the Switch Button of the MBSD is 1 and the variable “automatica” is 1. It will remain in this status as long as the hypotheses for irrigating the soil are met. Then, it will turn off, returning to 0.



This graph shows the trend of the pressure variable that varies over time based on the wind value. The control is carried out by the Matlab Function, which regulates the pressure based on fixed conditions explained in the previous paragraph (IV).

The analysis of the graphs shows that the system responds correctly to the simulated conditions, making the right decision and adjusting the pressure based on the detected environmental parameters.

Conclusions

The system illustrated, as discussed in Section II, offers numerous advantages over traditional irrigation systems by combining time control with soil conditions and climate factors.

Through a network of sensors, it integrates intelligent water management, preventing waste and skipping irrigation when it is not needed.

The possibility of remote management allows the end user to receive real-time notifications on the system's status, set customized parameters, and monitor water consumption, enabling a more efficient and conscious approach to irrigation.

The initial cost is higher compared to a traditional system, but given the benefits in terms of water savings and operational optimization, the investment can be easily recovered.

The smart irrigation system represents a significant step toward a more efficient and automated home.

For possible future developments, the irrigation system could be integrated with home automation solutions, allowing control through voice assistants and interaction with other smart devices, such as a robotic lawn mower.

Furthermore, the system could be further enhanced by integrating artificial intelligence algorithms capable of analyzing historical data and weather forecasts to optimize water usage even more. The adoption of solar-powered batteries would make the device completely autonomous and sustainable, reducing environmental impact and long-term operating costs.

References

[1] J.F. Kurose, K. W. Ross, “Reti di calcolatori e internet. Un approccio top-down”, Pearson, 7a edizione

[2] Andrei S. Tanenbaum, David J. Wetherall, “Reti di Calcolatori”, Pearson, 5 edizione

[3] https://www.rain.it/news-impianti-di-irrigazione/smart-irrigation-metodo-intelligente-irrigazione/

DETECTION OF SPACE OBJECTS ON VIDEO SEQUENCE OF THE STARRY SKY

Anton Shchukin
Department of Computer Technology and Software Engineering
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russian Federation

Abstract. *This paper proposes an algorithm for detection of space objects on video sequence of the starry sky. Spots are detected using the Lucas-Kanade Optical Flow and MOG2 background subtraction algorithms, object classification is performed using a Support Vector Machine (SVM) classifier.*
Keywords: *Calculating centroid of arbitrary shape, Lucas-Kanade Optical Flow detection algorithm, MOG2 sub tractor background, Support Vector Machine classificator.*

Data Description

This scientific paper uses several simulated videos of the starry sky created using the dynamic star field simulator from NASA Marshall Space Flight Centers Dynamic Star Field Simulator. The repository contains 4 video sequences with different video lengths, different numbers, positions, and velocities of stars and space objects. For each video, there is a MAT-object file that contains information about each frame of the video sequence, coordinates of the centroids of all figures in the video, divided into “Star” and “Object” classes. [1]

Practical section

The work consists of four subtasks: classifier training, features extract from the video sequence, object classification, and displaying the classification.

Two videos with pre-marked data were used to train the classifier: for each frame, the coordinates of the object centers and their classes were specified. All data is stored in MAT-object files. Based on this data, key features were extracted to distinguish stars from other objects. Since objects closer to the camera move differently than background stars, their trajectory and velocity are different from the total mass of points in the video. To enable the classifier to effectively distinguish objects in different videos with different object velocities, two meaningful features were defined:

- The difference between the velocity of a particular point and the average of the velocities of all objects in the video. (mag_diff)
- Difference in the trajectory of the object relative to other points. (angle_diff)

Fig. 1 shows the overall algorithm for computing the features for each video. To avoid undertraining and overtraining the model, 200 features from one video where objects have the fastest speed and 700 features from another video where objects have the slowest speed were selected. The implemented split features method splits the samples into balanced samples, so at the time of training the classes are balanced. Also the total sample of 900 points is divided into training and test sample, the test sample is 20% random samples which are not used in training the model.

The Support Vector Machine (SVM) [2-4] was chosen as the classification method for point classification. The main objective of the method is to find such a hyperplane that preserves the original distribution of points in the feature space to the greatest extent possible without additional assumptions. SVM has several important hyperparameters:

- Kernel – is used to denote the general algorithm of the classifier. There are several popular kernels: Liner kernel – the simplest kernel, it is often used in tasks with linearly separable data; Polynomial kernel – kernel that introduces polynomial function; RBF – the most common kernel that can separate data that are not linearly separable, it creates a decision boundary in the form of a radially symmetric bell; Sigmoid – kernel that is used for modeling neural networks.
- C – smooths the straight line and adjusts the classification accuracy. A low value of C makes the decision surface smooth, while a high C tends to classify all training examples more correctly.
- Gamma – the parameter indicates how much a particular example will influence the model.

Proper model selection directly affects the quality of object classification. The following figures show examples of feature space with different models.

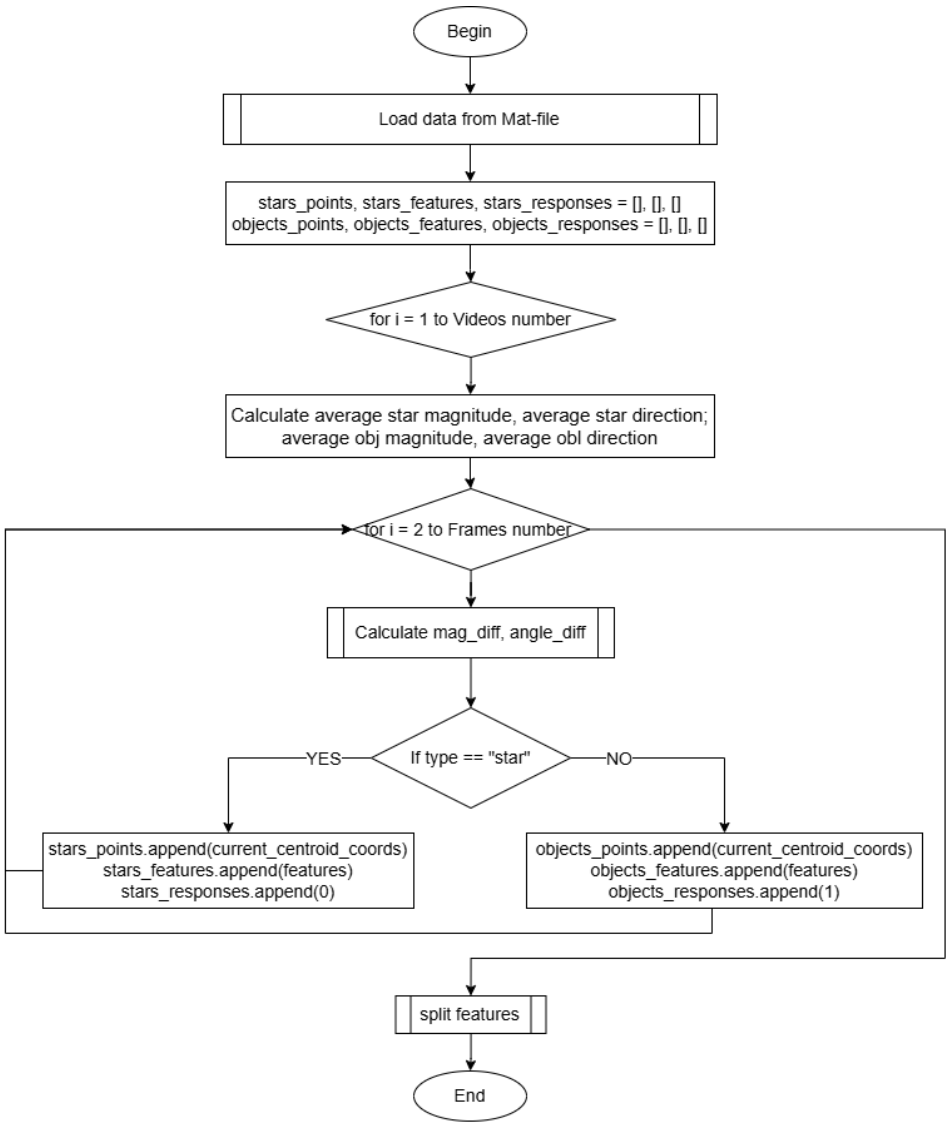


Fig. 1. Algorithm for creating features for the training sample

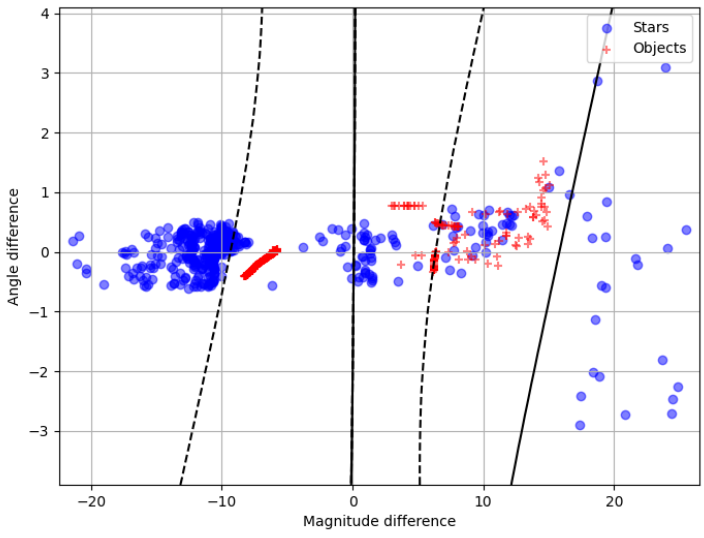


Fig. 2. Model 'C': 1, 'gamma': 'auto', 'kernel': 'sigmoid'

The model has the following metrics:

	precision	recall	f1-score	support
0 – Stars	0.54	0.87	0.67	97
1 – Space objects	0.48	0.14	0.22	83
accuracy			0.53	180

Judging by Fig. 2 and the metrics, we can say that the Sigmoid kernel is not suitable for solving this problem.

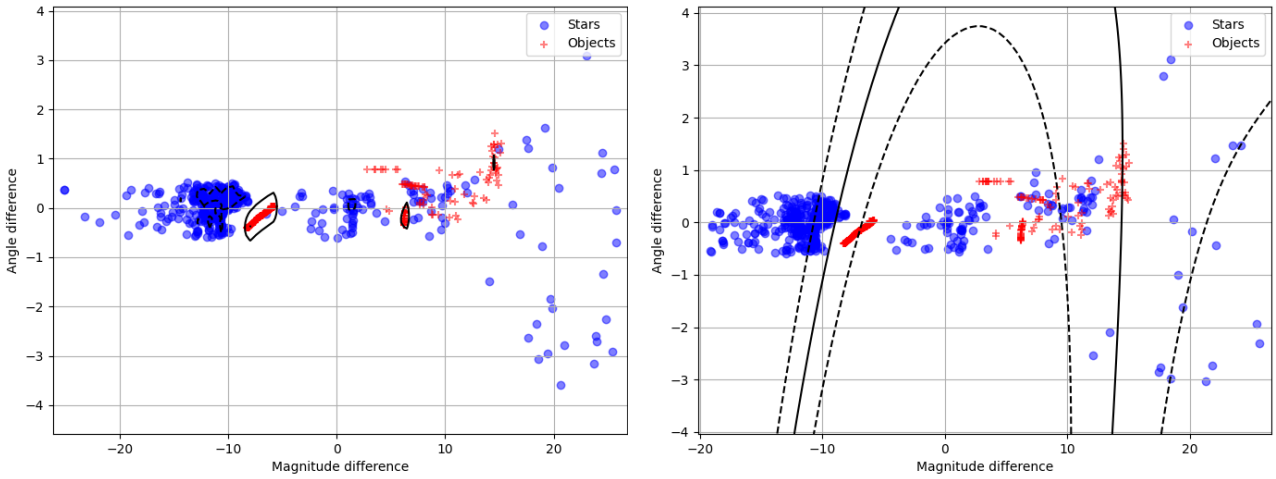


Fig. 3, 4. Models with RBF kernel, but misplaced hyperparameters C and gamma

Figures 3 and 4 show the results of overfitting and underfitting, respectively. The figure shows a small value of C and a large value of gamma, which led to the fact that the model tried to define classes as accurately as possible, without considering generalization. In contrast, Fig. 4 shows a too small gamma value and a large C value, which resulted in too smooth graph and a large number of errors.

