ITHEA

International Journal

INFORMATION THEORIES & APPLICATIONS

2022 Volume 29 Number 3

International Journal

INFORMATION THEORIES & APPLICATIONS
Volume 29 / 2022, Number 3

International Journal "INFORMATION THEORIES & APPLICATIONS" (IJ ITA)
is official publisher of the scientific papers of the members of
the ITHEA International Scientific Society

IJ ITA welcomes scientific papers connected with any information theory or its application. IJ ITA rules for preparing the manuscripts are compulsory.
The **rules for the papers** for IJ ITA are given on *www.ithea.org*.

Responsibility for papers *published in* IJ ITA belongs to authors.

International Journal "INFORMATION THEORIES & APPLICATIONS"
Vol. 29, Number 3, 2022

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with: University of Telecommunications and Posts, Bulgaria,
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,
Universidad Politécnica de Madrid, Spain,
Hasselt University, Belgium, University of Perugia, Italy,
Institute for Informatics and Automation Problems, NAS of the Republic of Armenia

Printed in Bulgaria
Publisher ITHEA®

Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: office@ithea.org
Technical editor: **Ina Markova, Vera Markova**

ISSN 1310-0513 (printed) **ISSN 1313-0463 (online)**

THE ROLE OF THINKER CONSCIOUSNESS IN MEASUREMENT ACCURACY: AN INFORMATIONAL APPROACH

Boris Menin

Abstract: *This article contends that understanding measurement accuracy requires considering the thinker's consciousness in the process. In contrast to modern statistical methods, which analyze sources of uncertainty related to variables and measurements, this study proposes an approach that also accounts for the thinker's role in storing, transmitting, processing, and utilizing information to formulate the model. By incorporating the finite amount of information in the model, the study proposes a method for estimating the limit of measurement accuracy and provides examples of its application in experimental physics and technological processes.*

Keywords: *Amount of Information; Finite Information Quantity; International System of Units; Modeling; Physical Constant; Speed of Sound*

ITHEA Keywords: *H.1 MODELS AND PRINCIPLES; I.6 SIMULATION AND MODELING*

DOI: *https://doi.org/10.54521/ijita29-03-p01*

Introduction

The remarkable progress in modern science and technology, from space exploration to the study of subatomic particles, would not have been possible without the development of measurement theory, sophisticated methods for processing experimental data, and powerful computers. To achieve even greater accuracy in scientific and technical calculations, researchers often rely on models that reflect their philosophical views, intuition, and experience. However, while scientists tend to focus on calculating known uncertainties when processing experimental data, they may overlook the importance of understanding the underlying uncertainties that arise during modeling, prior to conducting experiments.

Applying the concepts and mathematical apparatus of information theory can help us comprehend the physical reasons behind the uncertainties that surround observed physical objects [1]. One of the primary factors contributing to uncertainty during the model formulation stage (defining the qualitative-quantitative set of variables and functional relationships between them) is the amount of information contained within the model, a concept commonly used in physics and engineering fields to describe the complexity of theoretical models that explain physical phenomena.

In the article by Jaynes [2], the application of information theory to statistical mechanics and thermodynamics is explored, suggesting that entropy in thermodynamics can be understood as a measure of the amount of missing information about a system.

Kapur's book [3] discusses the use of maximum entropy models to extract the maximum amount of information from limited data and provides an overview of the role of information in physics.

Gell-Mann and Lloyd [4] argue that information is a fundamental aspect of the universe and explores the relationship between information and physical systems from an information theory perspective.

Allaire et al.'s article [5] proposes an information-theoretic metric for quantifying the complexity of engineering systems. The authors demonstrate the usefulness of their metric by applying it to a variety of engineering systems, including an aircraft wing, a gas turbine engine, and an electric power grid. They show that their metric can capture the key features of each system's complexity, including the number of components, the degree of interaction between components, and the level of redundancy built into the system.

Deffner and Jarzynski [6] present an information-theoretical approach to thermodynamics, characterizing the amount of information contained in the state of the system and obtaining a generalization of Kelvin-Planck, Clausius, and Carnot's provisions on the second law of thermodynamics for situations related to information processing. It is suggested that this approach can provide insight into the thermodynamic behavior of complex systems.

Fang and Ouellette's article [7] presents a novel approach for measuring the information content contained in fluid flow data. The proposed method employs information theory, a branch of mathematics that deals with the quantification and transmission of information, to analyze the patterns and structures present in fluid flow data. Authors suggest that the proposed approach has the potential to advance our understanding of fluid flow dynamics, providing crucial insights for addressing pressing engineering and biological challenges.

The examples represent just a small sample of the many works that have explored the concept of "the amount of information contained in the model" within the field of physics and technology.

This concept is derived from the necessity for researchers to use a system of units with a certain level of complexity. By using this idea, it is possible to calculate the value of the threshold discrepancy between the model and the observed object [8-10].

Physical and Mathematical Foundations of the Information Method

Every experiment or measurement is preceded by the researcher's decision to construct a model. Therefore, it is impossible to calculate the minimum achievable measurement accuracy without considering the modeling stage. In other words, *the accuracy limit for describing physical objects is not determined by the material measuring instrument employed, but by the design of the model* *constructed in accordance with the individual preferences, philosophical and theoretical positions, and will of the researchers.* Furthermore, this limit of accuracy far surpasses the limit imposed by quantum mechanics on measurement precision.

To calculate the limit of object reproduction accuracy, researchers can conceptualize the model as a "communication channel" between the observed phenomenon and the thinker. In doing so, they must select significant variables while disregarding potential connections between them, leading to the oversight of hidden effects. The selected variables are typically derived from a system of units, such as the International System of Units (SI).

According to [11], each variable is comprised of a scalar time parameter, a universal constant, an one-dimensional position or momentum component, and a dimensionless number that take real values from the set R. Additionally, each variable contains a "piece of information" [12] that is bounded from above and defined as a "finite information variable" or FIQ by Del Santo and Gisin [11]. The number of dimensionless FIQs in the SI system can be calculated as μ_{SI} = 38265 [13]. These concepts and formulas are applicable to models using different systems of units [14] and containing any FIQ, both dimensional and dimensionless [15].

The process of modeling physical phenomena involves representing the analyzed phenomenon as a communication channel [9,16]. However, because this model is subject to noise and various distortions introduced by the thinker, uncertainty is present even during the formulation stage, preceding computer calculations and experiments.

An important remark follows regarding the modeling process. To accurately capture the essence of this process, it is preferable to use the term "thinker" instead of the more commonly used term " observer". This choice of terminology emphasizes that during the act of modeling, the phenomenon being studied is not influenced by external factors, and that reality exists independently of the individual conducting the study. By using the term "thinker," we can better convey the idea that the modeling process involves a careful and deliberate consideration of the underlying principles governing the phenomenon, without the imposition of external biases, distortions, or perturbations.

The process of model coding involves mapping the original set of FIQs from the SI system to the model structure, using a specific mapping process designed to minimize the threshold for mismatches between the model and the object being studied. To accomplish this, the researcher selects the FIQs from the SI, which contains seven base quantities [17]: length (L), mass (M), time (T), thermodynamic temperature (Θ), electric current (I), luminous intensity (J), and quantity of substances (F). The selected base quantities and derived variables determine the group of phenomena (GoP) inherent in the model. A GoP refers to a set of natural phenomena or technological processes characterized by a qualitative and quantitative group of variables that reflect certain properties of the observed reality [18].

For instance, in experiments involving Faraday's law, reduced variables with dimensions, such as length (L), mass (M), time (T), and electric current (I), which belong to the GoP_{SI} ≡ LMTI, are typically used. Due to limited financial resources, time, computing power, and philosophical preferences, the researcher selects a very small number of variables in the model compared to the μ_{SI}, which was equal to 38.265 [8]. Consequently, by calculating the amount of information in both the model and SI, the comparative uncertainty of the model, ε, can be determined [9].

$$\varepsilon = \Delta_{\Sigma} / S = [(z' - \beta')/\mu_{SI} + (z'' - \beta'')/(z' - \beta')] \quad (1)$$

where Δ_{Σ} is the a priori general uncertainty of the model due to the choice of GoP and the number of FIQs considered, S is the observation interval of the main studied FIQ, z' is the number of FIQs in the selected GoP, β' is the number of base values in the selected GoP, z'' is the number of FIQs, recorded in the model, and β'' is the number of independent values recorded in the model. ε is an important element of information theory [19], although almost no attention has been paid to it in modern scientific and technical literature.

To determine the optimal number of FIQs for a specific GoP, it is necessary to calculate the derivative of ε with respect to z' - β' and set it equal to zero:

$$\left(z''-\beta''\right)=\left(z'-\beta'\right)^2/\mu_{SI} \tag{2}$$

By using equations (1) and (2), one can determine the characteristics of any model that impose certain requirements for the transmission of information from the observed phenomenon to the thinker and the number of variables needed in the model to achieve the optimal value of ε_{opt} specific to each GoP. ε_{opt} can be considered as the capacity of the communication channel (model) and is an important concept in the modeling process that has not been extensively covered in existing scientific literature. It can be used to describe the characteristics of the model that reproduces the studied process, as well as the preferred method for measuring certain objective functions, such as physical constants.

Table 1 presents the optimal values of ε_{opt} and the corresponding recommended amount of FIQs for each GoP, which were obtained using (1) and (2) [14].

The data in Table 1 reveal the boundaries that can be reached for source compression, which involves replacing the μ_{SI} of the object with the representation parameter $W(GoP_{SI})$ necessary to ensure minimal distortion (threshold mismatch) $D(\varepsilon_{opt})$ that can theoretically be achieved by reproducing the observed object. In this case, the bandwidth of the communication channel is determined as a function of *the a priori total amount of information obtained when formulating a model of a physical phenomenon*:

$$W(D)=\max_{\gamma_{CoP}\in Z_D}\Delta A(\mu_{CИ};\gamma_{CoP}) \tag{3}$$

where W(D) is the reproduction (or source distortion) parameter and Z_D is the set of all possible optimal FIQ variables corresponding to GoPs. The GoP corresponds to a distortion ε equal to $D(\varepsilon_{opt})$ when the source (object) is transformed into a model. Δ*A* is the a priori total amount of information in units of entropy obtained when formulating a physical phenomenon model.

Table 1. Comparative uncertainties and the optimal number of dimensionless criteria

GoP_{SI}	Comparative uncertainty, ε_{opt}	Number of FIQs attributable to GoP_{SI}, γ_{GoP} = z'-β'	Optimal FIQs number inherent in the model, γ_{mod} = z''-β''
LMT	0.0048	91	≈ 0.2 < 1
LMTF	0.0146	279	≈ 2
LMTI	0.0245	468	≈ 6
LMTθ	0.0442	846	≈ 19
LMTIF	0.0738	1412	≅ 52
LMTθF	0.1331	2546	≈ 169
LMTθI	0.2220	4247	≈ 471
LMTθFI	0.6665	12 751	≈ 4,249

The distortion resulting from compressing the source can be attributed to the researcher's preferences. The researcher selects a specific GoP based on his philosophical beliefs, which determines the number of FIQs included in the model. The difference between the number of selected FIQs and the optimal number of FIQs inherent in the model (γ_{mod}) is shown in Table 1. For example, let γ_1 and γ_2 be the number of FIQs in the first and second models, respectively, if the models refer to the same GoP, and $\gamma_1 < \gamma_{GoP} < \gamma_2$ and $|\gamma_1 - \gamma_{GoP}| < |\gamma_{GoP} - \gamma_2|$. By applying (1), we obtain the following equation:

$$|\varepsilon_1 \text{-} \varepsilon_{opt}| = |2 \cdot (\gamma_1 \text{-} \gamma_{CoP}) / \mu_{СИ}|, |\varepsilon_2 \text{-} \varepsilon_{opt}| = |2 \cdot (\gamma_2 \text{-} \gamma_{CoP}) / \mu_{СИ}|, \tag{4}$$

$$|\varepsilon_1 \text{-} \varepsilon_{opt}| / |\varepsilon_2 \text{-} \varepsilon_{opt}| = |\gamma_1 \text{-} \gamma_{CoP}| / |\gamma_2 \text{-} \gamma_{CoP}| < 1, \tag{5}$$

where ε_1 and ε_2 are the comparative uncertainties of the first and second model, respectively.

Although models with a large number of variables are generally preferred, this conclusion may not hold within the proposed approach. This is because the researcher can quickly determine a lower value when comparing $|\varepsilon_1 - \varepsilon_{opt}|$ with $|\varepsilon_{opt} - \varepsilon_2|$. Specifically, ε_1 uses a larger number of dimensionless criteria, which may be closer to the optimal γ_{GoP} and the optimal comparative uncertainty ε_{opt}. Therefore, a more informative model would utilize γ_1, which is closer to γ_{GoP}. Consequently, the informational approach enables the identification of the most suitable method for calculating the preferred model for the object under study.

The informational approach involves several key features during the modeling process. Firstly, researchers, guided by their own understanding of the observed phenomenon's physical nature (which is an arbitrary act), strive to create a model that closely represents the object in a reasonable amount of time and with limited financial costs. As a result, GoP_{SI} models with a small number of base quantities, such as LMT, LMTθ, and LMTI, are commonly used.

Models related to LMTθI or LMTθF are used much less frequently. This is largely due to the limited computing power of computers, the ease of establishing mathematical relationships between selected variables, and the lower complexity of calculating the variable under study's value and its measurement uncertainty using existing statistical methods.

Secondly, the use of a GoP with a limited number of base quantities and a low value of γ_{mod} (< 2), such as GoP_{SI} ≡ LM, GoP_{SI} ≡ LMT, and GoP_{SI} ≡ LMTF (refer to Table 1), may not achieve the required value of ε_{opt} in theory or practice. Conversely, the use of GoPs with many base quantities and high γ_{mod} values (>169, refer to Table 1) is hindered by the increased difficulty in calculating the model uncertainty due to the limited computing power of modern computers. This topic is discussed in detail in the chapter “Application examples of the information method”.

Thirdly, it is evident that the researcher's worldview leads to the initial and inevitable uncertainty of the model ε. It is not possible to calculate ε using statistical methods that rely on weighted factors or consistency criteria to process experimental data or computer calculations for a model that has already been constructed and implemented in the field. This fundamental difference highlights the disparity between the use of information as a physical substance and measurement theory [20].

Application examples of the information method

An analysis of scientific and technical works was conducted by *comparing the achieved comparative uncertainty of the model* ε_{mod} *with the theoretically substantiated* ε_{opt}, as shown in Table 1. When the uncertainties are close ($\varepsilon_{mod}/\varepsilon_{opt} \rightarrow 1$), it confirms the validity and usefulness of using either model to describe the process under study. On the other hand, a significant difference between these uncertainties ($\varepsilon_{mod}/\varepsilon_{opt}$<<1) indicates a high risk of applying a particular model.

This approach enables the use of concepts related to transmission, accumulation, and transformation of information in both theoretical research and applied problems.

The mass of information accumulated on the planet

There are two opposing perspectives regarding the nature of information. The first perspective asserts that information has no mass [12]. The second perspective, known as Landauer's principle, suggests that information is inherently physical [21]. For instance, [22] calculates the total mass of information (M_{info}) in kilograms that has been accumulated on the planet over n years, considering the annual growth rate (f %) of digital content creation on Earth:

$$M_{info}(n) = N_b \cdot k_b \cdot T \cdot \ln 2 \cdot ((f+1)^{n+1} - 1)/(f \cdot c^2) \approx 4.7 \cdot 10^{-16} \tag{6}$$

where N_b is the current annual rate of digital bit production on Earth, N_b = 7.3 1021 bits, T = 300 K is the temperature at which information is stored, f = 0.01, n = 1 year [22], $k_b = 1.3805 \cdot 10^{-23}$ m²·kg/(s²·K) - Boltzmann constant [23], ln2 = 0.6931, c = $2.99792458 \cdot 10^8$ m/s is the speed of light [23].

The accuracy of a calculated value for a physical variable is often assessed using relative uncertainty [24]. According to [15], the relative uncertainty (r_M) when calculating M_{info} (6) is found to be very small, approximately 4.2×10^{-6}. This suggests that Eq. (6) could be practically applicable. To further investigate the plausibility of the presented results, comparative uncertainty was used.

In [14], the authors used a model with GoP_{SI} ≡ LMTθ ($\gamma_{LMT\theta}$ = z'-β' = 846, Table 1) and γ_1 = z"-β" = 1, based on the π-theorem [25].

$$\varepsilon_M = (846/38,265 + 1/846) \approx 0.0233 \tag{7}$$

The ratio of ε_M to ε_{opt} is found to be approximately 0.53 << 1 (where ε_{opt} = 0.0442 and $GoP_{SI} \equiv LMT\theta$, Table 1). This is because the mechanical-thermal model used in [25] only considers one criterion (γ_1 = z''-β'' = 1), instead of the recommended γ_{mod} = z''-β'' = 19 (Table 1). As a result, the original model presented in [22] is overly simplified and requires further development to account for potential hidden influencing variables. Consequently, any predictive calculations within the framework of the FIQ-based method, which are deemed to have a clear physical content by the authors due to their philosophical ideas, should be accompanied by appropriate explanations of the possible limits of their applicability [26].

Measurement accuracy of physical constants

The approval of a new version of the International System of Units [17] stands as one of the important scientific achievements of the 21st century. This progress was made possible using of advanced experimental stands, as well as statistical methods for processing experimental data. The data for measuring physical constants were analyzed in accordance with the recommendations of the Committee on Data for Science and Technology (CODATA), using methods such as Bayesian linear regression and the method of least squares (LSA). However, it should be noted that the LSA approach can sometimes yield inadequate results [27], as the initial experimental values are often "corrected" to check the consistency of the results. Moreover, conflicting results may increase uncertainty, as noted by Pavese [28]. Additionally, statistical expert bias driven by personal beliefs or preferences [29] and the presence of subjective judgment [30] cannot be ignored.

The CODATA methodology was used to analyze the results of experiments based on models formulated by scientists, which overlooks the systematic effect resulting from the comparative uncertainty associated with the choice of GoP and the number of variables considered in the model. Therefore, this study proposes to analyze the data on measuring some physical constants using the FIQ-based method, which does not rely on tools such as consistency, asymptotic normality, or weighted estimates or coefficients.

An analysis of the results of measuring physical constants using various methods is presented in [9,10,14,15,31]. The data covers studies conducted by research centers from 2000-2019. Table 2 presents the results of applying measurement methods to obtain specific values of the experimental and optimal relative uncertainty, ε_{exp} and ε_{opt}, respectively.

The analysis of the results of measurements of physical constants using various methods from 2000-2019 revealed the following observations:

- A decrease in the ratio $\varepsilon_{exp}/\varepsilon_{opt}$ occurs when transitioning from models with a low number of base quantities (LMT) to models with a larger number of base quantities, such as LMTθ, LMTI, LMTθF, and LMTθI.

- All ratios of $\varepsilon_{exp}/\varepsilon_{opt}$ are greater than 1, which confirms the FIQ-based method's [18] fundamental proposition that any model's accuracy limit is determined by ε_{opt}.

