Vol. 23, no. 2, 2023 eISSN 2687-1653
PEER-REVIEWED SCIENTIFIC AND PRACTICAL JOURNAL

Advanced Engineering Research (Rostov-on-Don)

Mechanics

Machine Buiding and Machine Science

Information Technology, Computer Science and Management

www.vestnik-donstu.ru
DOI 10.23947/2687-1653

Advanced Engineering Research (Rostov-on-Don)

Peer-reviewed scientific and practical journal (published since 2000)

eISSN 2687-1653
DOI: 10.23947/2687-1653

Vol. 23, no. 2, 2023

The journal is aimed at informing the readership about the latest achievements and prospects in the field of mechanics, mechanical engineering, computer science and computer technology. The publication is a forum for cooperation between Russian and foreign scientists, it contributes to the convergence of the Russian and world scientific and information space.

The journal is included in the List of the leading peer-reviewed scientific publications (Higher Attestation Commission under the Ministry of Science and Higher Education of the Russian Federation), where basic scientific results of dissertations for the degrees of Doctor and Candidate of Sciences in scientific specialties and their respective branches of science should be published.

The journal publishes articles in the following fields of science:

– Theoretical Mechanics, Dynamics of Machines (Engineering Sciences)
– Deformable Solid Mechanics (Engineering, Physical and Mathematical Sciences)
– Mechanics of Liquid, Gas and Plasma (Engineering Sciences)
– Mathematical Simulation, Numerical Methods and Program Systems (Engineering Sciences)
– System Analysis, Information Management and Processing, Statistics (Engineering Sciences)
– Automation and Control of Technological Processes and Productions (Engineering Sciences)
– Software and Mathematical Support of Machines, Complexes and Computer Networks (Engineering Sciences)
– Computer Modeling and Design Automation (Engineering, Physical and Mathematical Sciences)
– Computer Science and Information Processes (Engineering Sciences)
– Machine Science (Engineering Sciences)
– Machine Friction and Wear (Engineering Sciences)
– Technology and Equipment of Mechanical and Physicotechnical Processing (Engineering Sciences)
– Engineering Technology (Engineering Sciences)
– Welding, Allied Processes and Technologies (Engineering Sciences)
– Methods and Devices for Monitoring and Diagnostics of Materials, Products, Substances and the Natural Environment (Engineering Sciences)
– Hydraulic Machines, Vacuum, Compressor Equipment, Hydraulic and Pneumatic Systems (Engineering Sciences)

Indexing and archiving	RSCI, CyberLeninka, EBSCO, Dimensions, DOAJ, Index Copernicus, Internet Archive, Google Scholar
Name of the body that registered the publication	Mass media registration certificate ЭЛ № ФС 77 – 78854 dated August 07, 2020, issued by the Federal Service for Supervision of Communications, Information Technology and Mass Media
Founder and publisher	Federal State Budgetary Educational Institution of Higher Education Don State Technical University (DSTU)
Periodicity	4 issues per year
Address of the founder and publisher	1, Gagarin sq., Rostov-on-Don, 344003, Russian Federation
E-mail	vestnik@donstu.ru
Telephone	+7 (863) 2–738–372
Website	http://vestnik-donstu.ru/
Date of publication	30.06.2023

Advanced Engineering Research (Rostov-on-Don)

Рецензируемый научно-практический журнал (издается с 2000 года)

eISSN 2687-1653
DOI: 10.23947/2687-1653

Том 23, № 2, 2023

Создан в целях информирования читательской аудитории о новейших достижениях и перспективах в области механики, машиностроения, информатики и вычислительной техники. Издание является форумом для сотрудничества российских и иностранных ученых, способствует сближению российского и мирового научно-информационного пространства.

Журнал включен в перечень рецензируемых научных изданий, в котором должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (Перечень ВАК) по следующим научным специальностям:

1.1.7 – Теоретическая механика, динамика машин (технические науки)
1.1.8 – Механика деформируемого твердого тела (технические, физико-математические науки)
1.1.9 – Механика жидкости, газа и плазмы (технические науки)
1.2.2 – Математическое моделирование, численные методы и комплексы программ (технические науки)
2.3.1 – Системный анализ, управление и обработка информации, статистика (технические науки)
2.3.3 – Автоматизация и управление технологическими процессами и производствами (технические науки)
2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки)
2.3.7 – Компьютерное моделирование и автоматизация проектирования (технические, физико-математические науки)
2.3.8 – Информатика и информационные процессы (технические науки)
2.5.2 – Машиноведение (технические науки)
2.5.3 – Трение и износ в машинах (технические науки)
2.5.5 – Технология и оборудование механической и физико-технической обработки (технические науки)
2.5.6 – Технология машиностроения (технические науки)
2.5.8 – Сварка, родственные процессы и технологии (технические науки)
2.5.9 – Методы и приборы контроля и диагностики материалов, изделий, веществ и природной среды (технические науки)
2.5.10 – Гидравлические машины, вакуумная, компрессорная техника, гидро- и пневмосистемы (технические науки)

Индексация и архивация:	РИНЦ, CyberLeninka, CrossRef, Dimensions, DOAJ, EBSCO, Index Copernicus, Internet Archive, Google Scholar
Наименование органа, зарегистрировавшего издание	Свидетельство о регистрации средства массовой информации ЭЛ № ФС 77 – 78854 от 07 августа 2020 г., выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций
Учредитель и издатель	Федеральное государственное бюджетное образовательное учреждение высшего образования «Донской государственный технический университет» (ДГТУ)
Периодичность	4 выпуска в год
Адрес учредителя и издателя	344003, Российская Федерация, г. Ростов-на-Дону, пл. Гагарина, 1
E-mail	vestnik@donstu.ru
Телефон	+7 (863) 2–738–372
Сайт	http://vestnik-donstu.ru/
Дата выхода в свет	30.06.2023

Contents

Содержание

MECHANICS
МЕХАНИКА

UDC 539.47

Original article

https://doi.org/10.23947/2687-1653-2023-23-2-113-120

Analysis of Stress-Strain State of a Cylinder with Variable Elasticity Moduli Based on Three-Dimensional Equations of Elasticity Theory

Jalala Ismayilova
Ganja State University, Ganja, Azerbaijan
celaleismayilova@mail.ru

Abstract

Introduction. Functionally graded materials are of great use, because heterogeneity of properties enables to control the strength and rigidity of structures. This has caused great interest in the topic in the world scientific literature. The construction of solutions to such problems depends significantly on the type of boundary conditions. In this paper, we consider the equilibrium of a thin-walled circular cylinder whose mechanical properties change along the radius. Homogeneous boundary conditions were set on cylindrical surfaces that had not been considered before, the effect was on the ends. The mathematical formulation of the problem was carried out in the linear theory of elasticity in the framework of axisymmetric deformation. Expressions were constructed for the components of the stress-strain state of the cylinder, in which some coefficients were found from the solution to the resulting system of linear algebraic equations.

Materials and Methods. The material of the cylinder was linearly elastic, the elastic modulus of which depended linearly on the radial coordinate. The basic research method was the asymptotic method, in which half the logarithm of the ratio of the outer and inner radii acted as a small parameter. Iterative processes were used to construct the characteristics of the stress-strain state of the cylinder.

Results. Homogeneous solutions to the boundary value problem were obtained for a linearly elastic functionally gradient hollow thin-walled cylinder. An analysis of these solutions made it possible to reveal the nature of the stress-strain state in the cylinder wall. For this purpose, an asymptotic analysis of the solutions was carried out, relations for displacements and stresses were obtained. It was determined that those solutions corresponded to the boundary layer, while their first terms determined Saint-Venant edge effect similar to the plate theory.

Discussion and Conclusion. The analytical solution to the equilibrium problem of a thin-walled cylinder inhomogeneous in radius constructed by asymptotic expansion can be used for numerical solution to a specific problem. For this, it is required to solve the obtained systems of linear algebraic equations and determine the corresponding coefficients. The resulting asymptotic representations provide analyzing the three-dimensional stress-strain state. The selection of the number of expansion terms makes it possible to calculate displacements and stresses with a given degree of accuracy. This analysis can be useful in assessing the adequacy of applied calculation methods used in engineering practice.

Keywords: linear theory of elasticity, functionally graded material, thin-walled hollow cylinder, homogeneous solutions, boundary layer, variational principle

Acknowledgements: the author would like to thank the editorial board and the reviewer for the attentive attitude to the article and the suggestions made that helped to improve its quality.

For citation. Ismayilova J. Analysis of Stress-Strain State of a Cylinder with Variable Elasticity Moduli Based on Three-Dimensional Equations of Elasticity Theory. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2):113–120. https://doi.org/10.23947/2687-1653-2023-23-2-113-120

Научная статья

Анализ напряженно-деформированного состояния цилиндра с переменными модулями упругости на основе трехмерных уравнений теории упругости

Д.Д. Исмайылова ✉

Гянджинский государственный университет, г. Гянджа, Азербайджан

✉ celaleismayilova@mail.ru

Аннотация

Введение. Функционально-градиентные материалы находят большое применение, т.к. неоднородность свойств позволяет управлять прочностью и жесткостью контрукций. Этим вызван большой интерес к данной теме в мировой научной литературе. Построение решения таких задач существенно зависит от типа граничных условий. В настоящей работе рассматривается равновесие тонкостенного кругового цилиндра, механические свойства которого заменяются вдоль радиуса. На цилиндрических поверхностях заданы однородные граничные условия, которые до этого не рассматривались, воздействие оказывается на торцах. Математическая постановка задачи осуществляется в линейной теории упругости в рамках осесимметричной деформации. В работе построены выражения для компонент напряженно-деформированного состояния цилиндра, в которых некоторые коэффициенты находятся из решения полученной системы линейных алгебраических уравнений.

Материалы и методы. Материал цилиндра является линейно упругим, модуль упругости которого линейно зависит от радиальной координаты. Основным методом исследования является ассимптотический метод, в котором в качестве малого параметра выступает половина логарифма отношения внешнего и внутреннего радиусов. Для построения характеристик напряженно-деформированного состояния цилиндра применены итерационные процессы.

Результаты исследования. Для линейно-упругого функционально-градиентного полого тонкостенного цилиндра получены однородные решения краевой задачи. Анализ этих решений позволяет раскрыть характер напряженно-деформированного состояния в стенке цилиндра. С этой целью проведен асимптотический анализ решений, получены соотношения для перемещений и напряжений. Установлено, что эти решения соответствуют пограничному слою, при этом их первые члены определяют краевой эффект Сен-Венана, аналогичный теории плит.

Обсуждение и заключение. Построенное с помощью асимптотического разложения аналитическое решение задачи о равновесии неоднородного по радиусу тонкостенного цилиндра может быть использовано для численного решения конкретной задачи. Для этого нужно решить полученные системы линейных алгебраических уравнений и определить соответствующие коэффициенты. Полученные асимптотические представления позволяют анализировать трехмерное напряженно-деформированное состояние. Выбор количества членов разложения позволяет рассчитать перемещения и напряжения с заданной степенью точности. Этот анализ может быть полезен при оценке адекватности прикладных методов расчета, применяемых в инженерной практике.

Ключевые слова: линейная теория упругости, функционально-градиентный материал, тонкостенный полый цилиндр, однородные решения, пограничный слой, вариационный принцип

Благодарности: авторы выражают благодарность редакции и рецензенту за внимательное отношение к статье и предложения, которые позволили повысить ее качество.

Для цитирования. Исмайылова Д.Д. Анализ напряженно-деформированного состояния цилиндра с переменными модулями упругости на основе трехмерных уравнений теории упругости. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2):113–120. https://doi.org/10.23947/2687-1653-2023-23-2-113-120

Introduction. Functionally graded materials are widely used in various constructions. Due to the dependence of mechanical properties on coordinates, it is possible to control the stress-strain state (SSS) of parts. An example of using such inhomogeneity is a cylinder whose mechanical properties depend on the radius. In this case, the cylinder may be of interest as a separate structure, or as being a division subring of a compound body, e.g., connecting two media with widely different properties. When calculating the SSS of a thin-walled cylinder, some applied theories can be used, an assessment of their adequacy. Especially in the case of inhomogeneous properties, it can be carried out using computer modeling or asymptotic analysis based on a three-dimensional formulation. The latter determines the relevance of this study.

A number of studies have been devoted to the investigation of the SSS of hollow cylindrical bodies within the framework of the linear theory of elasticity. In [1, 2], the mechanical behavior of a radially inhomogeneous cylinder in a three-dimensional formulation was studied on the basis of the spline collocation method and the finite element method. In [3], the SSS of a cylinder whose properties depended on the radius loaded with uniform internal pressure was described. In [4], an analytical study was carried out for a functionally graded piezoelectric cylinder. In [5], an exact solution was constructed to a radially inhomogeneous hollow cylinder with exponential Young's modulus, with constant Poisson ratio and power Young's modulus. In [6, 7], an analytical solution to the axisymmetric thermoelasticity problem for a continuous cylinder was obtained using the direct integration method when the coefficient of linear thermal expansion was an arbitrary function of the radius. In [8], a general asymptotic theory of a transversally isotropic homogeneous hollow cylinder was developed. New groups of solutions were obtained for a transversally isotropic homogeneous cylinder. The comparison of the constructed solutions to the solutions constructed using applied calculation methods was given. In [9, 10], some boundary value problems of elasticity theory were studied for a functionally gradient isotropic and transversally isotropic (the plane of isotropy was perpendicular to the axis) cylinder, in the case when the elastic modules were arbitrary continuous functions of the radius of the cylinder. In [11], an analysis of the bending deformation problem for a radially inhomogeneous cylinder was carried out. The analysis of the above papers shows that not all types of boundary conditions on cylindrical surfaces have asymptotic representations of solutions.

In this article, on the basis of an asymptotic analysis of three-dimensional equations of elasticity theory, the features of the SSS of a thin-walled cylinder whose properties vary linearly along the radius were studied. In this case, the inner border was fixed in the axial direction and was free in the radial direction.

This goal was achieved using several steps: asymptotic integration of differential equations and the construction of homogeneous solutions; derivation of formulas for the components of the displacement vector and stress tensor; consideration of boundary conditions on the face surfaces.

Materials and Methods. Radially inhomogeneous hollow thin-walled cylinder $\Gamma = \{r \in [r_1; r_2],\ \phi \in [0; 2\pi],\ z \in [-l_0; l_0]\}$ is considered in a cylindrical coordinate system with the origin on its axis. The problem of its equilibrium, in the case of fixing cylindrical surfaces along the axis and zero normal stresses, is solved in an axisymmetric formulation under the action of stresses at its faces.

The boundary value problem consists of the equilibrium equations [8]:

$$\frac{\partial \sigma_{rr}}{\partial r} + \frac{\partial \sigma_{rz}}{\partial z} + \frac{\sigma_{rr} - \sigma_{\phi\phi}}{r} = 0, \tag{1}$$

$$\frac{\partial \sigma_{rz}}{\partial r} + \frac{\partial \sigma_{zz}}{\partial z} + \frac{\sigma_{rz}}{r} = 0, \tag{2}$$

where $\sigma_{rr}, \sigma_{rz}, \sigma_{\phi\phi}, \sigma_{zz}$ — components of the stress tensor.

Defining relations [8]:

$$\sigma_{rr} = (2G + \lambda)\frac{\partial u_r}{\partial r} + \lambda\left(\frac{u_r}{r} + \frac{\partial u_z}{\partial z}\right), \tag{3}$$

$$\sigma_{zz} = (2G+\lambda)\frac{\partial u_z}{\partial z} + \lambda\left(\frac{\partial u_r}{\partial r} + \frac{u_r}{r}\right), \tag{4}$$

$$\sigma_{\phi\phi} = (2G+\lambda)\frac{u_r}{r} + \lambda\left(\frac{\partial u_r}{\partial r} + \frac{\partial u_z}{\partial z}\right), \tag{5}$$

$$\sigma_{rz} = G\left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right). \tag{6}$$

Here, $u_r = u_r(r,z)$, $u_z = u_z(r,z)$ — components of the displacement vector.

Lamé parameters vary linearly along the radius:

$$G(r) = G_* r, \quad \lambda(r) = \lambda_* r, \tag{7}$$

where G_*, λ_* — constants.

After substituting (3)–(7) into equations (1), (2), the dimensionless system of equations takes the form:

$$(2G_0+\lambda_0)\left(\frac{\partial^2 u_\rho}{\partial\rho^2} + \varepsilon\frac{\partial u_\rho}{\partial\rho}\right) + \varepsilon(G_0+\lambda_0)e^{\varepsilon\rho}\frac{\partial^2 u_\xi}{\partial\rho\partial\xi} + \lambda_0\varepsilon^2 e^{\varepsilon\rho}\frac{\partial u_\xi}{\partial\xi} + G_0\varepsilon^2 e^{2\varepsilon\rho}\frac{\partial^2 u_\rho}{\partial\xi^2} - 2G_0\varepsilon^2 u_\rho = 0, \tag{8}$$

$$G_0\left(\frac{\partial^2 u_\xi}{\partial\rho^2} + \varepsilon\frac{\partial u_\xi}{\partial\rho}\right) + (2G_0+\lambda_0)\varepsilon^2\left(e^{2\varepsilon\rho}\frac{\partial^2 u_\xi}{\partial\xi^2} + e^{\varepsilon\rho}\frac{\partial u_\rho}{\partial\xi}\right) + (G_0+\lambda_0)\varepsilon e^{\varepsilon\rho}\frac{\partial^2 u_\rho}{\partial\rho\partial\xi} = 0. \tag{9}$$

Here:

$\rho = \frac{1}{\varepsilon}\ln\left(\frac{r}{r_0}\right)$, $\xi = \frac{z}{r_0}$ — new dimensionless coordinates; $\varepsilon = \frac{1}{2}\ln\left(\frac{r_2}{r_1}\right)$ — in the case of wall thinness, small parameter;

$r_0 = \sqrt{r_1 r_2}$, $\rho \in [-1;1]$, $\xi \in [-l;l]$, $l = \frac{l_0}{r_0}$; $u_\rho = \frac{u_r}{r_0}$, $u_\xi = \frac{u_z}{r_0}$, $\lambda_0 = \frac{\lambda_* r_0}{G_1}$, $G_0 = \frac{G_* r_0}{G_1}$; G_1 — some parameter having the dimension of stress.

Let us consider a problem in which homogeneous boundary conditions are set on the lateral surfaces of the cylinder:

$$u_\xi\big|_{\rho=\pm1} = 0, \tag{10}$$

$$\sigma_{\rho\rho}\big|_{\rho=\pm1} = 0. \tag{11}$$

Stresses are applied to the faces of the cylinder:

$$\sigma_{\rho\xi}\big|_{\xi=\pm l} = t_{1s}(\rho), \tag{12}$$

$$\sigma_{\xi\xi}\big|_{\xi=\pm l} = t_{2s}(\rho), \tag{13}$$

$(s = 1;2)$.

$\sigma_{\rho\rho} = \frac{\sigma_{rr}}{G_1}$, $\sigma_{\rho\xi} = \frac{\sigma_{rz}}{G_1}$, $\sigma_{\xi\xi} = \frac{\sigma_{zz}}{G_1}$ — dimensionless stresses.

Stress vector components $t_{1s}(\rho), t_{2s}(\rho)$, $(s = 1;2)$ satisfy the equilibrium conditions.

To construct homogeneous solutions, we seek the components of the displacement vector in the form:

$$u_\rho(\rho;\xi) = u(\rho)e^{\alpha\xi}, \quad u_\xi(\rho;\xi) = w(\rho)e^{\alpha\xi}. \tag{14}$$

Substituting representations (14) into the system (8)–(11), we obtain:

$$(2G_0+\lambda_0)\left(u''(\rho) + \varepsilon u'(\rho)\right) + \varepsilon\alpha e^{\varepsilon\rho}\left((G_0+\lambda_0)w'(\rho) + \lambda_0\varepsilon w(\rho)\right) + \varepsilon^2 G_0\left(\alpha^2 e^{2\varepsilon\rho} - 2\right)u(\rho) = 0, \tag{15}$$

$$G_0\left(w''(\rho) + \varepsilon w'(\rho)\right) + (2G_0+\lambda_0)\varepsilon^2\left(\alpha e^{\varepsilon\rho}u(\rho) + \alpha^2 e^{2\varepsilon\rho}w(\rho)\right) + \varepsilon(G_0+\lambda_0)\alpha e^{\varepsilon\rho}u'(\rho) = 0, \tag{16}$$

$$w\big|_{\rho=\pm1} = 0, \tag{17}$$

$$\left[\left(2G_0+\lambda_0\right)u'(\rho)+\varepsilon\lambda_0\left(u(\rho)+\alpha e^{\varepsilon\rho}w(\rho)\right)\right]\Big|_{\rho=\pm1}=0. \tag{18}$$

We investigate boundary value problems (15)–(18) at $\varepsilon\to 0$. To solve (15)–(18) at $\varepsilon\to 0$, we use the asymptotic method [9–13].