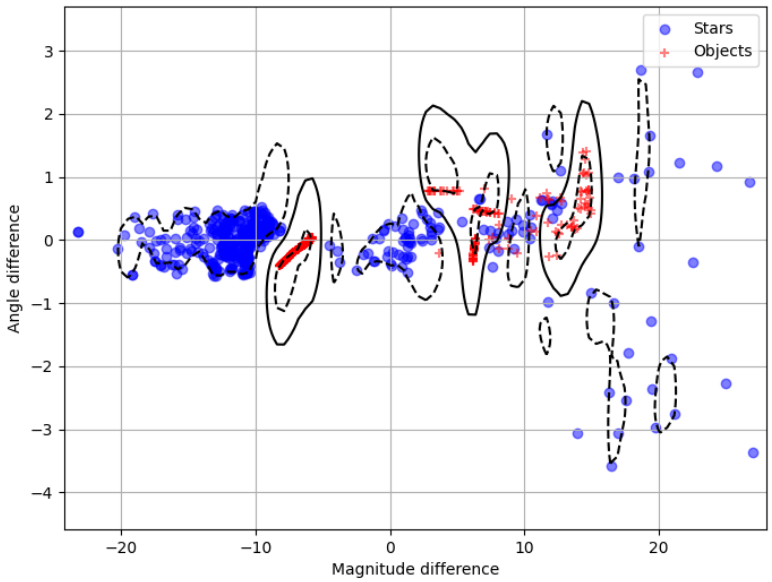


Fig. 5 – Model 'C': 1, 'gamma': 1, 'kernel': 'rbf'

In the process of searching for the optimal model, it was decided to focus on this model, which has better labels and better generalizing ability compared to other models.

	precision	recall	f1-score	support
0 – Stars	0.98	0.99	0.98	92
1 – Space objects	0.99	0.98	0.98	88
accuracy			0.98	180

Next, the classifier applies on a video sequence. To do this, the features first should be computed, then the class labels should be obtained from the classifier, and the results should be visualized on the video.

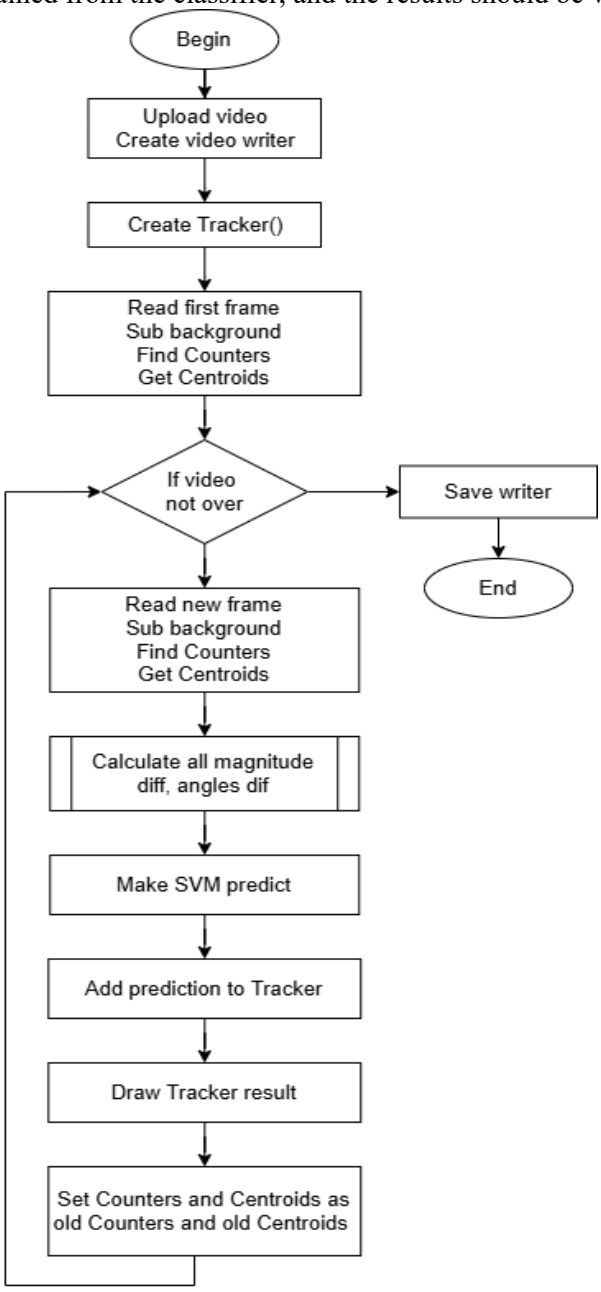


Fig. 6. Algorithm for generating features and recording classification results

In order to process the frame and identify the objects, the background must first be subtracted from the image, for this purposethe MOG2 background substrate algorithm is used, which determines and subtracts the background mask from the image based on luminance and pixie change rate, returning a layer of objects.

To analyze points on the video, the Lucas-Kanade optical flow detection algorithm algorithm is used [5], which allows using the coordinates of the centroids of points to track where the tracked point moved during the time period t, in this case, this is 1 frame. Knowing the previous and current coordinates, you can determine the velocity and angle of movement of the object.

After running the algorithm, a video of the classification of stars and objects can be obtained.

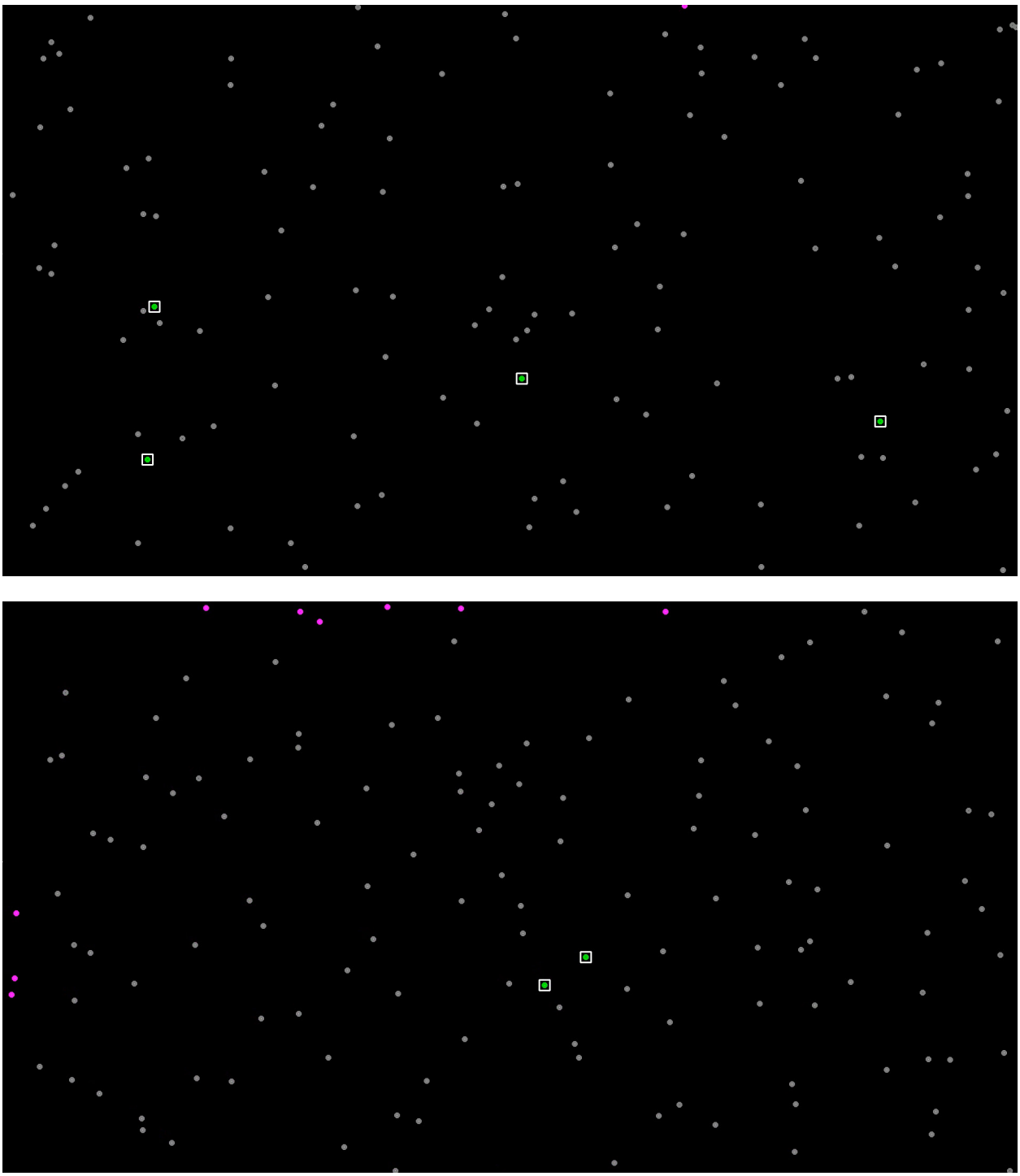


Fig. 6, 7. Marking of objects and stars in the video

Figures 6 and 7 show "Stars" marketed in grey and "Objects" marketed in green, with a white rectangle to make them more visible. Purple spots are objects that have not been classified yet.

Conclusions

For the classification of stars and space objects, an SVM classifier with an RBF kernel, and hyperparameters C = 1 and gamma = 1 was implemented, which in comparison with other models showed better classification accuracy = 0.98. Also, an algorithm that generates features from the video and using the classifier, creates a copy of the video with object labels was implemented. This development will be useful for scientific research of the near-Earth's orbit, as well as for adjusting the route of spacecraft, eliminating the large influence of human error.

Literature

1. A Robust Vision-based Algorithm for Detecting and Classifying Small Orbital Debris Using On-board Optical Cameras. Simulated environments generated using NASA Marshall Space Flight Center's Dynamic Star Field Simulator of on-board optical sensors and cameras. URL: https://github.com/nnategh/AMOS://github.com/nnategh/AMOS2019 /tree/master/assets/data (accessed: 24.02.2025)

2. Siji Joseph, Arun Pradeep – Object Tracking using HOG and SVM. Axis College of Engineering and Technology, 2017, 5 p.

3. Christopher M. Bishop. Pattern Recognition and Machine Learning, 2006, 758 p.

4. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong. Mathematics for Machine Learning, 2020, 417 p.

5. Saba Rabiee, Parisa Rangraz, Ahmad Shalbaf. Assessment of Carotid Artery Vibrations by Using Optical Flow Methods on Ultrasound Images and Optical Flow//FRONTIERS IN BIOMEDICAL TECHNOLOGIES. 2024, Vol. 11, No. 4, 631-641 p.

NETWORK SCENARIO FOR WEATHER PREDICTION AND ALERT

Alessio Salvatore Vetro, Giuseppe Alaimo
Computer Engineering and Networks Laboratory – Kore University of Enna – Italy
Alessiosalvatore.vetro@unikorestudent.it Giuseppe.alaimo002@unikorestudent.it

Abstract. *The increase of natural disasters due to climate change, needs to predict possible catastrophic phenomena with state-of-the-art systems that can detect natural hazards and their early manifestations in order to intervene as quickly as possible and be able to save the population of areas at risk.*
This study aims to simulate a network scenario between a radar, which provides values of atmospheric pressure and cloud speed; a processing system that takes care of analyzing the incoming data and processing it; a safety system where, based on the values processed by the previous systems, operations will be carried out to control the water level of a river and prevent possible flooding.

Introduction

Climate change and humans' transformations on the environment are affecting more frequently extreme weather events such as storms, hurricanes and subsequent flooding. These events mean a significant threat to communities located along waterways and coastlines. Effective weather forecast management and infrastructure protection require advanced monitoring and control systems that can detect hazardous conditions early and activate mitigation measures. In this context, this project proposes the development of a program based on Simulink TrueTime, which by monitoring various atmospheric and environmental parameters, such as water level and cloud speed, in real time and using fuzzy controllers and a MBDS for data analysis, is able to determine an alarm parameter indicating flood risk. Once the degree of alarm is determined, it is transmitted to the kernel of a security system that provides the activation of protective barriers in order to manage the flow of water and prevent flooding.

Related works

One of the studies, which this project is related to, is [1] main results of which are how temperature, humidity and pressure parameters processed by two fuzzy controllers, one using sugeno methodology and another using mamdani methodology, are more accurate with the last mentioned. This methodology is also studied in [2] where the values of temperature and velocity were analysed to predict the amount of rainfall.