- The use of BDL, BAO, and mechanical methods for measuring the Hubble constant and gravitational constant are not promising.

- In measuring the Boltzmann constant, k_b, researchers using DCGT achieved outstanding results compared to AGT ($\varepsilon_{exp}/\varepsilon_{opt} \approx 3.6$) and DBT ($\varepsilon_{exp}/\varepsilon_{opt} \approx 3.7$) at small values of $\varepsilon_{exp}/\varepsilon_{opt} \approx 2.3$.

Table 2. Summarized data for $\varepsilon_M / \varepsilon_{opt}$.

Ratio $\varepsilon_{exp} / \varepsilon_{opt}$	Physical Constant	Measurement method	GoP
2.3	Boltzmann constant, k_b	DCGT[1]	*LMTθI*

3.6	Boltzmann constant, k_b	AGT[2]	*LMTθF*
3.7	Boltzmann constant, k_b	DBT[3]	*LMTθF*
7.9	Gravitational constant, *G*	Electro-mechanical methods	*LMTI*
100	Gravitational constant, *G*	Mechanical methods	*LMT*
4.1	Hubble constant, H_0	CMB[4]	*LMTθ*
104	Hubble constant, H_0	BAO[5]	*LMT*
710	Hubble constant, H_0	BDL[6]	*LMT*
15.9	Planck constant, *h*	KB[7]	*LMTI*
32.6	Planck constant, *h*	XRCD[8]	*LMTθF*

[1]DCGT—dielectric constant gas thermometer,

[2]AGT—acoustic gas thermometer,

[3]DBT—Doppler broadening thermometer,

[4]CMB—cosmic microwave background,

[5]BAO— baryonic acoustic oscillations,

[6]BDL—brightness of distance ladder,

[7]KB — Kibble balance,

[8]XRCD—X-ray crystal density.

- Electromechanical methods provide a higher degree of confidence in measuring the gravitational constant, resulting in greater measurement accuracy (εexp/εopt ≈ 2.3) than mechanical methods (εexp/εopt ≈ 100).

- The BAO method (GoP_{SI} ≡ LMT) and BDL method (GoP_{SI} ≡ LMT) result in significantly higher ratios ($\varepsilon_{exp}/\varepsilon_{opt} \approx 100$ and $\varepsilon_{exp}/\varepsilon_{opt} \approx 710$, respectively) when measuring H_0 compared to $\varepsilon_{exp}/\varepsilon_{opt} \approx 4.1$ achieved using CMB. This suggests the possibility of missing hidden variables and indicates that considering all possible sources of uncertainty does not guarantee obtaining the true value of H_0 using the BAO method. The CMB method is theoretically substantiated, and the most reliable experimental data is achievable using this method.

- Of the two methods used to measure Planck's constant (KB, XRCD), the KB method ($\varepsilon_{exp}/\varepsilon_{opt} \approx 15.9$) is the more promising option for achieving higher accuracy in determining the value of *h*.

The results presented in this study provide strong support for the reliability of the FIQ-based method in identifying the preferred structure of the physical constant measurement method model. By utilizing comparative uncertainty as a criterion, the FIQ-based method has demonstrated its effectiveness as a reliable tool for this purpose, as demonstrated by the theoretically formulated and proven Equation 3.

It is important to note that the informational method approach differs fundamentally from the personalist Bayesian approach when it comes to evaluating the results of measurements of physical constants. While the Bayesian approach focuses on analyzing experimental data obtained after the model has been formulated, the informational method prioritizes the selection of the optimal model structure.

To achieve this, the FIQ-based method evaluates the comparative uncertainty of each candidate model structure, thereby enabling the identification of the most suitable model’s structure for the given data. In contrast, the Bayesian approach involves assembling an installation, testing the test bench, and implementing a series of measurements while simultaneously calculating all emerging uncertainties identified. Scientists consider both methods to be effective, but they differ in their approach and priorities.

While the CODATA method for processing experimental data on the measurement of physical constants has many advantages and is widely used, it is important to acknowledge some of the disadvantages of the Bayesian approach it relies on. These include:

- Limited experimental data: The CODATA method relies heavily on experimental data, and if there are only a limited number of high-quality measurements available, the resulting values for physical constants may be less precise or accurate. This can be particularly problematic for less well-studied physical constants or for those that are difficult to measure.

- Potential for bias: The CODATA method involves combining data from different experiments, each of which may have its own sources of bias or error. If these sources of bias are not properly accounted for, they may introduce additional bias into the result.

- Uncertainty estimation: Estimating the uncertainty associated with the measurements is an important part of the CODATA method. However, it can be challenging to accurately determine the uncertainty, particularly if there are multiple sources of error or if the uncertainty is non-uniform across the range of measurements.

- Dependence on assumptions: The CODATA method also relies on several assumptions, such as assuming that the underlying distribution of measurements is Gaussian or that the measurement errors are uncorrelated. These assumptions may not always be held in practice and can affect the accuracy of the result.

- Limited applicability: the CODATA method is primarily designed for processing experimental data on physical constants. It may not be as useful for other types of data or in situations where there are significant systematic errors or unknown sources of bias.

Application of the FIQ-based method for the analysis of measurements of the speed of sound

To analyze the results of measuring the speed of sound, three studies were selected, each with its own unique method of measurement, medium of sound propagation, and number of variables considered in the model. In [32], hydrogen chloride in the liquid and dense vapor phases was used as the medium. In [33], experiments were conducted on binary mixtures (N_2 + H_2), while in [34], a group of scientists used 36 elementary solids, including semiconductors and metals with high binding energies, as the propagation medium.

Each of the three studies presented novel and accurate experimental data, showing good agreement between theoretical calculations and experimental results, or with the available experimental literature data. In [33], three thermodynamic models were proposed based on the Helmholtz energy and their effectiveness was tested by comparing the experimental data on the speed of sound and acoustic virial coefficients with the results predicted by the reference thermodynamic models. In [34], the authors proved the existence of an upper limit on the speed of sound in condensed phases, which depends on the combination of two physical constants: the fine structure constant α and the ratio of the mass of an electron to the mass of a proton. According to the authors of [32], the proposed equation of state is applicable to the entire liquid region, making an important contribution to the accurate modeling of the thermal properties of hydrogen chloride.

However, none of the studies have compared the achieved relative uncertainty of the experiment with the difference between the theoretical predictions and the experimental data, which raises doubts about the legitimacy of using the formulated models. The application of these models in describing the propagation of sound in various media carries a significant risk. Moreover, in the case of relative uncertainty, the question of which model to prefer remains unanswered. To address this issue, we employed the FIQ-based method to determine the preference for the physical meanings of the compared models.

Table 3 presents the compressed data for the measurement of the speed of sound. After a careful analysis of the information provided in the articles, we have formulated the following comments:

Table 3. Comparison of research results

Variable/ Reference	Chosen GoP_{SI} of the model	Number of FIQs inherent in GoP_{SI}, γ_{GoP} = z'-β'	Optimal number of FIQs inherent in a model, γ_{modi} = z''-β'', i ☐ 1, 2, 3	The achieved experimental comparative uncertainty of the model, ε_i	The comparative uncertainty of the model, theoretically justified for the selected GoP, ε_{opti}	Ratio of $\varepsilon_i/\varepsilon_{opti}$
[32]	LMTθ	846	$\gamma_{mod1} \approx 19$	$\varepsilon_1 = 0.0233$	ε_{opt1} = 0.0442	≈ 0.53
[33]	LMTθ	846	$\gamma_{mod1} \approx 19$	$\varepsilon_2 = 0.0305$	ε_{opt2} = 0.0442	≈ 0.69
[34]	LMTIF	1412	$\gamma_{mod1} \approx 52$	$\varepsilon_3 = 0.0596$	ε_{opt3} = 0.0738	≈ 0.8

The researchers involved in the study of sound propagation may have their own philosophical biases that influence their choice of variables when modeling the process. However, it is important to note that the number of selected variables may be limited, potentially resulting in the omission of important, and underlying connections. This highlights the importance of considering a broader range of variables and potential interactions between them to model the complex process of sound propagation more accurately in different environments.

The ratios obtained from the experiments ($\varepsilon_1/\varepsilon_{opt1}$=0.53 < $\varepsilon_2/\varepsilon_{opt2}$=0.69 < $\varepsilon_3/\varepsilon_{opt3}$=0.8) provide evidence that the model proposed in [34] is the preferred choice. This is because the number of dimensionless variables considered in the [34] model is much closer to the recommended one. This is supported by the values of the ratio $\gamma_i\backslash\gamma_{opti}$:

$\gamma_1\backslash\gamma_{mod1}$ = 1\19 = 0.05 [32] < $\gamma_2\backslash\gamma_{mod2}$ = 4\19 = 0.21 [33] < $\gamma_3\backslash\gamma_{opt3}$ = 18\52 = 0.35 [34].

While acknowledging the scientific progress achieved in previous studies [32] and [33], the informational approach highlights the importance of considering many variables in a model. In the case of sound velocity, the model presented in [34] is preferred due to its inclusion of a greater number of variables. This approach deepens our understanding of the true nature of the phenomenon and opens up new avenues for research, revealing hidden connections and expanding our knowledge.

Discussion

This article presents a new approach to the problem of achievable measurement accuracy. Instead of focusing on technical improvements to statistical methods, computers, and experimental equipment, this approach considers the modeling process that precedes the experiment. It recognizes that model uncertainty, which is influenced by the consciousness of the thinker, introduces a new element that limits measurement accuracy.

The ε-equation represents a deep connection between the accuracy of measurement and the consciousness (will) of the thinker. While this equation may be difficult to understand, it highlights the importance of the model formulation stage in the measurement process. This approach shows that the existence of an additional limit on measurement accuracy is not mysterious; it is simply a consequence of the finite amount of information contained in the model used to understand the object under study.

By acknowledging the role of consciousness in the modeling process, this approach overcomes the difficulties of determining the causes of the initial "blurring" of the object under study and sets a possible limit on measurement accuracy. While this concept may be challenging to accept, it provides a clear and logical explanation for the limitations of measurement accuracy beyond the Heisenberg uncertainty principle.

The concept of ε represents an intrinsic conceptual uncertainty in any physical or mathematical model, independent of the measurement process. This uncertainty depends on the number of base and derived variables and the chosen GoP, which are determined by the thinker's will. It is important to note that ε is not a result of the measurement, but rather an inherent property of the model. While the overall uncertainty of the model's objective function is much larger than ε due to additional uncertainties arising from the structure of the model and processing of experimental data, ε provides a fundamental limit to the achievable accuracy of the measurement.

This proposal presents a new approach to assessing the achievable accuracy of measurements with a clear physical basis. While the implications of this approach for biological systems have not been fully discussed, we deliberately did not explore the application of the FIQ-based method to genetic processes, as variables from the system of units are not used in this area of research. Nonetheless, the use of information theory to represent the modeling of physical phenomena or technological processes is highly relevant and important in scientific and technical research. This informational approach provides a valuable tool for understanding the limitations and possibilities of modeling physical systems.

Our study has demonstrated the applicability of the information approach in selecting the optimal structure of the model for studying the behavior of a physically observable object. We have applied this method to calculate the speed of sound, measure physical constants, and investigate the potential mass of information on Earth. By adopting this paradigm, scientific research can go beyond the mere description of "how" a measurement is prepared and move towards constructing an optimal model from the thinker's perspective. Our findings provide a valuable framework for scientific research that can help researchers choose the best structure of the model for their studies.

Conclusions

The information method is not a direct solution for building an accurate model of an observed object. Instead, it provides researchers with a tool in the form of comparative uncertainty to construct a model that minimizes mismatches with the phenomenon or process under study.

The research presented the trade-off between model complexity and accuracy, as measured by the amount of information. It also explains why finding the optimal structure for a model that accurately describes a physical object can be challenging and ultimately depends on the subjective decisions of the researcher.

The informational approach can be useful in avoiding overly simplistic or complex assumptions, as real-world phenomena are often more complex than the models used to represent them. By striving for greater accuracy through more complex models, researchers can use the informational approach to better understand the world around them.

The informational approach also emphasizes the role of the intellect, will, consciousness, intuition, and life experience of the researcher as integral components of the reality being studied. This highlights the importance of considering the subjective experience of the researcher in modeling physical objects and processes.

The FIQ-based method is a promising approach that has the potential to improve our understanding of physical systems. It is based on the mathematical framework of information theory and provides a clear definition of comparative uncertainty, which can help us identify the most reliable structure for a model and interpret experimental results in a balanced manner. The method takes into account the finite amount of information available in the system of units and the model, as well as the subjective decisions of the researcher.

When building a model to represent a process, it's important to acknowledge that only a finite amount of information is available to accurately describe its state. This information is inherently uncertain, and this uncertainty cannot be avoided. This principle of finiteness, as described by [35], is not a result of mathematical or physical limitations, but rather a fundamental aspect of the model building process. Therefore, it's crucial to consider the limitations of the available information and the resulting uncertainty when using models to make predictions or draw conclusions.

Finally, while the FIQ-based method shows great potential, there is still much work to be done to fully develop and refine this approach. However, continued research into this method has the potential to deepen our understanding of physics and technology, and to provide new insights into the behavior of complex physical systems.

Bibliografy

[1] Grecis, F. On the meaning of measurement uncertainty. Measurement, Vol. 133, 2018, pp. 41-46. https://sci-hub.tw/10.1016/j.measurement.2018.09.073

[2] Jaynes, E.T. Information theory and statistical mechanics. Revs. Modern Physics, vol. 27, No. 289, 1955, pp. 620-630. https://bayes.wustl.edu/etj/articles/theory.1.pdf

[3] Kapur, J.N. Maximum-Entropy Models in Science and Engineering. Wiley eastern Limited, New Delhi, India, 1993. https://books.google.co.il/books?id=LuNIAp3QorUC&printsec=frontcover&vq=%22Maximum+Entropy+and+Bayesian+Methods%22&source=gbs_citations_module_r&cad=3#v=onepage&q=%22Maximum%20Entropy%20and%20Bayesian%20Methods%22&f=false

[4] Gell-Mann, M., Lloyd, S. (1996). Information measures, effective complexity and complete information. Difficulty, Vol. 2, N. 1, 1996, pp. 44–52. https://sci-hub.hkvisa.net/10.1002/(sici)1099-0526(199609/10)2:1%3C44::aid-cplx10%3E3.0.co;2-x

[5] Allaire, D., He, Q., Deyst, J., Willcox, K. An Information-Theoretic Metric of System Complexity With Application to Engineering System Design. J. Mech. Des., Vol. 134, No. 10, 100906, 2012. Pp. 1-10. https://doi.org/10.1115/1.4007587

[6] Deffner, S., Jarzynski, K. Information Processing and the Second Law of Thermodynamics: An Inclusive Hamiltonian Approach, 2013, pp. 1-14. https://arxiv.org/pdf/1308.5001.pdf

[7] Fang, L., Ouellette, N.T. (2021) Assessing the information content of complex flows. Phys. Rev. E, Vol. 103, 023301, 2021, pp. 1-7. https://link.aps.org/accepted/10.1103/PhysRevE.103.023301

[8] Menin B. 2017 Information Measure Approach for Calculating Model Uncertainty of Physical Phenomena. Amer. J Comput. Appl. Math.,

Vol. 7, No. 1, 2017, pp. 11–24. https://goo.gl/m3ukQi

[9] Menin, B. Construction of a model as an information channel between the physical phenomenon and thinker. J. Assoc. inf. sci. Technol., Vol. 72, No. 9, 2021, pp. 1198-1210. doi: 10.1002/asi.24473.p.

[10] Menin, B. Can Information Theory Help to Formulate an Optimal Model of a Physical Phenomenon? Journal of Applied Mathematics and Physics, Vol. 10, No. 10, 2022, pp. 2940-2950. https://www.scirp.org/pdf/jamp_2022101214331576.pdf

[11] Del Santo, F., Gisin, N. Physics without determinism: Alternative interpretations of classical physics. Phys. Rev. A, Gen. Phys., Vol. 100, No. 6, 2019, pp.1-9. doi: 10.1103/PhysRevA.100.062107.

[12] Burgin, M. Information theory: a multifaceted model of information. Entropy, Vol. 5, No. 2, 2003, pp. 146–160. https://sci-hub.tw/10.3390/e5020146

[13] Menin, B. Simplest Method for Calculating the Lowest Achievable Uncertainty of Model at Measurements of Fundamental Physical Constants. Journal of Applied Mathematics and Physics, Vol. 5, No. 11, 2017, pp. 2162-2171. https://goo.gl/qNyUC8

[14] Menin, B. *h*, *k*, N_A: Evaluating the Relative Uncertainty of Measurement. Am. J. Comp. Appl. Math., Vol. 8, No. 5, 2018, pp. 93–102.

http://article.sapub.org/10.5923.j.ajcam.20180805.02.html

[15] Menin, B. Hubble Constant Tension in Terms of Information Approach. Phys. Sci. Int. J., Vol. 23, No. 4, 2019, pp. 1-15. https://doi.org/10.9734/psij/2019/v23i430165

[16] Menin, B. Unleashing the Power of Information Theory: Enhancing Accuracy in Modeling Physical Phenomena. Journal of Applied Mathematics and Physics, Vol. 11, No. 3, 2023, pp. 760-779. Doi: 10.4236/jamp.2023.113051.

[17] Newell, D.B., Tiesinga, E. The International System of Units (SI). NIST Special Publication. Gaithersburg, MD, USA, Tech. Rep. 330, 2019, pp. 1–138. https://doi.org/10.6028/NIST.SP.330-2019

[18] Sedov, L.I. Similarity and Dimensional Methods in Mechanic. USA: CRC Press, Florida, 1993.

[19] Brillouin, L. Science and Information Theory. New York, NY, USA: Academic Press, 1953.