Nonzero solutions (15)–(18) correspond to the third iterative process, the components of the displacement vector are searched for in the form of expansions over a small parameter:

$$\begin{cases} u^{(3)}(\rho)=\varepsilon\left(u_{30}(\rho)+\varepsilon u_{31}(\rho)+\cdots\right), \\ w^{(3)}(\rho)=\varepsilon\left(w_{30}(\rho)+\varepsilon w_{31}(\rho)+\cdots\right), \\ \alpha=\varepsilon^{-1}\left(\beta_0+\varepsilon\beta_1+\cdots\right). \end{cases} \tag{19}$$

After substituting expansions (19) into equations (15)–(18) for terms of the first order, we have:

$$\left(2G_0+\lambda_0\right)u''_{30}(\rho)+\beta_0\left(G_0+\lambda_0\right)w'_{30}(\rho)+G_0\beta_0^2u_{30}(\rho)=0, \tag{20}$$

$$G_0w''_{30}(\rho)+\beta_0\left(G_0+\lambda_0\right)u'_{30}(\rho)+\left(2G_0+\lambda_0\right)\beta_0^2w_{30}(\rho)=0, \tag{21}$$

$$w_{30}(\rho)\big|_{\rho=\pm1}=0, \tag{22}$$

$$\left(\left(2G_0+\lambda_0\right)u'_{30}(\rho)+\lambda_0\beta_0w_{30}(\rho)\right)\Big|_{\rho=\pm1}=0. \tag{23}$$

Following [13], the spectral problem (20)–(23) corresponds to a potential solution for the plate.

Thus, the solutions are presented in the form:

$$\text{a)}\ u_\rho^{(3;1)}(\rho;\xi)=\varepsilon\sum_{k=1}^{\infty}T_k\left(-2G_0\beta_{0k}^2\sin\beta_{0k}\sin\left(\beta_{0k}\rho\right)+O(\varepsilon)\right)\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0k}+\varepsilon\beta_{1k}+\cdots\right)\xi\right), \tag{24}$$

$$u_\xi^{(3;1)}(\rho;\xi)=\varepsilon\sum_{k=1}^{\infty}T_k\left(2G_0\beta_{0k}^2\sin\beta_{0k}\cos\left(\beta_{0k}\rho\right)+O(\varepsilon)\right)\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0k}+\varepsilon\beta_{1k}+\cdots\right)\xi\right). \tag{25}$$

Here, β_{0k} is the solution to the equation:

$$\cos\beta_{0k}=0. \tag{26}$$

The stresses corresponding to solutions (24), (25) have the form:

$$\sigma_{\rho\rho}^{(3;1)}=\sum_{k=1}^{\infty}T_k\left(-4G_0^2\beta_{0k}^3\sin\beta_{0k}\cos\left(\beta_{0k}\rho\right)+O(\varepsilon)\right)\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0k}+\varepsilon\beta_{1k}+\cdots\right)\xi\right), \tag{27}$$

$$\sigma_{\rho\xi}^{(3;1)}=\sum_{k=1}^{\infty}T_k\left(-4G_0^2\beta_{0k}^3\sin\beta_{0k}\sin\left(\beta_{0k}\rho\right)+O(\varepsilon)\right)\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0k}+\varepsilon\beta_{1k}+\cdots\right)\xi\right), \tag{28}$$

$$\sigma_{\xi\xi}^{(3;1)}=\sum_{k=1}^{\infty}T_k\left(4G_0^2\beta_{0k}^3\sin\beta_{0k}\cos\left(\beta_{0k}\rho\right)+O(\varepsilon)\right)\exp\left(\frac{1}{\varepsilon}\left(\beta_{0k}+\varepsilon\beta_{1k}+\cdots\right)\xi\right), \tag{29}$$

$$\sigma_{\phi\phi}^{(3;1)}=O(\varepsilon). \tag{30}$$

$$\text{b)}\ u_\rho^{(3;2)}(\rho;\xi)=\varepsilon\sum_{i=1}^{\infty}F_i\left(2G_0\beta_{0i}^2\cos\beta_{0i}\cos\left(\beta_{0i}\rho\right)+O(\varepsilon)\right)\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0i}+\varepsilon\beta_{1i}+\cdots\right)\xi\right), \tag{31}$$

$$u_\xi^{(3;2)}(\rho;\xi)=\varepsilon\sum_{i=1}^{\infty}F_i\left(2G_0\beta_{0i}^2\cos\beta_{0i}\sin\left(\beta_{0i}\rho\right)+O(\varepsilon)\right)\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0i}+\varepsilon\beta_{1i}+\cdots\right)\xi\right). \tag{32}$$

Here, β_{0k} is the solution to the equation:

$$\sin\beta_{0i}=0. \tag{33}$$

The stresses corresponding to solutions (31), (32) have the form:

$$\sigma_{\rho\rho}^{(3;2)}=\sum_{i=1}^{\infty}F_i\left(-4G_0^2\beta_{0i}^3\cos\beta_{0i}\sin\left(\beta_{0i}\rho\right)+O(\varepsilon)\right)\exp\left(\frac{1}{\varepsilon}\left(\beta_{0i}+\varepsilon\beta_{1i}+\cdots\right)\xi\right), \tag{34}$$

$$\sigma_{\rho\xi}^{(3;2)}=\sum_{i=1}^{\infty}F_i\left(4G_0^2\beta_{0i}^3\cos\beta_{0i}\cos\left(\beta_{0i}\rho\right)+O(\varepsilon)\right)\exp\left(\frac{1}{\varepsilon}\left(\beta_{0i}+\varepsilon\beta_{1i}+\cdots\right)\xi\right), \tag{35}$$

$$\sigma_{\xi\xi}^{(3;2)} = \sum_{i=1}^{\infty} F_i \left(4G_0^2\beta_{0i}^3 \cos\beta_{0i} \sin\left(\beta_{0i}\rho\right) + O(\varepsilon)\right) \exp\left(\frac{1}{\varepsilon}\left(\beta_{0i} + \varepsilon\beta_{1i} + \cdots\right)\xi\right), \tag{36}$$

$$\sigma_{\phi\phi}^{(3;2)} = O(\varepsilon). \tag{37}$$

The general solution (15)–(18) will be a superposition of solutions (24), (25), (31), (32):

$$u_\rho(\rho;\xi) = \varepsilon\sum_{k=1}^{\infty} T_k \left(-2G_0\beta_{0k}^2 \sin\beta_{0k} \sin\left(\beta_{0k}\rho\right) + O(\varepsilon)\right)\times$$

$$\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0k} + \varepsilon\beta_{1k} + \cdots\right)\xi\right) + \varepsilon\sum_{i=1}^{\infty} F_i \left(2G_0\beta_{0i}^2 \cos\beta_{0i} \cos\left(\beta_{0i}\rho\right) + O(\varepsilon)\right)\times \exp\left(\frac{1}{\varepsilon}\left(\beta_{0i} + \varepsilon\beta_{1i} + \cdots\right)\xi\right), \tag{38}$$

$$u_\xi(\rho;\xi) = \varepsilon\sum_{k=1}^{\infty} T_k \left(2G_0\beta_{0k}^2 \sin\beta_{0k} \cos\left(\beta_{0k}\rho\right) + O(\varepsilon)\right)\times$$

$$\times\exp\left(\frac{1}{\varepsilon}\left(\beta_{0k} + \varepsilon\beta_{1k} + \cdots\right)\xi\right) + \varepsilon\sum_{i=1}^{\infty} F_i \left(2G_0\beta_{0i}^2 \cos\beta_{0i} \sin\left(\beta_{0i}\rho\right) + O(\varepsilon)\right)\times \exp\left(\frac{1}{\varepsilon}\left(\beta_{0i} + \varepsilon\beta_{1i} + \cdots\right)\xi\right), \tag{39}$$

Solutions (24), (25), (31), (32) have the character of a boundary layer. When moving away from the faces, solutions (24), (25), (31), (32) decrease exponentially.

To determine constants T_k, F_i, we use Lagrange variational principle. The variational principle takes the form [8]:

$$\sum_{s=1}^{2}\int_{-1}^{1}\left[\left(\sigma_{\rho\xi} - t_{1s}\right)\delta u_\rho + \left(\sigma_{\xi\xi} - t_{2s}\right)\delta u_\xi\right]\Big|_{\xi=\pm l} e^{2\varepsilon\rho} d\rho = 0. \tag{40}$$

Substituting (24–36) into (40), we have:

$$\sum_{k=1}^{\infty} M_{jk} T_{k0} = p_{0j}^{(1)}, \tag{41}$$

$$\sum_{i=1}^{\infty} Q_{ji} F_{i0} = p_{0j}^{(2)}. \tag{42}$$

Here:

$$M_{jk} = 16G_0^3\beta_{0j}^2\beta_{0k}^3\left(\beta_{0j} - \beta_{0k}\right)^{-1} \sin\beta_{0k} \sin\beta_{0j} \sin\left(\beta_{0j} - \beta_{0k}\right)\times\left(\exp\left(-\frac{\left(\beta_{0k} + \beta_{0j}\right)l}{\varepsilon}\right) + \exp\left(\frac{\left(\beta_{0k} + \beta_{0j}\right)l}{\varepsilon}\right)\right),\ \left(npu\ \ j \neq k\right)$$

$$M_{jj} = 16G_0^3\beta_{0j}^5 \sin^2\beta_{0j}\left(\exp\left(-\frac{2\beta_{0j}l}{\varepsilon}\right) + \exp\left(\frac{2\beta_{0j}l}{\varepsilon}\right)\right),\ \left(npu\ \ j = k\right)$$

$$p_{0j}^{(1)} = 2G\beta_{0j}^2 \sin\beta_{0j}\left[\int_{-1}^{1}\left(t_{21}(\rho)\cos\left(\beta_{0j}\rho\right) - t_{11}(\rho)\sin\left(\beta_{0j}\rho\right)\right)d\rho\cdot\exp\left(-\frac{\beta_{0j}l}{\varepsilon}\right) +\right.$$

$$\left. + \int_{-1}^{1}\left(t_{22}(\rho)\cos\left(\beta_{0j}\rho\right) - t_{12}(\rho)\sin\left(\beta_{0j}\rho\right)\right)d\rho\cdot\exp\left(\frac{\beta_{0j}l}{\varepsilon}\right)\right],$$

$$Q_{ji} = 16G_0^3\beta_{0j}^2\beta_{0i}^3\left(\beta_{0j} - \beta_{0i}\right)^{-1} \cos\beta_{0j} \cos\beta_{0i} \sin\left(\beta_{0j} - \beta_{0i}\right)\left(\exp\left(-\frac{\left(\beta_{0i} + \beta_{0j}\right)l}{\varepsilon}\right) + \exp\left(\frac{\left(\beta_{0i} + \beta_{0j}\right)l}{\varepsilon}\right)\right),\ \left(npu\ \ i \neq j\right)$$

$$Q_{jj} = 16G_0^3\beta_{0j}^5 \cos^2\beta_{0j}\left(\exp\left(-\frac{2\beta_{0j}l}{\varepsilon}\right) + \exp\left(\frac{2\beta_{0j}l}{\varepsilon}\right)\right),\ \left(npu\ \ i = j\right)$$

$$p_{0j}^{(2)} = 2G_0\beta_{0j}^2 \cos\beta_{0j}\left[\int_{-1}^{1}\left(t_{11}(\rho)\cos\left(\beta_{0j}\rho\right) + t_{21}(\rho)\sin\left(\beta_{0j}\rho\right)\right)d\rho\cdot\exp\left(-\frac{\beta_{0j}l}{\varepsilon}\right) +\right.$$

$$\left. + \int_{-1}^{1}\left(t_{12}(\rho)\cos\left(\beta_{0j}\rho\right) + t_{22}(\rho)\sin\left(\beta_{0j}\rho\right)\right)d\rho\cdot\exp\left(\frac{\beta_{0j}l}{\varepsilon}\right)\right],$$

$$T_k = T_{k0} + \varepsilon T_{k1} + \cdots,$$
$$F_i = F_{i0} + \varepsilon F_{i1} + \cdots.$$

Constants T_{kp}, F_{ip} $(p = 1, 2, ...)$ are found from systems of linear algebraic equations (41), (42), analogous to which are studied in [13].

Research Results. In the article, in an axisymmetric formulation, the solution to the problem of linear elasticity theory for a functionally graded hollow thin-walled cylinder, whose properties vary in thickness according to a linear law, was considered. Homogeneous cross boundary conditions were set on the lateral surfaces of the cylinder, and a stress vector was set at the faces. The constructed homogeneous solutions satisfied boundary conditions on cylindrical surfaces. For their construction, an asymptotic approach based on the expansion by a small parameter characterizing the relative thickness of the cylinder was used. To account for inhomogeneous boundary conditions at the faces, systems of linear algebraic equations similar to those studied in the literature were obtained. It was shown that the constructed SSS solutions had a boundary-layer character, which corresponded to the edge effect similar to the theory of inhomogeneous plates, which bears the name of Saint-Venant.

Discussion and Conclusion. Usually, when studying the SSS of thin-walled structures, applied calculation methods are built that reduce the dimension of the problem. In this regard, the task of determining the range of geometric and mechanical parameters in which these methods give acceptable accuracy is critical. The solutions to three-dimensional equations constructed in the work on the basis of asymptotic analysis make it possible to assess the adequacy of such applied theories with a predetermined accuracy threshold. In addition, these solutions can find application in the evaluation of numerical solutions to problems for structures with functionally graded materials.

References

1. Grigorenko AYa, Yaremchenko SN. Analysis of the Stress-Strain State of Inhomogeneous Hollow Cylinders. *International Applied Mechanics.* 2016;52(4):342–349. https://doi.org/10.1007/s10778-016-0757-3

2. Grigorenko AYa, Yaremchenko SN. Three-Dimensional Analysis of the Stress-Strain State of Inhomogeneous Hollow Cylinders Using Various Approaches. *International Applied Mechanics*. 2019;55(5):487–494. https://doi.org/10.1007/s10778-019-00970-2

3. Tutuncu N, Temel B. A Novel Approach to Stress Analysis of Pressurized FGM Cylinders, Disks and Spheres. *Composite Structures*. 2009;91(3):385–390. https://doi.org/10.1016/j.compstruct.2009.06.009

4. Hong-Liang Dai, Li Hong, Yi-Ming Fu, et al. Analytical Solution for Electromagnetothermoelastic Behaviors of a Functionally Graded Piezoelectric Hollow Cylinder. *Applied Mathematical Modelling*. 2010;34(2):343–357. https://doi.org/10.1016/j.apm.2009.04.008

5. Theotokoglou EE, Stampouloglou IH. The Radially Nonhomogeneous Elastic Axisymmentric Problem. *International Journal of Solids and Structures*. 2008;45(25–26):6535–6552. https://doi.org/10.1016/j.ijsolstr.2008.08.011

6. Tokovyy Yu, Chyzh A, Chien-Ching Ma. *Axisymmetric Thermal Stresses in a Radially-Inhomogeneous Elastic Cylinder Subjected to with-Respect-to-Length Varying Thermal Loadings*. In: Proc. 11th Int. Congress on Thermal Stresses. Palermo: Poguro edizioni; 2016. P. 263–266.

7. Tokovyy Yu, Chien-Ching Ma. Elastic Analysis of Inhomogeneous Solids: History and Development in brief. *Journal of Mechanics*. 2019;35(5):613–626. https://doi.org/10.1017/jmech.2018.57

8. Mekhtiev MF. *Asymptotic Analysis of Spatial Problems in Elasticity*. Singapore: Springer; 2019. 241 p. URL: https://link.springer.com/book/10.1007/978-981-13-3062-9 (accessed: 10.02.2023).

9. Akhmedov NK. Axisymmetric Problem of the Elasticity Theory for the Radially Inhomogeneous Cylinder with a Fixed Lateral Surface. *Applied and Computational Mechanics*. 2021;7(2):599–610.

10. Akhmedov NK, Akperova SB. Asymptotic Analysis of a 3D Elasticity Problem for a Radially Inhomogeneous Transversally Isotropic Hollow Cylinder. *Mechanics of Solids*. 2011;4:170–180. URL: https://mtt.ipmnet.ru/ru/Issues/2011/4/170 (accessed: 10.20.2023).

11. Akhmedov NK, Akbarova SB. Behavior of Solution of the Elasticity Problem for a Radial Inhomogeneous Cylinder with Small Thickness. *Eastern-European Journal of Enterprise Technologies Applied Mechanics*. 2021;6/7(114):29–42. https://doi.org/10.15587/1729-4061.2021.247500

12. Akhmedov NK, Sofiyev AH. Asymptotic Analysis of Three-Dimensional Problem of Elasticity Theory for Radially Inhomogeneous Transversally-Isotropic Thin Hollow Spheres. *Thin-Walled Structures*. 2019;139:232–241. https://doi.org/10.1016/j.tws.2019.03.022

13. Ustinov YuA. *Mathematical Theory of Transversely Inhomogeneous Slabs*. Rostov-on-Don: OOO «TSVVR»; 2006. 257 p. (In Russ.).

Received 01.03.2023
Revised 23.03.2023
Accepted 30.03.2023

About the Author:

Jalala Ismayilova, Postdoctoral student of the Department of General Engineering Disciplines and Technologies, Ganja State University (187, Khatai St., Ganja, AZ2003, Azerbaijan), ORCID, celaleismayilova@mail.ru

Conflict of interest statement: the author does not have any conflict of interest.

The author has read and approved the final manuscript.

Поступила в редакцию 01.03.2023
Поступила после рецензирования 23.03.2023
Принята к публикации 30.03.2023

Об авторе:

Джалала Джамшид кызы Исмайылова, докторант кафедры общих технических наук и технологий Гянджинского государственного университета (AZ2003, Азербайджан, г. Гянджа, ул. Гейдар Алиев, 187), ORCID, celaleismayilova@mail.ru

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Автор прочитал и одобрил окончательный вариант рукописи.

MECHANICS
МЕХАНИКА

UDC 519.688/534-16
https://doi.org/10.23947/2687-1653-2023-23-2-121-129

Original article

Implementation of Basic Operations for Sparse Matrices when Solving a Generalized Eigenvalue Problem in the ACELAN-COMPOS Complex

Pavel A. Oganesyan ✉, **Olga O. Shtein**
Southern Federal University, Rostov-on-Don, Russian Federation
✉ poganesyan@sfedu.ru

Abstract

Introduction. The widespread use of piezoelectric materials in various industries stimulates the study of their physical characteristics and determines the urgency of such research. In this case, modal analysis makes it possible to determine the operating frequency and the coefficient of electromechanical coupling of piezoelectric elements of various devices. These indicators are of serious theoretical and applied interest. The study was aimed at the development of numerical methods for solving the problem of determining resonance frequencies in a system of elastic bodies. To achieve this goal, we needed new approaches to the discretization of the problem based on the finite element method and the execution of the software implementation of the selected method in C# on the .net platform. Current solutions were created in the context of the ACELAN-COMPOS class library. The known methods of solving the generalized eigenvalue problem based on matrix inversion are not applicable to large-dimensional matrices. To overcome this limitation, the presented scientific work implemented the logic of constructing mass matrices and created software interfaces for exchanging data on eigenvalue problems with pre- and postprocessing modules.

Materials and Methods. A platform was used to implement numerical methods .net and the C# programming language. Validation of the research results was carried out through comparing the values found with solutions obtained in well-known SAE packages (computer-aided engineering). The created routines were evaluated in terms of performance and applicability for large-scale tasks. Numerical experiments were carried out to validate new algorithms in small-dimensional problems that were solved by known methods in MATLAB. Next, the approach was tested on tasks with a large number of unknowns and taking into account the parallelization of individual operations. To avoid finding the inverse matrix, a modified Lanczos method was programmatically implemented. We examined the formats for storing matrices in RAM: triplets, CSR, CSC, Skyline. To solve a system of linear algebraic equations (SLAE), an iterative symmetric LQ method adapted to these storage formats was used.

Results. New calculation modules integrated into the class library of the ACELAN-COMPOS complex were developed. Calculations were carried out to determine the applicability of various formats for storing sparse matrices in RAM and various methods for implementing operations with sparse matrices. The structure of stiffness matrices constructed for the same task, but with different renumbering of nodes of a finite element grid, was graphically visualized. In relation to the problem of the theory of electroelasticity, data on the time required to perform basic operations with stiffness matrices in various storage formats were summarized and presented in the form of a table. It has been established that the renumbering of grid nodes gives a significant increase in performance even without changing the internal structure of the matrix in memory. Taking into account the objectives of the study, the advantages and weaknesses of known matrix storage formats were named. Thus, CSR was optimal when multiplying a matrix by a vector, SKS was optimal when inverting a matrix. In problems with the number of unknowns of the order of 10^3, iterative methods for solving a generalized eigenvalue problem won in speed. The performance of the software implementation of the Lanczos method was evaluated. The

contribution of all operations to the total solution time was measured. It has been found that the operation of solving SLAE takes up to 95% of the total time of the algorithm. When solving the SLAE by symmetric LQ method, the greatest computational costs were needed to multiply the matrix by a vector. To increase the performance of the algorithm, parallelization with shared memory was resorted to. When using eight threads, the performance gain increased by 40–50%.

Discussion and Conclusion. The software modules obtained as part of the scientific work were implemented in the ACELAN-COMPOS package. Their performance for model problems with quasi-regular finite element grids was estimated. Taking into account the features of the structures of the stiffness and mass matrices obtained through solving the generalized eigenvalue problem for an electroelastic body, the preferred methods for their processing were determined.