Finally, the proposed paper takes into consideration [3] particularly with regard to the types of clouds and their dynamics. This analysis provides a more comprehensive understanding of cloud behavior and its implications for broader meteorological phenomena.

The proposed approach

The project begins with the activation of a radar system, which serves as the primary data acquisition source. This radar operates periodically, generating random values every 100 seconds for key meteorological parameters. These parameters include the speed of the clouds, the atmospheric pressure at the location of the clouds, and the atmospheric pressure at a river's location that is identified as being at risk of flooding. The collected data is then transmitted to a gateway and subsequently received by the wired network kernel of the processing system. The primary role of this kernel is to process the incoming data to generate an alert signal based on the analysis of the transmitted values. The data processing sequence starts with a fuzzy controller that implements the Mamdani methodology for its operations. This fuzzy controller produces an output that is further approximated into an integer value through a dedicated mathematical function. The integer value is then fed into the MBDS, a component designed to mitigate the risks of false alarms or transient critical conditions. The MBDS achieves this by filtering out unreliable or short-lived signals, ensuring that only verified and stable alerts are issued. The final step of data processing involves integrating the output of the MBDS with another parameter: the water level of the river, as received from the security system. This integrated dataset is analyzed by a secondary fuzzy system, also based on the Mamdani methodology, to calculate the final alarm value. This value represents the ultimate determination of the system's analysis and is transmitted through the alarm system kernel via a wired network to a second gateway, referred to as Gateway 2.

From here, it is distributed to two endpoints: the security system kernel for managing the river barriers and a parallel system responsible for issuing a weather alert notification. In addition, the security system continuously monitors the river's water level in real time and transmits this data back to Gateway 2. The alarm system subsequently receives this information and integrates it into the analysis conducted by the fuzzy controller. This feedback loop ensures a comprehensive and dynamic assessment of the environmental situation, enabling the system to respond effectively to changing conditions.

Scenario

The case above is based on a paper developed in Matlab/Simulink/TrueTime following the theoretical approach proposed in Section III. Specifically, with regard to radar, it consists of 3 random number generators:

-speed: indicates the speed at which the clouds are moving, it can take values from -

100 km/h to 100 km/h depending on the direction of displacement i.e. towards the river or in the opposite direction.

-pressure xi: indicates the atmospheric pressure in the area where the clouds were detected, values are taken from 950 hp (low atmospheric pressure area) to 1100 hp (high atmospheric pressure area).

-pressure xf: indicates the atmospheric pressure in the area where the river to be monitored is, values from 950 hp (low atmospheric pressure area) to 1100 hp (high atmospheric pressure area) are assumed.

The values are generated randomly every 100 seconds using a uniform random number generator. These values are then transmitted through a wired network kernel to a gateway. At the receiving end, the kernel of the processing system retrieves these values and processes them further. The processing system outputs the previously transmitted values, which are then analysed by a fuzzy controller employing the Mamdani methodology. This fuzzy controller calculates an output value ranging from 1 to 5, which is subsequently approximated using a matlab function. The approximate output from the fuzzy controller undergoes an additional verification step performed by the MBDS. This step ensures reliability by eliminating potential false meteorological alarms, specifically by disregarding alarm reports that persist for less than 50 seconds. This precaution minimizes the probability of transmitting false alarm. To determine the actual alarm value, the monitored river level must be evaluated along with the previous alarm value. The river level, is randomly generated every 100 seconds and can vary from 0 (indicating a normal level) to 20 (indicating a high-risk level). This approach ensures that the alarm system not only responds to meteorological data but also considers contextual environmental factors, enhancing the accuracy and robustness of the alarm mechanism. The security system is responsible for detecting the values of the river and transmitting them via wired network kernel to a gateway and forwarding them to the alarm system to generate the final alarm value via a fuzzy with mamdani methodology. The actual alarm value is transmitted via a wired network kernel, originating from the alarm system and directed to the gateway. From there, it is concurrently received by the kernel of the security system as well as the kernel of a weather alert system, enabling both systems to process the alarm data simultaneously.

Within the security system, a specialized matlab function has been designed to handle the alarm data. This function takes two critical inputs: the output received from the kernel, which represents the processed alarm value, and the previous status of the security barriers. The status of these barriers indicates whether they were activated or deactivated during the previous cycle. By integrating these inputs, the matlab function evaluates the current threat level and determining the necessary security response. This approach allows the system to dynamically adjust its actions based on real- time alarm data as well as historical barrier statuses.

Performance evaluation

Performance metrics are an important indicator of the quality and efficiency of a network scenario. The metrics measured in this paper are: response time, reaction time, and troughput. Response time refers to the time taken from when packets containing data of environmental parameters detected by the radar are sent to when they are actually received by the security system.

This value is found by the difference between the time recorded when the packets are sent in the radar kernel and the time when they arrive at the security system kernel. The response time for the river’s value is the same for the weather value because they are elaborated toghether. Reaction time is defined as the time between sending packets from a source and the start of processing those sent packets.

In this case, two reaction times are measured:

- The first is the time between the radar and the processing system, i.e., the difference between the time of sending data from the weather radar kernel and the time of arrival of those data to the processing system kernel.
- The second is the time between the security system and the warning system, i.e., the difference between sending river level values from the security system kernel and the arrival time to the warning system kernel.

	Reaction time	Response time
Weather value	0.050207	0.075514
River's value	0.025309	0.075514

The last metric measured is the troughput:

	Time	Value
1	2500.000	2.207e+04
2	7500.000	2.207e+04

Troughput is defined as the amount of bytes transmitted per second, that is, units of data transmitted in the unit of time. It can be seen from the figure shown that the value measured at two different instants is about 22,070 bytes/sec.

Conclusions

Climate change is a current phenomenon: temperatures are increasing, precipitation regimes are changing, glaciers and snow are melting, and the global standard sea level is rising.

To summarize, this paper presents a network scenario that analyzes meteorological values such as cloud speed and two atmospheric pressures in order to estimate a degree of warning, resulting in an alert and implementation of optimized protective measures.

References

[1] Setyanugraha, N.; Al Aziz, S.; Harmoko, I.W.; Fianti, F. Study of a Weather Prediction System Based on Fuzzy Logic Using Mamdani and Sugeno Methods. *Phys. Commun.* **2022**, *6*, 61–70.

[2] Janarthanan, R.; Balamurali, R.; Annapoorani, A.; Vimala, V. Prediction of rainfall using fuzzy logic. *Mater. Today Proc.* **2021**, *37*, 959–963.

[3] https://hdl.handle.net/10589/131695

[4] V. C. Gungor and G.P. Hancke "Industrial Wireless Sensor Network: Challenges, Design Principles, and Technical Approaches" IEEE Transaction on Industrial Electronics, Vol.56 No. 10, October 2009

REMOTE VENTILATION CONTROL SYSTEM USING ESP32

Dmitry Vinogradov
Saint Petersburg State University of Aerospace Instrumentation
e-mail: vinmitya@gmail.com

Abstract. *This article presents the development and analysis of a remote-controlled ventilation system based on the ESP32 microcontroller. It examines the hardware architecture, methods of integration with smart home platforms, and testing results. The proposed system provides automated fan on/off switching based on sensor readings (temperature, humidity, etc.) and remote control via a web interface or mobile application. Owing to its low cost and ease of implementation, this solution is suitable for both household and office spaces.*
Keywords: *ESP32, smart home, IoT, microcontroller, remote control, fan, sensors*

Introduction

In modern smart home systems, there is a growing trend toward implementing automated devices that provide comfort and energy savings. One of the key areas is climate control. To maintain optimal temperature and humidity, fans are often used, which can operate either manually or fully automatically when equipped with a control system.

The goal of this work is to create an affordable remote fan control system based on the ESP32 microcontroller. The proposed approach allows the device to respond to changes in environmental parameters (e.g., humidity, temperature) and to control the fan via a local network or the internet.

2. System Architecture

The proposed system consists of the following main modules.

2.1. ESP32 Module

- Controls the fan power via a relay or transistor switch.
- Reads data from various sensors (temperature, humidity, CO_2, etc.).
- Connects to a Wi-Fi network to communicate with a local server or a cloud platform.

2.2. Sensor Module

- Uses temperature and humidity sensors (DHT22, BME280) or CO_2 (MH-Z19B).
- Sends readings to the ESP32 to determine whether to turn the fan on or off.

2.3. Software

- Embedded firmware in C/C++ using Arduino IDE or PlatformIO.
- Implements a web server or MQTT client for remote control and data transmission.
- Automation logic is triggered when humidity or temperature exceeds a specified threshold.

2.4. Control Interface

- A web page or mobile application for manual fan control and sensor data monitoring.
- Possible integration into Home Assistant, OpenHAB, and other smart home platforms (via MQTT/HTTP).

3. Implementation Methods and Testing

3.1. Hardware

- ESP32 DevKit as the central microcontroller.
- A relay to switch the fan load (could be 220 V or 12 V).
- Sensors (DHT22/BME280) to gather environmental data.
- A power supply for the ESP32 and the relay.

3.2. Software Logic

- Initializing Wi-Fi to connect to the local network.
- Reading sensor data and processing it (calculating humidity, temperature).
- Providing a web interface or MQTT for remote monitoring and control.
- Triggering the fan when humidity/temperature exceeds a specified threshold.

```
/*
  Example firmware code for ESP32
  Fan control via relay and DHT22 sensor
  Author: Dmitry Fedorovich Vinogradov
  Version: 1.0
```

```
*/

// Include the required libraries
#include <WiFi.h>
#include <WebServer.h>
#include "DHT.h"

// Wi-Fi credentials (replace with your own)
const char* ssid     = "Your_SSID";
const char* password = "Your_PASSWORD";

// Pin settings
#define DHTPIN 4       // DHT22 sensor pin
#define DHTTYPE DHT22  // DHT22 sensor type
#define RELAY_PIN 2    // Relay pin for the fan

// Humidity threshold (can be adjusted as needed)
float humidityThreshold = 60.0;

// Create objects
DHT dht(DHTPIN, DHTTYPE);
WebServer server(80);

// Root page handler
void handleRoot() {
  float h = dht.readHumidity();
  float t = dht.readTemperature();

  // Create a simple HTML page
  String htmlPage = "<!DOCTYPE html><html><head><meta charset='UTF-8'><title>Ventilation
Control</title></head><body>";
  htmlPage += "<h2>Sensor Data</h2>";
  htmlPage += "<p>Humidity: " + String(h) + "%</p>";
  htmlPage += "<p>Temperature: " + String(t) + "C</p>";

  // Buttons for manual control
  htmlPage += "<h3>Fan Control</h3>";
  htmlPage += "<button onclick=\"location.href='/fanOn'\">Turn On</button>";
  htmlPage += "<button onclick=\"location.href='/fanOff'\">Turn Off</button>";

  htmlPage += "</body></html>";

  server.send(200, "text/html", htmlPage);
}

// Turn fan on
void handleFanOn() {
  digitalWrite(RELAY_PIN, LOW); // Active level may differ for the relay
  server.sendHeader("Location", "/");
  server.send(303);
}

// Turn fan off
void handleFanOff() {
  digitalWrite(RELAY_PIN, HIGH);
  server.sendHeader("Location", "/");
  server.send(303);
```

```
}

void setup() {
  Serial.begin(115200);

  // Initialize the relay and sensor
  pinMode(RELAY_PIN, OUTPUT);
  digitalWrite(RELAY_PIN, HIGH);
  dht.begin();

  // Connect to Wi-Fi
  WiFi.begin(ssid, password);
  Serial.print("Connecting to Wi-Fi ");
  while (WiFi.status() != WL_CONNECTED) {
    delay(500);
    Serial.print(".");
  }
  Serial.println("\nWiFi connected. IP address: ");
  Serial.println(WiFi.localIP());

  // Set up the web server
  server.on("/", handleRoot);
  server.on("/fanOn", handleFanOn);
  server.on("/fanOff", handleFanOff);

  server.begin();
  Serial.println("Web server started");
}

void loop() {
  server.handleClient();

  // Automatic fan activation based on humidity
  float currentHumidity = dht.readHumidity();
  if (!isnan(currentHumidity)) {
    if (currentHumidity > humidityThreshold) {
      digitalWrite(RELAY_PIN, LOW); // Turn the fan on
    } else {
      digitalWrite(RELAY_PIN, HIGH); // Turn the fan off
    }
  }

  delay(2000); // Update interval
}
```

3.3. Experimental Setup

Below is an approximate external view of the assembled system (Fig. 1).