[20] Menin, B. Applying Measurement Theory and Information-based Measure in Modeling Physical Phenomena and Technological Processes. Eur. J. Eng. Res. Sci., Vol. 3, No. 1, 2018, pp. 28-34. https://ej-eng.org/index.php/ejeng/article/view/594

[21] Landauer, R. Information is Physical. Phys. Today, Vol. 44, No. 5,

1991, pp. 23-29. https://sci-hub.se/10.1063/1.881299

[22] Vopson, M.M. The information catastrophe. AIP Adv., Vol. 10, 2020, pp. 1-5. https://sci-hub.se/10.1063/5.0019941

[23] Mohr, P.J., Newell, D.B., Taylor, B.N., Tiesinga, E. Data and analysis for the CODATA 2017 special fundamental constants adjustment. Metrologia, Vol. 55, 2018, pp. 125–146. https://iopscience.iop.org/article/10.1088/1681-7575/aa99bc/pdf

[24] Khruschov, V.V. 2021 On possible ways for future redefinition of the SI time unit, 2021, pp. 1-10. https://arxiv.org/ftp/arxiv/papers/2102/2102.05863.pdf

[25] Yarin, L. The Pi-Theorem. Berlin, Germany: Springer-Verlag, 2012. https://goo.gl/dtNq3D

[26] Menin, B. High accuracy when measuring physical constants: from the perspective of the information-theoretic approach. J. Appl. Math. Phys., Vol. 8, No. 5, 2020, pp. 861–867. DOI: 10.4236/jamp.2020.85067

[27] Wübbeler, G., Bodnar, O., Elster, C. 2017 Robust Bayesian Linear Regression with Application to an Analysis of the CODATA Values for the Planck Constant. Metrologia, Vol. 55, 2017, pp. 20-28. https://sci-hub.hkvisa.net/10.1088/1681-7575/aa98aa

[28] Pavese, F. The New SI and the CODATA Recommended Values of

the Fundamental Constants 2017 Compared with 2014, with a Comment to Possolo et al. Metrologia, Vol. 55, 2018, pp. 1-11. https://arxiv.org/ftp/arxiv/papers/1512/1512.03668.pdf

[29] Dodson, D. Quantum Physics and the Nature of Reality (QPNR) Survey: 2011. 2013. https://newatlas.com/confusion-basic-nature-quantum-mechanics/26216/

[30] Henrion, M., Fischhoff, B. 1986 Assessing Uncertainty in Physical Constants. Am. J. Phys., Vol. 54, 1986, pp. 791-798. https://sci-hub.se/10.1119/1.14447

[31] Menin, B. Simplicity of Physical Laws: Informational-Theoretical Limits. IEEE Access, Vol. 10, 2022, pp. 56711-56719. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9780195

[32] Thol, M., Dubberke, F.H., Baumhögger, E., Span, R., Vrabec, J. Speed of Sound Measurements and a Fundamental Equation of State for Hydrogen Chloride. J. Chem. Eng. data., Vol. 63, No. 7, 2018, pp. 2533–2547. https://doi.org/10.1021/acs.jced.7b01031

[33] Segovia, J.J., Lozano-Martin, D., Tuma, D., Moreau, A., Carmen Martín, M., Vega-Maza, D. Speed of sound data and acoustic virial coefficients of two binary (N2 + H2) mixtures at temperatures between (260 and 350) K and at pressures between (0.5 and 20) MPa.

The Journal of Chemical Thermodynamics, Vol. 171, 2022, pp. 1-13.
https://doi.org/10.1016/j.jct.2022.106791

[34] Trachenko, K. et al. 2020 Speed of sound from fundamental physical constants. Sci. Adv., Vol. 6, No. 41, 2020, pp. 1-6.
https://advances.sciencemag.org/content/advances/6/41/eabc8662.full.pdf

[35] Sternlieb, A. The principle of finiteness – a guideline for physical laws. J. Phys.: Conf. Ser., Vol. 437, 012010, 2013, pp. 1-12.
https://iopscience.iop.org/article/10.1088/1742-6596/437/1/012010/pdf

Authors' Information

Boris Menin – Mechanical & Refrigeration Expert, Beer-Sheva 8464209 Israel; e-mail: meninbm@gmail.com

Major Fields of Scientific Research: Modeling physical phenomena, Applying information theory for the model structure, Pumpable ice technology

DESIGN PRINCIPLES FOR SOFTWARE USING TECHNIQUES OF BIOLOGIC INFORMATION PROCESSING

Karl javorszky

Abstract**: We propose to use a technique of counting ‘in stereo’, by using **two** counting systems concurrently. Research in number theory has brought up a very slight inner discongruence within the numbering system. The proposal is to root the definition of the term ‘information’ in properties of natural numbers, with the refinement that the extent of information is equal to the extent of being otherwise. The property of being otherwise can be numerically defined on a collection of pairs of natural numbers. An etalon collection of pairs (a,b) is used to demonstrate the concept of being otherwise. The property of being otherwise is based on the numeric fact, presented in oeis.org/A242615, according to which there are **two** upper limits of the number of logical relations existing among members of collections, namely sentences which state the existence of **similarity, ‘n?’** contrasted to sentences that describe **differences, ‘n!’,** among the members of one and the same collection. The values f^{-1}**(number of different relations)** point back to slightly differing values of **n,** and this in the form of Δ (f^{-1}(n?), f^{-1}(n!)) {=, <, =, >, =, <, <<} for n {1, 11, 32, 66, 97, 136, 140}. The numeric incongruences appear to play important roles in Physics, Chemistry, Theoretical Biology. Here, in the context of software design, the practical relevance of being able to switch between **two** methods of describing one and the same state of a set, becomes apparent in the **increased efficiency** in the transmission of messages. Recognition of forms and patterns in (eg optical) multitudes becomes possible by using the properties of **cycles,** which are an artefact of periodic changes which sender and receiver subject the collection of tokens to, in a pre-agreed (or, in Nature, a-priori existing) algorithm. The theoretical foundations for the possibility of recognition of Gestalts are* *presented to be implications of the inner incongruence of the numbering system, as shown by oeis.org/A242615. There exist* ***shortcuts, *abbreviated names for* ***patterns.*** *This is equivalent to a condensation of information.*

Keywords*: Information, fundamental definition of; memory; learning; AI; enumerations of states of sets.*

ITHEA Keywords*: F.4.1, H.2.4, I, I.2.11, I.6.5, H.1, I.4.8, J.3, A.1, H.3.3, E.1, D.2.2, C.2.4, I.5.1, D.2, A.0*

Conference and topic*: SE 2022 SoftEngine International Conference on Software Engineering http://www.ithea.org/softengine/*

DOI*: https://doi.org/10.54521/ijita29-03-p02*

Introduction

We ask the reader to conceptualize an inbuilt quirk in the numbering system. The quirk itself is pointed out in oeis.org/A242615.

In essence, we make use of the inbuilt relative inexactitude to create a counting and enumerating system which is more detailed, more exact ***by orders of magnitudes*** *compared to the counting – enumerating system presently in use. One may hope that the fineness of units used in quantum-based models can be achieved by using the model presented here, which creates fractions of units.*

Example:

We use *two* eyes to perceive the world. Of the *two* pictures delivered by perception, cognition creates *one,* integrated picture. It is a culturally anchored, common sense approach that we speak about the form and content of that *one* picture, which has been assembled, if we talk about the world. As senders of messages, we rely that the receiver has a comparable neurological apparatus at their disposal. If we say "House H is located at the corner of Streets A, B", we think of *one* landscape or system of coordinates, where we can point out a specific location. We know that the recipient of this message will similarly employ *one* mental or physical map, on which they shall identify a location.

It would be quite unusual, if we said: “As seen by the ***left*** eye, House H is located at corner of Streets A, B, *displaced by the* ***perspective bias ‘left’*** *by λ %; a*s seen by the ***right*** eye; House H is located at corner of Streets A, B, *displaced by the* ***perspective bias ‘right’*** *by σ %.*” Aside of being unusual, there is nothing to say against this way of messaging the coordinates. The sentence is grammatical and we do have the technical means to establish exactly what we refer to by the terms *(λ, σ).*

The main hypothesis is that it is **worth while to refrain from integrating**, merging, fusing the two underlying pictures into one overriding picture, which we then by cultural tradition call The Picture. We do have the accounting tools and computer capacities to categorize and enumerate the slight deviations caused by factors (λ, σ). The slight fragments that distinguish the left picture from the right one are a nuisance in our current understanding. We do not question Nature’s ability to merge the two pictures into one within our brain. It can be done, and it is very practical that it is done, and common sense requires that if we talk about pictures we mean the net result, and not the underlying additions and subtractions of some small details that are artefacts of factors (λ, σ).

The question arises naturally, what benefit would it bring if we transmit the same coordinates in a more circuitous way, by using two reference points instead of one to build up the picture. There is no sense in describing a picture more exactly, if it is and remains the same picture about the same state of the world.

The surprise is that there are additional benefits to using two reference systems instead of one, further and beyond of being exact for the exactitude’s sake. In the example with the two optical pictures coming from two eyes, the focus is on that part of the two pictures which are congruent, translatable into each other and refer to one and the same state of the world.

Using the two metrics in a fashion that is not optimized towards congruence, one sees interdependences in the system which are best approached in a context of economics.

It is more *efficient* to transmit a message in batches of *66* tokens which can be in any order (think of beads in a sack) than have *66* tokens in a sequential order (think of beads forming a necklace).

It is more *efficient* to transmit a message in batches of *6 * 11* tokens which are sequenced than have *66* tokens which can be in any order. in a sequential order (think of beads forming a necklace).

The term *efficient* refers to the number of tokens needed to carry a given number of messages. The question: *“How many logical relations make up one object?”* can now be numerically approached.

The stereo approach to counting is logically solid, is nothing but an added chapter to arithmetic. Its basic concepts run against deep-seated neurological and cultural conventions and habitudes, but not against rules of computing.

Formal Introduction of the Model

1. What is the problem?

The meaning of the term ‘information’ needs to be rooted in the numbering system. Information is a phenomenon of Nature and also of our mental creations. It has to be understood as a fundamental property of our numbering system. The numeric facts shown in A242615 allow for using *two* reference points on **N.** Every measurement (statement, result) can be reformulated using principles and techniques known from interferences.

A second problem is, that in the theory in force until presently, it is not possible to *compress and unfold* information, like we observe Nature doing it in the genetic information transfer and in the memory. If one has only *one* **N**, one cannot collapse and recreate the rules that enforce properties of material assemblies. Having in addition to **N** further two variants of **N: Nλ, Nσ** allows protracted exchanges of *relation ↔ fragment of object.*

2. Why is it interesting and important?

It is high time to update the philosophy behind the fundamental axioms that restrict our ways of thinking about counting. In the 21^{st} century, in the midst of a social revolution regarding individuality, we should switch from the idea of the unit being of one standard kind, indistinguishable from any other units, to the idea of the unit being an individual, contained in a cohort of individuals, which are by their nature both similar and different compared with the other units of the cohort.

The time of infallibility and persuasion by the sword has passed her best days. It is time to allow for Truth to be not the one-and-only variant but to acknowledge the existence of alternative truths in the abstract fields, too. Using a model in which alternative truths exist requires multidimensional thinking, which the human brain is not as good at as are computers. A world view built by computers needs some getting used to, but its utter boring rationality gives one a solid halt in a world that is full of logical rivalries, contradictions, local breakdowns. The picture is not as unified and harmonic as previously, but is much closer to reality.

3. Why is it hard?

European culture, rooted in Judeo-Christian concepts of monotheism, has been maintaining that its concepts fit into each other seamlessly, perfectly, free of contradictions and ambiguities. The Cathars, who have evolved a system of concepts of duality, have been put to the sword as heretic in the 14^{th} century.

Their fate has been a strong warning for anyone experimenting with epistemological concepts of dialectic and tensions within the system model.

In our schools, children are taught, that *a+b=c* as a narrative has only one valid interpretation, namely that of *c,* where all differences that are present in the version *a+b* will **magically disappear** in the process of addition. While *a,b* can reasonably be assumed to be *two* ideas, *c* is imagined to be made up of *one* kind of units. The unification and regularization of basic units offers many advantages, yet the idealized version negates the existence of differences that were present in the form of *a,b.* Our children are offered a **heavily sanitized, family-friendly version** of the relations between the parts and the whole. In that version, taught up into our days, everything adds up finely, and there are neither rivalries nor conflicts among the parts of a whole.

We have been taught to use *one* sheet of millimeter paper as the background of any of our mental creations and explanations. This attitude may be practical and easy to administer, but it fails to take into account the actual reality of our living in a world that is based on *duality.* Whether it is *proton* ↔ *neutron, DNA* ↔ *RNA, female* ↔ *male, foreground* ↔ *background, sequential* ↔ *contemporary, breathing in* ↔ *breathing out, etc.,* we are confronted with empirical observations that call out our time-honored, traditional, idealized and simplified fundamental concept of *mono-perspectivism.*

We know of the practical advantages of using interferometry. Triangulation as a technique is a simple implication of using *two* sets of references (as opposed to the present basic system of ideas, in which all measurements are related to that one-and-only reference point of *0, Zero.*

As the examples with the picture in the brain, and in our communications among each other, show, having a convention that the *one, merged* picture is the only one to be referred to, brings forth advantages of uncomplicated and easy design principles for any system of thoughts that builds a model to depict Nature with. The unfortunate fact is, that Nature builds from the bottom up and not from the top down. Before the merged picture can evolve, there must exist *two* versions of it, which become merged by faculties of our brain. The *two, slightly deviating* pictures are ***logically prior to*** the *one, perfectly fitting picture.* Nature builds on the differences, uses the differences while integrating the two pictures into one. In the organizational level of adaptive, self-correcting feedback loops (our neurology, eg), the remainders and rests of the two pictures are recycled. Those fragments of the two basic pictures, which after having served their usefulness as constituents of the merged picture, remain as fragments (like snippets of pieces of cloth in a tailor's shop) and are recycled by the brain (we encounter them in the form of dreams); unrecycled these rests and remainders perform the procedure of *entropy*.

4. Why do naive approaches fail?

For psychologists, it is a natural and fundamental truth that there exist several versions of truth, each equally legitimate and valid. In the STEM environment, the concept of multicausal genesis is not rooted generally yet.

The a-priori existence of *patterns* is an undisputable fact of Nature. The classical mathematical philosophy does not foresee any meaningful relations among natural numbers as they stand. The importance of ***cycles*** as porters of a-priori logical relations between amounts, places and time has not yet been recognized.

The *upper limits* for the number of sentences that can be said about a collection of *n* members have only been dealt with, so far, in the form of counting sentences that describe differences among the members before a background of similarities. The parallel measure, the maximal number of distinct sentences that describe similarities among the members before a background of diversities, has not been addressed yet. *The existence* of upper limits of relations on objects *f(n)* paves the way to ideas of thresholds, quality transformations, saturation: these ideas are fundamental in biology, but not yet working day concepts in the technical sciences. *The extent* of upper limits – as we contrast the maximal similarity among objects with the maximal diversity among objects – opens up a completely new area of arithmetic, geometry and information transmission. The terms of trade in the *three-way Bazar* are **sliding *f(n).*** There exists a numeric table for the terms of trade among *{number of similarity relations, number of diversity relations, number of objects}.*

5. Why hasn't it been solved before?

Reason 1.: Ideologic, theological reluctance to depart from the unified concept. This is well reflected in the underlying idea of the Shannon concept of information: that messages transmission is identifying 1 of elements of **N** and communicating the algorithm to the receiver, who has at his disposal the same **N.**

The idea to use *two* measuring rods that are calibrated slightly differently, and use the immanent interference pattern between the two to identify 1 of elements of **N** as the background, and communicate the algorithm to the receiver, who has at his disposal the same interference pattern established by the slight relative mis-calibrations of the two measurement rods, this idea introduces concepts and techniques of *accounting*, with two different ledgers and one consolidated one. Maybe the idea is too much practical for number theory.

Reason 2.: Ehrenfels inhibition is the name for the observation, that humans' observational capacities are more challenged if the stimuli to be remembered are irritatingly close (similar) to each other. There appears to be a requirement of balance among similar and diverse so that our neurology functions best. We are, our neurology is optimized for an environment that contains in a relatively stable proportion input and impressions of a similar and of a different nature. We would not be optimized for a feature of Nature if that feature was not there. So there exists in Nature, by itself, a proportion between that what remains similar and that what changes. Tradition has it that one speaks about that, and uses as a background that, what remains and is the same. Relative to this background, processes can be observed to take place. Exchanging the perspectives and inquiring what the picture looks like if changes and rearrangements are axiomatic and are used as background, before which we can observe that what happens regularly, inevitably, predictably is a mental dexterity which salesmen, diplomats, lawyers and psychologist are trained in, but not those educated in the technical sciences. The background and the foreground are of course merged in our cognition, and they are irritatingly similar to each other. Dissecting their properties needs some returning attention.

Reason 3.: Availability of computers is a key factor. It is absolutely impossible to gain insight into the interactions of *(a,b)* in the adult version of the story of their romance, without a computer.

6. What are the key components of my approach and results?

The building blocks of the model are:

<u>a. the actors</u>

We conduct a mental experiment with a cohort of pairs of natural numbers *(a,b).* The cohort we prefer to use contains pairs of members which each can have *16* different variants, yielding a cohort size of *136* so-called *logical primitives. (© M. Abundis).*

<u>b. the habitat</u>

The logical primitives undergo ***periodic changes***. We use *aspects* of the members, like *a+b, a-b, b-2a* to create ***sequential orders.*** When 2 aspects are concurrently the case (two periodic changes happen concurrently, like the Earth’s rotation and the Moon’s position), there exists a coordinate for a ***planar place*** which is given by the two relevant sequential ranks, for each primitive. Using such planes that can be arranged perpendicularly, by having common axes, we find *two* Euclid-type spaces. The primitives have by this method found ***a spatial position***, actually *2* of such.

c. the paths in the habitat

The spatial grid itself is built up by ***cycles.*** Cycles appear as constituents of a procedure that is known as ‘reordering’. The topography of the habitat is rooted in the properties of the *similarity-related* cycles, the content of the habitat is determined by the cycles that are present due to the *diversities* among the members.

Using these building modules of the model, ***we observe:***

a. agglomerations

The collection being subject to more than two reorders at the same time, it is inevitable, that on specific spatial coordinates agglomerations of amounts come into existence. Like in traffic simulation: if sufficiently many cars are on the road traveling to sufficiently diverse and similar goals, pileups are unavoidable. The pileups come in ***types.*** (Like traffic jams can be differentiated on the number of cars involved, their momentum, composition of types of cars, etc.) As these agglomerations appear out of nowhere in our perception, although they had always been there, only we have not knowingly recognized them, like archetypes, we propose to name the various types of material agglomerations due to periodic changes that elementary logical symbols undergo, ***logical archetypes.*** The concept of logical archetypes should correspond to the concept of ***chemical elements.***

b. regularity

It is in the nature of periodic changes that they are a sequence of recurring states. The regularity allows a concept of ***evolutionary adaptability.*** Whether self-organization is an immanent property of Nature, can be decided by investigating the following hypothesis:

Any random arrangement of the logical primitives is closer (more similar, easier to transform into) to an order based on properties of logical primitives, than to a different arrangement which is not based on properties of logical primitives.

In our man-made habitat, it is easy to see that periodic changes impose a ***rhythm*** on the assembly.

c. predictability

Recurring states can be predicted.

A state can be identified/predicted by the appearance of such members of cycles which are known to be predecessors in cycles containing the constituent of the state to be predicted/identified. Thy cycles bring forth signs of change.

d. alternatives

The whole model is – but for the logical archetypes – a collection of alternatives. The alternatives have a numeric property of their extent of being otherwise. It is a huge accounting challenge, but it is theoretically possible to classify and tabulate all possible alternatives. Of this lot, only such need to be used, which are included in all ways of describing the collection: as a web of relations similarity-based, as a web of relations diversity based, as a content that is expressible in both descriptions, as a process that creates and reduces <*so-and-so much*> discongruence.

Closing remarks

The present paper offers an invitation. Why don't you explore the completely new landscape, which gives a formidable background for quite many of established concepts from Physics and Chemistry?

The ideas presented here make the impression of being no less than revolutionary. This reflects less on the novelty of the new algorithms, but rather on the conservativism and rigidity required in the technical sciences. There is nothing extraordinary in the idea to assemble an etalon cohort of simple symbols and subject them to periodic changes. The question is what one reads out of the observations and whether one dares believe that one has proceeded logically and rationally while attaining the observed results.