Keywords: piezoelectric materials, finite element method, sparse matrices, generalized eigenvalue problem, Lanczos method, Krylov subspace, preprocessing module, postprocessing module, triplets, coordinate storage format, compressed sparse row, CSR, compressed sparse column, CSC, Skyline

Acknowledgements: the authors would like to thank AN Solovyov and TS Martynova for their assistance in the development of numerical methods. Thanks are extended to the Russian Scientific Foundation for financial support of the research with grant no. 22–21–00318, https://rscf.ru/project/22-21-00318

For citation. Oganesyan PA, Shtein OO. Implementation of Basic Operations for Sparse Matrices when Solving a Generalized Eigenvalue Problem in the ACELAN-COMPOS Complex. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):121–129. https://doi.org/10.23947/2687-1653-2023-23-2-121-129

Научная статья

Реализация базовых операций для разреженных матриц при решении обобщенной задачи на собственные значения в комплексе ACELAN-COMPOS

П.А. Оганесян ✉, **О.О. Штейн**

Южный федеральный университет, г. Ростов-на-Дону, Российская Федерация

✉ poganesyan@sfedu.ru

Аннотация

Введение. Широкое использование пьезоматериалов в различных отраслях стимулирует изучение их физических характеристик и обусловливает актуальность таких изысканий. В рассматриваемом случае модальный анализ позволяет определить рабочую частоту и коэффициент электромеханической связи пьезоэлементов различных устройств. Эти индикаторы представляют серьезный теоретический и прикладной интерес. Цель исследования — разработка численных методов для решения задачи определения частот резонанса в системе упругих тел. Для достижения цели нужны новые подходы к дискретизации задачи на основе метода конечных элементов и выполнение программной реализации выбранного метода на языке C# на платформе .net. Актуальные решения созданы в контексте библиотеки классов комплекса ACELAN-COMPOS. Основанные на обращении матриц известные методы решения обобщенной задачи на собственные значения неприменимы к матрицам большой размерности. Для преодоления этого ограничения в представленной научной работе реализована логика построения матриц масс и созданы программные интерфейсы для обмена данными о задачах на собственные значения с модулями пре- и постпроцессинга.

Материалы и методы. Для реализации численных методов задействовали платформу .net и язык программирования C#. Валидация результатов исследования проводилась путем сравнения найденных значений с решениями, полученными в известных CAE-пакетах (англ. computer-aided engineering — компьютеризированная инженерия). Созданные подпрограммы оценивались с точки зрения производительности и применимости для задач большой размерности. Проводились численные эксперименты с целью валидации новых алгоритмов в задачах малой размерности, которые решаются известными методами в MATLAB. Далее подход тестировали на задачах с большим числом неизвестных и с учетом распараллеливания отдельных операций. Чтобы избежать нахождения обратной матрицы, программно реализовали модифицированный метод Ланцоша. Рассмотрели форматы хранения матриц в оперативной памяти: триплеты, CSR, CSC, SKyline. Для

решения системы линейных алгебраических уравнений (СЛАУ) задействовали итерационный симметричный метод LQ, адаптированный к этим форматам хранения.

Результаты исследования. Разработаны новые расчетные модули, интегрированные в библиотеку классов комплекса ACELAN-COMPOS. Проведены расчеты для определения применимости различных форматов хранения разреженных матриц в оперативной памяти и различных методов реализации операций с разреженными матрицами. Графически визуализирована структура матриц жесткости, построенных для одной и той же задачи, но с различной перенумерацией узлов конечноэлементной сетки. Применительно к задаче теории электроупругости обобщены и представлены в виде таблицы данные о времени, необходимом на выполнение базовых операций с матрицами жесткости в различных форматах хранения. Установлено, что перенумерация узлов сетки дает существенный прирост производительности даже без изменения внутренней структуры матрицы в памяти. С учетом поставленных задач исследования названы преимущества и слабые стороны известных форматов хранения матриц. Так, CSR оптимален при умножении матрицы на вектор, SKS — при обращении матрицы. В задачах с числом неизвестных порядка 10^3 выигрывают в скорости итерационные методы решения обобщенной задачи на собственные значения. Оценивалась производительность программной реализации метода Ланцоша. Измерялся вклад всех операций в общее время решения. Выяснилось, что операция решения СЛАУ занимает до 95 % от общего времени работы алгоритма. При решении СЛАУ симметричным методом LQ наибольшие вычислительные затраты нужны для умножения матрицы на вектор. Для увеличения производительности алгоритма прибегли к распараллеливанию с общей памятью. При использовании восьми потоков производительность выросла на 40–50 %.

Обсуждение и заключение. Полученные в рамках научной работы программные модули были внедрены в пакет ACELAN-COMPOS. Оценена их производительность для модельных задач с квазирегулярными конечноэлементными сетками. С учетом особенностей структур матриц жесткости и масс, получаемых при решении обобщенной задачи на собственные значения для электроупругого тела, определены предпочтительные методы для их обработки.

Ключевые слова: пьезоматериалы, метод конечных элементов, разреженные матрицы, обобщенная задача на собственные значения, метод Ланцоша, подпространство Крылова, модуль препроцессинга, модуль постпроцессинга, триплеты, координатный формат хранения, сжатый разреженный ряд, CSR, сжатый разреженный столбец, CSC, SKyline

Благодарности: авторы выражают благодарность А.Н. Соловьеву и Т.С. Мартыновой за помощь в разработке численных методов и Российскому научному фонду за финансовую поддержку исследования грантом № 22–21–00318, https://rscf.ru/project/22-21-00318

Для цитирования. Оганесян П.А., Штейн О.О. Реализация базовых операций для разреженных матриц в контексте решения обобщенной задачи на собственные значения в комплексе ACELAN-COMPOS. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):121–129. https://doi.org/10.23947/2687-1653-2023-23-2-121-129

Introduction. Devices made of piezoelectric materials have been widely used, actively studied and improved for a long time. Medical ultrasound devices (diagnostic equipment, ultrasonic scalpels) [1–4] and mobile energy generators [5] should be noted separately. Paper [6] described the combination of photo- and piezoelectric effects to create efficient compact energy sources. New materials designed for the application under specific conditions are being studied in science and industry. In [7], the creation of a lead-free piezo-active composition suitable for operation at various temperatures was considered.

In the study on piezoelectric elements, a key role is played by the modal analysis stage, which enables to establish the resonance and antiresonance frequencies of the device. These data:

– are needed to find out the operating frequency of the device;

– provide determining the electromechanical coupling factor — an important performance indicator of the device;

– are input information in numerical experiments for problems on forced oscillations.

The study was aimed at the creation of numerical methods for solving the problem of determining resonance frequencies for a system of elastic bodies. Achieving the stated goal required solving two tasks. The first was to develop

methods of discretization of the problem based on the finite element method (FEM). The second was to carry out a software implementation of the selected method in C# on the .net platform. All known programs take into account the context of the ACELAN-COMPOS class library [8]. When solving a generalized eigenvalue problem, methods based on matrix inversion are widely used. However, they are not applicable to large-dimensional matrices. In the presented research, this limitation was overcome as follows:

– the logic of constructing mass matrices was additionally implemented;

– software interfaces were created for exchanging data on eigenvalue tasks with pre- and postprocessing modules.

Materials and Methods. Principally, the proposed approach was designed to solve static problems of electroelasticity when implementing the averaging method [9], which was used to calculate the effective properties of piezo composites. In this regard, only stiffness matrices were presented at the stage of constructing global FEM matrices. In this study, we additionally implemented the logic of constructing mass matrices and developed software interfaces (application programming interface, API) for exchanging data on eigenvalue tasks with pre- and postprocessing modules. The developed routines were evaluated in terms of performance and applicability for large-scale tasks. Numerical experiments were carried out to validate the algorithms created for such small-dimensional problems that provide obtaining a solution by general methods in the MATLAB computing package. Next, testing was performed on tasks with a large number of unknowns and taking into account the parallelization of individual operations.

The mathematical model of the problem being solved consists of the defining relations [9]:

$$\rho_{\pi\kappa}\omega^2 u + \alpha_{\delta\varphi}\rho_{\varphi} i\omega u - \nabla\cdot\sigma = \phi_{\varphi}\,,\ \nabla\cdot D = 0\,, \tag{1}$$

$$\sigma = c_j^E\cdot(\varepsilon+\beta_{\delta\varphi}\dot{\varepsilon}) - e_j^T\cdot E\,,\ D+\varsigma_d\dot{D} = e_j\cdot(\varepsilon+\varsigma_d\dot{\varepsilon}) + э_j^S\cdot E\,, \tag{2}$$

$$\varepsilon = (\nabla u + \nabla u^T)/2\,,\ E = -\nabla\phi. \tag{3}$$

Here, σ — stress tensor; ρ_j — body density; ε — strain tensor; u — displacement vector; D — electric displacement vector; E — electric-field vector; f_j — body force vector; ϕ — electric potential; α_{dj}, β_{dj}, ς_d — damping coefficients; c_j^E, e_j^T, $э_j^S$ — tensors of elastic constants, piezoelectric modules and dielectric permittivity; index j — body number in the model.

The discretization is performed by replacing:

$$u(x,t) = N_u^T(x)\cdot U(t)\,,\ \ \phi(x,t) = N_\phi^T(x)\cdot\Phi(t)\,.$$

Here, N_u — shape function matrix for the displacement field; N_ϕ — shape function vector for electric potential; $U(t)$, $\Phi(t)$ — global vectors of the corresponding nodal degrees of freedom.

In this case, the original problem (1–3) takes the form:

$$M\cdot\ddot{a} + K\cdot a = F\,. \tag{4}$$

Here, matrices M and K are global matrices of mass and stiffness, respectively, and the vector is a general vector of unknowns:

$$a = [U, \Phi].$$

In the problem of the theory of electroelasticity:

$$M = \begin{pmatrix} M_{uu} & 0 \\ 0 & 0 \end{pmatrix},\ \ K = \begin{pmatrix} K_{uu} & K_{u\phi} \\ K_{u\phi}^T & -K_{\phi\phi} \end{pmatrix}. \tag{5}$$

Matrices M_{uu}, K_{uu} and $K_{\phi\phi}$ — symmetric. In the case of harmonic oscillations at natural frequency ω_i, it is possible to write:

$$a = v_i\sin(\omega_i t)\,,$$

denoting the corresponding eigenvector by v_i.

Consider free oscillations if $F = 0$. In this case, task (4) is represented as:

$$-\omega_i^2 M v_i + K\cdot v_i = 0\,. \tag{6}$$

Thus, the original problem is reduced to a generalized eigenvalue problem (6). For nonzero v_i, inequality (6) is solved by finding the matrix inverse to K. However, at the same time, the sparse matrix becomes full, i.e., the method is

unsuitable for large matrices. Therefore, it is needed to use other methods that do not require finding the inverse matrix. To solve this problem, a modified Lanczos method was programmatically implemented in this paper [10]. The author of this modification is T. S. Martynova. The description of the development is not given in this article. Of the operations used in the method, the most expensive from the point of view of computational resources was the solution to a system of linear algebraic equations (SLAE), needed for performing a spectral transformation.

Matrices M and K — sparse, with a small number of nonzero elements. Several formats are used to store such matrices in RAM:

– triplets, or coordinate format;

– CSR (compressed sparse row);

– CSC (compressed sparse column);

– Skyline storage format (SKS method).

The coordinate format involves storing triples (triplets) of values (i, j, k), representing coordinates (i, j) and values (k) of nonzero elements. CSR is sometimes referred to as CRS or Yale format. It involves storing a sparse matrix in the form of three arrays. Consider matrix of N size with NZ nonzero elements. We describe the possible organization of its storage. All nonzero elements must be placed in one array of NZ size. The positions of these elements in the columns should be placed in another array of NZ size, and the third array of N size should be used to store the indices of the first elements of the rows. Similarly, the storage in CSV format is implemented.

The SKS format assumes the storage of a variable-width matrix band that includes all nonzero elements. In this case, zeros are allowed. The efficiency of this format depends on the renumbering of the matrix rows. Methods for reducing the size of the tape are described in [11]; however, their applicability to the stiffness matrix obtained when solving a three-dimensional problem using FEM requires a separate study.

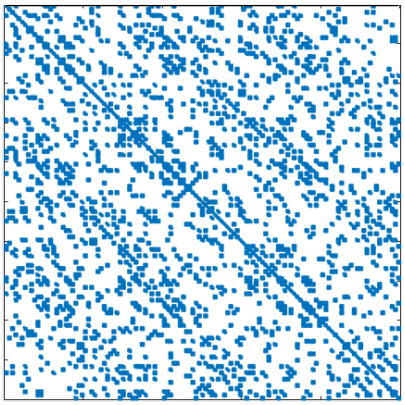
To solve the SLAE (system of linear algebraic equations), an iterative symmetric LQ method (SYMLQ [12]) adapted to the storage formats listed above was used.

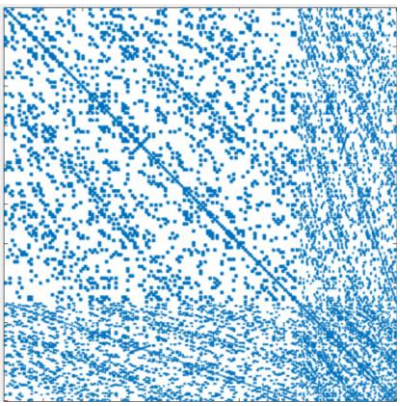
Research Results. At the beginning of the study, we chose the optimal storage format for sparse matrices. The coordinate format enabled to quickly add and change an element of the matrix. These operations were needed at the stage of assembling the global matrix and taking into account the boundary conditions. In addition, for ill-conditioned matrices, to which K refers, a preliminary transformation is often used for normalization. It is also convenient to perform it in the coordinate format. However, this format is ineffective when it comes to algebraic operations.

CSR is ill-adapted for changing the structure of the matrix: by adding a nonzero element, you need to insert into two arrays. In this case, the matrix is multiplied by a vector quite easily and efficiently.

SKS has similar problems with the addition of nonzero elements and is highly dependent on the renumbering of unknowns in the problem. We focus on the example of a quasi-regular grid, which is used in the ACELAN-COMPOS package to work with representative volumes of composites. The width of the band containing all nonzero elements can be predetermined and depends on the number of nodes and the type of final element. In the general case of an arbitrary finite element grid, it is difficult to estimate the size of the band in advance.

Four methods of numbering unknowns were used in numerical experiments. Figure 1 shows the structure of stiffness matrices constructed for the same task, but with different renumbering of nodes of a finite element grid.





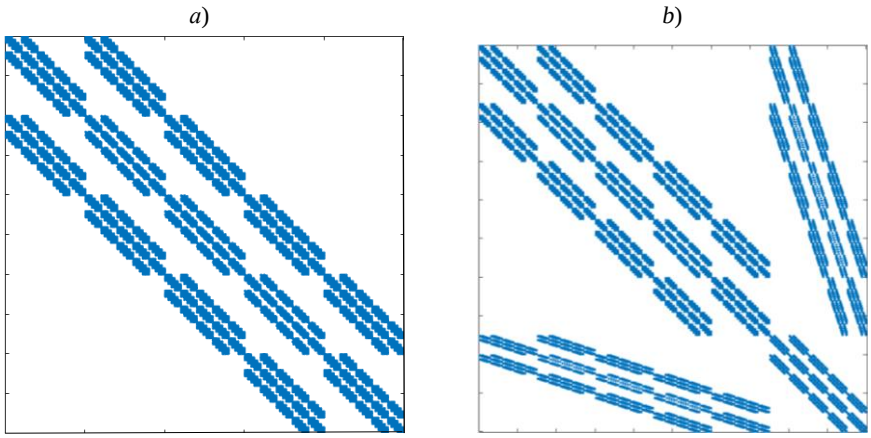


Fig. 1. Stiffness matrix structure with various methods of node numbering:
a — unknowns are ordered by nodes; *b* — first, displacement nodes are ordered,
then potentials; *c* — nodes are sorted by layers of the FE grid, and unknowns — by nodes;
b, *d* — nodes are sorted by layers of the FE grid, and unknowns — as in example

Thus, the grid was a cube with regular partitioning by eight-node finite elements. A model matrix of 500 lines was used for the illustrations. The matrices shown in Figures 1 *a* and 1 *b*, were not subjected to additional renumbering of nodes and differed only in the numbering of degrees of freedom. In 1 *a*:

$$a = [u_{1x}, u_{1y}, u_{1z}, \varphi_1, ..., u_{Nx}, u_{Ny}, u_{Nz}, \varphi_N].$$

In 1 *b*, the unknowns responsible for the potential distribution were collected at the end of the vector:

$$a = [u_{1x}, u_{1y}, u_{1z}, ..., u_{Nx}, u_{Ny}, u_{Nz}, \varphi_1, ..., \varphi_N].$$

The unknowns in matrices 1 *c* and 1 *d* were numbered similarly, but the nodes of the finite element grid were pre-numbered according to their coordinates through alternately sorting all nodes by each of the coordinates. This technique is widely used to build more efficient SLAE solution modules, as it enables to work with the matrix in a suitable band format, convenient for parallelization. Similar external modules were implemented for the ACELAN-COMPOS complex [13–15]; however, in this work, only formats for storing sparse matrices of a general form were used.

Table 1 summarizes the data on the time required to perform basic operations with matrices in various formats.

Table 1

Time to perform basic operations with the stiffness matrix in the problem of the theory of electroelasticity.
19,652 rows

Storage format	Operation	Elapsed time, ms			
		1 *a*	1 *b*	1 *c*	1 *d*
CSR	Conversion from coordinate format	123	132	97	117
CSR	Multiplication by vector, 100 operations	260	260	260	260
SKS	Conversion from coordinate format	690	703	124	268
SKS	Multiplication by vector, 100 operations	60,558	61,450	7,616	22,113

The experimental results showed that the conversion operation from a coordinate storage format to a compact one took little time. At the same time, the renumbering of grid nodes to form a block-tape matrix made it possible to get a noticeable increase in performance even without changing the internal structure of the matrix in memory. CSR format turned out to be optimal in terms of the efficiency of the matrix-vector multiplication operation. When the matrix was inverted, SKS format was more efficient, but for problems with the number of unknowns of the order of 103, iterative methods for solving a generalized eigenvalue problem worked noticeably faster.

Further, the performance of the software implementation of the Lanczos method was experimentally evaluated. The contribution of all operations to the total solution time was measured. As a result, it was found that the operation of solving SLAE took up to 95% of the total time of the algorithm. In the course of the algorithm, a Krylov subspace was

constructed, and depending on its dimension, the number of SLAE that needed to be solved changed. Note that the dimension of the Krylov subspace was chosen based on heuristics with respect to the number of desired eigenvalues. Here, the SLAE differed only in the right part, so that the requirements for allocated memory remained low. Among the basic operations used in the course of solving SLAE by the SYMLQ method, the greatest computational costs were needed for multiplying the matrix by the vector.

To increase the performance of the algorithm, the simplest parallelization with shared memory was implemented. Blocks of rows were allocated for the CRS format. They were transmitted to separate threads that calculated the corresponding components of the resulting vector. The performance gain was 40–50% when using 8 threads. At the same time, for matrices of the order of 10^3 elements, the increase was about 40%, and for matrices of the order of 10^4 — about 50%.

Discussion and Conclusions. Within the framework of this study, a method for solving a generalized eigenvalue problem for matrices obtained by modeling electroelastic bodies was implemented. Software modules were created in C# for constructing mass matrices by the finite element method and performing auxiliary operations within the framework of the Lanczos method (working with Krylov subspace vectors, reorthogonalization, finding eigenvectors). The computational complexity was mainly due to the operations of multiplying sparse matrices by a vector. In this regard, numerical experiments were carried out to determine the optimal formats for storing matrices, the optimal structure of the matrix obtained as a result of renumbering the nodes of the FE grid and degrees of freedom in the nodes. A version of the SYMMLQ iterative algorithm using parallel computing was developed. The final scheme of work included three points. First, global matrices were constructed in coordinate format with a renumbering algorithm (Fig. 1 *c*). Secondly, the data was converted to CRS format. Thirdly, the data was processed by the Lanczos method, which included the SYMLQ method for solving SLAE. The results of the work were included in the ACELAN-COMPOS software package.