For performance verification, a test model was assembled, which includes:

- ESP32 DevKit V1;
- 5 V relay;
- 12 V fan (for safety during testing);
- DHT22 sensor.

Testing was conducted in a room of about 12 m² at various humidity levels. The system was configured to activate the fan at humidity levels above 60%. Over the course of the five-day experiment, several activation cycles were recorded. Remote control was performed through the web page hosted on the ESP32. When integrated with Home Assistant, push notifications were sent each time the fan switched on.

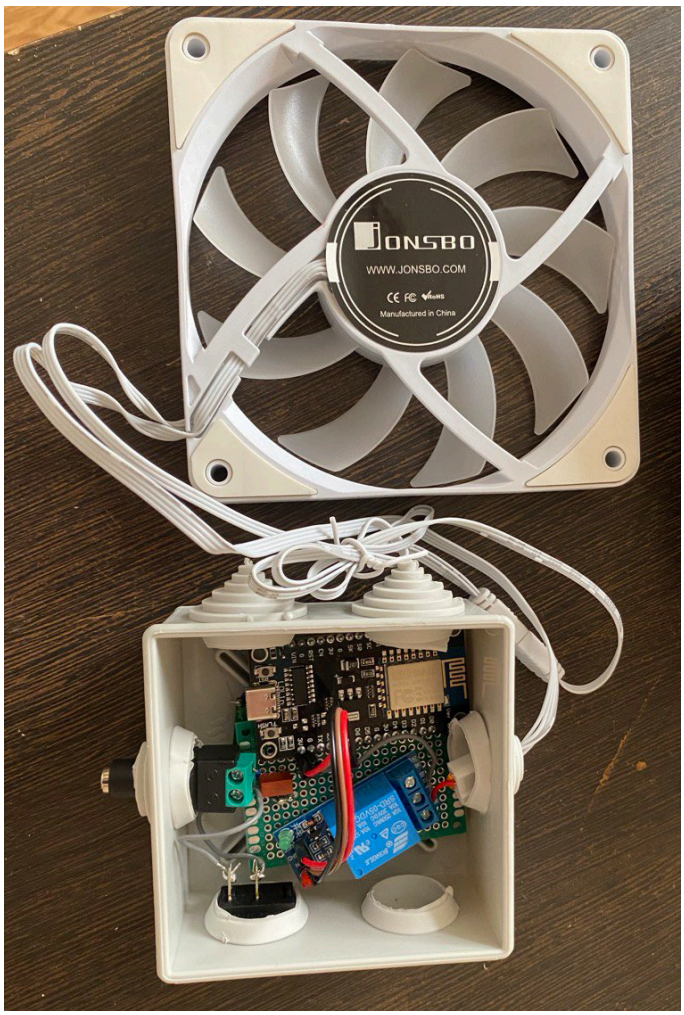


Fig. 1. View of the completed system

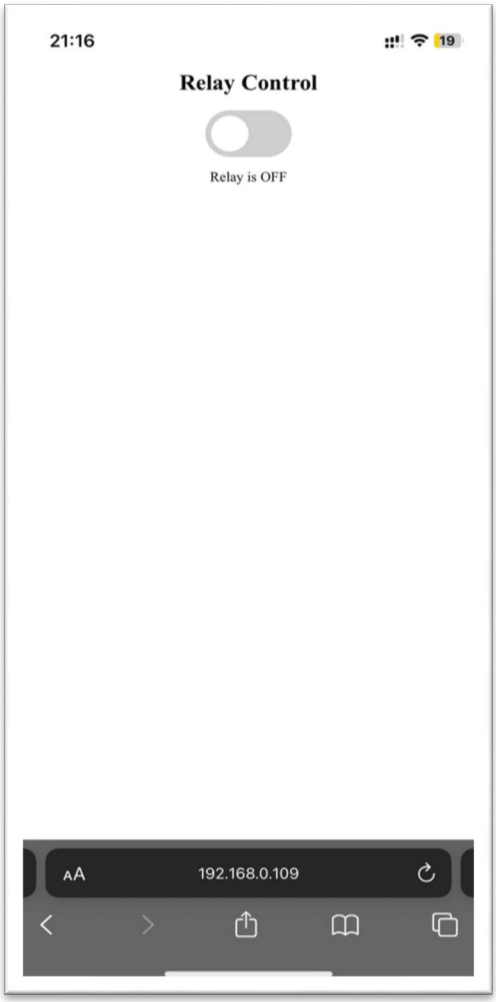


Fig. 2. Control interface

A comparative analysis of various methods for fan control (ranging from a simple timer and relay to an ESP32-based system with sensors) demonstrated that the proposed option offers extensive automation and control capabilities at a relatively low cost.

4. Integration into a Smart Home

Integrating the system into an existing smart home platform offers several advantages:

- Automatic adaptation: If humidity rises above the threshold, the fan switches on automatically.
- Expanded scenarios: Complex logic chains can be created (e.g., coordinating a kitchen exhaust fan and a bathroom fan).
- Parameter monitoring: Sensor readings are sent to the smart home server, allowing graphing and event notifications.
- Energy savings: A properly configured ventilation algorithm reduces heating/cooling costs.

Conclusion

The ESP32-based system developed here combines ease of implementation, low cost, and broad functionality for ventilation automation. Tests have shown that the device quickly responds to changes in the indoor environment and reliably transmits data over the network. Thanks to its modular design, the solution can be easily expanded with additional sensors and integrated into popular smart home platforms.

Key advantages:

- Low cost and component availability.
- Flexible configuration for both local and remote control.
- Automatic activation based on sensor data, as well as manual control.
- Full compatibility with Home Assistant, OpenHAB, and others.

In the future, the plan is to add more sensors (e.g., gas sensors or air quality sensors), implement energy-saving modes for the microcontroller, and develop a mobile application to simplify system configuration.

References

1. ESP32 Official Documentation [Electronic resource]. URL: https://docs.espressif.com/projects/esp-idf/en/latest/esp32 (accessed: 18.02.2025).

2. Arduino for ESP32 [Electronic resource]. URL: https://github.com/espressif/arduino-esp32 (accessed: 18.02.2025).

3. MQTT – Message Queuing Telemetry Transport [Electronic resource]. URL: https://mqtt.org/ (accessed: 18.02.2025).

4. Home Assistant Documentation [Electronic resource]. URL: https://www.home-assistant.io/docs (accessed: 18.02.2025).

AUTOMATED ANALYSIS OF TEXT DATA BASED ON MACHINE LEARNING ALGORITHMS

Roman Voronov
Saint Petersburg State University of Aerospace Instrumentation,
St. Petersburg, Russia
E-mail: rom.stud.spb@gmail.com

Abstract. *The paper deals with the method of automatic analysis of text data using machine learning algorithms. With the growing volume of unstructured data associated with the development of digital media, efficient processing of such data becomes an urgent task. Machine learning approaches for information retrieval, analyzing text tonality, creating chatbots and other methods of natural language processing are being investigated. The aim of the research is to develop methods and algorithms that provide high accuracy and performance in real-world conditions.*
Keywords: *machine learning, natural language processing, text data analysis, text classification, algorithms.*

Introduction

The amount of textual data is increasing every year, which creates challenges in various domains such as social media, scientific publications, user reviews and others. Processing this data manually requires significant time and human resources [1, 2]. Machine Learning (ML) provides efficient tools for analyzing large amounts of textual information by automating tasks related to Natural Language Processing (NLP) such as classification, tone detection and key aspect extraction.

These methods are actively used in search engines, chatbots, virtual assistants and other intelligent systems. The use of machine learning algorithms can improve the accuracy of analysis and optimize data processing, making it an important tool in the development of digital technologies. In response to growing data volumes, an approach (Fig. 1) has been developed that automates text analysis, minimizes errors, and speeds up decision-making.

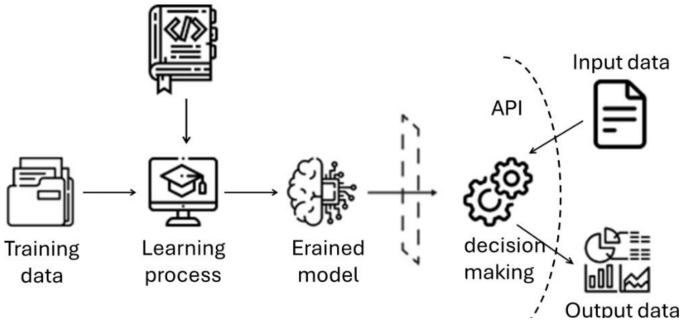


Fig. 1. Architecture of the algorithm for automated text analysis

Textual data preprocessing techniques

Basics of text data preparation

Effective analysis of textual data begins with its preparation (Fig. 2). Source text typically contains many unnecessary characters and redundant information. Basic steps include:

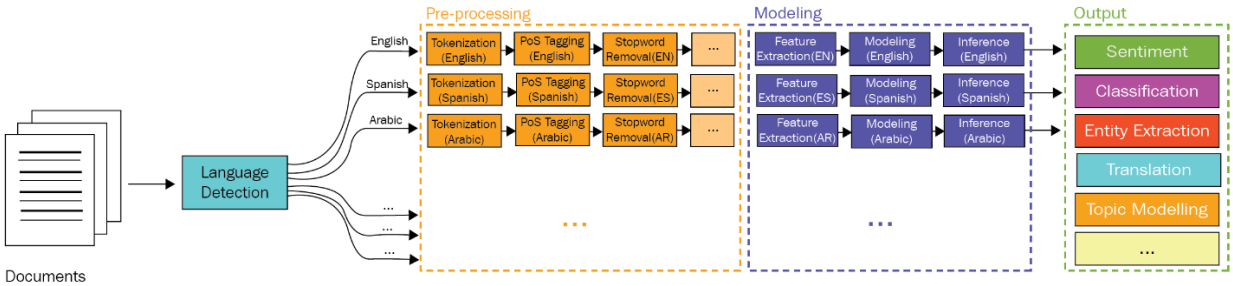


Fig. 2. A classic NLP flowchart: the pre-processing stage

1. **Text pre-processing**:
   - **Language detection –** analyzing text to identify language, which is important for selecting appropriate processing models.
   - **Tokenization –** breaking text into elements (words, phrases) to simplify analysis.
   - Part-of-speech – identifying parts of speech to account for grammatical context.**markup**
   - **Removing stop** words – eliminating frequently occurring but unimportant words.
2. **Text modeling**:
   - **Feature Extraction –** converting text to numeric format using methods such as TF-IDF and Word Embeddings.
   - **Model training and prediction –** text processing using machine learning models such as LSTM, Transformer, and BERT.
3. **Model selection**: Depending on the task, the following are used:
   - **Probabilistic methods** (Naive Bayes, HMM) – for simple classification problems.
   - **Classical models** (SVM, Random Forest) – to analyze tonality.
   - **Deep neural networks** (LSTM, Transformer, BERT) – for processing long texts.
4. **Training and Optimization**:
   - Splitting the data into training and test samples, training the model and optimizing it using methods such as Adam and SGD.
5. **Output**: after processing, the model can perform tasks such as:
   - Tonality analysis, classification, entity extraction, translation and topic modeling.

These steps ensure standardization of text data, making it suitable for efficient analysis and processing using machine learning algorithms (Fig. 3). Standardization includes eliminating heterogeneity and redundancy in the data, which is critical to achieving high performance models.