In the context of software, basic design principles of, one can rephrase the general concept in terms of search strategies. We can address 1 member of **N** by sequential methods or by methods of queries (contemporary searches, based on categories of commutative symbols). It has been shown (cf A242615) that query-based searches are ~ *350 %* **more efficient** relative to sequential searches if the batch of information carrying vehicles numbers ~ *66*.

Using the cycles as indexes of relationships, one finds a **lexicon** which is by its immanent features **already cross-referenced.**

Summary

It is unavoidable for the Reader, if they wish to continue investigations into cohesive counting, that they generate, program, set up their own *tautomat.* This extensive system of tables is comparable to the trigonometric tables, inasmuch as relations among numeric values are explicated. Please be aware that the model created by observing paths of strings of elements that belong to one cycle is in its complexity and versatility comparable to a **mother of all hybrids of Sudokus with Rubik cubes.** Instructions on how to build your own tautomat are eg in *Picturing Order.* The back-of-the-envelope laboratory prototype at www.tautomat.com is also available.

For a research institute which has insight, resources, circumstances to set up a public version of The Tautomat (like the Online Encyclopedia of Integer Sentences, a common resource), the advantages would be formidable. If you host the etalon laboratory animal, you will be the first to know, which variations and mutations that system of implications will more create, aside the few mentioned here. This could be a massive competitional advantage. One cannot lose a bet on the idea that ordering and reordering simple logical tokens will turn up typical patterns and that these manifold typical patterns will be of interest to Physics, Chemistry and Physiology.

Bibliography

Javorszky Karl. Natural Orders De ordinibus naturalibus, Morawa, Wien 2016, ISBN: 978-3-99057-137-8 https://www.facultas.at/item/32109801

Javorszky Karl. Picturing Order, https://www.researchgate.net/publication/329275837_Picturing_Order

Authors' Information

Karl Javorszky: *clinical psychologist; retired. Loeblichgasse 13/16, A-1090 Wien, e-mail: karl.javorszky@gmail.com*

Scientific Interest: Memory, How registering and reconstructing information is actually done, as a task of accounting and data processing.

MODEL OF SELECTION OF CHARACTERISTIC PARAMETERS OF A VOICE SIGNAL FOR ANNOUNCER IDENTIFICATION TASKS

Yana Bielozorova

Abstract: *A method for isolating the parameters of the voice signal characteristic of the speaker by selecting the features of the extremums of the surface of the scalogram of the voice signal, built on the basis of wavelet coefficients. The structural model of the speech signal is described, the sequence of mathematical transformations of the information channel of speech communication is considered and generalized, on the basis of which the approach to selection of features of language structures is offered. An algorithm for selecting the characteristics of speech signal structures has been developed. The selection uses a complex Morlet wavelet, which most effectively describes the structure of the speech signal at the first level of decomposition. By synthesizing the speech signal with the specified characteristics, the rational parameters of the wavelet transform settings for the implementation of the method are selected.*

Keywords: *voice signal, Morlet wavelet, special structures, fundamental frequency, speaker identification.*

DOI: https://doi.org/10.54521/ijita29-03-p03

Introduction

The accuracy of speaker identification is based on the correct choice of identification features and the correct presentation of the voice signal. One of the most important characteristics and often determining is the frequency of the fundamental tone. Currently, there is no single approach to determining the fundamental frequency, so effective identification of the announcer is possible only on the basis of a comparison of the frequencies of the fundamental tone, performed on the basis of the same method of determining the frequency of the fundamental tone.

The basis for improving the accuracy of identification of the speaker by the voice signals left by him is a qualitative determination of the frequency of the fundamental tone, which in turn depends on the correct interpretation of the characteristics of the speaker in the voice signal.

Thus, the construction of a model for the selection of characteristic parameters of the voice signal, providing high accuracy in describing the characteristics of the speaker in the voice signal is an important scientific task.

Related works

Nowadays, it is well known that vocalized fragments of a person's speech signal maintain the periodicity of voice oscillations. The speech signal should be considered in the form of pulsating oscillations of the air flow, where the period of repetition of the pulses of the speech signal is called the frequency of the fundamental tone. It has been studied that the fundamental tone determines the structure of the speech signal, and is also the main parameter of the speech signal [Mallat, 1999]. The intonation contour of a person is a trajectory of changing the melody of the fundamental tone frequency.

Prosody of speech formation, one of the components of which is intonation, significantly distinguishes the acoustic signal of speech from written language. Therefore, the fundamental frequency carries a significant amount of information contained in the speech signal of the person. From the point of view of presenting this information, the process of allocating the fundamental frequency is of independent interest [Mallat, 1999, Muzy et al., 1991]. When calculating the fundamental frequency, it is also necessary to take into account both slow changes in the trajectory of the fundamental frequency, and its rapid changes and the very moments of inclusion / deactivation of the speech signal. However, when constructing systems for analysis, synthesis, compression or identification of a speech signal, it should be borne in mind that the fundamental frequency is usually used as one of the main features needed for improved description of the speech signal. When it is necessary to perform calculations of the fundamental frequency from a real speech signal, it is necessary to pay attention to the following points [Muzy et al., 1991]:

1) the ability to work at different noise levels;

2) estimation of the fundamental frequency should be performed with minimal error;

3) the minimization of error in a wide range of changes in the frequency of the fundamental tone, emotional and other changes in the language of the person, leading to variations in the frequency of the fundamental tone;

4) must perform all previous signal transformations;

5) effectively divide the signal into vocalized/unvocalized fragments.

To get the maximum amount of information contained in the fundamental frequency loop, you should also pay attention to the additional point when you need to perform instantaneous calculations of the fundamental frequency value.

Materials and methods

Model of speech signal representation based on wavelet analysis and determination of characteristics of its self-similar structures.

Let us present a speech signal in the form of components with different structure:

$$f(t) = x_1 f_1(t) + x_2 f_2(t) + \cdots + x_s f_s(t) \quad (1)$$

In the case of correlation of coefficients $x_1, x_2, \ldots, x_s$ it is difficult to draw a conclusion about the type of approximation functions. The main way to solve this version of the signal is to present it as components

$$\sum_i g_i f_j(t_i) f_k(t_i) = \delta_{jk} \quad (2)$$

where $g_i = 1/\sigma_i^2$.

Given that the functions f_i in (1) have a different structure, which is subject to changes at random moments of time, the most effective way to describe them is to use approximation methods based on basic decomposition

$$f_i(t) = \sum_n c_{in} \varphi_{in}(t) \quad (3)$$

where $f_i \in L^2(R), \varphi_{in}$ – basic spatial functions $L^2(R)$.

To create models that adapt to the structure of the signal, it is proposed to use nonlinear approximation schemes. In this case, the approximation f is performed by M vectors depending on the structure of the signal

$$f_M = \sum_{m \in I_M} \langle f, \varphi_m \rangle \varphi_m \quad (4)$$

where I_M –a set of indices that is determined by the properties of the function f.

The mathematical construction (1) taking into account the introduced properties (1) - (4) in relation to the description of the signal structure is called the structural model of the speech signal.

Given the search for self-similar structures in speech signals, their different shape and length, the most appropriate space for their representation is the space of wavelet bases. Wavelet coefficients $c_{j,n} = \langle f, \psi_{j,n} \rangle, where \{\psi_{j,n}\}_{(j,n)\in Z^2}$, are considered as a result of mapping the function f into a space with a resolution j.

Consider the process of selection of self-similar structures in the speech signal. Naturally, to obtain the characteristics of the signal and the selection of any structures in the signal, wavelet transform algorithms are used, which allow to decompose the signal by the operation of shift and tension of the wavelet ψ. The peculiarity of the wavelet is that its average value is zero, and its integral has the form

$$\mathrm{Wf}(a,b) = \int f(t)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)\mathrm{dt}\# \quad (5)$$

And allows us to estimate the behavior of f in the vicinity of point b, which is directly related to *a*. A feature of the wavelet coefficients obtained from (5) is the effective representation of the property of the function f at the proximity of the scale a to zero in the vicinity of point b. We write the Taylor polynomial for some neighborhood v, if the condition of differentiability of the function fm times in the interval $[v-l; v+l]$

$$\rho_v(t) = \sum_{k=0}^{m-1}\frac{f^{(k)}(v)}{k!}\# \quad (6)$$

With this kind of polynomial and evaluation of the quality of its description $\varepsilon_v(t) = f(t) - \rho_v(t)$ the following condition must be fulfilled

$$\forall t \in [v-l; v+l] |\varepsilon_v(t)| \leq \frac{|t-v|^m}{m!} \sup_{u \in [v-l; v+l]} |f^m(u)| \# \quad (7)$$

The estimate of the maximum error $\varepsilon_v(t)$ in the pursuit of t to v is determined by the order of differentiation f in the vicinity of v. To clarify the maximum value of the error, we use Lipschitz smoothness by adding the following indicator [Pavlov A.V. et al., 2007]:

The function f satisfies the Lipschitz condition $\alpha \geq 0$ at point v, if there exists $K > 0$ and the polynomial ρ_v of degree $m = \lfloor \alpha \rfloor$ such that

$$\forall t \in R; |f(t) - \rho_v(t)| \leq K|t-v|^{\alpha} \# \quad (8)$$

If the function f satisfies (4) for all $v \in [a, b]$ with a constant K, independent of v, then it is assumed that it satisfies the Lipschitz condition α on $[a, b]$. It is known that Lipschitz indicators give the most accurate representation of smoothness, which is used throughout the range.

In [Lardiès J. et al., 2004] it is shown if the wavelet ψ has n zero moments, i.e.

$$\int_{-\infty}^{+\infty} l^k \psi(t) \mathrm{dt} = 0, k = \overline{0; n-1} \# \quad (9)$$

and n derivatives, then there is $A > 0$ for $f \in L^2(R)$, which satisfies the uniform exponent Lipschitz $\alpha, \alpha \leq n$ on $[a, b]$

$$\forall (s, u) \in R^{+ \times [a,b] |\mathrm{Wf}(s,u)| \leq \mathrm{As}^{\alpha + 1/2}} \# \quad (10)$$

We can conclude that f will satisfy the Lipschitz condition α on $[a+\varepsilon, b-\varepsilon]$ for any $\varepsilon > 0$, if f is bounded, and wavelet coefficients Wf(s,u) will satisfy (10) for $\alpha < n$.

Thus, condition (10) makes it possible to say that the decrease in the amplitude of the wavelet transform of the speech signal depending on the scale is associated with the uniform and point smoothness of Lipschitz.

Estimating the self-similarity of f at point v on the basis of Lipschitz indicators can be quite difficult, due to the fact that they can change arbitrarily in the vicinity of point v. We use Jaffar's theorem [Xuedong, 2001] to impose the sufficiency condition on the wavelet transform for estimating the Lipschitz smoothness of the function f at the point v.

Let the wavelet ψ have n zero moments and n derivatives. If $f \in L^2(R)$ satisfies the Lipschitz condition $\alpha \leq n$ at point v, then there exists A such that

$$\forall(s,u) \in R^{+\times R|\mathrm{Wf}(s,u)|\leq \mathrm{As}^{\alpha+1/2}\left(1+\left|\frac{u-v}{s}\right|^{\alpha}\right)}\# \qquad (11)$$

Conversely, if $\alpha < n$ — not whole, but exists $A, \alpha < \alpha$ such that

$$\forall(s,u) \in R^{+\times R|\mathrm{Wf}(s,u)|\leq \mathrm{As}^{\alpha+1/2}\left(1+\left|\frac{u-v}{s}\right|^{\alpha}\right)}\# \qquad (12)$$

then f satisfies the Lipschitz condition α at the point v.

Therefore, when killing the scale s, the calculated amplitudes of the wavelet coefficients based on speech signals have a rapid decrease to zero in areas where the signal is smooth and has no self-similar structures.

Therefore, if |Wf(s,u)| has no local maxima on a small scale, it is assumed that the function f describing the speech signal is locally smooth and the process of isolating self-similar structures of the speech signal function f can be constructed by determining the maximum values of the functions $|\mathrm{Wf}(s,u)|$ on a small scale. It is taken into account that the scale parameters are limited by the parameters of the segmentation of the speech signal and its step.

After the selection of a self-similar structure, the next task is its classification. We use the approach proposed in [Chui, 1992, SounGen]. Expression (6) can be written in a similar form

$$\log_2|\mathrm{Wf}(s,u)| \leq \log_2 A + \left(\alpha + \frac{1}{2}\right)\log_2 s\# \quad (13)$$

Therefore, the smoothness parameters at point v are determined by the slope of the function $\log_2 s$ (and accordingly $\log_2|\mathrm{Wf}(s,u)|$) along the maxima line. The peculiarity of the lines of maxima is its construction on the basis of the points of maxima of the module, which is the curve $s(u)$ in the coordinates (s,u).

The classification of self-similar structures of the speech signal will be performed using (13) as follows.

We introduce the notation $O_v(s,u)$ – as a line of maxima of the wavelet transform of the speech signal converging to the point $u = v$, at $s \to 0$. For each point v we define the slope $\log_2 O_v(s,u)$ as a function $\log_2 s$ for $s \to 0$:

$$\log_2 O_v(s,u) = \log_2 A + \left(\alpha' + \frac{1}{2}\right)\log_2 s\# \quad (14)$$

We will consider that at the point $u = v$ we have a self-similar structure α'.

An effective solution to the problem of classifying the self-similar structure of the speech signal depends on the characteristics of the basic function of the wavelet transform ψ.

For example, if the wavelet basis ψ has n zero moments, then there is a function θ such that

$$\psi = (-1)^n \theta^{(n)} \int_{-\infty}^{+\infty} \theta(t)\mathrm{dt} \neq 0 \# \tag{15}$$

The wavelet transform is defined in the form

$$\mathrm{Wf}(s,u) = s^n \frac{d^n}{\mathrm{du}^n}\left(f * \overline{\theta}_n\right)(u)\# \tag{16}$$

where $\theta_s(t) = s^{-1/2}\theta\left(\frac{-t}{s}\right)$ [Chui, 1992]. Given this, if the wavelet ψ is represented by only one zero moment, then

$$\psi = -\theta`, \mathrm{Wf}(s,u) = s\frac{d}{\mathrm{du}}\left(f * \overline{\theta}_s\right)(u)\# \tag{17}$$

In (17) the maxima $|\mathrm{Wf}(s,u)|$ – represent the smoothed by the function $\overline{\theta}_s$ maxima of the first derivative function of the speech signal f. These multiscale maxima determine the location of the breakpoints and differences of the function of the speech signal f, and, accordingly, allow us to describe the location of self-similar structures. If the wavelet ψ has two zero moments, the representation of the maxima of the modulus of the function of the speech signal f will look like

$$W_2 f(s,u) = s^2 \frac{d^2}{\mathrm{du}^2}\left(f * \overline{\theta}_s\right)(u)\# \tag{18}$$

And they themselves will correspond to the local features of self-similar structures of the speech signal. Thus, it was found that performing a wavelet transform allows to obtain a set of wavelet coefficients that describe the speech signal. The most popular method of presenting the results of wavelet transform is a scalogram (Fig. 1), which allows you to visually represent and assess the location of the surface extremums, built on the basis of wavelet coefficients $W(a,b)$.

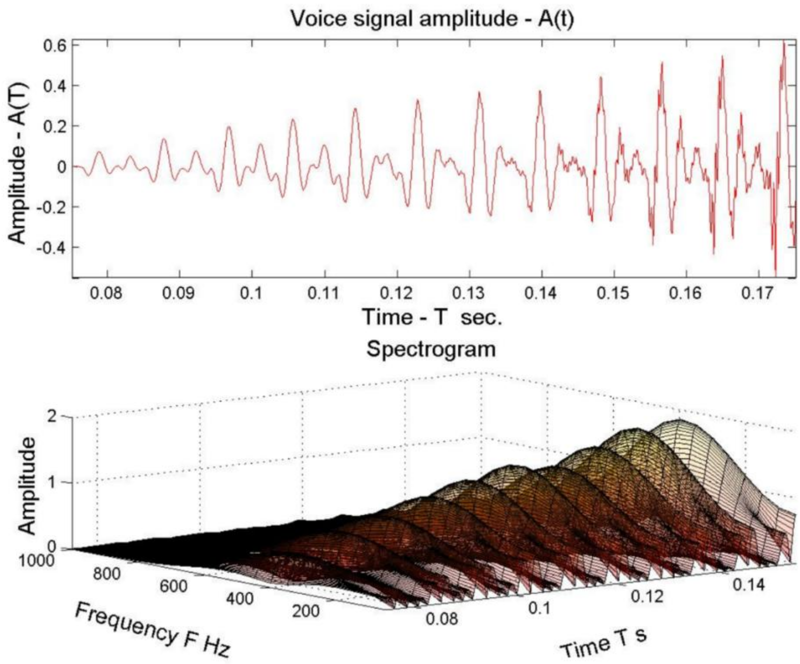

Figure 1: Scalogram of a speech fragment

Of particular interest are the local extrema of the surface coefficients. Theoretically, the analysis of self-similar structures can be performed on the basis of scalogram parameters, but there are a number of statistical functions that allow the evaluation of spectrum characteristics more efficiently.

The general view of such statistical measures of the measure can be represented as

$$M(q,a) = \sum_{l \in L(a)} \left| W\left(a, t_l(a)\right) \right|^q \# \qquad (19)$$

where l – local maximum line, L – a set of lines of maxima of wavelet coefficient modules, $t_l(a)$ – maxima of wavelet coefficients related to lines l scale a.

According to [Turiel and al., 2006], the dependence is executed

$$M(q,a) \sim a^{\tau(q)} \# \tag{20}$$

where $\tau(q)$ is determined for the value of q by calculating the slope $ln\big(M(q,a)\big)$ from lna, which is called the scaling exponential. Setting the value of q in (9) we obtain the dependence $\tau(q)$. The dependence $\tau(q)$ allows to obtain a multifractal spectrum of the speech signal based on the wavelet transform [Oświęcimka and al., 2020], which allows to describe the main characteristics of self-similar structures. The following dependence is used to obtain the multifractal spectrum

$$\begin{cases} D(h) = \min\limits_{q}[qh - \tau(q)] \\ h = \partial\tau/\partial q \end{cases} \# \tag{21}$$

The stability of this method of obtaining the characteristics of self-similar structures is to use the frequency-time window, which automatically performs averaging operations, as well as to obtain modules of wavelet coefficients in the calculation. It should be noted that from the point of view of the energy approach to the analysis of self-similar structures, the maxima of the wavelet transform coefficients at different levels of decomposition are the most significant.

The presented approach allows to expand the possibilities of speech signal analysis through the use of fractal and wavelet analysis. Unlike existing methods, this allows the analysis of non-stationary and short-lived signals.

Experiments

Selection of parameters and study of the effectiveness of the method of increasing the informativeness of the fundamental frequency.