References

1. Lisong Deng, Mingxiang Ling. Design and Integrated Stroke Sensing of a High-Response Piezoelectric Direct Drive Valve Enhanced by Push-Pull Compliant Mechanisms. *Review of Scientific Instruments.* 2022;93(3):035008. https://doi.org/10.1063/5.0067483

2. Urtnasan Erdenebayar, Jong-Uk Park, Pilsoo Jeong, et al. Obstructive Sleep Apnea Screening Using a Piezo-Electric Sensor. *Journal of Korean Medical Science*. 2017;32(6):893–899. https://doi.org/10.3346/jkms.2017.32.6.893

3. Skaliukh AS, Gerasimenko TE, Oganesyan PA, Solovieva AA. Effect of Geometric and Physical Parameters on Resonant Frequencies of Ultrasonic Vibrations of Elastic and Piezoelectric Element System. *Vestnik of Don State Technical University*. 2017;17(4):5–13. https://doi.org/10.23947/1992-5980-2017-17-4-5-13

4. Bulletti A, Capineri L, Floridia D. Automatic System to Measure the Impedance of Piezoelectric Actuators Used in Ultrasonic Scalpels. In book: *Sensors and Microsystems.* Cham: Springer; 2014. Vol. 268. P. 71–74. https://doi.org/10.1007/978-3-319-00684-0_14

5. Keli Li, Qisheng He, Jiachou Wang, et al. Wearable Energy Harvesters Generating Electricity from Low-Frequency Human Limb Movement. *Microsystems & Nanoengineering*. 2018;4:24. https://doi.org/10.1038/s41378-018-0024-3

6. Wenbo Peng, Chenhong Wang, Fangpei Li, et al. Piezo- and Photo-Voltage Field-Effect Transistor. *Nano Energy*. 2022;105:108025. https://doi.org/10.1016/j.nanoen.2022.108025

7. Tangyuan Li, Chang Liu, Peng Shi, et al. High-Performance Strain of Lead-Free Relaxor-Ferroelectric Piezoceramics by the Morphotropic Phase Boundary Modification. *Advanced Functional Materials*. 2022;32(32):2202307. https://doi.org/10.1002/adfm.202270184

8. Kurbatova NV, Nadolin DK, Nasedkin AV, et al. Finite Element Approach for Composite Magneto-Piezoelectric Materials Modeling in ACELAN-COMPOS Package. In book: *Analysis and Modelling of Advanced Structures and Smart Systems*. Singapore: Springer; 2018. Vol. 81. P. 69–88. https://doi.org/10.1007/978-981-10-6895-9_5

9. Belokon AV, Nasedkin AV, Soloviev AN. New Diagrams of Finite Element Dynamic Analysis of Piezoelectric Devices. *Journal of Applied Mathematics and Mechanics*. 2002;66(3):491–501. (In Russ.)

10. Zhongming Teng, Lei-Hong Zhang. A Block Lanczos Method for the Linear Response Eigenvalue Problem. *Electronic Transactions on Numerical Analysis*. 2017;46:505–523. https://doi.org/10.13140/RG.2.2.16369.68962

11. Chagas G, Oliveira SLGD. Metaheuristic-Based Heuristics for Symmetric-Matrix Bandwidth Reduction: A Systematic Review. *Procedia Computer Science*. 2015;51:211–220. https://doi.org/10.1016/j.procs.2015.05.229

12. Paige CC, Saunders MA. Solution of Sparse Indefinite Systems of Linear Equations. *SIAM Journal on Numerical Analysis.* 1975;12(4):617–629. https://doi.org/10.1137/0712047

13. Fassbender H, Ikramov K. SYMMLQ-like Procedure of *Ax* = *b* where *A* is a Special Normal Matrix. *Calcolo*. 2006;43(1):17–37. https://doi.org/10.1007/s10092-006-0112-x

14. Vasilenko A, Veselovskiy V, Metelitsa E, et al. *Precompiler for the ACELAN-COMPOS Package Solvers*. In: Proc. 16th Int. Conf.: Parallel Computing Technologies. Cham: Springer; 2021. Vol. 12942. P. 103–116. https://doi.org/10.1007/978-3-030-86359-3_8

15. Steinberg BYa, Vasilenko AA, Veselovskiy VV, et al. Solvers for Systems of Linear Algebraic Equations with Block-Band Matrices. *Bulletin of the South Ural State University. Series "Mathematical Modeling, Programming & Computer Software".* 2021;14(3):106–112. https://doi.org/10.14529/mmp210309

Received 03.04.2023
Revised 28.04.2023
Accepted 04.05.2023

About the Authors:

Pavel A. Oganesyan, Cand.Sci. (Phys.-Math.), Associate Professor of the Mathematical Modeling Department, Vorovich Institute for Mathematics, Mechanics, and Computer Science, Southern Federal University (105/42, Bolshaya Sadovaya St., Rostov-on-Don, 344006, RF), ResearcherID, ScopusID, ORCID, AuthorID, poganesyan@sfedu.ru

Olga O. Shtein, Teaching assistant of the Applied Mathematics and Programming Department, Vorovich Institute for Mathematics, Mechanics, and Computer Science, Southern Federal University (105/42, Bolshaya Sadovaya St., Rostov-on-Don, 344006, RF), ORCID, shteyn@sfedu.ru

Claimed contributorship:

PA Oganesyan: basic concept formulation; research objectives and tasks; software implementation; text preparation; formulation of conclusions.

OO Shtein: software implementation; performing of numerical experiments; preparation of illustrations.

Conflict of interest statement: the authors do not have any conflict of interest.

All authors have read and approved the final manuscript.

Поступила в редакцию 03.04.2023
Поступила после рецензирования 28.04.2023
Принята к публикации 04.05.2023

Об авторах:

Павел Артурович Оганесян, кандидат физико-математических наук, доцент кафедры математического моделирования Института математики, механики и компьютерных наук «Южный федеральный университет» (344006, РФ, г. Ростов-на-Дону, ул. Большая Садовая, 105/42), ResearcherID, ScopusID, ORCID, AuthorID, poganesyan@sfedu.ru

Ольга Олеговна Штейн, ассистент кафедры прикладной математики и программирования Института математики, механики и компьютерных наук «Южный федеральный университет» (344006, РФ, г. Ростов-на-Дону, ул. Большая Садовая, 105/42), ORCID, shteyn@sfedu.ru

Заявленный вклад соавторов:

П.А. Оганесян — формирование основной концепции, цели и задачи исследования, программная реализация, подготовка текста, формулирование выводов.

О.О. Штейн — программная реализация, проведение численных экспериментов, подготовка иллюстраций.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Все авторы прочитали и одобрили окончательный вариант рукописи.

MACHINE BUILDING AND MACHINE SCIENCE
МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ

UDC 62-52
https://doi.org/10.23947/2687-1653-2023-23-2-130-139

Original article

Surface Quality Forming under Parts Finishing and Strengthening Treatment with an Eccentric Hardener

Mikhail A. Tamarkin, Elina E. Tishchenko ✉, Omar C.A. Hashash
Don State Technical University, Rostov-on-Don, Russian Federation
✉ lina_tishenko@mail.ru

Abstract

Introduction. The formation of the quality parameters of the surface layer and the operational properties of the parts occurs throughout all stages of their manufacture. However, the decisive impact is most often exerted by the stages of finishing. Therefore, in modern digital engineering, the task of process support of high quality of the surface layer of the part is one of the challenges in solving the problem of improving the quality and reliability and increasing the life cycle of manufactured machines. Surface plastic deformation treatment is instrumental in improving the performance characteristics of machine parts. Its essence is that the required quality parameters of parts are obtained not by removing a layer of material, but by plastic deformation. During the processing, both the dimensions of the parts and the physical and mechanical properties of the surface layers are changed. In this case, the technologist has the opportunity to significantly increase the life cycle of the manufactured products through controlling the process. These studies are aimed at providing the required quality parameters of the surface layer under processing with an eccentric hardener.

Materials and Methods. The article presents the results of research on a new method of surface plastic deformation treatment — with an oscillating eccentric hardener. The considered processing method enables to obtain high quality of the treated surface, to process large-sized parts in places that are stress concentrators, to process welds, small areas of surfaces, whose hardening is needed for the part to fulfill its intended service. A set of theoretical studies was carried out; their results provided determining the parameters of a single interaction of the indenter and the surface of the part, the diameter of the plastic imprint and its depth.

Results. Dependences for determining the surface roughness, the depth of the hardened layer and the degree of deformation were obtained. The resulting formulas were tested for adequacy by experimental studies.

Discussion and Conclusion. The obtained research results can be used in the technological design of surface plastic deformation treatment processes. Further tasks for the study of the considered processing method are determined.

Keywords: oscillating tool, eccentric hardener, surface roughness, hardened layer depth, degree of deformation

Acknowledgements: the authors would like to thank the editorial board of the journal and the reviewers for their attentive attitude to the article.

For citation. Tamarkin MA, Tishchenko EE, Hashash O. Surface Quality Forming under Parts Finishing and Strengthening Treatment with an Eccentric Hardener. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):130–139. https://doi.org/10.23947/2687-1653-2023-23-2-130-139

Научная статья

Формирование качества поверхностного слоя при отделочно-упрочняющей обработке деталей эксцентриковым упрочнителем

М.А. Тамаркин, Э.Э. Тищенко ✉, Омар Хашаш
Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация
✉ lina_tishenko@mail.ru

Аннотация

Введение. Формирование параметров качества поверхностного слоя и эксплуатационных свойств деталей происходит на протяжении всех этапов их изготовления. Однако решающее влияние чаще всего оказывают этапы финишной обработки. Поэтому в современном цифровом машиностроении задача технологического обеспечения высокого качества поверхностного слоя детали является одной из важнейших при решении проблемы повышения качества, надежности и увеличения жизненного цикла производимых машин. Ведущую роль в повышении эксплуатационных характеристик деталей машин играет обработка поверхностным пластическим деформированием, сущность которой заключается в том, что требуемые параметры качества деталей достигаются не удалением слоя материала, а его пластическим деформированием. В процессе обработки производится изменение как размеров деталей, так и физико-механических характеристик поверхностных слоев, управляя которыми технолог имеет возможность значительно увеличивать жизненный цикл производимой продукции. Целью настоящих исследований является обеспечение необходимых параметров качества поверхностного слоя при обработке эксцентриковым упрочнителем.

Материалы и методы. В статье представлены результаты исследований нового метода обработки поверхностным пластическим деформированием — осциллирующим эксцентриковым упрочнителем. Рассматриваемый метод обработки позволяет получать высокое качество обработанной поверхности, осуществлять обработку крупногабаритных деталей в местах, являющихся концентраторами напряжений, обрабатывать сварные швы, небольшие участки поверхностей, упрочнение которых необходимо для выполнения деталью своего служебного назначения. Выполнен комплекс теоретических исследований, по результатам которых определены параметры единичного взаимодействия индентора с поверхностью детали, диаметр пластического отпечатка и его глубина.

Результаты исследования. Получены зависимости для определения шероховатости поверхности, глубины упрочненного слоя и степени деформации. Полученные формулы прошли проверку адекватности экспериментальными исследованиями.

Обсуждение и заключение. Полученные результаты исследований могут быть использованы при технологическом проектировании процессов обработки поверхностным пластическим деформированием. Определены дальнейшие задачи по исследованию рассматриваемого метода обработки.

Ключевые слова: осциллирующий инструмент, эксцентриковый упрочнитель, шероховатость поверхности, глубина упрочненного слоя, степень деформации

Благодарности: авторы выражают благодарность редакции журнала и рецензентам за внимательное отношение к статье.

Для цитирования. Тамаркин М.А., Тамаркин Э.Э., Хашаш О. Формирование качества поверхностного слоя при отделочно-упрочняющей обработке деталей эксцентриковым упрочнителем. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):130–139. https://doi.org/10.23947/2687-1653-2023-23-2-130-139

Introduction. The reliability and durability of machine parts largely depend on the quality of their surface layer. From numerous works on mechanical engineering technology, it is known that the formation of the quality parameters of the surface layer occurs at all stages of their manufacture. However, the dramatic impact is most often exerted by the stages of finishing. Therefore, in modern digital engineering, increased attention is paid to the production design of highly efficient finishing operations of parts, which enables to respond to a challenge of increasing their life cycle. Surface plastic deformation (SPD) treatment is critical in improving the performance of machine parts carried out under finishing operations. In contrast to traditional cutting methods, the quality parameters of the surface layer in SPD are obtained

through performing plastic deformation with special tools or working media. In the course of processing, simultaneously with the change in the size of the processed parts, the physical and mechanical properties of the surface layers are changed. In this case, the technologist has the opportunity to significantly increase the life cycle of the manufactured products through controlling the process.

It should be noted that the widespread use of many SPD methods in industry is hindered by poor knowledge of their basic laws, difficulties arising in the process of designing optimal combinations of processing modes, and design parameters of the tooling. On numerous occasions, treatment modes are assigned on the assumption of the results of private experimental studies, which provides low processing efficiency [1–6].

The objective of these studies is to provide the required quality parameters of the surface layer during processing with an eccentric hardener.

Materials and Methods. The need for applying SPD under the conditions of modern machine-building industries results in the creation of new processing methods. One of such methods is the processing of SPD with an oscillating tool — with an eccentric hardener.

Figure 1 shows a kinematic diagram of an eccentric hardener consisting of vibrating body 1 suspended on flat springs 2. Vibrations of vibrating body 1, acting normally to the treated surface, are excited by the rotation of eccentric mass (unbalance) 3 around the vertical axis. The rotation axis of the eccentric mass is restricted from moving relative to vibrating housing 1. The rotational motion is transmitted to the eccentric from electric motor 5 through flexible shaft 6. Tool head 4 with an instrument of the appropriate geometric form is attached to housing 1. The motion of tool 4 is limited by limiter 7 (a workpiece). In this case, the tool is an indenter with a spherical sharpening. It can be made in the form of a roller or a ball. The vibration system in the eccentric hardeners can be represented as a single-mass system with two degrees of freedom, which is under the action of a force varying according to the harmonic law. To study the system dynamics, we consider the features of its free oscillation under the action of centrifugal vibration excitation and the nature of the movement of the system hitting the limiter (a part).

Under free oscillation, the vibrating system fixed at the end of flat springs 2 (Fig. 1) performs harmonic oscillations, which are excited by the rotation of eccentric 3 with a constant angular velocity.

The proposed device can be effective when processing formed parts of not the most complex profile, and in some cases — when processing simple surfaces, such as planes or bodies of rotation.

First, it is required to check the possibility of providing a wide range of the energy of the impact of the tool head on the surface of the workpiece in combination with relatively low altitude characteristics of surface roughness.

Due to the lower rigidity of a flat spring (in our case, two springs) in the X direction, in comparison to the stiffness in the Y direction, the system describes a trajectory close to an ellipse with a larger semi-axis in the X direction. To analyze the law of motion of the system, we decompose the trajectories along the X and Y axes.

The equation of motion of the center of gravity is nothing more than a mathematical expression of Newton's second law.

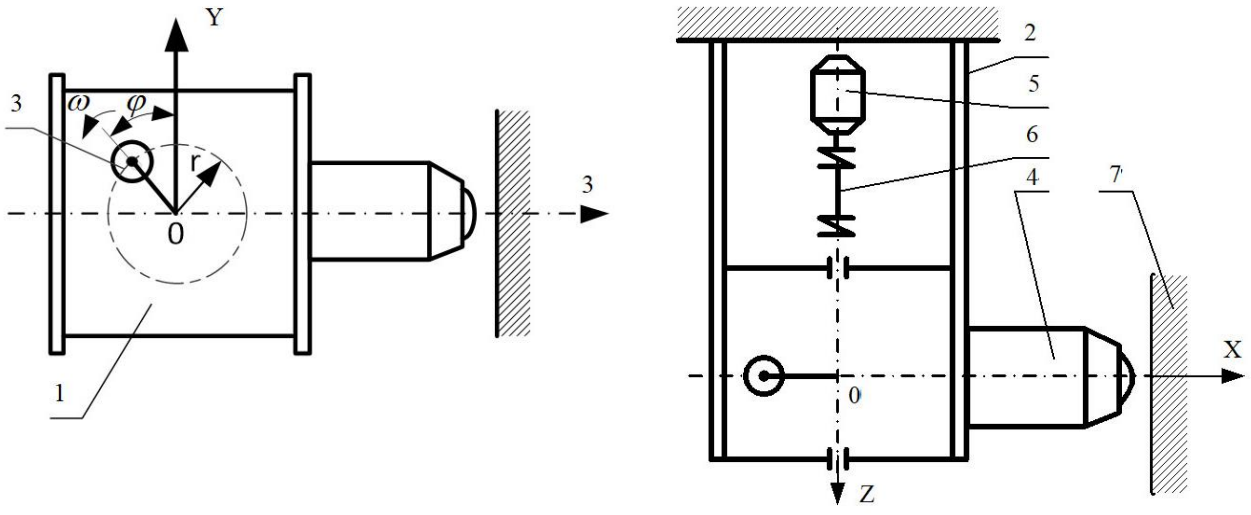


Fig. 1. Scheme of eccentric hardener: 1 — housing; 2 — flat spring; 3 — eccentric mass; 4 — tool head; 5 — electric motor; 6 — flexible shaft; 7 — limiter (workpiece)

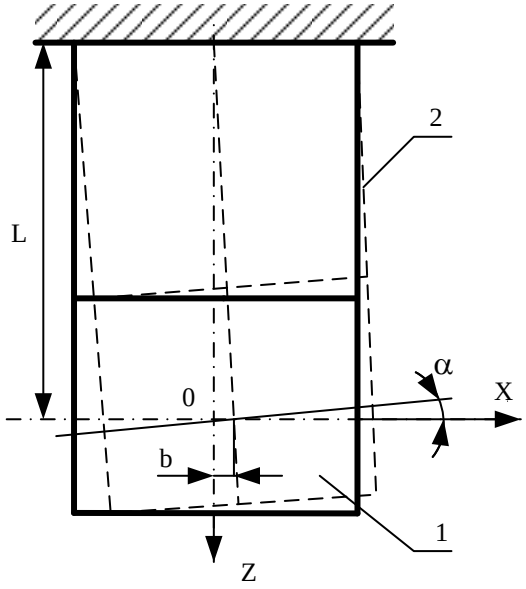


Fig. 2. Rotations of the housing of the hardener during vibrations

Research Results. To establish the main regularities of the process variables effect on the interaction of the oscillating indenter and the treated surface, it is necessary to take into account the kinetic energy of the indenter, the number of indenters, the radius of the indenter, the efficiency of the device, physical and mechanical properties of the workpiece [1–6].

Taking into account all the forces acting on the moving indenter, it is possible to write the equation of motion of the tool head in the Y direction:

$$m_c \frac{d^2y}{dt^2} = -c_1 y - \mu \frac{dy}{dt} + m_{cam} r\omega^2 \cos\omega t - m_c g, \quad (1)$$

in the X direction:

$$m_c \frac{d^2x}{dt^2} = -cx - \mu \frac{dx}{dt} + m_{cam} r\omega^2 \sin\omega t, \quad (2)$$

where $m_c \frac{d^2y}{dt^2}$, $m_c \frac{d^2x}{dt^2}$ — projections of the inertia forces of the system on Y and X axes, respectively; c_1, y, cx — projections of spring resistance forces on Y and X axes; $\mu \frac{dx}{dt}$, $\mu \frac{dy}{dt}$ — projections of the resistance forces of the medium on Y and X axes;; $m_{cam} r\omega^2 \cos\omega t$, $m_{cam} r\omega^2 \sin\omega t$ — projections of the perturbing force on Y and X axes; $m_c g$ — gravity (weight) of the vibrating system; m_c — mass of the vibrating system; c_1, c — spring stiffness in the Y and X direction; μ — resistance of the medium; m_{cam} — mass of the eccentric; r— distance from the axis of rotation of the eccentric to its center of gravity; ω — angular velocity; t — current time value; y, x — current coordinate value.

Due to the significantly greater stiffness of the springs in the Y direction relative to the stiffness in the X direction, the amplitude of the indenter movement in the Y direction is noticeably less than the amplitude in the X direction. Therefore, we assume that the system performs harmonic oscillations only in the X direction, i.e., we consider an indenter with only one degree of freedom.

If we neglect the damped oscillation, then the equation of motion will have the form:

$$x = b\sin(\omega t + \beta), \quad (3)$$

where b — amplitude of the oscillations; β — phase difference between the exciting force and the movements of the center of gravity of the indenter.

Substituting this expression into equation (2), we find:

$$b = \frac{m_{cam} r\omega^2}{\sqrt{(c - \omega^2 m_c)^2 + \omega^2 \mu^2}}, \quad (4)$$

$$tg\beta = \frac{\omega\mu}{c - \omega^2 m_c}. \quad (5)$$

Value μ is determined from the expressions:

$$\mu = \frac{\omega_1 m_c \delta}{2\pi}, \quad (6)$$

where ω_1 — frequency of natural oscillations; δ – logarithmic decrement of attenuation.

After differentiating equation (4) in time and examining the function at the extremum, we obtain an expression for determining the maximum speed of the indenter:

$$V_X = \frac{m_{cam} r \omega^3}{\sqrt{(c - \omega^2 m_c)^2 + \omega^2 \mu^2}}. \quad (7)$$

The largest kinetic energy of the indenter is determined from the equation:

$$T = \frac{m_c V_X^2}{2} = \frac{m_c m_{cam}^2 r^2 \omega^6}{2[(c - \omega^2 m_c)^2 + \omega^2 \mu^2]}. \quad (8)$$

The analysis of the interaction of spherical indenters and a deformable half-space (the surface layer of the workpiece) is described in the classical works of I. V. Kudryavtsev [1, 2, 4, 5]. The diameter of the plastic print can be determined from the dependence

$$d = \sqrt[4]{\frac{D_i \cdot T \cdot \eta}{M \cdot HD}}. \quad (9)$$

In this case, the depth of the plastic print can be defined as:

$$h = \frac{1}{4}\sqrt{\frac{T \cdot \eta}{M \cdot D_i \cdot HD}}, \quad (10)$$

where T — kinetic energy of the tool head; HD — dynamic hardness of the part material (ratio of the impact energy of the spherical indenter to the volume of the displaced material upon impact); D_i — diameter of the indenter; η — efficiency of the device; M — number of indenters.