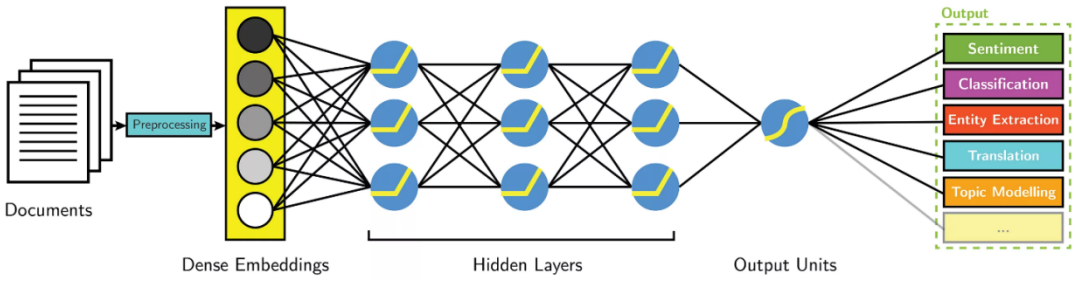


Fig. 3. Structure of neural network for text analysis

1. **Vectorization (Dense Embeddings)**: Text is converted into multidimensional numeric vectors, allowing models to preserve meaning and context. Methods include:
   - **TF-IDF –** determines the importance of words based on their frequency.
   - **Word Embeddings** (Word2Vec, GloVe, FastText) – create dense vector representations by considering semantic relations.
   - **Bag of Words –** represents text as a set of frequency vectors without considering word order.
   - **Contextual Embeddings** (BERT, GPT, T5) – create context-dependent embeddings where a single word can have different vectors depending on the environment.
2. **Processing by neural network model (Hidden Layers)**: Vector data is fed into deep neural networks that reveal hidden patterns.
   - **LSTM** is a recurrent neural network for processing text sequences.
   - **Transformer** (BERT, GPT-3, T5) – Uses the attention mechanism to effectively analyze text.
   - **Feedforward Neural Networks –** applied for final text classification.

These methods provide an efficient way to encode text, transforming it from an unstructured format to a numerical form that can be analyzed by machine learning algorithms. This transformation allows models to access the key data contained in the text and extract its key properties.

Evaluation of algorithm results

Comparative performance analysis of algorithms

The choice of algorithm depends on the characteristics of the data and the problem. The following metrics were used to objectively assess the quality of the work [6]:

- **Accuracy –** the proportion of correctly classified objects.
- **Completeness (Recall) –** important to minimize missing positive examples.
- **F1-measure –** balances accuracy and completeness, especially for unbalanced data.

Experimental results:

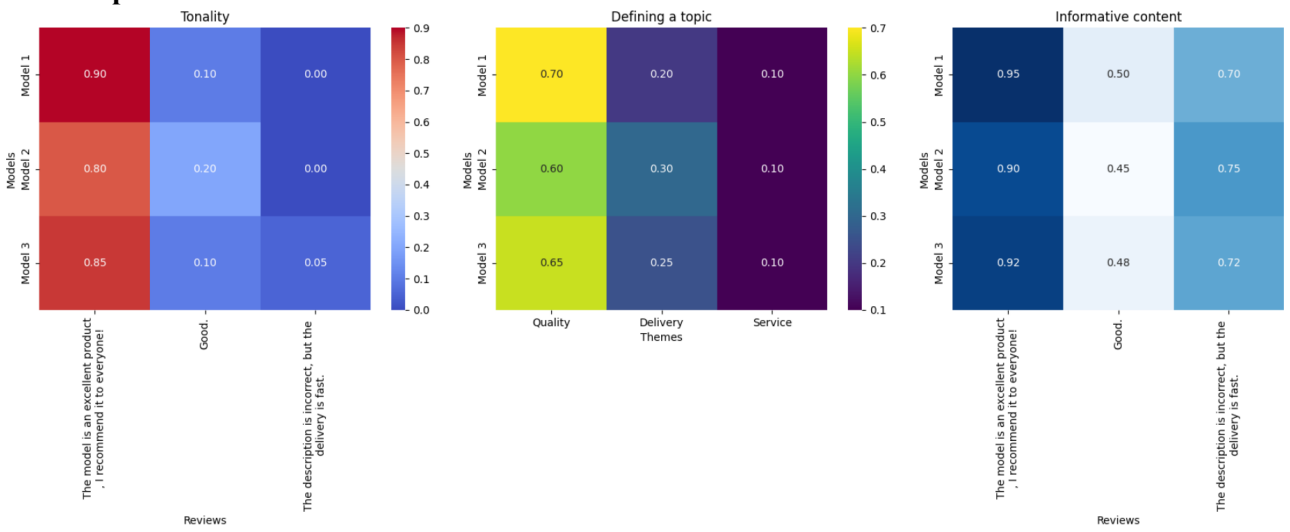


Fig. 3. Metrics for evaluating NLP models

1. **Text Classification**:
   - **SVM –** 92% accuracy on small structured data.
   - **Naive Bayes –** 88% accuracy on big data.
   - **Random Forest –** 90% accuracy on complex data.
2. **Tonality Analysis**:
   - **BERT** is 94% accurate due to contextualization.
   - **LSTM –** inferior to Transformers.
3. **Text clustering**:
   - **K-means –** effective on clearly separated clusters, but performs poorly with overlapping clusters.
   - **DBSCAN –** better on non-uniform data, but requires careful tuning of hyperparameters.

Experimental results showed that the choice of model depends on the specifics of the data. Traditional methods such as SVM and Random Forest are effective for structured data, while transformers such as BERT are better at analyzing context. Noise tolerant algorithms such as DBSCAN are suitable for clustering heterogeneous data [7]. Optimal model selection is critical for accurate and efficient text analysis.

Conclusion

Each of the considered machine learning algorithms has its own advantages and limitations, which influence their choice depending on the specifics of the task and data. For example, SVM is efficient on small structured data but requires large computational resources when processing voluminous data, while Naive Bayes is simple and efficient on large arrays but is limited in accounting for complex dependencies. Random Forest is versatile but can suffer from overfitting on noisy data, and BERTs (transformers) provide accurate context analysis, but their use is associated with high computational cost. LSTM performs well on sequential data but is inferior to transformers in terms of accuracy and speed. Clustering methods such as K-means and DBSCAN perform well on specific data, but require careful tuning for complex scenarios.

In the course of the research, a method for automated analysis of text data using these algorithms was developed, which demonstrated high accuracy in the tasks of classification, tone analysis and clustering. The results of the research allow automating the processing of large amounts of text data, reducing labor costs and providing practical applications in various fields such as information retrieval, user data analysis and document automation.

References

1. Brink, H. Machine learning [Text] / Richards Joseph, Feverolf Mark / Brink Henrik, Richards Joseph, Feverolf Mark – SPb.: Peter. 2017. p. 336.

2. Hobson, L. Natural Language Processing in Action [Text] / Hobson Lane, Hannes Hapke, Cole Howard. 2020. p.42.

3. Ivanko, A. F., Neural networks: general technological characteristics [Text] / Ivanko, A. F., Sizova, Y. A. // Scientific Review. Technical Sciences. 2019. No. 2. pp. 17-23.

4. Petrov, I. V. Application of neural networks in natural language processing / I. V. Petrov, E. N. Smirnova // Bulletin of Computer and Information Technologies. 2021. № 5. – pp. 67 72

5. Bolshakova E.I., Klyshinsky E.S., Lande D.V., Noskov A.A., Peskova O.V., Yagunova E.V. Automatic processing of texts in natural language and computer linguistics. M.: MIEM, 2011. 272 p. 522

6. Kyunghyun Cho. Introduction to Neural Machine Translation with GPUs (part 3). // NVIDIA Developer Blog [Electronic recource]. URL: https://devblogs.nvidia.com/introduction-neural-machine translation-gpus-part-3/ (22.12.2018).

7. Lewis, David D, 1998, Naive (Bayes) at forty: The independence assumption in information retrieval. In Proe, of the European Conference on Machine Learning (ECML) p. 4-15.

POSSIBILITIES OF THEORETICAL ESTIMATION OF OPTICAL RADIATION PENETRATION DEPTH IN MULTICOMPONENT BIOLOGICAL ENVIRONMENTS

Maria Zanevskaya
Saint Petersburg State University of Aerospace Instrumentation,
67 Bolshaya Morskaya St., Saint Petersburg, 190000, Russia
mnevskaya1@gmail.com

Abstract. *The paper presents theoretical calculation of the penetration depth of optical radiation into different types of biological tissue. Experimental study of the penetration depth of radiation for human biological tissues with different modification of the position of the radiation source and receiver is carried out. The obtained theoretical and experimental results can be further used for the development of effective monitoring systems of microcirculatory-tissue systems and confirm the prospects of non-invasive spectrophotometric diagnostics application in clinical practice.*
Keywords: *biological tissue, optical radiation, absorption, hardware-software complex.*

Introduction

The microcirculation system is designed to carry out metabolic processes between blood vessels and tissues, as well as the transport of nutrients contained in them to each cell of the human body [1]. It is on these processes that human health and performance depend. Evaluation of the microcirculatory channel activity allows us to judge the functional state of physiological systems of the human body, which is an indicator of the body's protection against diseases of the cardiovascular and respiratory systems. The main marker indicating deterioration of the functional state is deviation from the normal functions of the cardiorespiratory system, the main task of which is to provide the body with oxygen [2]. That is why monitoring the level of oxygenation of human biological tissues is a task that requires accurate and modern solutions.

Materials and methods

Nowadays, optical non-invasive methods are widespread, thanks to which it is possible to create accurate, compact and safe for human functional state assessment systems. Techniques such as optical reflectance spectroscopy, near infrared spectroscopy have significantly advanced the understanding of the haemodynamics of cutaneous, muscular, cerebral and gastrointestinal microcirculation. Among the promising non-invasive methods for analysing the state of biological tissues is the spectrophotometric method. It is based on the ability of various chemical compounds to interact with radiation by absorption. For further development of sensors for assessing the functional state of the organism, it is necessary to study and conduct theoretical studies of the interaction of optical radiation with human biological tissues. The processes of interaction of radiation with the biological environment are shown in Fig. 1.

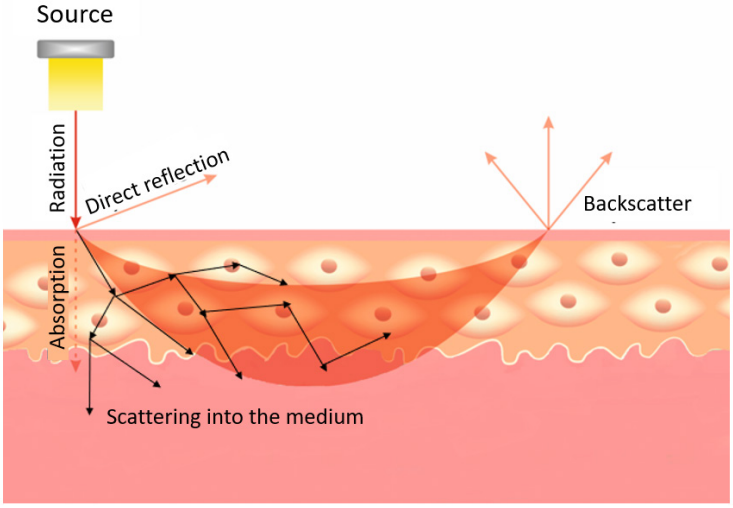


Fig. № 1. Interaction of optical radiation with biological tissue

The ability of substances to absorb electromagnetic radiation is determined by such factors as substance concentration, temperature (such a temperature difference as from 30 ^{o}C to 90 ^{o}C corresponds to a decrease in the absorption coefficient of water by 0.885 cm^{-1} [3]), radiation wavelength and thickness of the absorbing layer. In biological tissues, the main absorbing centres – chromophores – are water molecules, fat, haemoglobin and other structural components.

An important criterion that needs to be studied is the depth of penetration of radiation into biotissue, as it allows determining the optimal dose (sufficient to obtain the necessary information, but not harmful to the patient's health) and type of radiation during medical procedures.

Thus, as an object of research we should consider the process of interaction of optical radiation with tissues of human body, how the type of absorbing medium affects the depth of penetration of radiation and the possibility of using the results of this research in medical practice.

At the beginning it is necessary to clarify that biological tissue is a system including cells and intercellular substance, which are united by common origin and functions [4]. For the purpose of further practical application, this study considered those tissues in which microcirculation is most likely to be traced by spectrophotometry: dermis (epithelial tissue type), skin and blood vessels (connective tissue type)

The dermis is the middle layer of the skin, which in turn is formed by the reticular and papillary layers (lower and upper dermis respectively) Collagen and elastin are the main structural components of the dermis and provide it with strength and elasticity [5].

Blood is a liquid connective tissue consisting of 55% plasma and 45% blood cells. Formed blood elements are cells and postcellular elements, the most important function of which is the transfer of oxygen to internal tissues and organs due to red blood cells containing haemoglobin [6].

Blood artery includes three layers: intima, media, adventitia Their width is: 10^{-5} – 10^{-4} cm, $3 \cdot 10^{-3}$ – $2 \cdot 10^{-2}$ cm, $3 \cdot 10^{-3}$ – $3 \cdot 10^{-2}$ cm respectively.