As mentioned earlier, we will use the complex Morlet wavelet as a basic wavelet to isolate self-similar structures. An important feature of this wavelet is the ability to determine the instantaneous frequency using the analytical form of the wavelet [SoundGen]. It is believed that the Morlet is wavelet is a wavelet of small oscillations, ie it provides a center frequency of about 1 Hz, has good localization in time and frequency resolution. It is this set of properties that makes the Morlet wavelet one of the best for speech signal analysis. An analytical representation of the Morlet wavelet can be recorded as

$$\psi(t) = \left[exp\left(\frac{-t^2}{2\sigma_t^2}\right) - \sqrt{2}exp - \left(\frac{\omega_c^2\sigma_t^2}{4} + \frac{t^2}{\sigma_t^2}\right)\right] exp(j\omega_c t)\# \quad (22)$$

where $\omega_c = 2\pi F_c$, F_c – central wavelet frequency.

$$\sigma_t = \frac{1}{2\pi\sigma_f}\# \quad (23)$$

where σ_t – standard Gaussian deviation ($4\sigma_f$ – width of the wavelet).

The product $\omega_c\sigma_t$ relates the width of the Gaussian wavelet to the frequency of its oscillations. For the Morlet wavelet, this product must take on fairly large values ($\omega_c\sigma_t \geq 5$). The most frequently used range is $5 \leq \omega_c\sigma_t \leq 10$, when $0{,}8 \leq F_c \leq 1$ [Markel and al., 1982].

Given that $\omega_c\sigma_t$ provides a link between the width of the wavelet and the frequency of its oscillations, to determine the rational parameters of the wavelet, it is necessary to analyze the combination of parameters $(F_c, \omega_c\sigma_t)$, o more accurately identify features and structures speech signal.

It is known [Solovyov and al., 2014] that the activities of phonetics as an analysis of vowel sounds, the frequency range of which is quite wide, but it can be divided into the following subbands depending on the specific sound or set.

In view of this, the total frequency range of the study was divided into subbands of 100-500 Hz, 500-1500 Hz, 1500-4000 Hz, from 4000 Hz. These ranges will be used to fix the scales of the wavelet transform settings.

In order to select the parameters of the wavelet, a speech signal with predefined characteristics was synthesized. The open source software SoundGen [SoundGen] was used for the synthesis. This is a library in the R programming language that allows you to synthesize a language with specified parameters. The estimateVTL function was used to generate the formant frequencies, and the Soundgen function was used to generate the speech signal. A signal with frequencies of 450 Hz, 1400 Hz, 2800 Hz and 4300 Hz was created (Fig. 2).

Results

During the selection of parameters, the central frequency of the wavelet changed in 0.5 Hz increments in the range of 0.8-1 Hz. For each value of the center frequency, the parameter $\omega_c\sigma_t$ changed in steps of 0,5 in the range from 5 to 10. For each combination of parameters, the wavelet transform modulus was calculated, the scales corresponding to the maximum of the coefficient modulus were found, and the frequencies were calculated, deriving the phase wavelet transform coefficients corresponding to these scales. The obtained parameters of the wavelet transform settings are presented in table 1.

The next step of testing the developed methods was to study the effectiveness of determining the frequency of the fundamental tone. Comparative tests to determine the frequency of the fundamental tone on the basis of the proposed method to determine the frequency of the fundamental ion with the Pitch method for pure signal, signal with added noise, signal limited by the telephone channel band.

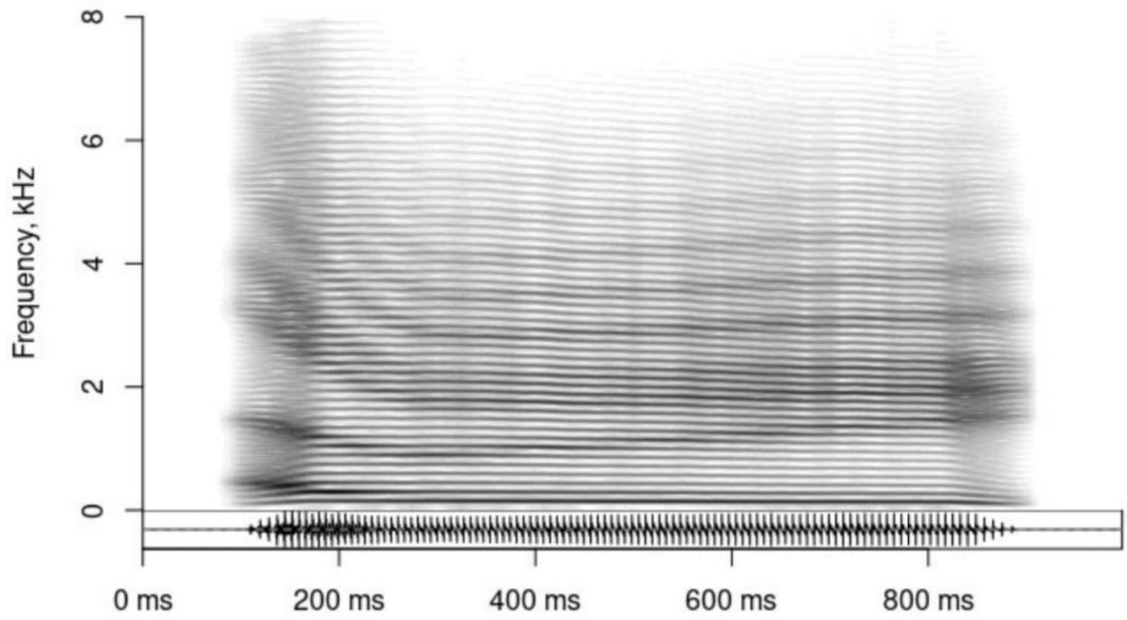

Figure 2: Synthesized speech signal with specified frequency parameters (picture representation is given without changes - as obtained in the software)

Table 1

Configuration parameters $(F_c, \omega_c\sigma_t)$ for selected frequency bands

Frequency range, Гц	F_c	$\omega_c\sigma_t$
100-500	0,8	5
500-1500	1	7
1500-4000	0,9	8
>4000	1	9

Adjustment of method parameters was performed by selection so that the generalized error of allocation of the frequency of the fundamental tone for the test speech material would be minimal. The total test results by estimating the generalized error are presented in Fig. 3, ten men, ten women, for 300 language fragments for each person. The generalized error was calculated from the normalized correlation coefficient with a single delay for calculations by each method, followed by summation.

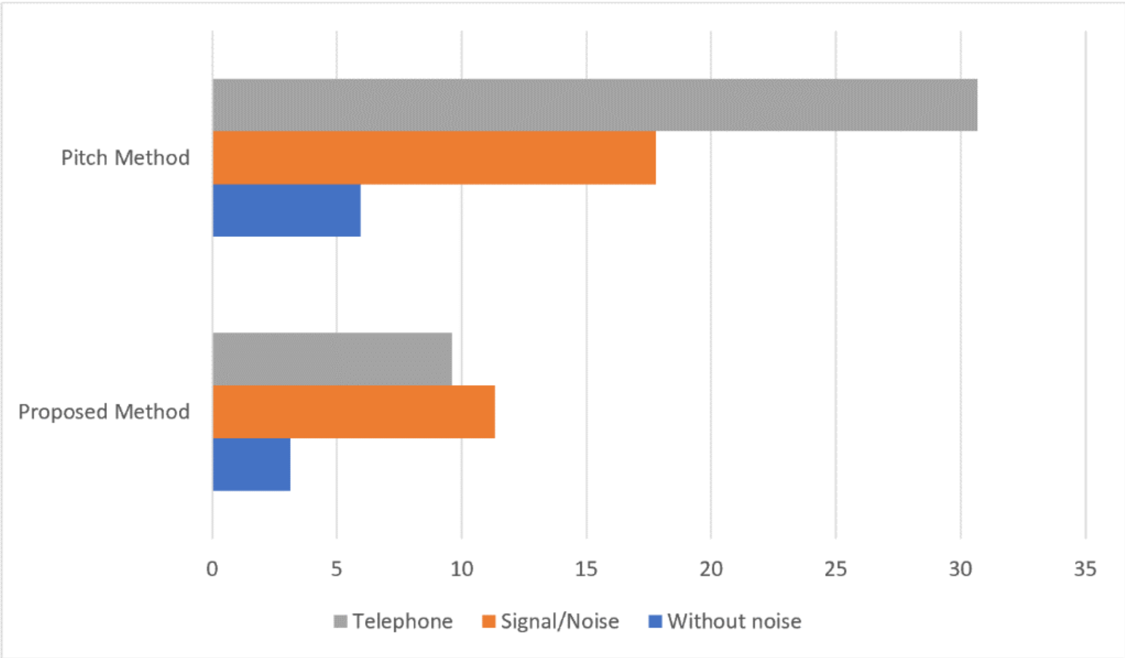

Figure 3: Test results of fundamental tone selection methods (%)

The developed method of determining the fundamental frequency showed the best average results in determining the frequency of the fundamental tone and provided proper tracking of the trajectory of the fundamental frequency throughout the proclamation, even at the ratio Signal/noise = 5 dB. Therefore it is possible to consider that the method is competitive in comparison with other considered methods of allocation of the basic tone which demand manual adjustment of parameters.

Conclusion

An algorithm for selecting special structures in a speech signal has been developed and an informative feature for linguistic identification of a person who, unlike existing ones, uses the values of wavelet transform coefficients of speech signal on segments where extremes of fundamental frequency correlation are observed. who are responsible for the individuality of the speech signal, and achieve high accuracy of identification. An experiment on the selection of rational parameters of the wavelet transform for the implementation of the method is performed, the coefficients of the wavelet transform adjustment for the frequency bands are obtained.

As an assessment of the effectiveness of the study, a comparative experiment was performed to determine the frequency of the fundamental tone based on the proposed approach and the Pitch method. The option to determine the frequency of the fundamental tone, based on the proposed method, reduces the error of determination in all analyzed variants of the voice signal.

Thus, the Model of allocation of characteristic parameters of a voice signal for speaker identification tasks presented in the article solves the problem of increasing the accuracy of speaker identification based on the fundamental frequency by improving the accuracy of determining the structural features of the voice signal.

As a continuation of the study, it is proposed to consider the relationship between the structural elements of the voice signal with its fractal characteristics.

Bibliography

[Mallat, 1999]. Mallat S. A wavelet tour of signal processing, Courant Institute, New York University, 1999, 671 pp.

[Muzy et al., 1991]. Muzy J.F., Bacry E., Arneodo A. Wavelets and multifractal formalism for singular signals: application to turbulence data // Phys. Rev. Lett. 1991. V.67. P.3515−3518.

[Pavlov A.V. et al., 2007]. Pavlov A.V., Anishenko V.S. Multifractal signal analysis based on wavelet transform //Saratov University, 2007. №1. 3-25 pp.

[Lardiès J. et al., 2004]. Lardiès J., Ta M.N., Berthillier M. Modal parameter estimation from output-only data using the wavelet transform, Archive of Applied Mechanics, Vol. 73. 2004. 718-733 pp.

[Xuedong, 2001]. Huang Xuedong. Spoken language processing: a guide to theory, algorithm and system development. – New Jersey: Prentice Hall PTR, 2001. 910 p.

[Chui, 1992]. Chui C. An introduction to wavelets. Academic Press. 1992. p. 278

[SoundGen]. A set of tools with a clear code for voice synthesis, manipulation and analysis. http://cogsci.se/soundgen.html

[Markel and al., 1982]. Markel J.E., Gray A.H. Linear Prediction of Speech. New York, NY: Springer. (1982).

[Turiel and al., 2006]. Turiel, A. & Pérez-Vicente, Conrad & Grazzini, Jacopo. (2006). Numerical methods for the estimation of multifractal singularity spectra on sampled data: A comparative study. Journal of Computational Physics. 216. 362-390. 10.1016/j.jcp.2005.12.004.

[Oświęcimka and al., 2020]. Oświęcimka, P., Drożdż, S., Frasca, M. et al. Wavelet-based discrimination of isolated singularities masquerading as multifractals in detrended fluctuation analyses. NonlinearDyn 100, 1689–1704 (2020). https://doi.org/10.1007/s11071-020-05581-y

[Solovyov and al., 2014] Solovyov Victor, Byelozorova Yana. Multifractal approach in pattern recognition of an announcer's voice / Victor Solovyov, Yana Byelozorova // TEKA. Commission of motorization and energetics in agricultire. – 2014. – Vol. 14, № 2.– pp. 164 – 170.

Authors' Information

Yana Bielozorova *– PhD, Associate Professor of Software Engineering Department, National Aviation University, Kyiv, Ukraine.*

E-mail: *bryukhanova.ya@gmail.com*

***Major Fields of Scientific Research**: Speech Recognition Models, Wavelet analysis, Software Architecture*

ABOUT THE ISSUE OF JOINT PROCESSING OF IMAGES OF THE EARTH'S SURFACE, OBTAINED FROM DIFFERENT TYPES OF MEDIA

Valeriy Pysarenko, Julia Pisarenko, Ivan Varava,Olga Gulchak

***Abstract**: The problem of joint processing of images of the Earth's surface obtained from spacecraft and unmanned aerial vehicles, which is associated with the need to combine frames of Earth sections during image preprocessing, is considered. A procedure has been developed for processing images of the Earth's surface to monitor the area in order to identify long-term changes, and a procedure has been developed for processing images of the earth's surface to monitor the area in order to identify long-term changes/*

***Keywords**: spacecraft, unmanned aerial vehicles, Earth's surface, images.*

***DOI**: https://doi.org/10.54521/ijita29-03-p04*

Introduction

The procedure for remote sensing of the Earth (ERS) from a spacecraft is widely used to assess the state of the terrain, for example, to detect the state of vegetation [1] or, very importantly, the consequences of destruction after technological and environmental or military events. But these methods do not provide detailed information about the state of the infrastructure. In this case, there is a need to support technologies for space survey of territories with additional methods of photographing a given surface of a land plot from heights from 10 m to 200 m using telephoto equipment placed on board unmanned aerial vehicles (UAVs).

Thus, the problem arises of joint processing of video and photographic materials obtained by UAV and spacecraft (SC) equipment. Joint processing of photographs or video materials from the spacecraft and UAV boards provides the most comprehensive information about the state of the studied territory of the Earth (Fig.1)

Improvement of scientific approaches to the creation of software and hardware tools for the use of unmanned aerial vehicles (UAVs) and spacecraft for remote monitoring from the air of territories is a modern urgent problem. This is due to the problem of searching, detecting and identifying objects in a given area, with determining the coordinates of these objects. It is also important to track changes in the state of objects over time with the subsequent transmission of photo or video information in real time [1, 2, 3, 4]. A comprehensive assessment of damage from damage and destruction as a result of technical and environmental disasters is extremely important, because only after that it is possible to plan the restoration of the infrastructure of the territory. As for precision farming, the development of this direction requires accurate and promptly updated information about the state of vegetation and soil. Obtaining such information is possible only when using remote sensing of agricultural fields due to the large size of their areas. The ability to quickly identify diseased areas of vegetation is essential. This allows at the early stages of development with minimal costs and quickly localize and "cure the disease." [1, 5].

Thus, the problem arises of joint processing of video and photographic materials obtained by UAV and spacecraft (SC) equipment. Joint processing of photographs or video materials from the spacecraft and UAV boards provides the most comprehensive information about the state of the studied territory of the Earth (Fig.1)

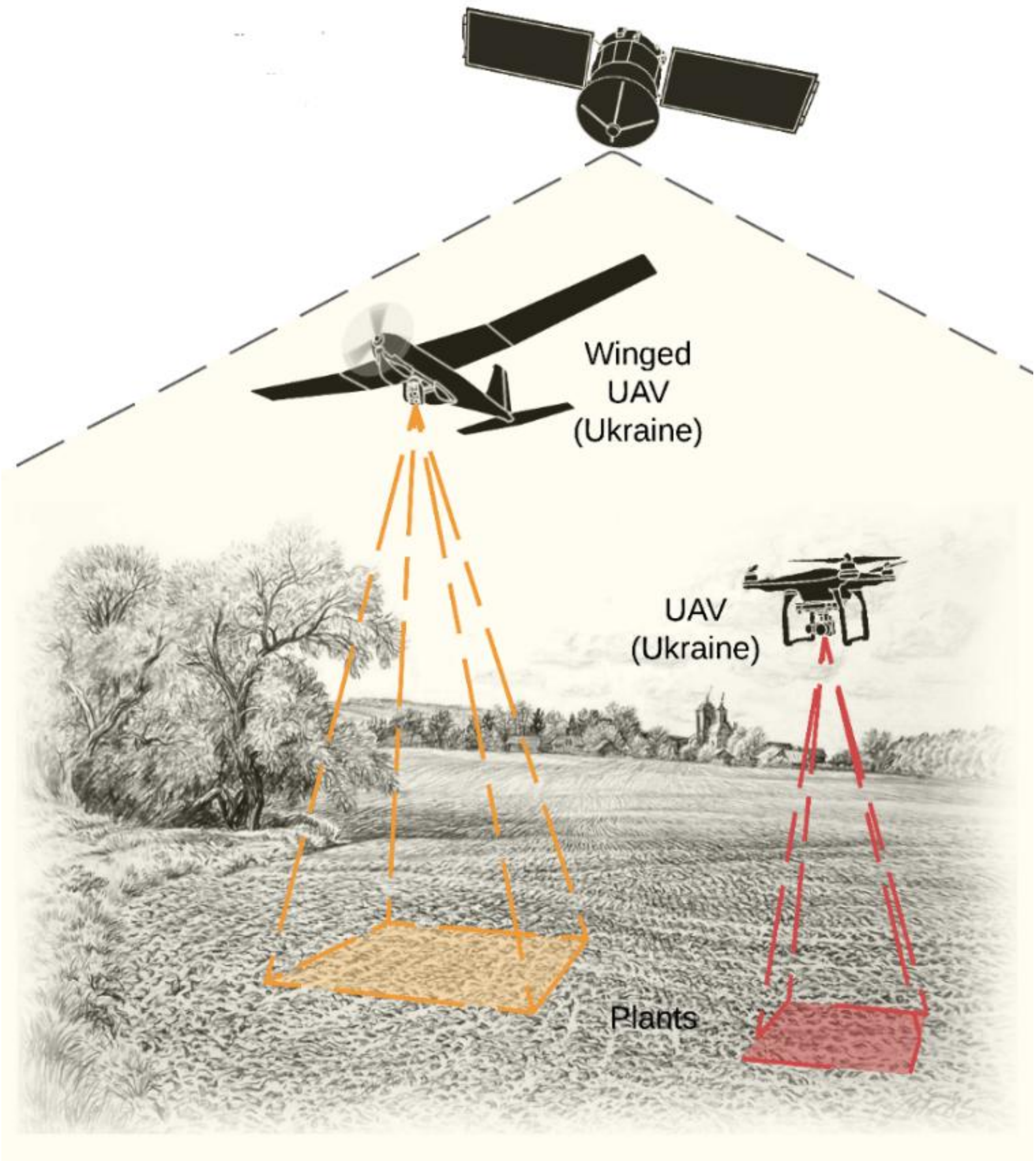

Fig.1 Illustration for the procedure of remote sensing of the earth's surface from various types of media

II. Formulation of the problem

The task of the authors' research was to improve the quality of environmental monitoring using neural network identification and classification of objects on multi-zone satellite images and multi-zone images obtained from UAVs.