When processing with an eccentric hardener, the roughness parameters of the treated surface can receive a constant (steady-state) value, which is reproduced during further processing of the surface of the part. The relief of the resulting surface can be both isotropic and anisotropic and is formed by repeatedly superimposing traces of a single interaction.

When the oscillating indenter interacts with the initial protrusions of the microasperity, its height decreases with a simultaneous decrease in the depth of the cavities of the microasperity. With increasing processing time, the initial surface roughness profile is completely reshaped. As a result, a new microrelief is formed, and it has a specific character for each SPD method [7–19].

The finally formed roughness of the treated surface is called “established”. As a rule, its altitude parameters do not depend on the initial one. They are formed under the specific conditions of each processing method and depend on its process variables. Based on the methodology of papers [3, 4], a dependence was obtained for determining the steady-state surface roughness during treatment with an eccentric hardener:

$$Ra = 0.0075\sqrt{\frac{T \cdot \eta}{D_i \cdot M \cdot HD}}. \quad (11)$$

The parameters of the surface layer hardening, which include the depth of the hardened layer and the degree of deformation, have a major impact on increasing the life cycle of the machined parts. As a result of theoretical studies, analytical dependences were obtained for their calculation during processing with an eccentric hardener:

$$h_n = \sqrt[8]{\frac{\left(\frac{T \cdot \eta}{D_i \cdot M \cdot HD}\right)^3}{D_i}}, \quad (12)$$

$$\varepsilon = 1.13\sqrt[4]{\frac{T \cdot \eta}{D_i^3 \cdot M \cdot HD}}. \quad (13)$$

The above dependences correspond to the physical meaning of the phenomena occurring under processing, and have been verified during complex experimental studies.

In the course of the experimental studies, samples from various materials often utilized for the manufacture of machine parts were used: high-quality and alloy steels (steel 45, HVG, steel 30, steel 30HGSA, etc.), aluminum alloys (AL1, AVT, D16, etc.). Flat samples were treated with an eccentric hardener under different modes. Ball and roller indenters were used.

According to the theoretical dependences, graphs of the dependences of the roughness of the treated surface, the depth of the hardened layer and the degree of deformation on the processing modes, characteristics of working media and processed materials were constructed.

In the graphs (Fig. 3–6), a solid line shows curves constructed according to theoretical formulas, and the dots show the results of experimental studies. The construction of confidence intervals with a confidence factor of 95% has been performed.

There is high convergence of the results, which indicates that the theoretical dependence reflects correctly the phenomena occurring during the processing of SPD with an oscillating tool — an eccentric hardener.

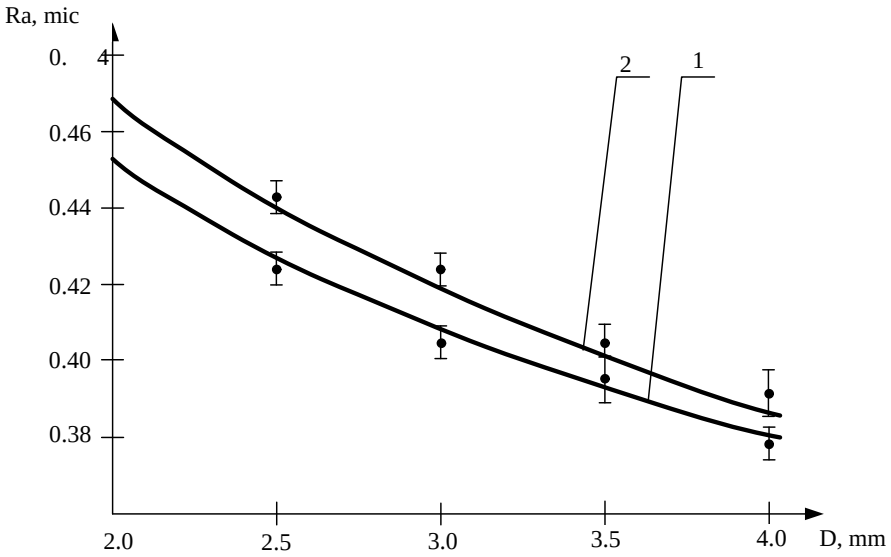


Fig. 3. Dependence of the surface roughness on the radius of the indenter:
1 — steel 45 sample material; 2 — HVG sample material

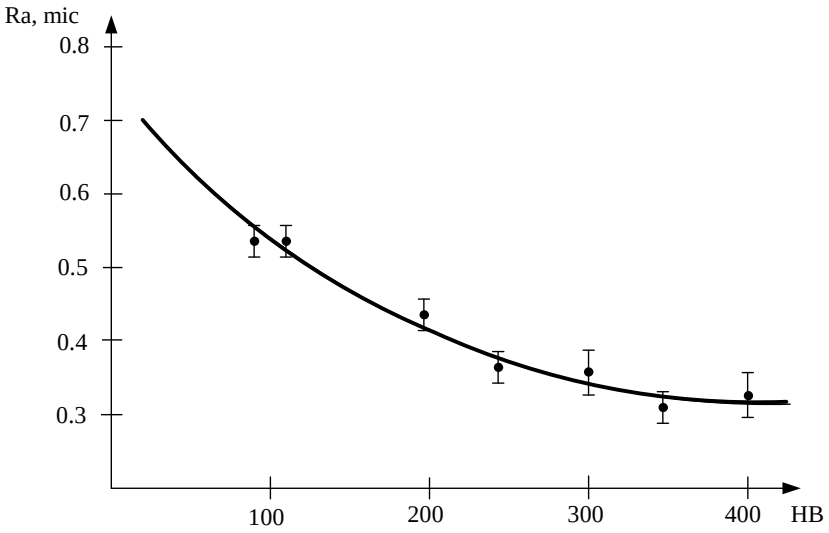
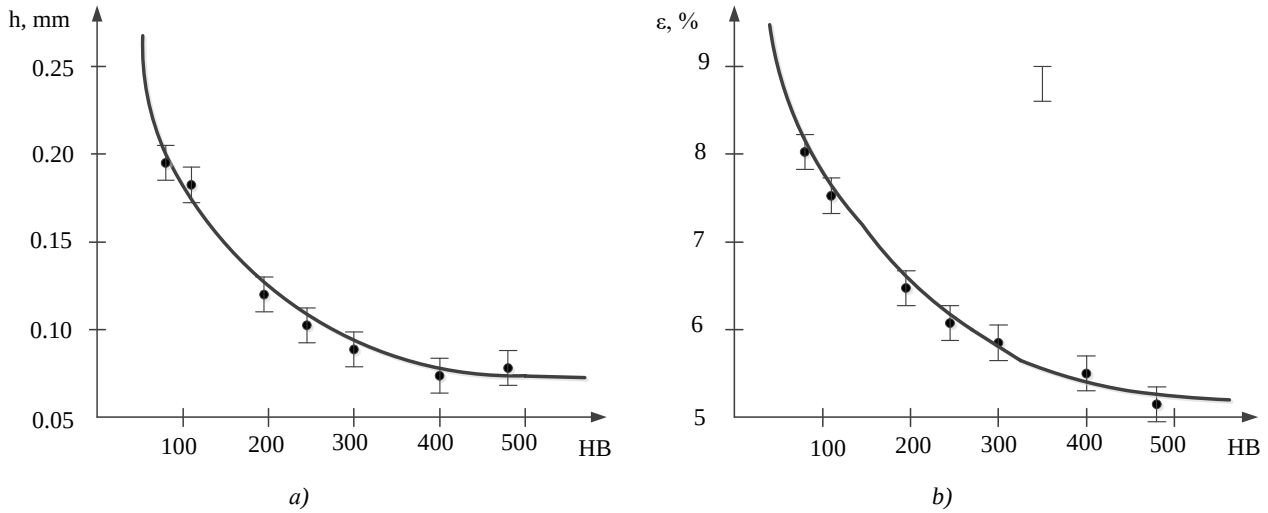
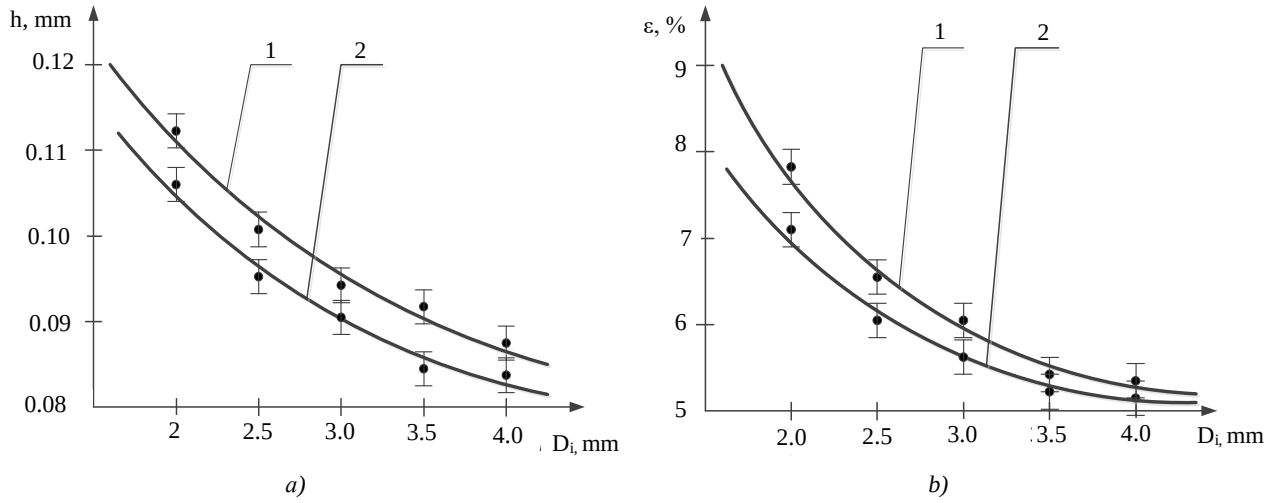


Fig. 4. Dependence of the surface roughness on the Brinell hardness of the part



a) *b)*

Fig. 5. Dependence of hardening parameters on Brinell hardness for various materials: *a* — on the depth of the hardened layer; *b* — on the degree of deformation



a) *b)*

Fig. 6. Dependence of hardening parameters on diameter of the indenter: *a* — depth of the hardened layer; *b* — degree of deformation. 1 — steel 45 sample material; 2 — HVG steel sample material

Discussion and Conclusion. Based on the results of the conducted research, the following conclusions can be drawn:

1. A theoretical dependence has been obtained that provides determining the kinetic energy of the indenter under processing with an oscillating tool — an eccentric hardener.

2. Dependences have been obtained for determining the diameter and depth of the plastic print, as well as the surface roughness according to parameter Ra, the depth of the hardened layer and the degree of deformation that provide predicting the quality of the treated surface.

3. The results of theoretical and experimental studies of the treatment process with an eccentric hardener were compared. The discrepancy in the results did not exceed 15%.

4. The process under study is subject to further investigation in order to determine other parameters of the quality of the treated surface, e.g., the magnitude of residual stresses in the surface layer, and [5, 6, 20] also to expand the range of treatment modes and design parameters in order to determine their optimal range.

5. The dependences obtained for determining the key parameters of the surface layer quality make it possible to predict the results of processing and can be used to design processes with an eccentric hardener.

References

1. Tamarkin M, Tishchenko E, Astashkin A, et al. Module System Developing of Computer-Aided Engineering for Process Technologies with Ball-Shaft Hardener. In book: A. Guda (ed.) *Lecture Notes in Networks and Systemfor Connected and Automated Vehicles.* Cham: Springer; 2023. Vol. 509. P. 1605–1613. https://doi.org/10.1007/978-3-031-11058-0_163

2. Tamarkin M, Tishchenko E, Murugova E, et al. Surface Quality Assurance and Process Reliability in the Processing with a Ball-Rod Hardener. *E3S Web of Conferences*. 2020;175:05008. https://doi.org/10.1051/e3sconf/202017505008

3. Beskopylny A, Meskhi B, Beskopylny N, et al. Strengthening of Welded Joints of Load-Bearing Structures of Robotic Systems with Ball-Rod Hardening. In book: *Robotics, Machinery and Engineering Technology for Precision Agriculture. Proceedings of XIV International Scientific Conference "INTERAGROMASH 2021".* Singapore: Springer; 2022. P. 1–12. https://doi.org/10.1007/978-981-16-3844-2_1

4. Beskopylny A, Meskhi B, Veremeenko A, et al. Influence of Boundary Conditions on the Strengthening Technology of a Welded Joint with a Ball-Rod Hardener. *IOP Conference Series: Materials Science and Engineering*. 2020;1001:012047. https://doi.org/10.1088/1757-899X/1001/1/012047

5. Tamarkin M, Tishchenko E, Fedorov V. Theoretical Bases of the Surface Layer Formation in the Finishing and Hardening Treatment of Details by SPD in Flexible Granular Environment. *IOP Conference Series: Materials Science and Engineering.* 2016;124:012169. https://doi.org/10.1088/1757-899X/124/1/012169

6. Tamarkin MA, Tishchenko EE, Tishchenko RG. Technological Support of Digital Production when Processing Parts Using a Ball-Rod Hardener. *iPolytech Journal*. 2022;26(2):184–196. https://doi.org/10.21285/1814-3520-2022-2-184-196

7. Lebedev V, Sokolov V, Davydova I. Prediction of Physical and Mechanical Characteristics of Quality of Surface Layer, Modified by Surface Plastic Deformation. *Strengthening Technologies and Coatings*. 2018;14(2):54–58.

8. Lebedev V, Kirichek A, Chunakhova L. *Effectiveness of Application of Additional Strengthening Processing of Surface Plastic Deformation on Increase in Fatigue Life of Parts*. In: Proc. 5th Int. Conf. on Industrial Engineering (ICIE 2019). Cham: Springer; 2019. Vol. 2. P. 17–25. https://doi.org/10.1007/978-3-030-22063-1_3

9. Blumenstein V, Makhalov M. The Metal Surface Layer Mechanical Condition Transformation in Machining Processes. *MATEC Web of Conferences.* 2019;297:05001. https://doi.org/10.1051/matecconf/201929705001

10. Smolentsev V, Kuzovkin A, Safonov S. Nano-Transformations in the Surface Layer of Materials under Combined Processing by Unbound Granules. *Materials Today: Proceedings*. 2019;11(1(26)):20–25. https://doi.org/10.1016/j.matpr.2018.12.100

11. Smolentsev V, Safonov S. The Technological Methods of Surface Layer Modification in Construction Materials. *MATEC Web of Conferences*. 2017;129:01077. https://doi.org/10.1051/matecconf/201712901077

12. Makhalov MS, Blumenstein VYu. The Residual Stress Modeling in Surface Plastic Deformation Machining Processes with the Metal Hardening Effect Consideration. *Solid State Phenomena*. 2022;328:27–37. https://doi.org/10.4028/p-z92o0e

13. Makhalov MS, Blumenstein VYu. The Surface Layer Mechanical Condition and Residual Stress Forming Model in Surface Plastic Deformation Process with the Hardened Body Effect Consideration. *IOP Conference Series: Materials Science and Engineering.* 2017;253:012009. https://doi.org/10.1088/1757-899X/253/1/012009

14. Blumenstein V, Makhalov M. The Metal Surface Layer Mechanical Condition Transformation in Machining Processes. *MATEC Web of Conferences.* 2019;297:05001. https://doi.org/10.1051/matecconf/201929705001

15. Makhalov MS, Blumenstein VYu. Finite Element Surface layer inheritable condition residual stresses model in surface plastic deformation processes. *IOP Conference Series: Materials Science and Engineering.* 2016;126:012004. http://dx.doi.org/10.1088/1757-899X/126/1/012004

16. Chigirinskii YuL. Surface Quality after Different Treatments. *Russian Engineering Research*. 2011;31(8):816–819. https://doi.org/10.3103/S1068798X11080065

17. Plotnikov AL, Chigirinskii YuL, Frolov EM, et al. Formulating CAD/CAM Modules for Calculating the Cutting Conditions in Machining. *Russian Engineering Research*. 2009;29(5):512–517. https://doi.org/10.3103/S1068798X09050207

18. Chigirinskii YuL. Formalized Approaches in Technological Design. *Russian Engineering Research*. 2010;30(3):305–307. https://doi.org/10.3103/S1068798X10030251

19. Chigirinskii YuL, Firsov IV, Chigirinskaya NV. Information System for the Design of Machining Processes. *Russian Engineering Research*. 2014;34(1):49–51. https://doi.org/10.3103/S1068798X14010031

20. Tamarkin MA, Tishchenko MA, Tishchenko EE, et al. Development of Design Methodology of Technological Process of Ball-Rod Hardening with Account for Formation of Compressive Residual Stresses. *Advanced Engineering Research (Rostov-on-Don).* 2020;20(2):143–149. https://doi.org/10.23947/1992-5980-2020-20-2-143-149

About the Authors:

Mikhail A. Tamarkin, Dr.Sci. (Eng.), Professor, Head of the Mechanical Engineering Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), ScopusID, ORCID, AuthorID, tehn_rostov@mail.ru

Elina E. Tishchenko, Cand.Sci. (Eng.), Associate Professor of the Mechanical Engineering Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), ScopusID, ORCID, AuthorID, lina_tishenko@mail.ru

Omar C.A. Hashash, Postgraduate of the Mechanical Engineering Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), ORCID, omar-hashash@mail.ru

Claimed contributorship:

MA Tamarkin: basic concept formulation; research objectives and tasks; academic advising; research results analysis.

EE Tishchenko: calculation analysis; text preparation; formulation of the conclusions.

OCA Hashash: conducting experimental studies; processing the results.

Received 01.04.2023
Revised 21.04.2023
Accepted 29.04.2023

Conflict of interest statement: the authors do not have any conflict of interest.

All authors have read and approved the final manuscript.

Об авторах:

Михаил Аркадьевич Тамаркин, доктор технических наук, профессор, заведующий кафедрой технологии машиностроения Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), ScopusID, ORCID, AuthorID, tehn_rostov@mail.ru

Элина Эдуардовна Тищенко, кандидат технических наук, доцент кафедры технологии машиностроения Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), ScopusID, ORCID, AuthorID, lina_tishenko@mail.ru

Омар С.А. Хашаш, аспирант кафедры технологии машиностроения, Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), ORCID, omar-hashash@mail.ru

Заявленный вклад соавторов:

М.А. Тамаркин — формирование основной концепции, цели и задачи исследования, научное руководство, анализ результатов исследований.

Э.Э. Тищенко — проведение расчетов, подготовка текста, формирование выводов.

О.С.А. Хашаш — проведение экспериментальных исследований, обработка результатов.

Поступила в редакцию 01.04.2023
Поступила после рецензирования 21.04.2023
Принята к публикации 29.04.2023

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Все авторы прочитали и одобрили окончательный вариант рукописи.

MACHINE BUILDING AND MACHINE SCIENCE
МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ

UDC 669.058.7(620.18+620.193)
https://doi.org/10.23947/2687-1653-2023-23-2-140-154

Original article

Aspects of Thermal Protection of Machine-Building and Power Equipment: Application of Oxidation-Resistant Combined Nickel-Based Coatings

Valeriy N. Varavka, Oleg V. Kudryakov, Vyacheslav I. Grishchenko
Don State Technical University, Rostov-on-Don, Russian Federation
✉ varavkavn@gmail.com

Abstract

Introduction. In the areas of power engineering where the thermal energy of superheated steam is used, an important aspect of providing the reliability and safety of equipment is the heat resistance of the materials employed. In the manufacture of induction superheaters, the optimal material for the steam pipe (coil) is copper. However, its ultimate resistance to oxidation does not exceed 400 °C, which significantly limits the efficiency of steam generators. Therefore, the objective of the work was to study the kinetics of oxidation of the combined galvanic coating of the Mo-Ni-Cr system applied to copper tubular samples and intended for thermal protection of steam generator coils.

Materials and Methods. A combined electroplating of the Mo-Ni-Cr system with a total thickness of 12–35 μm was formed on the experimental copper tubular samples. A Mo sublayer with a thickness of about 1.5 μm on the surface of the copper tube was formed to prevent the diffusion of Cu into the Ni coating. A 1.5 μm thick chromium layer on the coating surface acted as an indicator of the oxidation process. A comparative analysis of the oxidation processes of the copper surface and the combined coating of the Mo-Ni-Cr system on a copper substrate was carried out using the methods of optical and electron microscopy, energy dispersive analysis, and precision determination of the growth parameters of oxide films.

Results. The intervals of thermal stability of the copper substrate and nickel coating were experimentally determined. The obtained experimental dependences characterized the parabolic law of copper oxidation with the formation of a single-phase diffusion zone of CuO at temperatures above 350 °C, and nickel at temperatures above 750 °C, when the transition of NiO monoxide into oxide Ni_2O_3 began. The growth of oxide films according to quadratic laws provided a rapid increase in the thickness of the films, the accumulation of stresses in them, cracking, and chipping.

Discussion and Conclusion. It is shown that the Mo-Ni-Cr electroplating is resistant to heating during long-term operation up to temperatures of 750–800 °C. The functional roles of Mo and Cr in the coating architecture were described. The work focused on the applied aspect of using the coating under study to increase the thermal stability of the steam pipelines of industrial induction superheaters with low and medium power.