Results and discussion

Theoretical calculations of the depth of penetration of radiation into tissue are made using the Bouguer-Lambert-Bera law, which states that the intensity of radiation passing through a medium decreases exponentially with distance:

$$I(z) = I_0 \cdot \exp(-\mu_a \cdot z), \quad (1)$$

where I_0 is the initial intensity of the incident radiation; I(z) – intensity of radiation at depth z; μ_a – absorption coefficient; z – depth of penetration of radiation (cm); d – width of the absorbing layer.

On the basis of this law it is possible to derive a formula for the penetration depth of radiation [7]:

$$h = \frac{1}{\mu_a}. \quad (2)$$

As a result of theoretical calculations of the depth of penetration of radiation into blood (at the volume water content in oxygenated blood of 90% [8] and values of blood oxygen saturation: 60%, 80%, 90%, 99%) we can conclude that the best absorption of radiation is blood with oxygen saturation value of 99% at a wavelength of 700 nm, where the depth of penetration reached 1.4 cm. At the same wavelength, blood with an oxygenation value of 80% absorbs radiation best, where the depth has reached a value of 1 cm and model 3 (blood oxygenation value of 90%), where the depth is 1.2 cm. At 99%, blood absorbs radiation best at a wavelength of 940 nm, where the depth of radiation penetration is 0.8 cm.

When considering different values of blood volume content in the dermis (0%, 5%, 10%, 15%) and a fixed value of water volume content (65%), the results of theoretical calculations of radiation penetration depth showed the fact that as the saturation of the dermis with blood increases, optical radiation loses its ability to penetrate into its deep layers. Also with increasing wavelength the depth of penetration of radiation into biological tissue increases.

In the following calculations the value of the volume content of water in the arterial vessel wall was taken as 80% and as a result of the calculation we can conclude about the ability to penetrate only the upper layer of the arterial vessel wall – the adventitia – regardless of the wavelength (the range of 400 – 940 nm was considered). At the same time, the light penetrates deepest into the arterial wall at a wavelength of 400 nm.

The results of the calculations below demonstrate the ability of biological tissues to absorb light of the optical range at different depths, which gives optical methods the opportunity to be used to assess properties such as saturation.

As mentioned earlier, there are many methods and instruments for measuring saturation. However, it is possible to combine non-invasiveness, compactness and patient safety by using wearable optical sensors based on spectrophotometry. By illuminating the sample with radiation sources of different wavelength ranges (from visible to infrared), the radiation would be backscattered back to the integrated photosensitive element. As a result of this process, it is possible to observe the spectral characteristics of backscattered radiation from different tissue depths [9]. It is possible to control the depth of penetration of radiation into the tissue by fabricating the device in several configurations that differ in the distance between the source and the radiation detector. The configurations perform measurements in the 'on reflection' mode, in which the associated sources and detectors are located on the same side of the biological tissue at a certain distance from each other. With such an arrangement of the optical system elements, the character of the optical path from the source-detector pair corresponds to an arc-shaped curve [10]. At that, the penetration depth of radiation into biological tissue is approximately half of the distance between the transmitter and the optical detector (figure 2) [11].

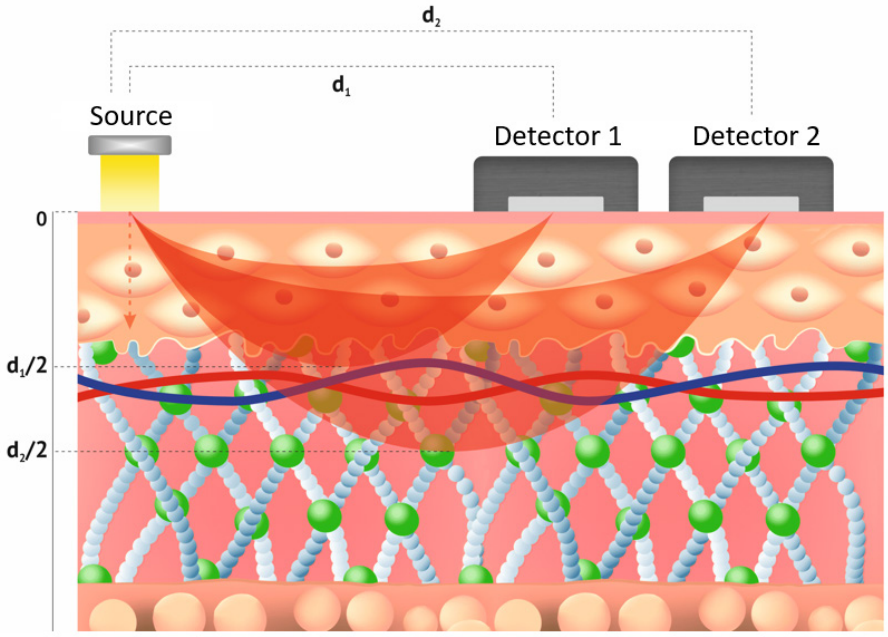


Fig. 2. Schematic of the hardware-software complex operation in the 'on reflection' mode with the trajectories of the detected radiation propagation at different source-detector distances

This development was tested in an experimental study aimed at investigating the spectral characteristics of biological tissues at different penetration depths of radiation. The study was conducted with the participation of 14 subjects, who were subjected to physical exercise with the above described device fixed on the forearm. To make the result more visible, the physical activity was carried out in stages: warm-up, load, recovery. According to the results of the experimental study, it became possible to obtain spectral characteristics of backscattered radiation in the visible, near-infrared wavelength ranges in the surface layer of biological tissue of the forearm and deeper layers (2.5 mm and 5 mm, respectively). The following trends were observed: the maximum value of the radiation intensity falls at the wavelength of 535 nm and reached the value of 65 thousand conventional units. For the majority of the subjects, among the three stages of physical activity, a greater intensity of radiation was observed during the warm-up stage. This can be explained by a decrease in the density and diameter of functioning capillaries [12]. Fig. 3 shows the spread of intensity values for each wavelength.

Experimentally obtained backscattered radiation intensities in deeper layers have slightly different values. The maximum intensity of radiation is traced at a wavelength of 535 nm, and reaches 20 thousand conventional units. Also at each wavelength, a regular decrease in the intensity value with each subsequent stage of physical activity can be seen. As in the case of the surface layer, this pattern may be due to changes in the density and diameter of blood vessels, as well as changes in the rheological properties of blood, morphofunctional restructuring of the cardiovascular system and changes in the parameters of systemic haemodynamics due to exercise [13]. Fig. 4 shows the spread of intensity values for each wavelength.

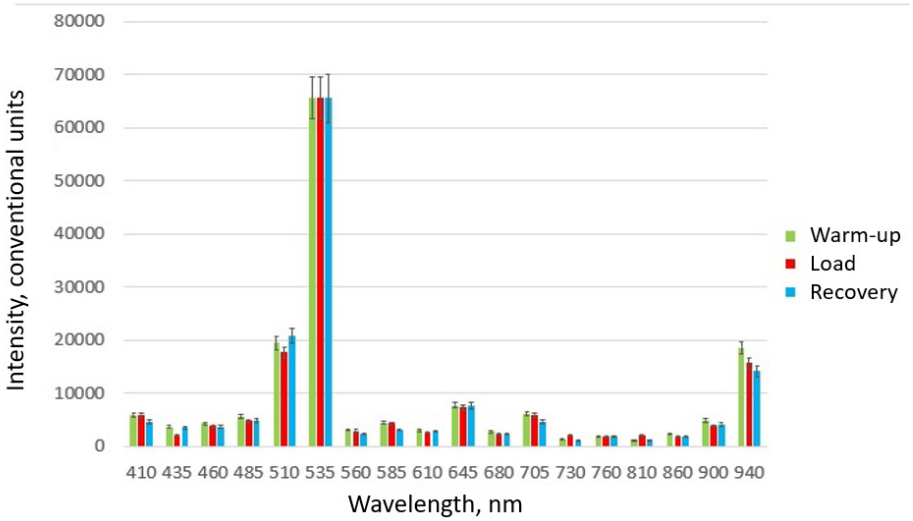


Fig. 3. Measured spectral response during loading and recovery for a typical subject in the superficial layer of forearm biological tissue

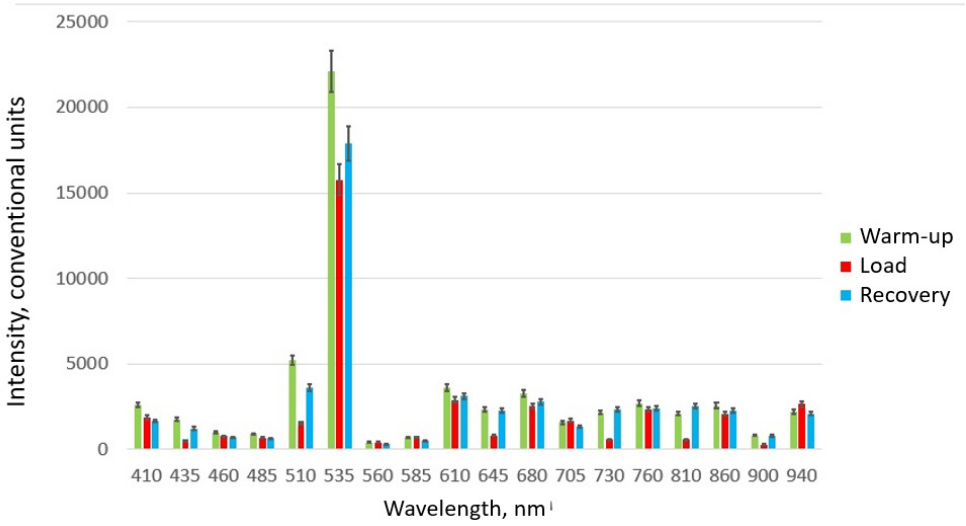


Fig. 4. Spectral characteristic of backscattered and reflected radiation in the visible, near-infrared wavelength range in deeper layers

Conclusion

A noticeable difference in the behaviour of signals at different depths carries valuable information about the state of the human microcirculatory channel, since in the superficial layers of the skin microcirculation provides heat exchange and cell nutrition, and also participates in the regulation of body temperature. In deeper tissues, such as muscles and internal organs, microcirculation plays a key role in the delivery of oxygen and nutrients to cells. At different depths in tissues, the microcirculation may be more or less active, depending on the oxygen and nutrient needs of the cells. During exercise, microcirculation in muscles increases to provide additional oxygen and energy [14].

List of references

[1] Thorn C. E. et al. Is mean blood saturation a useful marker of tissue oxygenation? //American Journal of Physiology-Heart and Circulatory Physiology. – 2009. – T. 296. – №. 5. – P. H1289-H1295.

[2] Tishutin, N.A. Approach to the assessment of the functional state of the organism / N.A. Tishutin, E.S. Pitkevich // Science – education, production, economy: materials of the 72nd Reg. scientific-practical

conference of teachers, research workers and graduate students, Vitebsk: VGU named after P.M. Masherov, 2020. – C. 329-331.

[3] Pushkareva A. E. Methods of mathematical modelling in biotissue optics: textbook // SPb: SPbSU ITMO. – 2008. – T. 103.

[4] Guimarães C. F. et al. The stiffness of living tissues and its implications for tissue engineering //Nature Reviews Materials. – 2020. – T. 5. – №. 5. – C. 351-370.

[5] Brown T. M., Krishnamurthy K. Histology, dermis //StatPearls [Internet]. – StatPearls Publishing, 2022.

[6] Almezhghwi K., Serte S. Improved classification of white blood cells with the generative adversarial network and deep convolutional neural network //Computational Intelligence and Neuroscience. – 2020. – T. 2020.

[7] Zanevskaya M. Yu. S., Zaitseva A. Yu. Investigation of the optical radiation penetration depth on theoretical models of inhomogeneous biological media and using a wearable hardware complex for medical monitoring / / Zhurnal tekhnicheskogo physika. – 2024. – T. 94. – №. 9. – C. 1524-1531.

Cambridge University Press, 2016.

[8] Firago V. A., Radchikova V. S. Determination of hydration of human surface tissues and parameters of their microcirculatory channel of the circulatory system. – 2023.