The research reported on included both theoretical research and the solution of applied problems. The mathematical basis of image processing is the computational apparatus for pattern recognition. Practical research is based on experiments, software implementation and testing of algorithms and technology.

Thus, the task of the authors' research is to increase the accuracy of recognition of objects on the earth's surface (infrastructure objects or the vegetation state of agricultural lands). The authors have experience in this area. In particular, one of the topical research areas of Department 265 of the V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine (NASU) is the creation of technologies and intelligence systems after technical and environmental events, which are most often accompanied by the destruction of infrastructure. The team of authors has long been working on improving unmanned aerial vehicles and adapting their capabilities to perform many important and relevant tasks for the industry and agriculture of Ukraine. For example, in 2017, an employee of the 265th department Prokopchuk V.V. took part in the manufacture and testing of these UAV models, and on August 25, 2017 presented a cargo drone that set a Ukrainian record - lifted a weight of 16 kg into the air.

This article discusses the use of a scientific idea, which consists in developing a method for joint computer processing of digital and analog images obtained from UAVs, as well as quasi-simultaneous and reusable multi-zone space images. Currently, this direction is being developed by Department 265 of the Institute of Cybernetics of the National Academy of Sciences of Ukraine

III Problems in the collaborative image processing

Below is a list of the most typical problems and tasks that arise in the development of methods for such joint processing of images of the earth's surface, including multi-temporal ones (which is sometimes necessary to identify the dynamics of changes in individual parts of the earth's surface). The solution of these problems occurs during the so-called pre-processing of images, after which it becomes possible to further process the prepared images according to the adopted program:

- The problem of compensating for the mismatch of different-time frames of remote sensing images.
- The problem of combining images obtained from different aircraft.
- The problem of the impact of atmospheric phenomena.
- The problem of filtering noise in the image.
- The problem of joint processing of images of the earth's surface with different resolutions.

Below is an overview of the current state of research on these problems of joint image processing.

1. The problem of compensating for misalignment of different-time frames of remote sensing images.

Images obtained from an unmanned aerial vehicle (UAV) are misaligned relative to each other at different times. This problem also arises when processing images from spacecraft is required, namely, an automated search for changes in space images of different times. Then it becomes necessary at the pre-processing stage to use methods for eliminating the misalignment of different-time frames of remote sensing images. Among the methods for eliminating such an obstacle in the processing of remote sensing image frames, there are methods based on the search for key points and phase correlation. The algorithm for compensating the misalignment of multi-temporal image frames using key points is based on a preliminary search for an image of key (special) points. Such procedures include, for example, methods: Speeded Up Robust Features (SURF), Scale-Invariant Feature Transform (SIFT), as well as such more modern approaches as Oriented FAST and Rotated BRIEF (ORB) and KAZE/AKAZE. [2]. Such methods (in general, the SIFT method and their modifications) are often used to solve the problems of processing images obtained from both air and space carriers. In the case when monitoring requires long-term observations (for example, for mapping and monitoring glacier velocities, in the practice of searching for changes in terrain using satellite images or using aviation), images obtained by different systems, under different conditions (different angle and direction of sighting) are often used, different times of the day, etc.). For the correlation of multi-temporal images, the algorithms FFT (Fast Fourier Transform), NCC (Normalized Cross-Spatial Correlation), Cross-corre - cross-correlation on orientation images or the joint registration and correlation algorithm can then be used. (CO-Registration of Optically Sensed Images and Correlation - COSI-Corr) [3]

2. The problem of combining images obtained from different aircraft.

This implies an exact comparison of the pixels of two images corresponding to the same points on the earth's surface. This gives rise to data on the spatial resolution of images. Distortions can occur when gluing microframes and diverging axes of devices, shooting angles that form multispectral images.

A common practice for assessing the quality of combined images is visual analysis. It is obviously subjective and poorly suited for mass image processing. Combining images with different spatial resolutions and in different spectral ranges also causes difficulties [4].

It happens that the spatial resolution of images differs by several times, and for a number of tasks one-to-one correspondence of pixels of different spectral channels is required with the highest possible accuracy of the spatial resolution of the panchromatic channel. That is, at the stage of imaging, there should be a one-to-one correspondence of pixels corresponding to the same piece of terrain and belonging to different spectral channels. Devices that form the image are often not coaxial, which makes the task of combining their images relevant. In addition, already building an image from a set of microframes is a complex procedure that requires the use of high-precision technologies. Therefore, at the first stage of image processing, it is necessary to strictly control the spatial correspondence of images. Some of the algorithms are described in the publication [6].

3. The problem of the influence of atmospheric phenomena

Unfavorable weather events, such as rain, fog, snow, can lead to inaccurate performance of object identification tasks using video frames. Many scientific studies are devoted to the use of improved vision systems that reduce the impact of precipitation on the quality of video information [7].

Some well-known algorithms for eliminating the effect of precipitation on the quality of photographs are used: the spatial median filtering algorithm, algorithms based on photometric and dynamic rain models, a two-stage algorithm for improving the quality of video images. Each of these approaches has its own advantages and disadvantages; therefore, the task of developing an algorithm that compensates for the effect of precipitation on the images obtained during remote sensing is relevant. Often the elimination of the described shortcomings is carried out in two stages: the procedure for detecting and eliminating particles of precipitation and the procedure for increasing the contrast are performed separately in order to reduce the foggy effect on the distant plan [7].

4. The problem of filtering noise in the image

In the process of obtaining remote sensing data, there are images that have been distorted by noise generated during the generation or transmission stage. The reasons that influenced the appearance of noise in the image may be a malfunction in the communication channels, errors in the operation of the signal recording device, and others. Satellite imagery and UAV imagery are an effective way to obtain data, but the quality of data depends not only on the hardware component of the remote sensing process. Ultimately, this leads to a deterioration in the quality of visual perception and a decrease in the likelihood of decisions that will be made based on the analysis of such images. To successfully solve the problems of searching and identifying objects, determining various kinds of quantitative characteristics, it is necessary that the primary images be characterized by a high indicator of visual quality, which is lost due to unsatisfactory conditions for obtaining images, imperfection of information transmission systems and their display.

5. The problem of joint processing of images of the earth's surface with different resolutions.

The problem of comparing two images over the past two decades has been an active research topic in the field of processing remote sensing images of the earth's surface. Most existing methods consider images of the same area, where viewpoints differ by small shifts in position, orientation, and viewing parameters such as focal length. Image elements associated with different shots have a comparatively close resolution. For this reason, elements of the scene that appear in different images will be approximately the same scale. When images from spacecraft and UAVs of the same area are compared, new problems arise in comparing two images with different spatial resolutions. This issue has received little attention in the past [7].

Obviously, the resolution with which a 3D object is observed in an image mainly depends on two factors: the distance d from the camera to the object and the focal length f associated with the camera lens. Image resolution increases as f increases and decreases as d increases. Therefore, $r = f/d$ is a first order approximation of the image resolution measurement. Therefore, there is a need to develop methods for matching a pair of images with completely different resolutions.

At the same time, we have input image data for processing with low resolution (image from the spacecraft) and images with high resolution (image from the UAV).

In the research on the topic of the report, the development of an algorithm for pre-processing images was provided and carried out, taking into account the features of digital images from a spacecraft and UAVs.

The images received from the spacecraft have a mismatch in colors, which further complicates their joint processing. At the pre-processing stage, it is proposed to use an algorithm for eliminating the mismatch of color channels of Earth remote sensing (ERS) images, based on the search for key points and differing by their additional check for each channel.

IV Complex processing of multi-temporal images of the Earth's surface.

Development of a procedure for complex neural network processing of multi-temporal images of the earth's surface obtained by remote sensing equipment located on spacecraft and UAVs.

Consider the procedure for processing images of the earth's surface for the following task: monitoring of the area should be carried out in order to determine long-term changes in the earth's surface. A typical task is the assessment of deforestation areas with the possibility of assessment and spontaneous deforestation. This is necessary to calculate the damage caused by unauthorized logging, to calculate the funds that should be provided for the restoration of natural resources.

At the same time, it is necessary to conduct video monitoring of the study area periodically in the long term, in fact, every time to map the area in automatic mode. The idea is to obtain multi-temporal images of some areas of the tested forest massifs, bordering each other, with a clear spatial reference. In this case, changes in the brightness of the earth's surface must correspond to real changes in the earth's surface. To do this, it is better to use pictures taken with one sensor. Then the same shooting technology will be used, so the differences in brightness in the images will be guaranteed to be associated with changes in the earth's surface. In addition, if you need to compare the state of the area for different years, then it is desirable that the pictures were taken approximately in the same months and at the same time of day. This ensures that long-term changes are detected, rather than seasonal dynamics of the area. After selecting images for comparison, it is necessary to perform their radiometric calibration and atmospheric correction, since it is important that the difference in the state of the atmosphere does not interfere with the detection of real changes on the ground.

Image processing will consist of the following steps:

1) The array under study (the area from which images have already been taken) with the coordinates (two-dimensional) of each “frame” is considered a template.

2) Selection of geometric spatial references, key points that are used to define the boundaries of each frame. And when the UAV passes the next time after a certain time, a picture of the same area is obtained. The time interval between successive studies of the territory is determined from the characteristics and objectives of the task.

3) The operation of combining the boundaries of these multi-temporal images is carried out (determining the "common" frame for two snapshots). (Images

received from an unmanned aerial vehicle (UAV) in different periods of time have discrepancies relative to each other. The same problem arises when processing images from spacecraft is required, namely an automated search for changes in satellite images at different times. Then it becomes necessary at the stage pre-processing, use methods for eliminating the discrepancy between frames of remote sensing images at different times).

4) Mathematical processing to detect the correlation of these two images is possible after the following preparatory work:

a) elimination of errors due to scale mismatch for different terrain images;

b) elimination of errors due to differences in the spatial resolution of images of the same area of the surface;

c) elimination of the influence of differences in the angles of inclination of the shooting cameras. Distortions arise when gluing microframes and divergent axes of devices, shooting angles that form multispectral images;

d) elimination of the influence of atmospheric changes [2];

e) creation of a panchromatic image (black and white) of a specific area using multi-zone imaging, if it is decided to use the methods of correlation analysis of panchomatic images. Otherwise, methods of applying the correlation processing of multispectral images can be applied;

f) orthorectification (orthocorrection) of images if necessary, mathematical rigorous transformation of the original image (image) into an orthogonal projection and elimination of distortions caused by terrain, shooting conditions and camera type.

Note: not all of the above preparatory work may be needed, it all depends on the specific requirements for monitoring land masses.

Below in fig.2-4 one of the operations of preliminary processing of images of the earth's surface - image alignment - is illustrated for the purpose of the next step of the technique: comparison of images at different times. Various satellite images of one of the areas were processed using OpenCV, an open source computer vision library.

Fig. 2,3 - satellite images of one area of the territory, made with an interval of several days.

Fig. 2

Fig. 3

Fig. 4 Photo from the UAV superimposed on a satellite the area photo as an example of the program operation for combining two photos using OpenCV (an open source computer vision library available for free use for scientific, academic and commercial purposes).

After combining the images, it becomes possible to separate 2 frames of images at different times for processing them by the methods of correlation analysis [3] using computer processing programs and identify changes in the state of the terrain [4,5].

Conclusion

1.The article analyzes the current state of scientific research on the problem of joint processing of images from different media or multi-temporal images of the same area of the earth's surface, which carry information about the state of land use areas.

2. A procedure has been developed for processing images of the earth's surface to monitor the area in order to identify long-term changes. At the same time, it is necessary to conduct video monitoring of the study area periodically in the long term, in fact, every time to map the area in automatic mode. The idea is to obtain multi-temporal images of some areas of the test territories bordering each other with a clear spatial reference and subsequent compensation in the general case for the placement of multi-temporal frames of remote sensing images, combining images obtained from various aircraft of atmospheric phenomena and noise filtering in image.

3. A procedure has been developed for processing images of the earth's surface to monitor the area in order to identify long-term changes After combining images, it becomes possible to select 2 frames of images at different times for their processing by methods of correlation analysis using computer processing programs and detecting changes in the state of the terrain [6,7].

Bibliography

Cobeña Cevallos, J.P., Atiencia Villagomez, J.M., Andryshchenko, I.S. (2019) Convolutional Neural Network in the Recognition of Spatial Images of Sugarcane Crops in the Troncal Region of the Coast of Ecuador. 13th International Symposium “Intelligent Systems” (INTELS’18), 757-863. doi: 10.1016/j.procs.2019.02.001. [Online]. Available: https://doi.org/10.1016/j.procs.2019.02.001.

[Bickel et al, 2018] V.T. Bickel, A. Manconi, and F. Amann, Quantitative Assessment of Digital Image Correlation Methods to Detect and Monitor

Surface Displacements of Large Slope Instabilities, Remote Sens, 2018, № 10, pp. 865-883 DOI: https://doi.org/10.3390/rs10060865.

[Leprince et al, 2007] S. Leprince, S. Barbot, F. Ayoub, and J.-P. Avouac, Automatic and Precise Orthorectification, Coregistration, and Subpixel Correlation of Satellite Images, Application to Ground Deformation Measurements, IEEE Transactions on Geoscience and Remote Sensing, 2007, Vol. 45, No. 6, pp. 1529-1558, July 2007 DOI: https://doi.org/10.1109/TGRS.2006.888937.

[Huete and Lium 1994] A. Huete, H. Liu, An error and sensitivity analysis of the atmospheric and soil correcting variants of the NDVI for the MODIS- EOS, IEEE Transactions on Geoscience and Remote Sensing, July 1994, vol. 32, pp. 897-905. Doi: https://doi.org/10.1109/36.298018.

[Huang et al, 2018] Huang, H., Deng, J., Lan, Y., Yang, A., Deng, X., Zhang, L. (2018) A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery. PLoS ONE 13(4): e0196302. doi: 10.1371/journal.pone.0196302. [Online]. Available: https://doi.org/10.1371/journal.pone.0196302.

A. I. Aleksanin, M. A. Morozov, E. V. Fomin, Problems of image alignment with pixel accuracy. Modern problems of remote sensing of the Earth from space. 2019. V. 16. No. 1. pp. 9–16

[Leprince et al, 2007] S. Leprince, S. Barbot, F. Ayoub, and J.-P. Avouac, Automatic and Precise Orthorectification, Coregistration, and Subpixel Correlation of Satellite Images, Application to Ground Deformation Measurements, IEEE Transactions on Geoscience and Remote Sensing, 2007, Vol. 45, No. 6, pp. 1529-1558, July 2007 DOI: https://doi.org/10.1109/TGRS.2006.888937.

Authors' Information

Valeriy Pysarenko, *V.M.Glushkov Institute of Cybernetics of the NAS of Ukraine, chief of dept. №265 of Mathematical Problems of Applied Informatics V.M.Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine,Kyiv, Ukraine, e-mail: pisarenkojv@gmail.com*

Major Fields of Scientific Research: General theoretical information research, Multi-dimensional information systems

Julia Pysarenko, *V.M.Glushkov Institute of Cybernetics of the NAS of Ukraine, Senior Researcher lab. №236 of Virtual Environment Systems for Organization of Scientific Research, Kyiv, Ukraine, jvpisarenko@gmail.com*

Major Fields of Scientific Research: Multi-dimensional information systems, robototechnics

Ivan Varava, *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Associate Professor at the Department of Automation of Power Processes and Systems Engineering National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" Kyiv, Ukraine, ivan.varava@ukr.net*

Major Fields of Scientific Research: Information systems, robototechnics, image processing, programming

Olga Gulchak, *V.M.Glushkov Institute of Cybernetics of the NAS of Ukraine, engineer of dept. №265 of Mathematical Problems of Applied Informatics V.M.Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine, e-mail: olga-gulchak@ukr.net.Major Fields of Scientific Research: robototechnics, image processing*

COMPUTER SIMULATION OF A TRANSPORT RELOADING COMPLEX

Mykola Dekhtyaruk

Abstract. *The article considers the reloading complex, in which the machines input flow is served by the rule of relative priorities. Algorithms functioning of a separate elements are constructed, on the basis of which its computer model developed, by means of object-oriented programming system - Embarcadero RAD Studio XE10 Seattle. Simulation of the reloading complex is performed, and the calculation of the main parameters of the complex – total and average waiting time of machines in the queue, total and average length of the queue, etc., depending on the intensity of the machines input flow and service channel performance, to analyze and select optimal modes work of the complex.*

Keywords: *Reloading complex, Algorithm simulation model, Service channel, Object-oriented programming system.*

ITHEA Keywords*: H. Information Systems; H.0 General; H.1 Models and Principles; H.4 Information Systems Applications.*

DOI*: https://doi.org/10.54521/ijita29-03-p05*

1. Introduction

Transport is one of the most important sectors of the state's economy and belongs to the sphere of material services production. The main task of transport is timely, high-quality and full satisfaction of a national economy needs and the population in transportations. The most important element in the development of freight technology is the choice of transport and technological system. Each transport and technological system can be represented as a set of standard operations, the main of which are: cargo; transport network; rolling stock; cargo concentration points; loading and unloading means; participants in logistics processes; packaging materials [Christopher, 2016].

The main factors determining the choice of transport and technological systems are local technological processes that take place in all parts of the transport logistics system, which have a number of features and depend on the type of cargo, mode of transport and its structure, industry characteristics, state of logistics process elements [Shortle, 2018].

The main reserves for improving the transport and logistics process are in the rational organization of interaction of participants in the supply chain, in the coordination of their interests and the search for mutually beneficial and suitable solutions. Progress in the field of information technology makes it possible to significantly increase the efficiency of transport logistics, information and computer support has a proper place among the key logistics functions [Trivedi, 2016].

The purpose of this work is to develop a simulation model of the reloading complex, in which the input flow of machines is served by the rule of relative priorities, in the form of a computer software package. based on the use of the latest information technologies - Embarcadero RAD Studio XE10 Seattle - object-oriented programming system [RAD, 2019].

Using a computer model to simulate the operation of the reloading complex, and calculate the basic parameters of the complex – total and average waiting time of machines in the queue, the total and average length of the queue, as well as their variance, depending on the intensity of input flow and service channel performance, in order to analyze and optimal modes selection the operation of the reloading complex.

2. The Primary Research Materials

The theory and practice development of the freight transportation organization is given considerable attention in scientific publications. Thus, the articles [Khorshidian, 2019; Kumar, 2016; Escobar, 2020; Heidari, 2018] present the results of the study of routing optimization by the criterion of minimum delivery time. In [Nasiri, 2018; Soeanu, 2020; Lai, 2015] methods of route selection based on alternative sampling are considered.

Publications [Sun, 2020; Luisa, 2015] use routing modeling methods in transport systems based on fuzzy logic. The heuristic model of route selection is considered in articles [Tokgeoz, 2015; Manley, 2015]. In works [Yu, 2020; Rodado, 2017; Havelock, 2018] the statistical model of a routing choice in transport systems is considered. The method of cargo transportation using genetic algorithms is described in publications [Tintu, 2020; Mohammed, 2017; Chiappone, 2016; Lu, 2017; Yang, 2016].