Keywords*:* superheaters, heat resistance, oxidation process, electroplating, microstructure, electron microscopy, gravimetric analysis

Acknowledgements: the authors would like to thank the editorial board of the journal and the reviewers for their attentive attitude to the article.

For citation. Varavka VN, Kudryakov OV, Grishchenko VI. Aspects of Thermal Protection of Machine-Building and Power Equipment: Application of Oxidation-Resistant Combined Nickel-Based Coatings. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):140–154. https://doi.org/10.23947/2687-1653-2023-23-2-140-154

Научная статья

Аспекты теплозащиты машиностроительного и энергетического оборудования: применение стойких к оксидированию комбинированных покрытий на основе никеля

В.Н. Варавка✉, **О.В. Кудряков**, **В.И. Грищенко**
Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация
✉ varavkavn@gmail.com

Аннотация

Введение. В тех областях энергетического машиностроения, где используется тепловая энергия перегретого пара, важным аспектом обеспечения надежности и безопасности оборудования является теплостойкость используемых материалов. При изготовлении индукционных пароперегревателей оптимальным материалом для паропровода (змеевика) является медь. Однако её предельная стойкость к оксидированию не превышает 400 °С, что существенно ограничивает эффективность работы парогенераторов. Поэтому целью работы было исследование кинетики окисления комбинированного гальванического покрытия системы Mo-Ni-Cr, нанесенного на медные трубчатые образцы и предназначенного для теплозащиты змеевиков парогенераторов.

Материалы и методы. На опытных медных трубчатых образцах было сформировано комбинированное гальваническое покрытие системы Mo-Ni-Cr с общей толщиной 12–35 мкм. Подслой Mo толщиной около 1,5 мкм на поверхности медной трубки был сформирован для предотвращения диффузии Cu в Ni-покрытие. Слой хрома толщиной 1,5 мкм на поверхности покрытия выполнял роль индикатора процесса окисления. Сравнительный анализ процессов окисления поверхности меди и комбинированного покрытия системы Mo-Ni-Cr на медной подложке выполнен с использованием методик оптической и электронной микроскопии, энергодисперсионного анализа, а также прецизионного определения параметров роста оксидных пленок.

Результаты исследования. Экспериментально определены интервалы термической устойчивости медной подложки и никелевого покрытия. Полученные экспериментальные зависимости характеризуют параболический закон окисления меди с образованием однофазной диффузионной зоны CuO при температурах выше 350 °С и никеля при температурах выше 750 °С, когда начинается переход монооксида NiO и в оксид Ni_2O_3. Рост оксидных пленок по квадратичным законам приводит к быстрому увеличению толщины пленок, накоплению в них напряжений, растрескиванию и скалыванию.

Обсуждение и заключение. Показано, что гальваническое покрытие Mo-Ni-Cr устойчиво к нагреву при длительной эксплуатации вплоть до температур 750–800 °С. Описаны функциональные роли Mo и Cr в архитектуре покрытия. Работа акцентирована на прикладном аспекте использования исследуемого покрытия для повышения термической устойчивости змеевика-паропровода промышленных индукционных пароперегревателей малой и средней мощности.

Ключевые слова: парогенераторы, теплостойкость, окислительный процесс, гальванические покрытия, микроструктура, электронная микроскопия, гравиметрический анализ

Благодарности: авторы выражают благодарность редакции журнала и рецензентам за внимательное отношение к статье.

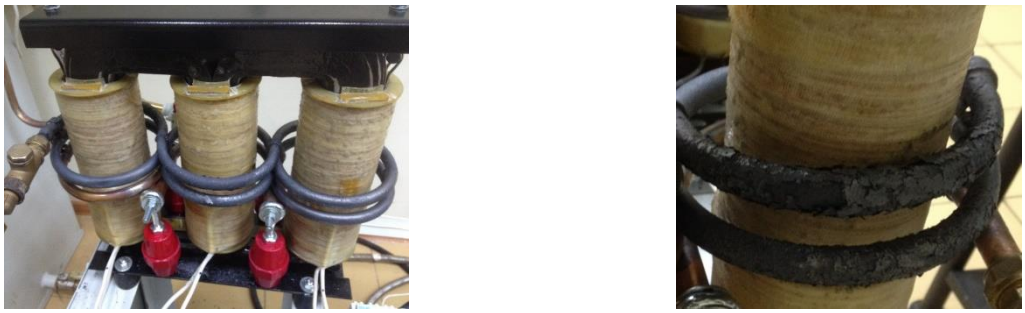
Для цитирования. Варавка В.Н., Кудряков О.В., Грищенко В.И. Аспекты теплозащиты машиностроительного и энергетического оборудования: применение стойких к оксидированию комбинированных покрытий на основе никеля. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):140–154. https://doi.org/10.23947/2687-1653-2023-23-2-140-154

Introduction. Electrochemical deposition of metals is widespread in industry, being the basis of electroplating. One of the features of the development of this branch of science is usually attributed to the fact that its development took place “almost exclusively empirically” [1], starting primarily from the needs of various industries. Despite the fact that at present, the theoretical foundations of electrochemistry have been worked out quite deeply [2–6], the applied aspect is

still of priority importance here and determines most of the scientific tasks being solved, mainly related to the special conditions for the use of electroplating coatings. Under such particular conditions, superheaters are operated, for example, whose performance is associated with a significant change in the composition and temperature of steam along the length of the steam coil [7]. Depending on the power of the steam generator, the temperature of the coil along its length can vary from 150 to 650 °C, and for super-critical steam parameters in high-capacity modern steam turbines — even higher [8–10]. This work studies the possibility of using electroplating coatings to protect the steam pipe of an induction superheater from oxidation at high temperatures. On the totality of physical and technological properties (electrical conductivity, thermal conductivity, ability to plastic deformation, machinability by cutting, etc.), copper is currently an indispensable structural material for the manufacture of coils for household and low-power industrial steam generators. This is the reason for the interest in heat-protective coatings. However, the oxidizing ability of copper is also high, and its oxidation resistance does not exceed 400 °C. Based on the operating conditions of the superheaters under consideration, this circumstance offers a challenge of using coatings whose oxidation resistance level is above the thermal barrier of 600 °C.

Materials and Methods. Taking into account the complex configuration of the coil, the presence of a large length of curved surfaces and its considerable overall dimensions, electrochemical deposition was chosen as the most technologically advanced method of applying a heat-protective coating.

To select the composition of such a coating, accurate data on its operating modes were required. For this purpose, a thermal imaging analysis of the thermal operating conditions of an experimental induction three-coil six-turn steam generator with a capacity of 10 kW was carried out (Fig. 1) [11]. The steam pipe was made of profiled copper tube ∅25×1.5 mm of technical copper M2 grade according to GOST 617–2006.



a) *b)*

Fig. 1. Experimental induction superheater with copper steam pipe: *a* — general view; *b* — the most heavily oxidized sections of the steam pipe (coil) at the steam outlet

Quantitative thermal analysis of the operating conditions of the steam pipeline, including forced operating modes, was performed using a no-contact thermal imager of Fluke Ti401 PRO model (manufactured by Fluke Corp., USA) [12] with the main technical characteristics:

– infrared spectral range: from 7.5 to 14 μm (long-wavelength);

– thermal sensitivity: ≤ 0.075 °C at an object temperature of 30 °C (75 mK);

– error: ±2 °C (at low temperatures) or 2%;

– degree of protection: according to GOST 14254-96 (IEC 60529): IP54.

The heat capacity of water vapor (c_p = 33.6 J/(mol·K) under normal conditions) is approximately twice lower than the heat capacity of water (c_p = 75.3 J/mol·K), which changes significantly the conditions of heat removal in the coil and contributes to the intensification of oxidation of the surface of the steam pipe; therefore, substantial heterogeneity of the degree of oxidation is observed along its length (Fig. 1 *b*). The thermal analysis results (Fig. 2) showed that the maximum heating temperatures were fixed at the output half-turn of the coil (Fig. 1 *b*). The range of their values was 530–540 °C with an absolute maximum of 541.38 °C. The average temperature values of most superheated (oxidized) half-turns were at the level of 420–460 °C.

Taking into account the results obtained, in further studies on heat-resistant coatings to protect against oxidation of the coil surface, it is needed to focus on the maximum temperature load of 600 °C. In this regard, it is reasonable to use

nickel-based coatings. Ni forms the basis of most modern heat-resistant superalloys used in thermal power engineering [13–15], and the technology of electroplating Ni is quite well developed [15–18].

When applying experimental electroplating coatings to samples of copper tubes of technical copper M2, standard deposition modes and compositions of electrolytes containing Ni and Cr recommended by GOST 9.305 and 9.306 were used. During the operation of the steam generator, the coating is practically not subjected to mechanical action; therefore, it should not demonstrate outstanding mechanical properties during performance. At the same time, when applied to a curved convex surface, internal tensile stresses are formed in the coating. As the number of working heat shifts increases, their level grows. In this regard, the thickness of the investigated coating on the steam line should not be too large. It was taken as the average of the recommended in the literature ranges of values for nickel electroplating coatings performing protective and decorative functions, and was an approximate level of 20 μm.

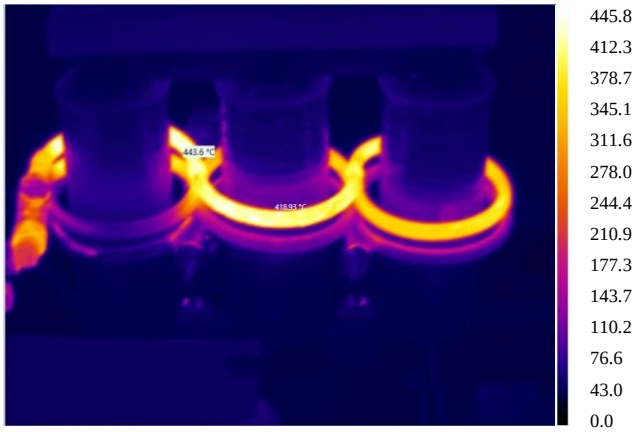


Fig. 2. IR image of a general view of the thermal fields of a working superheater (the window of the Fluke Connect program, the location of the coil is similar to Fig. 1 *a*, all values are in degrees Celsius)

To study the microstructure of coatings and the topography of their surface, dual beam (electron/ion) scanning electron microscope ZEISS CrossBeam 340 (SEM) was used, which provides using a focused ion beam (FIB) to etch and perform cross-sections (sections of a given configuration) of samples directly in the vacuum chamber of the microscope with high positioning accuracy. The elemental composition of the studied surfaces was monitored using energy dispersive X-ray detector (EDAX) of X-Max 50N (Oxford Instruments) model, built into an electron microscope.

The study of the oxidation kinetics of coatings was carried out through measuring the mass index of high-temperature gas corrosion (weight gain of the sample as a result of heating and oxidation). To determine the degree of oxidation of copper samples, both coated and uncoated, all samples were weighed before and after experiments at different stages of interaction. Gravimetric studies were carried out on analytical scales of "VLR-20" brand with a weighing accuracy of 10^{-5} g.

Results and Discussion

1. Qualitative analysis of oxidation kinetics. Properties of the chemical interaction of nickel and oxygen are manifested in the fact that Ni forms two modifications of monoxide: α-NiO with a hexagonal lattice (below 252 °C) and β-NiO with a face-centered cubic lattice. The transition occurs under continuous heating in the range of 250–300 °C. It was experimentally established that when heated to 630 °C, a diffusion process ran through a thin film of NiO monoxide, above 640 °C, a chemical process of NiO formation was established, which, when heated above temperatures of 800 °C, could cause the formation of Ni_2O_3 oxide [19].

The applied aspect of using a heat-resistant nickel coating on a copper substrate was complicated by two circumstances: the unlimited solubility of the Cu-Ni system components (Fig. 3) and the possibility of the Kirkendall interaction effect [21, 22] at the "Ni-coating — Cu-substrate" boundary. These phenomena reduced the resistance of the coating to oxidation due to the dissolution of copper in the coating.

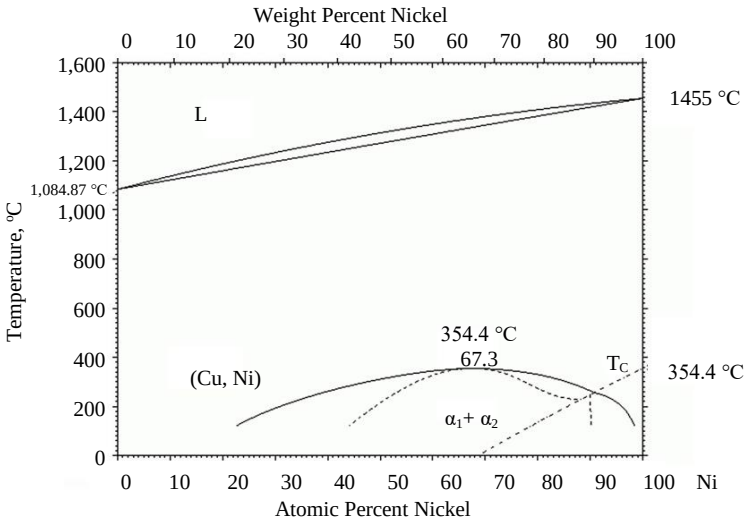
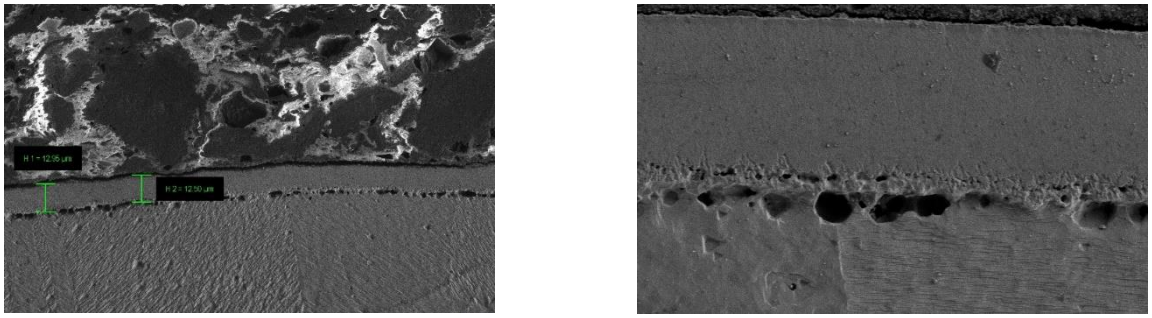


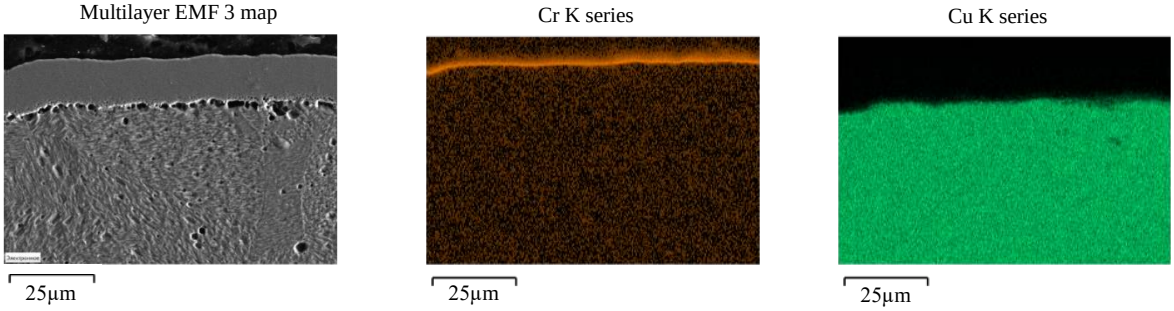
Fig. 3. Diagram of the phase states of the copper-nickel system [20]

To exclude the negative impact of these circumstances on the oxidation resistance of the coating, a combined electroplating of the Mo-Ni-Cr system was formed on experimental copper tubular samples (рис. 4). Mo sublayer with a thickness of about 1.5 μm on the surface of the copper tube (Fig. 4 *b*) was formed to prevent the diffusion of Cu into the Ni coating under long-term operation of the steam pipe due to the practically insoluble system of Cu-Mo components and the limited solubility of Ni-Mo [20]. A layer of chromium 1.5 μm thick on the coating surface (Fig. 4 *b*) served as an indicator of the oxidation process (for more information, see below). The total thickness of the coating on the experimental samples with coatings was 12–35 μm. The elemental distribution in the cross-section of the initial Mo-Ni-Cr coating (before the experiment with sample heating) is shown in Fig. 5.



a) *b)*

Fig. 4. Initial MoNiC coating in cross section, SEM: *a* — coating with thickness markers; *b* — homogeneous microstructure of the coating and hydrogen porosity at the boundary with the substrate



a) *b)* *c)*

d) *e)*

Fig. 5. Color maps of the distribution of chemical elements by the depth of Mo-Ni-Cr coating, EDAX: *a* — general view of the coating in cross-section (SEM); *b*–*e* — distribution of elements in general view image: Cr (*b*), Cu (*c*), Mo (*d*), Ni (*e*)

For the experimental study of the oxidation kinetics under conditions as close as possible to the operating conditions of the steam pipe, the samples for the study were made of copper tubes with the corresponding wall thickness and diameter (Fig. 6).

After plating with the technology that included heat treatment elements [11], the coating acquired a greenish hue characteristic of nickel monoxide NiO. When operating the steam pipe, the coating was applied only to the outer (convex) surface of the samples. However, in order to correctly determine the weight gain of the coating on the experimental samples, the coating was applied on both sides.

Simultaneous heating of pure copper samples and coated samples was carried out at a fixed temperature in the range of 350–1000 °C in SNOL 6.7/1300 (2.4 kW) furnace in air. Exposure at a given temperature was 30 minutes. For the statistical picture of the experiment, heating at each set temperature was carried out for 5–7 samples with their separate loading into the furnace. The selective results of the experiment are visualized in Figure 7.

From the experimental data obtained, it follows that under the conditions of the conducted heating, copper is relatively thermally-resistant to a temperature of 300-350 °C. At these temperatures, a dense thin oxide film of brown color is formed on the copper surface, regardless of its curvature (Fig. 7).