[9] Mazing M. C., Zaitseva A. Y. POSSIBILITIES OF APPLICATION OF OPTICAL Tissue Oximetry METHOD FOR DIAGNOSTICS OF FUNCTIONAL STATUS OF MICROCIRCULATION // GEORGIEVSKY READINGS: INherited and Acquired PATHOLOGIES OF BLOOD SWERTHING-TROMBOSIS AND BLOOD BLOWING (DIAGNOSTICS, PROPHYLACTICS, TREATMENT). – 2022. – C. 37-41.

[10] Alekseev V. A., Perminov A. S., Yuran S. I. Mutual location of the source and the receiver of the radiation of the sensor for photoplethysmography // Instruments and methods of measurements. – 2011. – №. 1 (2). – C. 5-9.

[11] Hamaoka T., McCully K. K. Review of early development of near-infrared spectroscopy and recent advances in studies on muscle oxygenation and oxidative metabolism //The journal of physiological sciences. – 2019. – T. 69. – C. 799-811.

[12] Mikhailov P. V. et al. Changes in the parameters of the microcirculation system in response to physical exercise of different intensity // Yaroslavl Pedagogical Bulletin. – 2012. – T. 3. – №. 1. – C. 121-124.

[13] Grigorieva M. E. et al. Blood flow at physical loads of different types //Sports medicine: science and practice. – 2023. – T. 12. – №. 4. – C. 45-58.

[14] Borisevich S. A. Features of microcirculation and transcutaneous partial pressure of oxygen in athletes with different orientation of physical loads // Modern problems of science and education. – 2012. – №. 3. – C. 305-305.

LAN TESTER DEVELOPMENT

Artem Zhilka
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
e-mail: 2165563@gmail.com

Abstract. *The paper describes the development of a LAN tester – a twisted pair diagnostic device. A complex methodology combining voltage measurements, resistance calculations and analysis of charging/discharging process time characteristics for determining the line break length is presented.*
Keywords: *LAN tester, twisted pair, crimp order, capacitive method, calibration, analog measurements.*

As the volume of information transmitted and the complexity of network infrastructures grows, the importance of reliable cable diagnostics increases. Crimping errors, short circuits and line breaks can lead to significant network disruptions and significant financial losses. In this paper we propose a diagnostic technique to detect crimping faults in RJ-45 connector.

Technical implementation of the tester.

The basis of the tester is the Arduino Nano [1] microcontroller, which has the following characteristics:

- Eight analog inputs (ADC) for reading voltages from twisted pair lines.
- Sufficient digital outputs to control a passive response part consisting of resistors with different ratings.
- I2C interface for connection of a character display showing diagnostic results.

The device includes:

- RJ-45 connectors – provide the physical connection of the cable under test;
- Passive response part – a set of resistors, with selected ratings, allowing comparative analysis of measured resistances;
- The power supply is a 9V 6F22 type battery, which provides a stable power supply for the whole circuit;
- LCD display – with 20×4 character resolution, connected via I2C interface for convenient visual control.

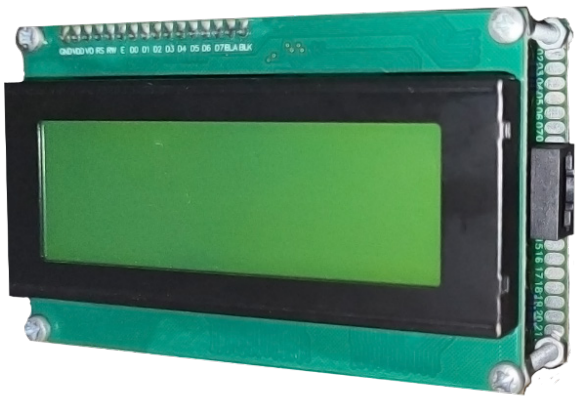


Fig. 1. Layout of the device

Software

The program part of the tester is written in Arduino C language and consists of the following main modules:

- System initialization and calibration of input parameters;
- Serial measurement of voltages on analog inputs and setting of test lines;
- Calculation of inter-contact resistances and restoration of crimp order by solving a system of equations;
- Application of the capacitive method for determining the line length to break;
- Error handling, logging and display of results.

The tester software is implemented in Arduino C language and consists of several stages: initial voltage check, diagnostics of line integrity, resistance measurement, determination of length to break by capacitive method and display of results.

To restore the crimp order, the voltages U_i are measured on each of the eight lines. Based on the difference between the values of U_i and $U_{(j),}$ the pairwise resistances are calculated using the formula [2]:

$$R_{ij} = \frac{(U_i - U_j) \cdot R_p}{U_j - U_{\min}}$$

where $R_p = \frac{R_1 \cdot R_2}{R_1 + R_2}$ – is the equivalent resistance of the resistive network, and U_{min} – is the minimum voltage registered on the line.

The table of nominal values for each line is set experimentally and includes the following parameters: R_{nom}, R_{min}, R_{max}.

The main task is to compare the measured resistances with reference values, which allows you to restore the sequence of pins in the connector.

Consider the algorithm:

1. **Pairwise comparison**: for each pair of lines, the values of R_{ij} are calculated using the above formula.
2. Minimum pair: a pair of lines with the minimum resistance value is selected.
3. **Solution of the system of equations**: after determining the minimum pair, the system of linear equations is solved to clarify the values of resistances for all lines [3]:

$$\begin{cases} R_1 + R_2 = R_{12} \\ R_1 + R_3 = R_{13} \\ R_2 + R_3 = R_{23} \end{cases} \rightarrow R_1 = \frac{R_{12} + R_{13} + R_{23}}{2}$$

$$R_2 = R_{12} - R_1$$

$$R_3 = R_{13} - R_1$$

4. **Comparison with a reference: The values obtained are compared with the nominal parameters. If the error does not exceed the specified tolerance, the line is considered to be correctly crimped, otherwise an error is recorded.**

When a line break is detected, the capacitive method is used to estimate the length of the line to the fault. The initial U1 and final U2 voltages on the broken line are measured, and then the formula is applied [4]:

$$L = \frac{-1}{\ln\left(\frac{U_2}{U_1}\right) - k} \cdot c,$$

where k and c are tuning coefficients.

Fig. 2. The result of the device operation. Direct crimp cable connected

Fig. 3. The result of the device layout. Cross-crimped cable connected

To test the developed methodology, cable samples with different crimping modes were prepared:

- Straight crimp – standard connection according to T568A/B;
- Cross-crimping – with an intentional error that simulates cross-crimping;
- Cables with breaks – artificially damaged to test the capacitive method.

Based on the results obtained, a statistical model was built to show the distribution of errors when compared with reference values. The results showed that after calibration the average error of crimp determination was less than 5%, and the accuracy of break length determination was up to 15% under stable temperature conditions.

The developed layout and algorithm are able to determine the order of LAN-cable crimping with high accuracy. The considered complex approach provides high accuracy of diagnostics, reduces the influence of errors and allows to detect both crossing and line breaks reliably.

References

1. Arduino Nano // Documentation and Specifications [website]. – URL: https://docs.arduino.cc/

2. Matvienko V. A. (2016) Fundamentals of circuit theory: textbook for universities. – Ekaterinburg : UMC UPI,. – p. 162.

3. Pestrikov V. M. (2023) Mathematical Methods in Engineering: textbook. – St. Petersburg: HSTE SPbSUITD. – p. 158.

4. Anodina-Andrievskaya E. M., Zhilka A. A., Kulikov V. V. Implementation of capacitive analysis methods in the environment of computer mathematics MATLAB // Mathematical methods and models in high-tech production. – St. Petersburg: SUAI, 2024. – pp. 226-228.

EVALUATION OF METHODS FOR CALCULATING THE CENTER OF GRAVITY OF A SYSTEM OF POINTS WITH TWO PARAMETERS OF WEIGHTING COEFFICIENTS

Alexander Zhmurin
Saint Petersburg State University of Aerospace Instrumentation,
Saint Petersburg, Russia
E-mail: qqwwww2006@gmail.com

Abstract. *The results of calculations of the center of gravity of a system of points are compared with the weighting coefficients calculated by various methods. The following methods for calculating weight coefficients are considered: the classical method, the Fishburne method and the proportional method.*
Keywords: *weighting factor, Fishburne method, method evaluation, center of gravity*

Introduction

The result of calculating the center of gravity of a system of points with certain weights may vary depending on the method of calculating the weights. This article provides an assessment of methods for calculating weighting coefficients for a system of points, which is a matrix with dimensions N by M.

The principle of calculating weight coefficients

There are two parameters determining the weighting factor – a random estimate of the significance of the point and the distance of the point to the geometric center, the further the point is from the geometric center, the smaller its weighting factor.

The point system is represented by an N by M matrix, where the cell index denotes the coordinates of the point (x; y), respectively. The geometric center is a cell with indexes (N/2; M/2) corresponding to a point with coordinates (N/2; M/2). The cells of the matrix contain estimates of points by degree of significance, randomly distributed in a discrete normal distribution in a certain range [1].

The differences in calculating the center of gravity point will depend on the method of calculating the weighting coefficients [2]. The classical calculation of the weighting coefficients is as follows: there is a matrix of expert estimates of the significance of Ae points with dimensions N by M. A matrix of distance weights Ad is calculated relative to the geometric center of the system (matrix cells (N/2; M/2)), where the cell of the geometric center has the maximum distance weight Wd. Relative to the geometric center, the distances to the following cells are calculated and their weight is calculated as the quotient of the maximum weight of the geometric center and the distance to the geometric center increased by one, so that when the distance is zero, division by 0 does not occur:

$$w^{d}{}_{i,j} = \frac{W_d}{(1+\sqrt{(i-\frac{N}{2})^2+(j-\frac{M}{2})^2})} \tag{1}$$

After obtaining the matrix of distance weights Ad, in the cells of which the weights wdi,j are located, the operation of obtaining the matrix of total weights is performed. The operation of obtaining total weights can also have different variants [3-4]. In our case, the product operation is accepted:

$$A_p = A_e \times A_d \tag{2}$$

Next, the center of gravity of the resulting system is calculated:

$$x_{цm} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} i \cdot w_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{i,j}},\ y_{цm} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} j \cdot w_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{i,j}}. \tag{3}$$

To calculate the weight coefficient matrix using the Fishburne method, the coefficient matrix Ap is multiplied by the rank matrix Ar, which is a matrix of "isolines" along the edges of which the values of the

minimum ranks are located and the values of the ranks increase towards the center, reaching the maximum value in the geometric center of the matrix [5]:

$$A_f = A_p \times A_r \tag{4}$$

The point of the center of gravity of such a system of points is calculated by the formula (3).

Calculating the matrix of weights using the proportional method also changes the matrix of total weights in expression (2). The proportional method multiplies the matrix Ap by the proportional matrix Aep, which is calculated by dividing the values in the cells of the matrix of Ae scores by the total sum of the scores:

$$A_{ep}(i;j) = \frac{A_e(i;j)}{\sum_{i=1}^{n}\sum_{j=1}^{m} A_e(i;j)} \tag{5}$$

And then the final matrix of weighting coefficients has the form:

$$A_a = A_p \times A_{ep} \tag{6}$$

The center of gravity of such a system is also calculated using the formula (3)

The results of calculations of the center of gravity points using three methods

Below are the results of a study evaluating the calculation of gravity points of the above-described point system. The sizes of all the matrices were assumed to be 10 by 10. The geometric center of such a matrix is (5; 5). The matrix of expert estimates had a discrete distribution in the range from 0 to 10. The maximum weight of the geometric center Wd was 10. The expected results are as follows: the points of the centers of gravity will have coordinates close to the geometric center, since the matrix of expert estimates has a normal distribution. The results are shown in Fig. 1.