The modeling results of the routes choice using data from the Global Positioning System (GPS), focused on trucks that perform long trips, are given in [Hess, 2015]. The publication [Nyrkov, 2015] describes a differential method for determining the location of land vehicles, using the GLONASS/GPS system, using special processing algorithms. Probabilistic methods of controlling the movement of ground objects, using satellite navigation systems, are considered in [Zhilenkov, 2015].

Analysis of the results of these studies shows that these works use various analytical methods for organizing the movement of cargo flows and modes of operation of individual elements and parts of logistics systems.

At the same time, the use of the latest information technologies and systems helps to increase the efficiency of transportation. Information systems of logistic processes automation allow to automate all information and technological activity of the transport enterprises participating in processes of the freight transportations organization. Automation of transport logistics is necessary to increase the efficiency and optimization of transport. Due to computer data processing, introduction of information systems of routing, accounting and planning at the transport enterprise, transport logistics reaches a qualitatively new level [Shortle, 2018; Trivedi, 2016].

Therefore, the development of a simulation model, by means of object-oriented programming, and modeling of the reloading complex, in order to analyze and select the optimal modes of its operation is relevant.

3. Development of a mathematical model and algorithmization of the reloading complex

The block diagram of the reloading complex is shown in fig. 1. The operation of the system is described as follows. Two types of machines come to the reloading complex: the first consists of machines arriving for unloading (service time of such machines is τ_1; the second - of machines arriving for loading (service time of such machines is equal to τ_2.

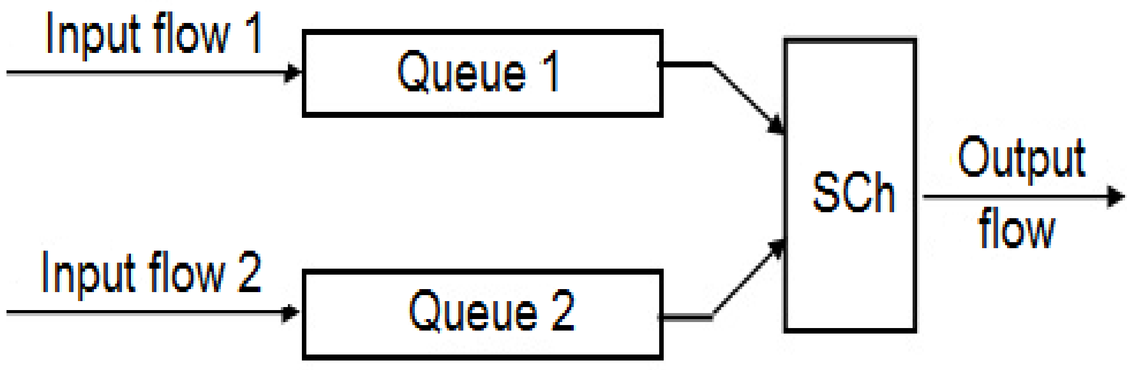

Fig. 1. Block diagram of the reloading complex

where SCh - servicing channel of machines input flows.

.

The reloading complex can serve only one machine at a time. If at the time of arrival of a new car the service channel is occupied, it becomes into one of two queues: one consists of cars arriving for unloading, the other - of cars arriving for loading. Let the service of the machines be carried out in the following order: at the moment of release of the service channel the service of the machine standing first in line for unloading begins. Only if this queue is empty, the first machine is serviced from the queue for load (expressed in the language of queue theory, we can say that the service of machines in the system are performed according to the rule of relative priorities). In this case, the reloading complex is considered as a queuing system (QS).

Because the moments of machines arrival for service are random variables, the length of the queue and the waiting time for service will also be random, so as a result of modeling it is necessary to determine their statistical characteristics - the average value and variance.

Consider the average waiting time of machines in the queue for streams 1 and 2. For each *i*-th machine, the waiting time in the queue is equal to the difference between the time when it began to be serviced and the time when it came into the system. The average waiting time is:

$$\omega_{1ave} = \frac{1}{n_1}\sum_{i=1}^{n_1}\omega_{1i}; \quad \omega_{2ave} = \frac{1}{n_2}\sum_{i=1}^{n_2}\omega_{2i} \tag{1}$$

where n_1 and n_2 are the number of machines, respectively, streams 1 and 2, which were serviced by the system during T.

Summarizing the number of machines in the queue at short intervals and dividing the amount by the number of summations, we obtain the average value of the queue length:

$$L_{1ave} = \frac{1}{N}\sum_{i=1}^{N} m_{1i}; \quad L_{2ave} = \frac{1}{N}\sum_{i=1}^{N} m_{2i} \tag{2}$$

where m_{1i} and m_{2i} is the number of machines in the first and second queues at the time of observation *i*;

N - the number of observation moments (moments of receiving statistics) for time *T*.

The variance of the values of *ω* and *L*:

$$D\omega_1 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1} (\omega_{1i} - \omega_{1ave})^2 \tag{3}$$

$$D\omega_2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2} (\omega_{2i} - \omega_{2ave})^2 \tag{4}$$

$$DL_1 = \frac{1}{N - 1}\sum_{i=1}^{N} (m_{1i} - L_{1ave})^2 \tag{5}$$

$$DL_2 = \frac{1}{N - 1}\sum_{i=1}^{N} (m_{2i} - L_{2ave})^2 \tag{6}$$

To build a simulation model of the reloading complex, you must first be able to model the moments of machines arrival in the queue of each flow. Both streams of machines coming into the service channel can be described by the function of distribution *A*(*t*) of time intervals between the moments of their arrival in the queue (this function will be different for each stream), that is $A(t)=P(Q<t)$.

The practice of queuing systems research shows that in many cases the approximation of the intervals distribution function between the moments of machines entry into system by exponential function appears satisfactory:

$$A(t) = \begin{cases} 1 - e^{-\lambda t} & for\, t \geq 0; \\ 0 & for\, t < 0, \end{cases} \tag{7}$$

where λ is the intensity of the input flow (a value inverse to the average time interval between the arrival of two adjacent machines).

In this case, the requirements at the input of the system form the s-called Poisson flow. Poisson flows of requirements are often found in practice, because they are in a sense boundary for different flows. For most applications similar to logistics tasks, the replacement of non-Poisson flows by Poisson with the same intensities leads to a solution that differs little from the real one, and sometimes does not differ at all. In this case, the errors of the solution, as a rule, are within the accuracy of the original data, which are often known approximately [Kirpichnikov, 2018; Gnedenko, 2018].

For the first stream $Q_{1i+1} = t_{1i} + Q_{1i}$, where the values of Q_{1i} are distributed by the law $A(t)$. It is easy to prove that if there is a random variable R distributed evenly in the interval (0, 1), then to obtain a quantity Q having a distribution function $A(t)$, we must solve the equation $A(Q)=R$ relatively to Q. In particular, for $A(t) = 1 - e^{-\lambda t}, t \geq 0$, it is necessary to solve the equation $1 - e^{-\lambda Q} = R$.

From here:

$$Q = -(1/\lambda)\ln(1-R) \tag{8}$$

Therefore, obtaining Q is reduced to finding R, which is subject to a uniform distribution, and calculating Q by (7).

All programming systems have special standard programs that produce so-called pseudo-random numbers, the sequence of which is subject to a uniform distribution in the interval (0, 1). Obtaining one such number requires one reference to this standard program.

Thus, the algorithm for obtaining the arrival time of the next machine in the queue can be represented in the form of a diagram shown in Fig. 2 for stream 1, assuming that this stream is Poisson.

Next, you need to make an algorithm that describes the logic of the queue (it is clear that the algorithm will be identical for both queues). Assume that the queue has a maximum length of *LM* (number of waiting places). One-dimensional array *P*, consisting of elements (cells) P_1, P_2, ... P_{LM}, simulates the "places" of this queue. Each of these cells can be either free or occupied by a machine. As an equivalent of the car, it is convenient to take the moment of its arrival in the queue - this information will be needed to determine the waiting time of the car in the queue. Cell *X* is used as a working cell when writing the next machine in the queue or when selecting from the queue.

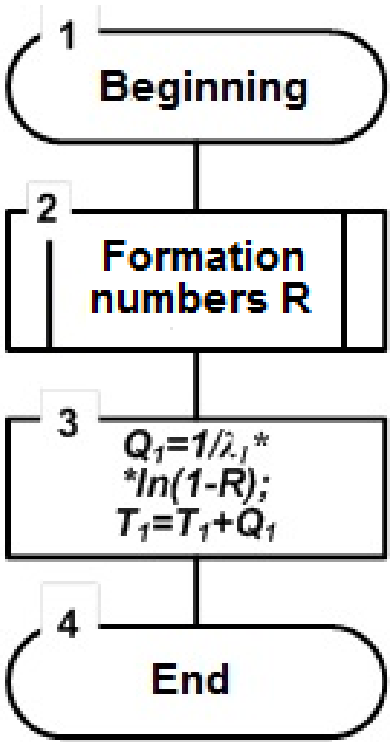

Fig. 2. Scheme of the algorithm for modeling the input stream 1

PER and *POS* cells contain information that allows you to determine the presence of the first and last machines in the queue, respectively. This information is the number (index) of the corresponding cell of the array *R*.

When the machine is selected from the queue, the contents of the *PER* cell are increased by one according to the queue selection rule in the order of arrival. The contents of the *POS* cell increase by one when writing to the queue of a new machine. The maximum value of the variables *PER* and *POS* is equal to *LM*, so their change occurs according to the formulas:

$$PER = \begin{cases} PER+1, & if\ PER < LM; \\ PER+1-LM, & if\ PER \geq LM. \end{cases} \quad (9)$$

$$POS = \begin{cases} POS+1, & if\ POS < LM; \\ POS+1-LM, & if\ POS \geq LM. \end{cases} \quad (10)$$

Variable *NP* - the number of machines in the queue; when selecting from the queue, the *NP* decreases by one, and when writing to the queue, it increases by one. The variable *PUST* is equal to one if there are no machines in the queue, and is equal to zero, in otherwise case. The *POLN* variable is equal to one if there are no free spaces in the queue (in which case the new machine cannot be written to the queue). The *WYB* variable must be set to 1, if the reference to the algorithm is performed for the purpose of sampling. If the reference to this algorithm is performed for writing, it is necessary to set *WYB*=0. The scheme of the algorithm modeling the work of queue 1 is shown in Fig. 3.

The operation of the service channel can be described by the sign "free - busy", which changes over time. As a sign, you can take the binary variable *KSW* (*KSW*=1 if the service channel is free and *KSW*=0 if the service channel is busy).

Let the previous releasing moment of the service channel be T_3. Then, according to the logic of the system, the request to queue 1 is executed, and if it is not empty, then the next releasing moment of the service channel will be:

$$T_3 = T_3 + \tau_1 \quad (11)$$

If queue 1 is empty, then there is an appeal to queue 2. There are two options:

1) if this queue is not empty then the next releasing moment of the service channel will be:

$$T_3 = T_3 + \tau_2 ; \quad (12)$$

2) if queue 2 is empty, the value KSW=1 is set.

At the time of machines selection from the queue for service, you can calculate the waiting time in the queue according to the formula:

$$W = T_3 + X . \quad (13)$$

To determine the average waiting time and variance of this value, it is necessary to perform the accumulation (summation) of the waiting time and the square of this value:

$$SW = SW + W ; \quad (14)$$

$$SW2 = SW2 + W^2 , \quad (15)$$

while N=N+1.

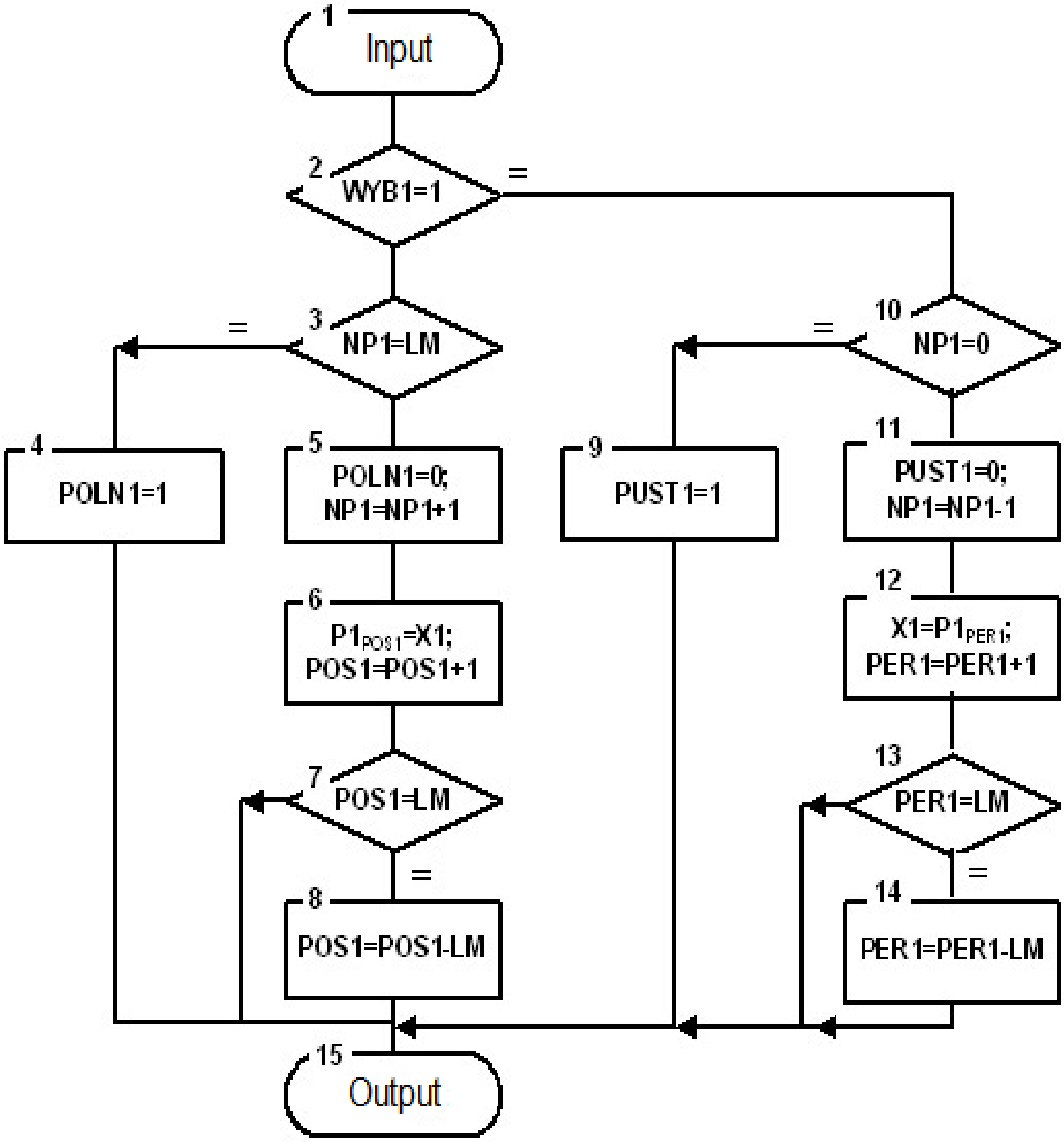

Fig. 3. Scheme of the algorithm for modeling the work of the queue 1

The scheme of the algorithm, which simulates the logic of the service channel and simultaneously performs the accumulation of statistical data, is shown in Fig. 4. At certain points in time (with a period of *DT*) it is necessary to calculate and print the following statistical characteristics:

$$SWR = SW / N \, , \tag{16}$$

$$DW = SW2/(N-1) - N/(N-1) \times WSR^2 \, . \tag{17}$$

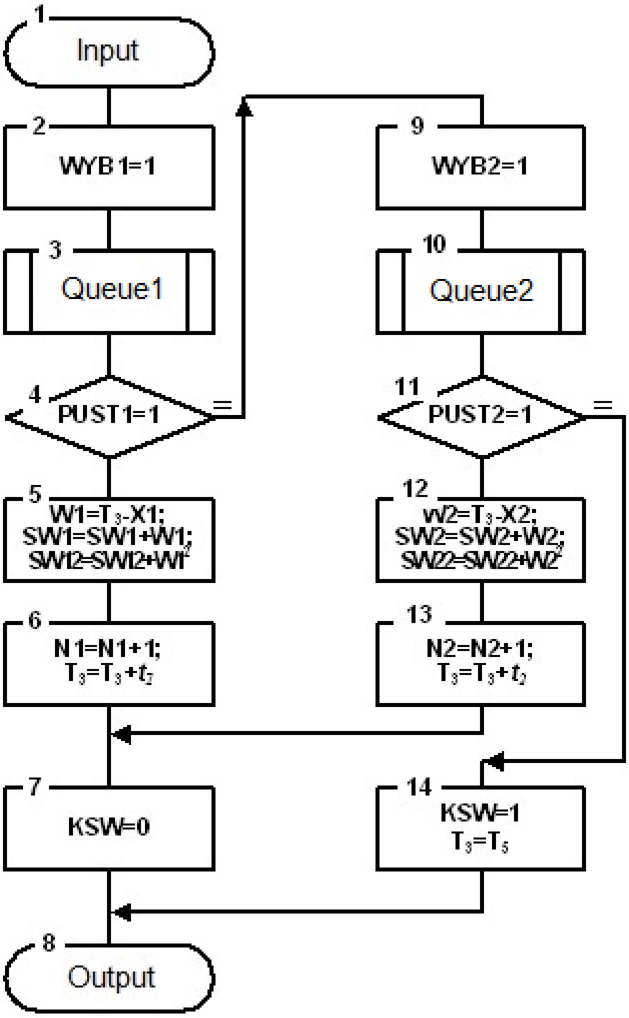

Fig. 4. Scheme of the algorithm for modeling the operation of the service channel

It remains to make an algorithm that provides the correct sequence of events in the system model. Such events are the arrival in the system of the machine of the first or second stream, the release of the service channel, the printing of statistical characteristics. The operation of this algorithm is based on the introduction of the system time concept, which is used to represent time-ordered events.

System time changes discretely, passing sequentially through all moments of events. The role of the algorithm, that controls the sequence of events, is now to determine the time of the next event and transfer control to those algorithms, that simulates this event. The moment of occurrence of the next event can be determined by finding the minimum element in the so-called list of future events, which represents the nearest moments of requirements arriving of the first or second streams, the nearest dismissal moment of the service channel, the nearest moment of printing statistical characteristics. This minimum element will determine the new value of the system time and what event should be performed at this time.

After this event is performed, the list of future events is restored by calculating a new point in time, that corresponds to the type of event, that occurred. For example, if the minimum list item is the arrival time of the first flow machine, you must refer to the subroutine, that implements the algorithm of queue 1, to write the machine in this queue, and then generate the value of the arrival time of the next first flow machine (new future event), according to the algorithm shown in Fig. 2.

The algorithm scheme of the whole system model is shown in Fig. 5. In this scheme, the list of future T_{IKR} events (block 3 in Fig. 5) also includes the end of the simulation.