Fig. 6. Samples for the study of oxidation kinetics; external surfaces of reference samples of pure copper (right) and copper coated with Mo-Ni-Cr (left), prepared for experiments

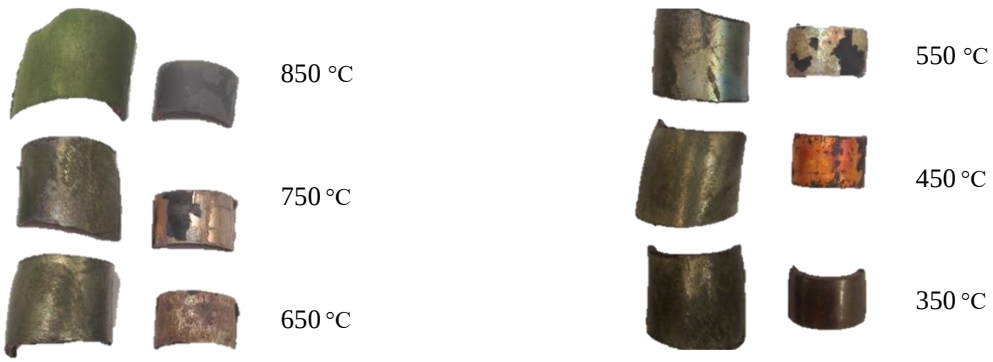
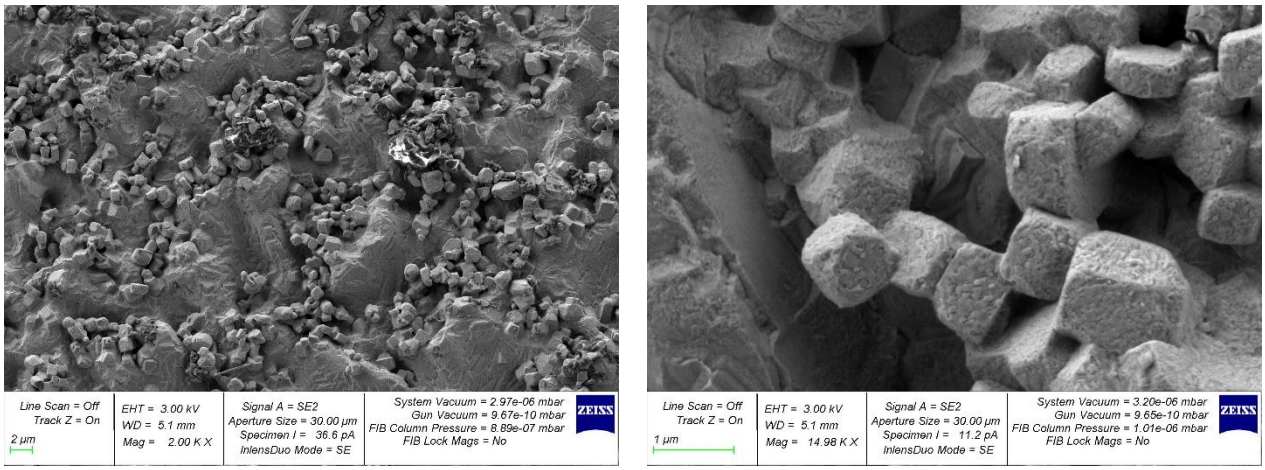


Fig. 7. Comparison of the outer surface of the samples after heating to the specified temperatures: on the left — coated samples, on the right — pure copper samples

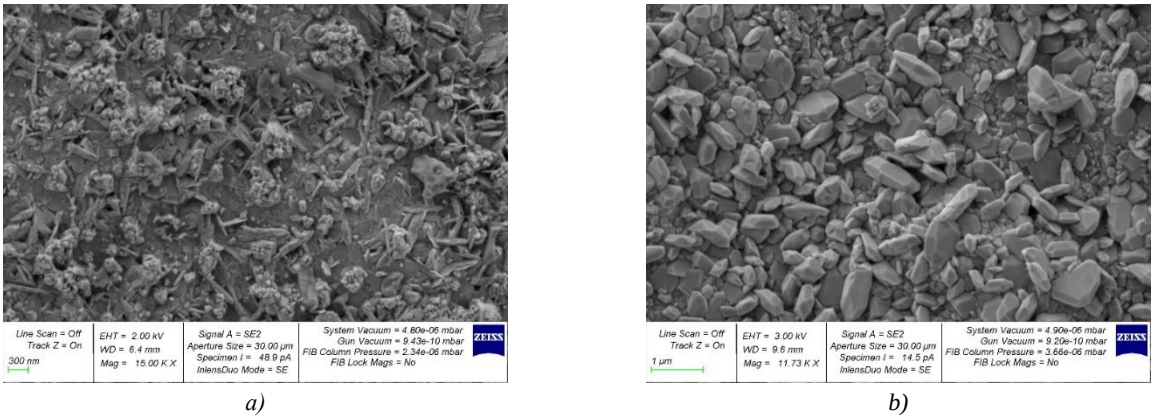


a) *b)*

Fig. 8. Copper surface of the sample oxidized at a temperature of 650 °C, SEM: *a* — location of CuO crystals on the copper surface; *b* — morphology of CuO crystallites

According to the literature data [23, 24], it appears to be two-layered: a thin sublayer of Cu_2O is located on the surface of the sample, and a layer of CuO is located outside. Due to the small thickness of the latter, the internal stresses in the film are small. It has good adhesion to the substrate, low roughness, it does not loosen, and does not chip off the surface. As the heating temperature increases, the growth of the outer oxide layer CuO accelerates. Already at a temperature of 450 °C, it becomes very loose and crumbles from the surface (Fig. 1 *b*). At the same time, a Cu_2O sublayer of a characteristic reddish hue is found under it (Fig. 7). During further heating, peeling and shedding of the CuO oxide layer progresses — it practically does not stay on the surface up to temperatures of 650–700 °C. This, apparently, is due to the nature of crystallization of copper monoxide: its crystallites have a strict cut close to cubic (Fig. 8 *b*), weak conjugation with each other, and, most importantly, high heterogeneity of the nucleation sites (Fig. 8 *a*). Starting from the heating temperatures of ≈750 °C, the copper oxide film is compacted, the strength of its adhesion to the surface increases. On the concave surface of the samples, due to compressive configuration stresses, the CuO film is strong enough or can peel off completely from the entire surface of the sample without crumbling. At temperatures of 800–900 °C, the CuO oxide film behaves similarly on the outer (convex) surface of the samples.

The surface of samples coated with Mo-Ni-Cr practically does not change up to a heating temperature of 750 °C. Upon further heating above 800 °C, the coating is first covered with a film of Cr_2O_3 oxide having a characteristic bright green color (Fig. 7). Then, at heating temperatures above 900 °C, the deeper layers of the coating are oxidized. Amphoteric chromium oxide Cr_2O_3 (Fig. 9 *b*) is fundamentally different in its morphology and crystallization character from CuO copper monoxide (Fig. 8). Cr_2O_3 oxide crystals have a characteristic polyhedron shape with a predominance of prismatic crystallites. Due to the large dispersion of crystallites in size, they, unlike copper monoxide, form a layer on the surface with a high packing density of crystallites.



a) *b)*

Fig. 9. Surface of samples coated with Mo-Ni-Cr, SEM:
a — after heating to a temperature of 650 °C; *b* — after heating to a temperature of 850 °C

If, prior to the oxidation of the chromium layer (below 800 °C), the coating surface has a very weak crystallinity character (Fig. 9 *a*), then the appearance of Cr_2O_3 oxide gives the surface a well-known polycrystalline appearance (Fig. 9 *b*). In general, the nickel coating does not change its composition and structure by cross-section up to a temperature of 850 °C (Fig. 10). The chemical composition of the coating surface at this temperature indicates the initial degree of oxidation of the chromium layer. The presence of a thin layer of chromium oxide on the surface is confirmed by the data of energy dispersion analysis (EDAX) both by the depth of the coating and by the surface. Figure 11 shows that oxygen is concentrated in a much narrower surface layer (~1 μm) than the chromium layer (~3 μm), which characterizes the initial stage of oxidation.

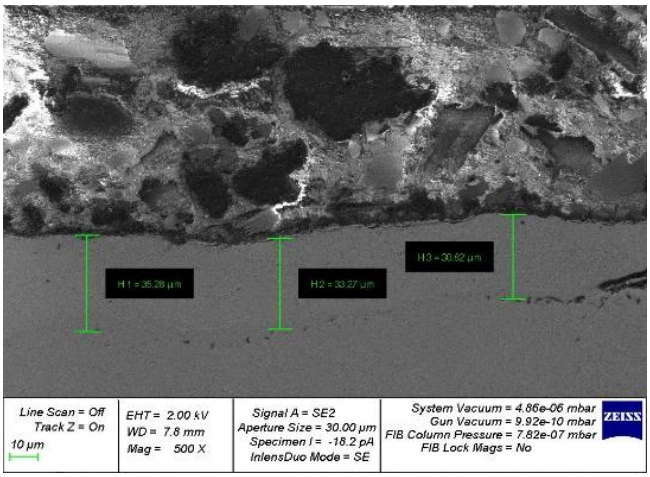
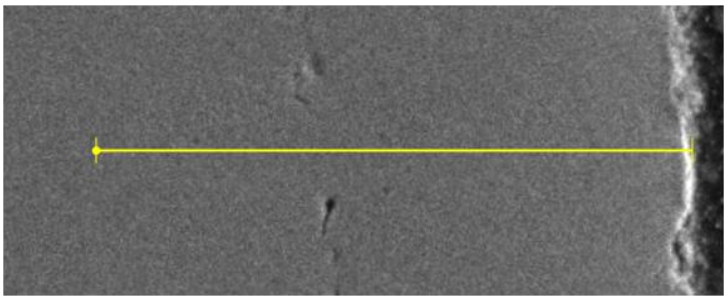
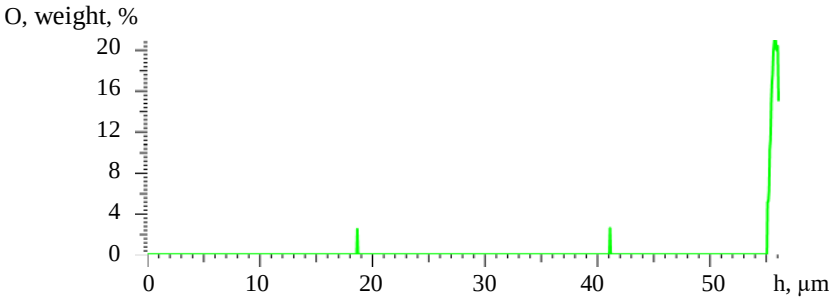


Fig. 10. Thickness and structure of Mo-Ni-Cr coating in cross section after heating up to 850 °C, SEM

To reduce the volume of the article, the EDAX data is not provided in full. Its results show that the amount of oxygen on the surface increases from 30 to 50 at. % due to a similar decrease in chromium concentration, which indicates the oxidation of chromium (since the concentration of Ni does not change when heated from 650 °C to 850 °C, and the presence of Ni in the detection results is caused by the penetration of X-ray radiation through a thin layer of chromium into the nickel base of the coating during EDAX analysis). The composition of the oxides corresponds to the compound Cr_2O_3.



a)



b)

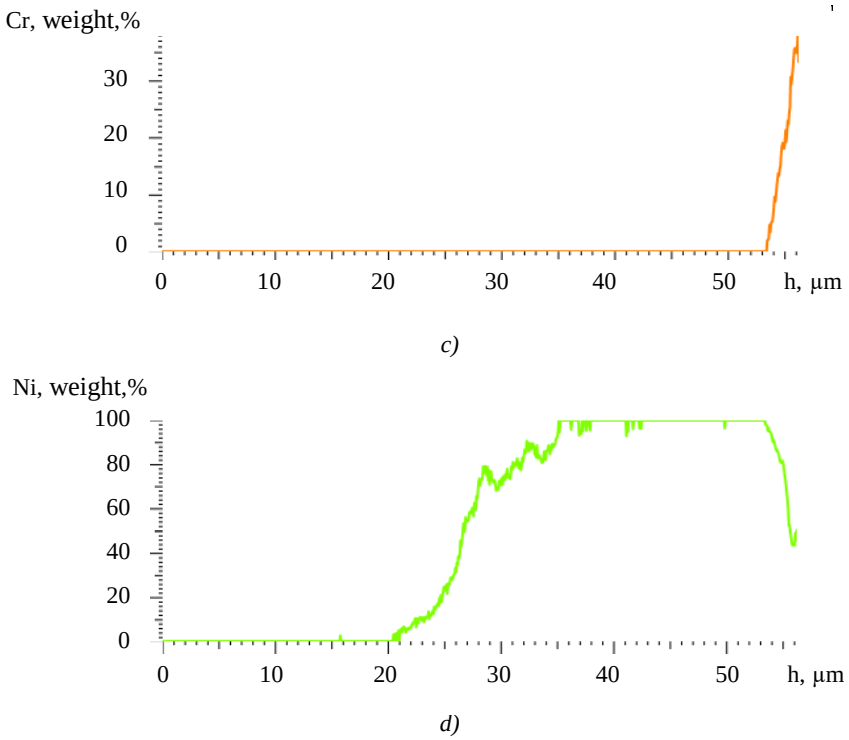
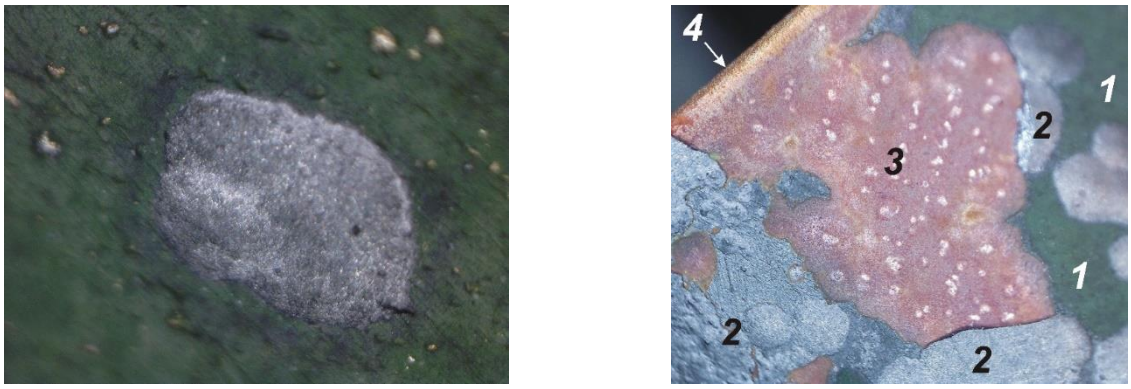


Fig. 11. Distribution of the main chemical elements in Mo-Ni-Cr coating by depth h after heating to 850 °C, EDAX: *a* — coating in cross section, the scanning direction (SEM) is shown; *b–d* — content of elements in the scanning direction; *b* — oxygen; *c* — chromium; *d* — nickel

All stages of the oxidation process of the coating are shown in Figure 12. At temperatures above 800 °C, the surface of the coating starts to oxidize, as indicated by a change in its color — the coating acquires the Kensington Green color.



a) *b)*

Fig. 12. Successive stages of oxidation of nickel coating, optical microscopy, ×100: *a* — initial stage: germination of nickel oxide (dark-gray precipitates) on the surface of chromium oxide (green field) at 850 °C; *b* — final stage: collective phase pattern in the area of coating chip at 1,000 °C, where: 1 — chromium oxide Cr_2O_3; 2 — nickel oxide Ni_2O_3; 3 — copper surface (coating chip); 4 — sample edge

This color corresponds to chromium oxide Cr_2O_3. Already at a temperature of 850 °C, rare single formations of Ni_2O_3 nickel oxide can be found on the surface of the coating (Fig. 12 *a*, gray color), which, under further heating, gradually increase their area occupied on the surface. As they grow, which mainly spreads tangentially to the surface, under the impact of internal stresses, the Ni_2O_3 oxide film cracks and then chips off, exposing the surface of the substrate — pure copper of a reddish hue (Fig. 12 *b*).

Thus, the basic result of this part of the research should be considered experimentally established temperature ranges of permissible use of materials for the manufacture of steam pipe of steam generator. Thus, a copper steam pipe without

coatings is operable up to a temperature of 300 °C and can only be used to generate wet (not overheated) steam. Starting from a temperature of ~400 °C and up to ~700 °C, the CuO oxide film formed is very loose and easily crumbles from the copper surface. At higher temperatures, the film becomes denser, thicker, and its adhesion to the copper substrate increases. However, it is still prone to chipping during heat exchange. Its presence on the surface of the steam pipe significantly slows down the heat removal, and the chemical reactions of high-temperature gas corrosion that continue during heating, work in the direction of reducing the thickness of the pipe. Due to the heterogeneity of the ongoing processes, the operation of the steam pipe under these conditions becomes unpredictable from the point of view of emergency situations. The use of a combined electroplating of the Mo-Ni-Cr system increases the efficiency of the steam pipe to a temperature of 750–800 °C. When the temperature reaches 850 °C, the coating starts to oxidize along with copper. At 950 °C and above, the oxidized coating is prone to chipping, and its operation is subject to the same risks as the copper pipe. A distinctive feature of the investigated coating is self-testing: if the heating temperature exceeds 800 °C during operation, the surface layer of chromium turns the coating bright green and signals the danger of overheating. The indicator layer of galvanic chrome after oxidation can be easily restored, and the operation of the steam pipe then continues.

Quantitative analysis of oxidation kinetics

As part of the performed studies, a quantitative analysis of the kinetics of oxidation of pure copper samples and Mo-Ni-Cr coated samples was carried out. The specific mass gain $M = \Delta m/S$, observed during the heating process, was used as a measured parameter, where Δm — increase in the mass of the sample, g; S — area of the oxidized surface of the sample, cm^2. According to the qualitative analysis method described above, the experimental data of M values were obtained during the heating of tubular samples made of pure copper and coated samples. They are presented in Tables 1 and 2.

The Tables show the spread intervals of the data obtained for fixed values of heating temperatures (Table 1) or the holding time in the furnace (Table 2), as well as the average value of M from each interval.

Statistical processing of the data shown in Tables 1 and 2, performed using the MathCAD application software package, which included interpolation procedures, allowed us to obtain kinetic dependences shown in Figures 13 and 14. The rectilinear graphs of the obtained dependences, shown in Figure 13 *b* in Arrhenius coordinates ($-ln\ M - 1{,}000/T$), characterized the parabolic law of oxidation of copper at temperatures above 350 °C and nickel — at temperatures above 750 °C [19, 25–27].

Table 1

Experimental data on the specific mass gain of samples under furnace heating in air for 30 min

No. of experiment	Furnace temperature, °C	Specific weight gain M, 10^{-5} g/cm^2	
		Uncoated copper tube	Ni-coated copper tube *
1	350	1.35 ± 0.31	–
2	450	6.23 ± 1.37	–
3	550	24.73 ± 3.60	–
4	650	58.51 ± 9.02	1.52 ± 0.24
5	750	138.88 ± 17.85	4.25 ± 0.58
6	850	241.03 ± 25.25	13.12 ± 1.08
7	1,000	564.70 ± 49.76	35.71 ± 2.78

* minimum weight gain equal to 10^{-5} g, measured on the analytical scales, was taken for the criterion of the absence of oxidation (dash in the Table)

Table 2

Experimental data on specific mass gain of samples at different exposure time in the air of the furnace

No. of experiment	Holding time in the furnace, min	Specific weight gain M, 10^{-5} g/cm^2	
		Uncoated copper tube at 600 °C	Ni-coated copper tube at 800 °C
1	5	13.68 ± 2.02	3.03 ± 0.41
2	15	29.43 ± 4.15	5.93 ± 0.54
3	30	35.76 ± 4.84	7.11 ± 0.67

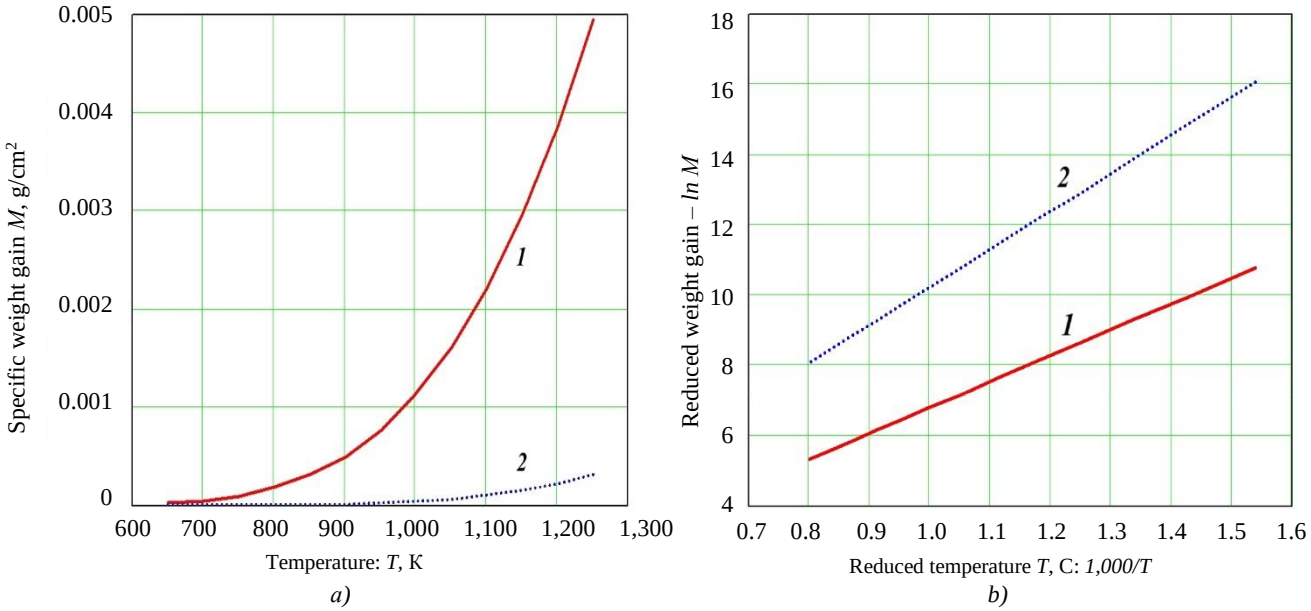


Fig. 13. Temperature dependences of mass gain M of pure copper samples (1) and samples coated with Mo-Ni-Cr (2): *a* — in absolute units; *b* — in relative coordinate system

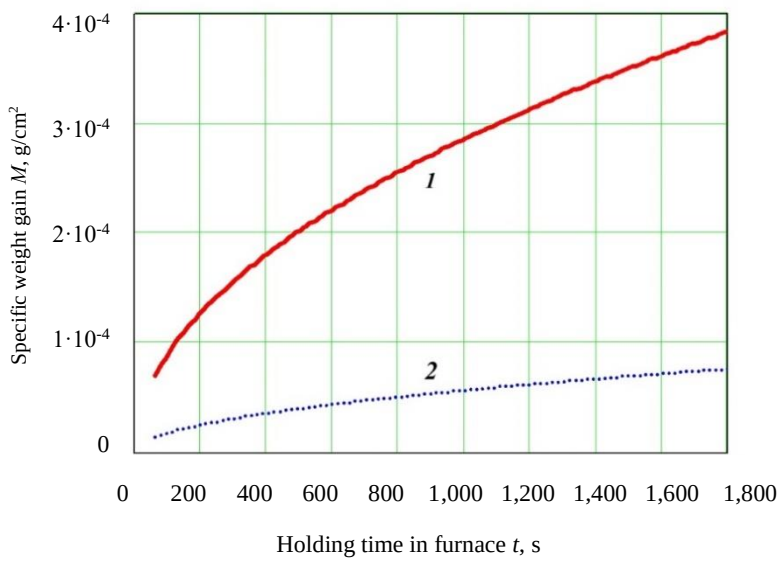


Fig. 14. Kinetics of time variation of mass gain M of pure copper samples at 600 °C (1) and samples coated with Mo-Ni-Cr at 800 °C (2)

The growth of oxide films according to quadratic laws occurs with the formation of single-phase diffused zones, in this case consisting of CuO and NiO oxides, respectively. It causes a rapid increase in the thickness of the films, the accumulation of stresses in them, cracking and chipping. An additional contribution to the acceleration of this process is made by the curved outer surface of the copper tube [28, 29].