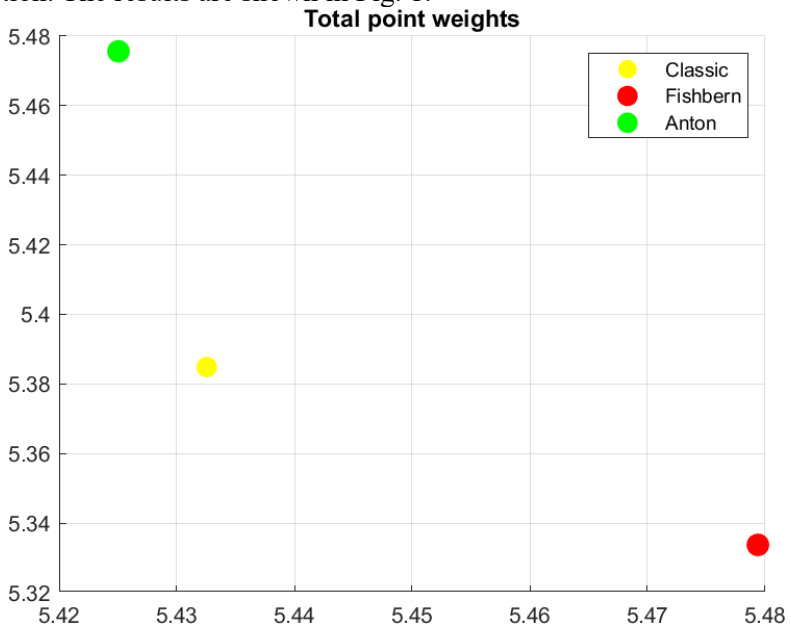


Fig. 1. Points of the centers of gravity

In Fig. 1, the x and y coordinates are located along the abscissa and ordinate axes, respectively. The points of the centers of gravity calculated using three methods have coordinates close to the geometric center of the system, and the deviations of the coordinates of the points of the centers of gravity do not exceed 0.1, which we can conclude that all three methods have a fairly good calculation accuracy.

Conclusion

This article considers the evaluation of the calculations of the coordinates of the center of gravity points of a system of points using three different methods using weighting coefficients. As a result, the

evaluation of the three methods for calculating the weighting coefficients showed that all three methods have approximately the same calculation accuracy.

References

1. Analysis of approaches to the selection of weighting coefficients of criteria by the method of paired comparison of criteria. Spiridonov S.B., Bulatova I.G., Postnikov V.M. // Online Journal of Science, Volume 9, No. 6 (2017)

2. The method of estimating the weighting coefficients of the elements of organizational and technical systems. Boyko A. A., Degtyarev I. S. // Control, communication and security systems. 2018. No. 2. pp. 245-266.

3. Methods and models of statistical decision-making in conditions of uncertainty. Artisan E.S. // Dissertation for the degree of Candidate of Economic Sciences // Moscow, 2020

4. The theory of decision-making: a textbook / S. M. Borodachev. Yekaterinburg : Ural Publishing House, University, 2014. 124 p.

5. Estimation of probability distribution by Fishburn sequences: necessity and advantages // Sigal A.V., Artisan E.S. // Management of large systems. Issue 70, KFU named after Vernadsky. Simferopol, 2020

CONTENTS

GREETINGS

Scott Reynolds, 2025 ISA SOCIETY PRESIDENT 3

Gerald W. Cockrell, ISA SOCIETY FORMER PRESIDENT..... 5

PROFESSIONALS SPEAKING

Chabanenko A. PRODUCT QUALITY ASSESSMENT USING VIDEOMETRIC SYSTEMS..... 7

Kriachko M. GENERALIZED ATOMIC WAVELETS..... 11

Silina A., Gaiduk E. AUTOMATING THE DECISION-MAKING PROCESS BY DEVELOPING A MODEL FOR THE EVOLUTION OF A SEA PASSENGER PORT..... 16

THE TWENTY-FIRST ISA EMEA&PAKISTAN STUDENT PAPER COMPETITION

Alekseev K. AUTOMATION IN GAMEDEV: TOOLS FOR AUTOMATIC TRANSLATION AND LOCALIZATION 22

Bagaeva A. CONSTRUCTION OF A SMALL-SIZED AIRCRAFT IMPLEMENTED ON AN ARDUINO BOARD WITH REMOTE CONTROL FUNCTIONALITY VIA A MOBILE DEVICE INTERFACE 25

Banukov P. THE ROLE OF AUTOMATION IN MODERN METROLOGY 29

Belova M. AUTOMATION OF PROCESSES OF TECHNOLOGICAL PREPARATION OF PRODUCTION BY APPLICATION OF NEURAL NETWORK TECHNOLOGIES..... 31

Bobovich A. WHY LINEAR ACTIVATION FUNCTIONS SHOULDN'T BE USED IN MACHINE LEARNING 36

Bobryshov A. DEVELOPMENT OF THE SYSTEM OF STRESS-TESTING OF ELECTRICAL CONTROL AND MEASURING DEVICES FOR DETERMINATION OF ESTIMATED FUNCTIONS OF DEGRADATION OF ACCURACY OF DEVICES 39

Bukhvalova P., Rachkovskaya E. SEMANTIC SEGMENTATION OF AIRBORNE LIDAR DATA USING RANDLANET NERUAL NETWORK MODEL 47

Casadio D. THE ETHICS OF USING ARTIFICIAL INTELLIGENCE IN DECISION-MAKING: BALANCING AUTOMATION AND HUMAN CONTROL..... 52

Chabanenko G. USING A CHATBOT SYSTEM TO AUTOMATE MARKETING PROCESSES..... 55

Chembarisova R. FORMATION AND PROCESSING OF SIGNAL-CODE STRUCTURES IN SMALL-SIZED AIRBORNE RADAR STATIONS 58

Dianov V. COMPARISON OF CONVOLUTIONAL NEURAL NETWORK MODELS FOR THE POSSIBILITY OF HUMAN DETECTION IN THERMAL IMAGING DATA USING KERAS AND TENSORFLOW 61

Galeeva E. GENERAL ARTIFICIAL INTELLIGENCE: OPPORTUNITIES, RISKS AND THE FUTURE OF DIFFERENT INDUSTRIES..... 65

Golovkin M. THE PROCESS OF COLLECTING A DATASET FOR TRAINING A NEURAL NETWORK FOR CLASSIFICATION BASIC ALGORITHMS AND DATA STRUCTURES WRITTEN IN THE PYTHON PROGRAMMING LANGUAGE 67

Grigoriev I. MODELING OF A TRAFFIC LIGHT OBJECT..... 69

Gromysh Y. RESEARCH OF THE NOISE IMMUNITY OF IMAGES MASKED BY HADAMARD MATRICES 75

Interlice Emanuele Maria, Nunzio Vincenzo Di Bella DEVELOPMENT OF AN INTELLIGENT ALARM SYSTEM FOR HOME SECURITY 84

Ivanov K. DEVELOPMENT OF A LIGHT MONITORING SYSTEM BASED ON A WI-FI MICROCONTROLLER..... 89

Kalinichenko M. AUTOMATIC SPACE MAPPING USING LIDAR..... 92

Karabaeva D. RESEARCH OF THE AUDIO DATA NOISE-IMMUNITY MASKED BY HADAMARD MATRICES 96

Kazakevich T. ANALYSIS OF THE QUALITY OF MATRIX MASKING OF VISUAL INFORMATION WITH CONSIDERATION OF HUMAN PERCEPTION..... 104

Khudyakov V. MAIN NON-GAUSSIAN PROBABILITY DENSITIES USED IN MODELING TECHNICAL SYSTEMS..... 110

Kitaev V. OPTIMIZATION OF DIFFRACTION GRATING PARAMETERS FOR MULTI-CRITERIA SPECTRAL ANALYSIS TASKS..... 114

Kleshnin B. IMPLEMENTATION OF A CONVOLUTIONAL NEURAL NETWORK ALGORITHM IN MATLAB 119

Komarov T. DEVELOPMENT OF A RELIABILITY MANAGEMENT SYSTEM FOR THE MACHINE-BUILDING INDUSTRY BASED ON MODERNIZATION OF THE TECHNICAL READINESS COEFFICIENT CALCULATION 122

Komarova V. DEVELOPMENT OF A METHODOLOGY FOR ASSESSING THE NEED FOR IMPLEMENTING CORRECTIVE OR PREVENTIVE ACTIONS..... 125

Kosmynin D. STUDY OF THE AUDIO SIGNAL ENTROPY CALCULATION ALGORITHM FOR SYNTHESIZED VOICE RECOGNITION 134

Kreyzo M. LIBRARY SYSTEM IMPLEMENTED IN NODE.JS AND MYSQL..... 139

Kulikov V. DEVELOPMENT AND ANALYSIS OF MATHEMATICAL MODEL OF ULTRASONIC RANGE FINDER 142

Kuzmenko Yu. DEVELOPMENT OF A REAL-TIME LIGHT MONITORING ALGORITHM..... 145

Leontyev V. IOT IMPLEMENTATION IN RAILWAY ROLLING STOCK DIAGNOSTICS: ANALYSIS OF MEASUREMENT ACCURACY AND DATA CORRECTION METHODS..... 150

Lisovenko S. MODELING OF PLANT GROWTH PROCESS..... 155

Simone Giovanni Matraxia, Paolo Francesco Garrasi. REDUCING TRAFFIC JAMS IN HISTORIC URBAN INTERSECTIONS USING FUZZY NETWORKS AND SMART TRAFFIC MANAGEMENT: A PRACTICAL IMPLEMENTATION..... 158

Matveev A. RECOGNITION OF OBJECTS IN A POINT CLOUD USING DEEP LEARNING..... 164

Pisklenov T. DEVELOPMENT OF A HARDWARE-SOFTWARE SYSTEM USING A SINGLE-BOARD COMPUTER AND NEURAL NETWORK MODELS FOR THE TASK OF EARTH SURFACE OBJECT RECOGNITION..... 170

Pitomets M. INVESTIGATION OF A METHOD FOR DETECTING UNMANNED AIRCRAFT LANDING SPOTS USING A POINT CLOUD 174

Pogodina D. PSYCHOLOGY AND AI: ALTERNATIVES FOR SELF-HELP AND THE FUTURE OF PSYCHOTHERAPY WITH LANGUAGE MODELS..... 179

Rassykhaeva M. AUTOMATION OF QUALITY MANAGEMENT PROCESSES 182

Rybkin I. COMPARATIVE CHARACTERISTICS OF LEAST SQUARES AND LEAST MODULI METHODS FOR APPROXIMATING REGRESSION RELATIONSHIPS IN THE PRESENCE OF ANOMALOUS OUTLIERS 185

Ryndina K. AUTOMATION OF MEASURING PROCESSES WITH A ROUND-ROBIN METER 188

Ryvkina Ya. THE INFLUENCE OF THERMAL FLOW TURBULENCE ON THE SPATIAL CHARACTERISTICS OF A LASER BEAM 191

Davide Sambito, Erica Schembri. PROPOSAL FOR AN INTELLIGENT IRRAGATION SYSTEM 198

Shchukin A. DETECTION OF SPACE OBJECTS ON VIDEO SEQUENCE OF THE STARRY SKY 202

Alessio Salvatore Vetro, Giuseppe Alaimo. NETWORK SCENARIO FOR WEATHER PREDICTION AND ALERT 208

Vinogradov D. REMOTE VENTILATION CONTROL SYSTEM USING ESP32 211

Voronov R. AUTOMATED ANALYSIS OF TEXT DATA BASED ON MACHINE LEARNING ALGORITHMS 216

Zanevskaya M. POSSIBILITIES OF THEORETICAL ESTIMATION OF OPTICAL RADIATION PENETRATION DEPTH IN MULTICOMPONENT BIOLOGICAL ENVIRONMENTS 220

Zhilka A. LAN TESTER DEVELOPMENT 225

Zhmurin A. EVALUATION OF METHODS FOR CALCULATING THE CENTER OF GRAVITY OF A SYSTEM OF POINTS WITH TWO PARAMETERS OF WEIGHTING COEFFICIENTS 228

The scientific edition

ИЗВЕСТИЯ КАФЕДРЫ UNESCO ГУАП «ДИСТАНЦИОННОЕ ИНЖЕНЕРНОЕ ОБРАЗОВАНИЕ»

Сборник статей

Выпуск 10

BULLETIN OF THE UNESCO CHAIR “DISTANCE EDUCATION IN ENGINEERING” OF THE SUAI

Collection of the papers

Issue 10

ISBN: 978-5-8088-2026-5

9 785808 820265

Computer imposition *A. N. Koleshko*
Papers are publish in author's edition

Подписано в печать 07.04.2025. Дата выхода в свет: 11.04.2025. Формат 60×84 1/8.
Усл. печ. л. 27,5. Тираж 150 экз. Заказ № 79.

Редакционно-издательский центр ГУАП
190000, г. Санкт-Петербург, ул. Большая Морская, 67, лит. А
Распространяется бесплатно

Submitted for publication 07.04.2025. Passed for printing 11.04.2025.
Format 60×84 1/8.
Department of operative polygraphy SUAI
67A, B. Morskaia, 190000, Saint Petersburg, Russia