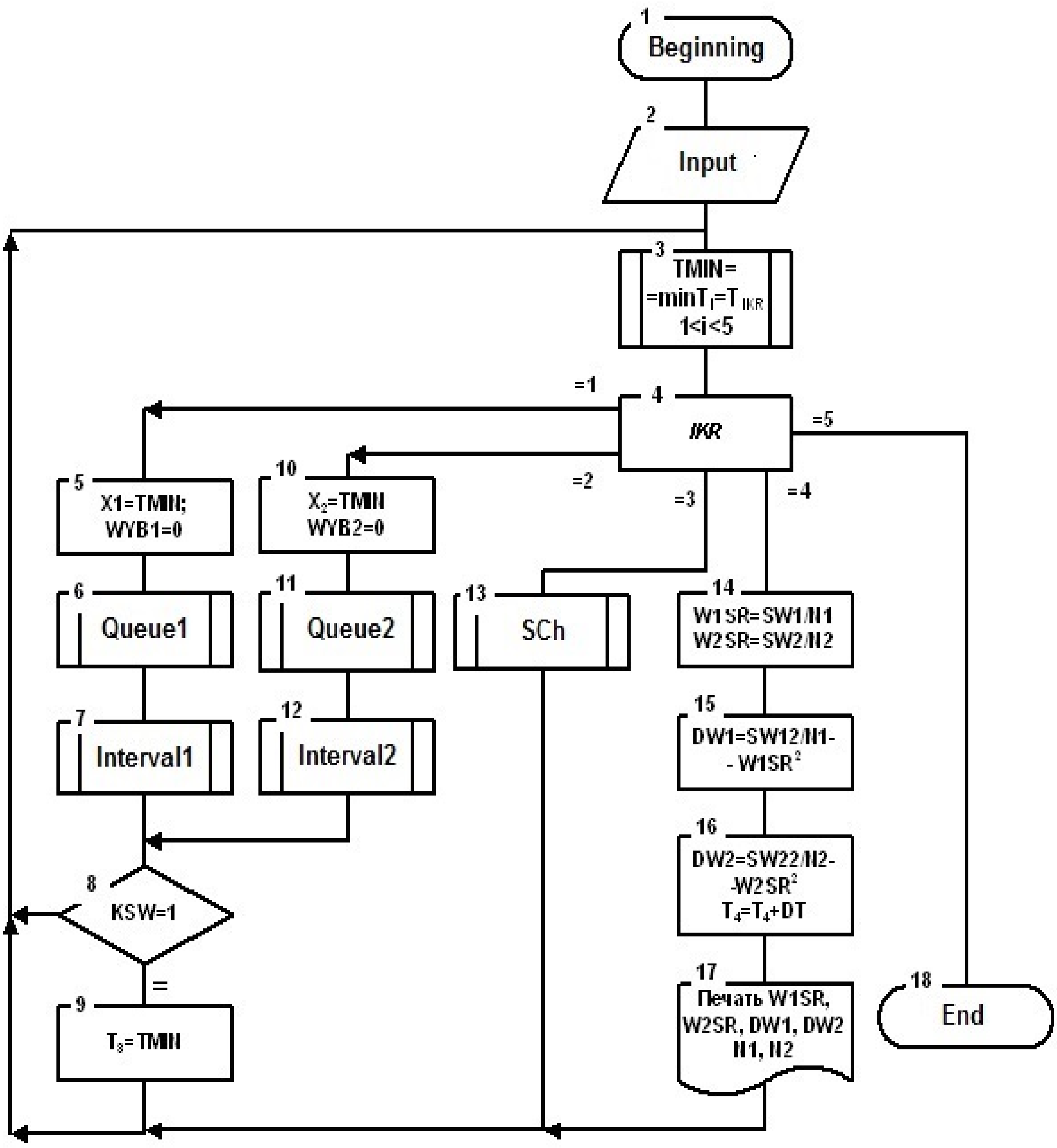

Fig. 5. Scheme of the general modeling algorithm reloading complex

4. Simulation model development of the reloading complex and conducting experimental research

Based on the developed structural scheme of the reloading complex (Fig. 1), built algorithms for modeling individual elements (Fig. 2 - Fig. 4) and the general algorithm of the complex (Fig. 5), developed a computer simulation model of the reloading complex.

The Embarcadero RAD Studio XE10 Seattle object-oriented programming system was used to develop a simulation model of the reloading complex, in which the input flow of machines is serviced according to the rule of relative priorities [RAD, 2019].

In Fig. 6 shows the main window of the computer simulation model of the reloading complex.

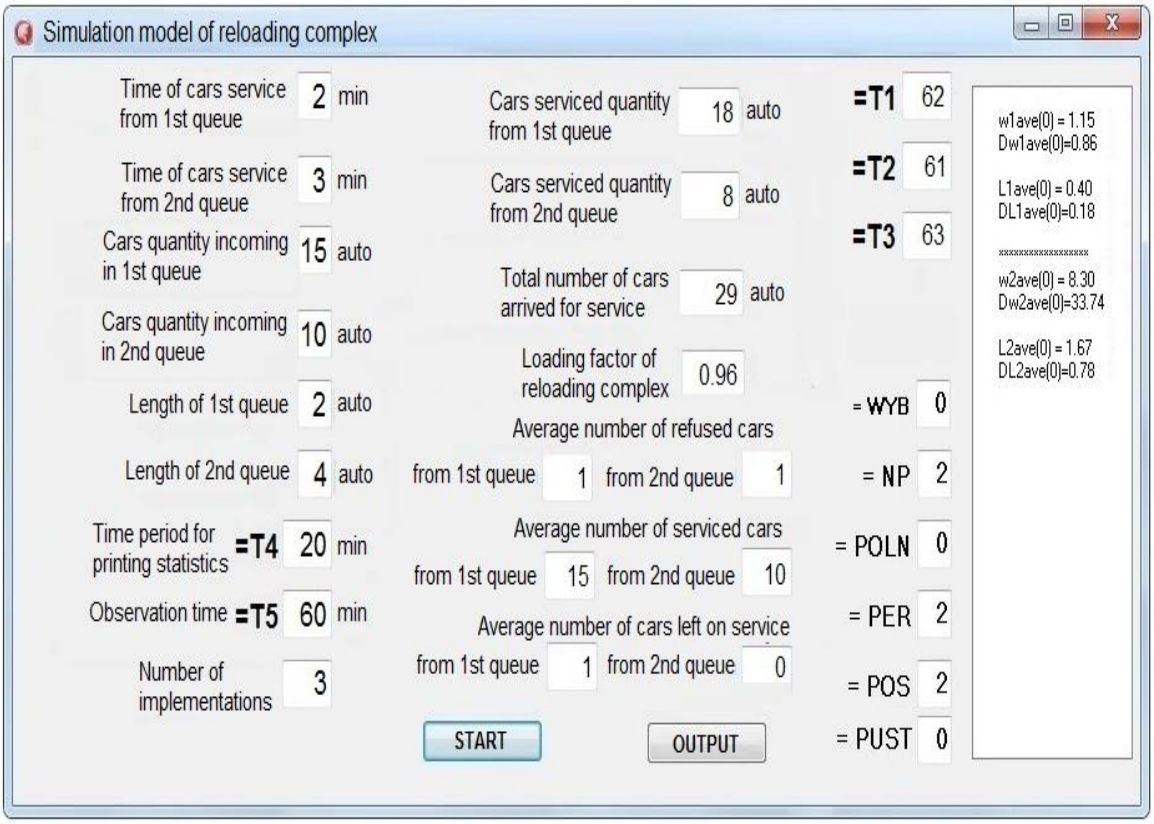

Fig. 6. Simulation model window of reloading complex.

.

The input parameters were:

• machine service time in the first and second queues - τ_1 =2, τ_2 =3;

• the number of machines arrival in the first and second queues for 1 hour of system operation - λ_1 =15, λ_2 =10;

• length of the first queue L_1 =2 and the second queue L_2 =4 (maximum = 5);

• period of printing statistics time T_4 =20 min;

• observation time T_5 =60 min;

• number of system implementations = 1, 2, 3, ... , *n*.

Simulation of the machines input flow was carried out in accordance with the scheme of the algorithm shown in Fig. 2.

The entry of machines into the system was carried out using the built-in Embarcadero RAD Studio XE10 Seattle standard program - a generator of pseudo-random numbers, the sequence of which is subject to uniform distribution in the interval (0,1).

Simulation of the first and second queues was performed according to the scheme of the algorithm shown in Fig. 3. The total number of machines entering the first and second queues for 1 hour of system operation was calculated:

$$\lambda_{tot} = \lambda_1 + \lambda_2 . \quad (18)$$

Simulation of the reloading complex service channel was performed in accordance with the scheme of the algorithm shown in Fig. 4. At the same time, the waiting time of machines in the queue and the square of this value were summed, for the subsequent determination of the average waiting time and the average value of the queue length, as well as the variance of the values ω and L. The following system parameters were also calculated:

• the number of serviced cars, received separately from the first and second queues for 1 hour of its operation:

$$\mu_{1ave} = 60/\tau_1 \times \frac{\lambda_1}{\lambda_1 + \lambda_2}; \tag{19}$$

$$\mu_{2ave} = 60/\tau_2 \times \frac{\lambda_2}{\lambda_1 + \lambda_2}; \tag{20}$$

• the total number of serviced machines received from the first and second queues for 1 hour of system operation:

$$\mu_{tot} = \mu_{1ave} + \mu_{2ave}. \tag{21}$$

In addition, the calculation of the following parameters was also performed:

- the average number of cars that were denied service from the first and second queues;

- the average number of serviced cars from the first and second queues;

- the average number of cars remaining in the first and second queues.

Simulation of the entire reloading complex was performed according to the scheme of the algorithm shown in Fig. 5, which provides the correct sequence of events in the system model. The following events are (block 3 in Fig. 5):

1. Work of the first queue;

2. Work of the second queue;

3. Operation of the service channel;

4. Calculation and printing of statistical characteristics;

5. End of system work.

5. Discussion and Outlook

The results of calculating the average waiting time in the first (ω_{1ave}) and second (ω_{2ave}) queues, the average length of the first (L_{1ave}) and second (L_{2ave}) queues and the variance of ω and L, which were calculated by formulas (1)-(6), were displayed in the window of the ListBox component (Fig. 6), and for each implementation of the system separately.

From the simulation results (Fig. 6) it is seen, that for the given of the input flow intensity, coming in the first and second queues $\tau_1 = 2$ and $\tau_2 = 3$, and the performance of the cars service channel, coming from the first and second queues $\lambda_1 = 15$, $\lambda_2 = 10$, give the following values of the main parameters reloading complex:

• average waiting time in the first queue $\omega_{1ave} = 1.15$ min, and in the second queue $\omega_{2ave} = 8.3$ min;

• the length average value of the first queue $L_{1ave} = 0.40$ cars and the second queue $L_{2ave} = 1.67$ cars;

• variance of the average waiting time in the first queue $D\omega_{1ave} = 0.86$, and in the second queue $D\omega_{2ave} = 33.74$;

• variance of the length average value of the first queue $DL_{1ave} = 0.18$ and the second queue $DL_{2ave} = 0.78$.

Thus, simulation and obtaining the basic parameters of the reloading complex, depending on the intensity of the machines input flow and performance of the service channel, makes it possible to analyze and select the optimal modes of the complex working – determine the optimal number of machines or service channels required for cargo delivery and execution reloading works, which ensures optimal use of vehicles and reloading machines.

In addition, the results of modeling the reloading complex show, that simulation models of transport logistics systems are convenient to build on a block basis, and each block is subject to its own logic, which simulates a particular process or element of the system being modeled. Some of these blocks can be implemented as standard, by including them in libraries of standard programs.

6. Conclusions

The article considers the reloading complex, in which the incoming flow of machines is served by the rule of relative priorities. In this case, the reloading complex is considered as a queuing system.

Modeling algorithms schemes of separate elements (input stream 1 and queue 1, service channel) and the scheme of the general modeling algorithm of a reloading complex are constructed, on the basis of which its computer model developed, by means of programming system - Embarcadero RAD Studio XE10.

The calculation of the main characteristics reloading complex, such as the total and average queue length, total and average waiting time in the queue, as well as their variance, depending on the intensity of the machines input flow and performance channel, which allows analysis and selection of optimal modes.

Bibliography

[Christopher, 2016] Christopher M. Logistics & supply chain management, 5th ed, p. 328. Pearson UK (2016)

[Shortle, 2018] Shortle J., Thompson J., Gross D., Harris C. Fundamentals of queueing theory, 5th ed, p. 576. John Wiley & Sons (2018). https://doi.org/10.1002/9781119453765

[Trivedi, 2016] Trivedi K. Probability and Statistics with Reliability, Queueing, and Computer Science Applications, 2nd ed, p. 830. Hoboken, New York: John Wiley & Sons (2016)

[RAD, 2019] RAD Studio XE10 Seattle Product Reviewer's Guide, (2019). http://docwiki.embarcadero.com/RADStudio/XE10.

[Khorshidian, 2019] Khorshidian H., Shirazi M., Ghomi S. An intelligent truck scheduling and transportation planning optimization model for product portfolio in a crossdock. Journal of Intelligent Manufacturing, vol. 30(1), pp. 163–184, (2019)

[Kumar, 2016] Kumar R., Kondapaneni K., Dixit V., Goswami A. Multi-objective modeling of production and pollution routing problem with time window: a self-learning particle swarm optimization approach. Computers & Industrial Engineering, vol. 99, pp. 29–40, (2016)

[Escobar, 2020] Escobar J., Marin A., Lince J. Multi-objective mathematical model for the redesign of supply chains considering financial criteria optimisation and scenarios. International Journal of Mathematics in Operational Research, vol. 16(2), pp. 238–256, (2020)

[Heidari, 2018] Heidari F., Zegordi S., Tavakkoli-Moghaddam R. Modeling truck scheduling problem at a cross-dock facility through a bi-objective bi-level optimization approach. Journal of Intelligent Manufacturing, vol. 29(5), pp. 1155–1170, 2018

[Nasiri, 2018] Nasiri M., Rahbari A., Werner F., Karimi R. Incorporating supplier selection and order allocation into the vehicle routing and multi-cross-dock scheduling problem. International Journal of Production Research, vol. 56(19), pp. 6527–6552, (2018)

[Soeanu, 2020] Soeanu A., Ray S., Berger J., Boukhtouta A., Debbabi M. Multi-depot vehicle routing problem with risk mitigation: model and solution algorithm. Expert Systems with Applications, vol. 145, pp. 135–147, (2020).

[Lai, 2015] Lai X., Bierlaire M. Specification of the cross-nested logist model with sampling of alternatives for route choice models. Transp. Res. Part B Methodol, vol. 80, pp. 220–234, (2015)

[Sun, 2020] Sun Y. Green and reliable freight routing problem in the road-rail intermodal transportation network with uncertain parameters: a fuzzy goal programming approach. Journal of Advanced Transportation, vol. 2020, pp. 1–21, (2020)

[Luisa, 2015] Luisa M., Vitetta A. Route choice on road transport system: a fuzzy approach. J. Intell. Fuzzy Syst, vol. 28(5), pp. 2015–2027, (2015)

[Tokgeoz, 2015] Tokgeoz E., Alwazzi S., Trafalis T. A heuristic algorithm to solve the single-facility location routing problem on Riemannian surfaces. Computational Management Science Journal, vol. 12(3), pp. 297–415, (2015)

[Manley, 2015] Manley E., Orr S., Cheng T. A heuristic model of bounded route choice in urban areas. Transp. Res. Part C Emerg. Technol, vol. 56, pp. 195–209, (2015)

[Yu, 2020] Yu Y., Zhou J., Xu W. Regret-based multi-objective route choice models and stochastic user equilibrium: a noncompensatory approach. Transportmetrica A: Transport Science, vol. 16(3), pp. 473–500, (2020)

[Rodado, 2017] Rodado D., Escobar J., Garcrıa-Crаceres R., Atencio F. A mathematical model for the product mixing and lot-sizing problem by considering stochastic demand. International Journal of Industrial Engineering Computations, vol. 8(2), pp. 237–250, (2017)

[Havelock, 2018] Havelock J., Ommen B., Granmo O. Novel distance estimation methods using “stochastic learning on the line” strategies. IEEE Access, vol. 6, pp. 48438–48454, (2018)

[Tintu, 2020] Tintu G., Amudha T. Genetic algorithm based multiobjective optimization framework to solve traveling salesman problem. Advances in

Computing and Intelligent Systems, pp. 141–151, Springer, Singapore, (2020)

[Mohammed, 2017] Mohammed M., Abd-Ghani M., Hamed R., Mostafa S. Solving vehicle routing problem by using improved genetic algorithm for optimal solution. Journal of Computational Science, vol. 21, pp. 255–262, (2017)

[Chiappone, 2016] Chiappone S., Giuffre O., Grana A., Mauro R. Traffic simulation models calibration using speed-density relationship: an automated procedure based on genetic algorithm. Expert Systems with Applications, vol. 44, pp. 147–155, (2016)

[Lu, 2017] Lu F., Bi H., Huang L., Bo W. Improved genetic algorithm based on delivery time control for Fourth Party Logistics. In Proceedings of the IEEE Conference on Automation Science and Engineering (CASE), pp. 390–393, Xinan, China, August (2017)

[Yang, 2016] Yang Y., Dong H., Qin Y., Zhang Q. Parameter calibration method of microscopic traffic flow simulation models based on orthogonal genetic algorithm. In Proceedings of the22nd International Conference on Distributed Multimedia Systems, Colchester, UK, November (2016)

[Hess, 2015] Hess S., Quddus M., Rieser-Schüssler N., Daly A. Developing advanced route choice models for heavy goods vehicles using GPS data. Transp. Res. Part E Logist. Transp. Rev, vol. 77, pp. 29–44, (2015)

[Nyrkov, 2015] Nyrkov A., Sokolov S., Belousov A. Algorithmic support of optimization of multicast data transission in networks with dynamic routing. J Modern Applied Science, vol. 9(5), pp. 162–176, (2015)

[Zhilenkov, 2015] Zhilenkov A., Nyrkov A., Cherniy S. Evaluation of reliability and efficiency of distributed systems rigs. Automation in the industry, vol. 6, pp. 50–52, (2015) [Kirpichnikov, 2018] Kirpichnikov A., Yelizarov A. Methods of Applied Queuing Theory, 2nd ed, p. 223. Moscow: LENAND, (2018)

[Gnedenko, 2018] Gnedenko B. Theory of probability, p. 520. Routledge, (2018). https://doi.org/10.1201/ 9780203718964

Author’s information

Mykola T. Dekhtyaruk *– Ph.D. Phys.- Mat. Sciences, Associate Professor, Senior Researcher, Professor of the Information Technologies and Programming Department, University "Ukraine", st. Lvivska, 23, Kyiv, Ukraine, 03115; e-mail: dekh@ukr.net.*

Major Fields of Scientific Research: Computer modeling, Transport logistics systems, Queuing systems.

TABLE OF CONTENTS