Conclusions

1. The performed set of studies has shown that the combined electroplating of the Mo-Ni-Cr system is a sufficiently effective protection of the copper steam pipe from oxidation. The coating is able to provide a long-term operation of the steam generator up to heating temperatures of 750–800 °C.

2. Long-term heat resistance of the coating is provided by an internal Ni layer with a recommended thickness of 20–30 μm. The study of the oxidation kinetics of the coating, performed by optical and electron microscopy, energy dispersion analysis, as well as using precision methods for determining the growth parameters of oxide films, has shown that the nickel coating is indifferent to heating up to temperatures of 600–650 °C. In the temperature range 700–900 °C, the oxidation of the coating occurs with the formation of NiO monoxide according to the parabolic law. At higher temperatures, oxidation progresses due to the formation of Ni_2O_3, oxide film, which quickly causes its growth, cracking and chipping.

3. The combined architecture of the investigated nickel coating includes two thin layers of Mo and Cr. The Mo sublayer with a thickness of about 1.5 μm is located on the surface of the copper tube (substrate). Its function is to prevent the mutual diffusion of Ni and Cu at the coating – substrate interface during long-term operation of the steam generator, since the dissolution of copper reduces the heat resistance of nickel and disrupts the performance of the coating. The outer layer of chromium with a thickness of 2–3 μm serves as an indicator of the degree of oxidation of the coating. The first sign of excessive oxidation of the coating is the appearance of a bright green hue on the surface of the coating, which is associated with the formation of chromium oxide Cr_2O_3 at temperatures ≥800 °C. The overheating indicator — a layer of chromium — is easily updated and contributes to the prolongation of the life cycle of the steam generator.

References

1. Knauschner A. *Oberflächenveredeln und Plattieren von Metallen*. Leipzig: VEB Deutscher Verlag Grundstoffindustrie; 1982. 416 p.

2. Stock JT, Orna MV (eds.). *Electrochemistry, Past and Present*. Washington, DC: American Chemical Society; 1989. 606 p.

3. Gamburg YuD. *Electroplating Coatings.* Moscow: Tekhnosfera; 2006. 216 p. (In Russ.) https://www.elec.ru/files/2019/09/23/Galvanicheskie_pokrytia_Spravochnik_po_primeneniyu.PDF (accessed: 10.02.2023).

4. Andreev YuYa. *Electrochemistry of Metals and Alloys.* Moscow: MISiS; 2011. 256 p. (In Russ.)

5. Ujjal Kumat Sur (ed.). *Recent Trend in Electrochemical Science and Technology*. London: IntechOpen; 2012. 318 p. https://doi.org/10.5772/1891

6. Djokic SS (ed.). *Modern Aspects of Electrochemistry.* New York, NY: Springer; 2014. P. 1–84.

7. Song Liu, Yaping Wei, Shiqiang Chen, et al. Development and Application of an Ultrahigh-Temperature Steam Generator. *Advances in Materials Science and Engineering*. 2020;2020:4243170. https://doi.org/10.1155/2020/4243170

8. Weitzel PS, Tanzosh JM, Boring B, et al. *Advanced Ultra-Supercritical Power Plant (700 to 760 ℃) Design for Indian Coal*. In: Proc. POWER-GEN Asia Conference and Exhibition. Bangkok, 2012. P. 281–291.

9. Wang Chongbin, Xu Xueyuan, Zhu Yufeng, et al. Research on the Heating Surface Material Properties for 700 °C USC Boiler. In book: *Energy Materials*. Cham: Springer; 2014. P. 151–159. URL: https://link.springer.com/chapter/10.1007/978-3-319-48765-6_16 (accessed: 10.02.2023).

10. Booras GS, Powers JM, Riley C, et al. *Engineering and Economic Analysis of an Advanced Ultra-Supercritical Pulverized Coal Power Plant with and without Post-Combustion Carbon Capture*. Palo Alto, CA: Electric Power Research Institute; 2015. 111 p.

11. Shipilov VM, et al. *Electric Steam Superheater.* RF Patent, No. 2736270. 2020. 11 p. (In Russ.)

12. *Thermovizor Fluke Ti401 PRO*. URL: https://fluke-russia.ru/catalog/teplovizor-fluke-ti401-pro (In Russ.) (accessed: 02.02.2023).

13. Sims ChN, et al. (eds.). *Superalloys II: High-Temperature Materials for Aerospace and Industrial Power*. New York, NY: John Wiley & Sons; 1987. 640 p.

14. Reed RC. *The Superalloys. Fundamentals and Applications*. Cambridge: Cambridge University Press; 2006. 372 p. https://doi.org/10.1017/CBO9780511541285

15. Toshio Narita, et al. Advanced Coatings on High Temperature Applications. *Materials Science Forum*. 2006;522–523:1–14. https://doi.org/10.4028/www.scientific.net/MSF.522-523.1

16. Yuebo Zhou, Peng X, Wang F. Oxidation of a Novel Electrodeposited Ni–Al Nanocomposite Film at 1050 °C. *Scripta Materialia*. 2004;50(12):1429–1433. https://doi.org/10.1016/j.scriptamat.2004.03.014

17. Karimzadeh A, Aliofkhazraei M, Walsh FC. A Review of Electrodeposited Ni–Co Alloy and Composite Coatings: Microstructure, Properties and Applications. *Surface Coatings Technology*. 2019;372:463–498. https://doi.org/10.1016/j.surfcoat.2019.04.079

18. Morteza Alizadeh, Abbas Cheshmpish. Electrodeposition of Ni–Mo/Al2O3 Nano-Composite Coatings at Various Deposition Current Densities. *Applied Surface Science*. 2019;466:433–440. https://doi.org/10.1016/j.apsusc.2018.10.073

19. Ryabukhin AG, Novoselova EG, Samarin II. Air Oxidation of Nickel with the Formation of Thin Films. *Bulletin of the South Ural State University. Series "Metallurgy".* 2005;10:34–40. (In Russ.)

20. *Phase Diagrams of Binary Systems*. URL: https://himikatus.ru/art/phase-diagr1/diagrams.php (In Russ.) (accessed: 10.02.2023).

21. Bhadeshia HKDH. *Kirkendall Effect*. University of Cambridge. URL: http://www.phase-trans.msm.cam.ac.uk/kirkendall.html (accessed: 10.02.2023).

22. Hideo Nakajima. The Discovery and Acceptance of the Kirkendall Effect. *Journal of The Minerals, Metals & Materials Society (JOM).* 1997;49(6):15–19. https://doi.org/10.1007%2Fbf02914706

23. Yao Zhi Hu, Rahul Sharangpani, Sing-Pin Tay. Kinetic Investigation of Copper Film Oxidation by Spectroscopic Ellipsometry and Reflectometry. *Journal of Vacuum Science & Technology A*. 2000;18(5):2527–2532. https://doi.org/10.1116/1.1287156

24. Avetisyan AG, Chatilyan HA, Kharatyan SL. Kinetic Features of the Initial Stages of High-Temperature Oxidation of Copper. *Chemical Journal of Armenia*. 2013;66(3):407–415. URL: https://www.researchgate.net/publication/271459151 (In Russ.) (accessed: 10.02.2023).

25. Haugsrud R. On the High-Temperature Oxidation of Nickel. *Corrosion Science*. 2003;45(1):211–235. https://doi.org/10.1016/S0010-938X(02)00085-9

26. Mrowec S, Grzesik Z. Oxidation of Nickel and Transport Properties of Nickel Oxide. *Journal of Physics and Chemistry of Solids*. 2004;65(10):1651–1657. https://doi.org/10.1016/j.jpcs.2004.03.011

27. Avetisyan A, Chatilyan NA, Kharatyan SL. Kinetic Features of the Initial Stages of High-Temperature Oxidation of Nickel. *Chemical Journal of Armenia*. 2014;67(1):27–35. URL: https://www.researchgate.net/publication/271458661 (accessed: 10.02.2023).

28. Karmhag R, Tesfamichael T, Wäckelgard E, et al. Oxidation Kinetics of Nickel Particles: Comparison Between Free Particles and Particles in an Oxide Matrix. *Solar Energy*. 2000;68(4):329–333. https://doi.org/10.1016/S0038-092X(00)00025-6

29. Lei Zhou, Ashish Rai, Nicholas Piekiel, et al. Ion-Mobility Spectrometry of Nickel Nanoparticle Oxidation Kinetics: Application to Energetic Materials. *Journal of Physical Chemistry C*. 2008;112(42):16209–16218. https://doi.org/10.1021/jp711235a

Received 03.03.2023
Revised 29.03.2023
Accepted 05.04.2023

About the Authors:

Valeriy N. Varavka, Dr.Sci. (Eng.), Professor of the Materials Science and Technology of Metals Department, Head of "Materials" Research and Educational Center, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF) ResearcherID, ScopusID, ORCID, AuthorID, varavkavn@gmail.com

Oleg V. Kudryakov, Dr.Sci. (Eng.), Professor of the Materials Science and Technology of Metals Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF) ResearcherID, ScopusID, ORCID, AuthorID, kudryakov@mail.ru

Vyacheslav I. Grishchenko, Cand.Sci. (Eng.), Associate Professor, Head of the Hydraulics, Hydropneumoautomatics and Heat Management Department, Head of "Spektr" Technical Engineer Center, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), , ResearcherID, ScopusID, ORCID, AuthorID, vig84@yandex.ru

Claimed contributorship:

VN Varavka: academic advising; basic concept formulation; research objectives and tasks; discussion of the results; text preparation; formulation of conclusions.

OV Kudryakov: planning and organization of experiments; conducting metal-physical studies; analysis of the results obtained; correction of the conclusions.

VI Grishchenko: preparation of the experimental base, samples and equipment; technical management of the experimental research process; computational analysis; discussion of the results.

Conflict of interest statement: the authors do not have any conflict of interest.

All authors have read and approved the final manuscript.

Поступила в редакцию 03.03.2023
Поступила после рецензирования 29.03.2023
Принята к публикации 05.04.2023

Об авторах:

Валерий Николаевич Варавка, доктор технических наук, профессор кафедры материаловедения и технологии металлов, директор НОЦ «Материалы» Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1) ResearcherID, ScopusID, ORCID, AuthorID, varavkavn@gmail.com

Олег Вячеславович Кудряков, доктор технических наук, профессор кафедры материаловедения и технологии металлов Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1) ResearcherID, ScopusID, ORCID, AuthorID, kudryakov@mail.ru

Вячеслав Игоревич Грищенко, кандидат технических наук, доцент, заведующий кафедрой гидравлики, гидропневмоавтоматики и тепловых процессов, руководитель ИТЦ «Спектр», Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1) ResearcherID, ScopusID, ORCID, AuthorID, vig84@yandex.ru

Заявленный вклад соавторов:

В.Н. Варавка — научное руководство, формирование основной концепции, цели и задачи исследования, обсуждение результатов, подготовка текста, формирование выводов.

О.В. Кудряков — планирование и организация экспериментов, проведение металлофизических исследований, анализ полученных результатов, корректировка выводов.

В.И. Грищенко — подготовка экспериментальной базы, образцов и оборудования, техническое руководство процессом экспериментальных исследований, выполнение расчетов, обсуждение результатов.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Все авторы прочитали и одобрили окончательный вариант рукописи.

MACHINE BUILDING AND MACHINE SCIENCE
МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ

UDC 621.9
https://doi.org/10.23947/2687-1653-2023-23-2-155-168

Original article

Comprehensive Assessment of the Manufacturability of Products

Peter Yu. Bochkaryov[1, 2] ✉, **Richard D. Korolev**[1, 3], **Larisa G. Bokova**[4]

[1] Kamyshin Technological Institute, VSTU branch, Kamyshin, Russian Federation

[2] Saratov State Vavilov Agrarian University, Saratov, Russian Federation

[3] OOO PCF "Ex-Form", Saratov, Russian Federation

[4] Yuri Gagarin State Technical University of Saratov, Saratov, Russian Federation

✉ bpy@mail.ru

Abstract

Introduction. The assessment of the manufacturability of products — as a stage of production planning and a key aspect of the development of modern industrial machining systems — is an urgent task of modern mechanical engineering. In this regard, theoretical and practical research on the development of methodological approaches to determining the weight significance of quantitative indicators in assessing the manufacturability of parts is highly relevant. The objective of the presented work was to develop an evaluation method aimed at improving the quality of part processing and the effectiveness of the performance of multiproduct manufacturing systems based on the development of additional quantitative indicators for assessing production manufacturability.

Materials and Methods. To assess the impact of quantitative production indicators associated with time spent during equipment downtime, a model was created. It was aimed at predicting event flows of delivery of batches of parts for manufacturing for a specific operation and flows of processed parts using the queuing theory apparatus. This approach makes it possible to take into account both the design-engineering characteristics of parts, the features of a particular production system, and the emerging manufacturing situation.

Results. The degree of influence of the manufacturability indicators at the level of the process operation was determined by assessing the possible impact on the components when calculating piece-calculation time ($T_{шт.к.}$). The interrelations between the manufacturability indicators and expenses for all items of the production cost of part processing ($C_{ОП}$), as well as costs associated with organizational downtime of equipment ($C_{пр.о.i}$) were established. The degree of influence of the indicators of manufacturability relative to other indicators was determined by using the apparatus of paired comparisons in decision-making in relation to all structural elements of production costs.

Discussion and Conclusion. The approach to the implementation of this design procedure was described, which provided taking into account the composition and capabilities of processing equipment of a particular production and the actual production situation. The developed formalized models make it possible to comprehensively predict the impact of the manufacturability indicators of parts on the performance effectiveness of machining systems during their manufacture.

Keywords: production planning, product manufacturability assessment, quantitative indicators of manufacturability, machining production systems, production efficiency

Acknowledgements: the authors would like to thank the editorial board of the journal and the reviewer for attentive attitude to the article and suggestions made that helped to improve its quality.

For citation. Bochkaryov PYu, Korolev RD, Bokova LG. Comprehensive Assessment of the Manufacturability of Products. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):155–168. https://doi.org/10.23947/2687-1653-2023-23-2-155-168

Научная статья

Комплексная оценка производственной технологичности изделий

П.Ю. Бочкарев[1,2], Р.Д. Королев[1,3], Л.Г. Бокова[4]

[1] Камышинский технологический институт — филиал Волгоградского государственного технического университета, г. Камышин, Российская Федерация

[2] Вавиловский университет, г. Саратов, Российская Федерация

[3] ООО «ЭКС-ФОРМА», г. Саратов, Российская Федерация

[4] Саратовский государственный технический университет имени Ю. А. Гагарина, г. Саратов, Российская Федерация

✉ bpy@mail.ru

Аннотация

Введение. Оценка производственной технологичности изготавливаемых изделий — этап технологической подготовки и ключевой аспект развития современных производственных механообрабатывающих систем — является актуальной задачей современного машиностроения. В этой связи теоретические и практические исследования по разработке методических подходов к определению весовой значимости количественных показателей при оценке производственной технологичности деталей являются весьма актуальными. Целью представленной работы явилась разработка метода оценки, нацеленного на повышение качества обработки деталей и эффективности функционирования многономенклатурных производственных систем на основе разработки дополнительных количественных показателей оценки производственной технологичности.

Материалы и методы. Для оценки влияния количественных производственных показателей, связанных с затратами времени при простое оборудования, создана модель прогнозирования потоков событий поступления партий деталей на изготовление на определенную операцию и потоков обработанных деталей с использованием аппарата теории массового обслуживания. Такой подход позволяет учесть, как конструкторско-технологические характеристики деталей, особенности конкретной производственной системы, так и складывающуюся производственную ситуацию.

Результаты исследования. Посредством оценки возможного влияния на составляющие при расчете штучно-калькуляционного времени ($T_{шт.к.}$) на уровне технологической операции была определена степень влияния показателей технологичности. Установлены взаимосвязи между показателями технологичности и затратами по всем статьям технологической себестоимости обработки заготовки ($C_{ОП}$), а также затратами, связанными с организационными простоями оборудования ($C_{пр.о.i}$). С помощью применения аппарата парных сравнений при принятии решений применительно ко всем структурным элементам производственных затрат определена степень влияния показателей технологичности относительно других показателей.

Обсуждение и заключение. Описан подход к выполнению данной проектной процедуры, позволяющий учитывать состав и возможности технологического оборудования конкретного производства и реально складывающуюся производственную ситуацию. Разработанные формализованные модели позволяют комплексно спрогнозировать влияние показателей технологичности деталей на эффективность функционирования механообрабатывающих систем при их изготовлении.

Ключевые слова: технологическая подготовка производства, оценка технологичности изделий, количественные показатели производственной технологичности, механообрабатывающие производственные системы, эффективность функционирования производства

Благодарности: авторы выражают благодарность редакции и рецензенту за внимательное отношение к статье и высказанные предложения, которые позволили повысить ее качество.

Для цитирования. Бочкарев П.Ю. Королев Р.Д., Бокова Л.Г. Комплексная оценка производственной технологичности изделий. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):155–168. https://doi.org/10.23947/2687-1653-2023-23-2-155-168

Introduction. The development of machine-building production under modern conditions is impossible without a serious increase in scientific research related to the development of theory and methodological principles of formalization of all stages of product production, which are the basis of future intelligent support systems for the creation and manufacture of technical objects. In this aspect, the solution to the tasks of design and process planning is a challenge [1, 2]. Despite numerous works in this area, automated systems that provide for even minor functional actions of designers and technologists related to the implementation of creative design solutions have not been created yet.

A prerequisite for the production planning of the effective functioning of machining systems is monitoring and analysis of the current production situation, as well as information about the condition of equipment and engineering support. Rational production decisions can be made only based on full knowledge of the above. Even an experienced technologist is not able to collect and analyze such a large amount of information. Therefore, decisions are often made subjectively and unreasonably, the design of processes and their implementation are spaced out in time, and the use of computing systems is hindered by the lack of models describing the process of production planning.

R&D works on the creation of a system for planning multiproduct processes are devoted to solving the tasks formulated [3]. They are based on a conceptual approach to the formalization of all design procedures for providing the process planning of machining industries, taking into account specific features, capabilities of equipment and tooling. One of such design procedures is the assessment of the manufacturability of products, which is traditionally given insufficient attention. The role of this stage is significantly underestimated.

All performance indicators of the production system operation are determined by the complexity of the products and the degree of production capacity. There is often inconsistency between these two indicators, which causes the inability to meet the requirements for the quality of products, downtime, and irrational use of equipment. Objective data on the feasibility of manufacturing products in a specific production system, along with known tasks and methods of solving them, should be obtained precisely when evaluating the manufacturability of products.

Materials and Methods. Scientific studies on creating formalized models for establishing links between engineering and design tasks for the preparation of industrial machine-building systems are of great importance. Due to the increasing global rivalry in the manufacturing sector, the primary task is to increase the efficient operation of equipment during the implementation of production processes, taking into account compliance with the specified requirements for the quality of parts, which, in turn, are installed during the design process.

The challenges of the modern conditions of the operation of industrial complexes involve providing the manufacturability of products. Currently, methods for assessing the manufacturability of products, taking into account the need for compliance with the requirements of standards, directly depend on the qualification of the technologist (designer) and their knowledge. This approach does not guarantee making reasonable engineering decisions and hinders the automation of project procedures.

The assessment of manufacturability as a stage of pre-production is carried out to establish the relationship between the costs of manufacturing the product and its design features. The results of such an assessment are often contradictory, there is no complete mathematical description of the procedure for its implementation.

To resolve the current situation, it seems appropriate to implement the following steps in practice [4–6]:

– establishment of relative weight characteristics of manufacturability indicators based on the parameters of products. The solution to this problem at the stages of development of working design documentation, when there are no engineering solutions for manufacturing, is difficult to implement, but paramount;

– development of the existing range of quantitative indicators for the implementation of the procedure for assessing manufacturability; they should provide taking into account specific approaches to the production planning for particular industrial complexes.

The creation of methodological support for the design procedure of assessing the manufacturability of products should be based on an extensive design and engineering database that takes into account its structure and the relationship between the elements of models used in the design and implementation of the processes. The planning system of multiproduct processes meets these requirements and enables, along with the possibility of evaluating known and used quantitative indicators in production, to offer new ones [7].

In accordance with the principles laid down when creating a process planning system, the key performance criterion is the operating time of the production system for manufacturing the whole set of products. It includes all the costs of the production cycle and is directly related to the cost of production of parts. Given this situation, the authors propose an approach that provides a conclusion about the significance of these indicators for specific production conditions. The approach is based on the establishment of relationships between the elements included in the estimation of the cost of manufacturing parts, and quantitative indicators of manufacturability.

Research Results. The sequence of implementation of the developed approach includes several design procedures that take into account both the design features of the parts being processed and the organizational and technological features of the production system, including the composition and capabilities of the equipment, as well as the specifics of the program of manufactured products.

Initially, at the level of the process operation, the degree of influence of the manufacturability indicators was established by assessing the possible impact on the components when making the time per piece calculation ($T_{шт.к.}$). Figure 1 shows the structure detail ($T_{шт.к.}$) for the turning operation, through which the analysis was carried out and the possibilities of the impact of production performance indicators on each individual value in the calculations ($T_{шт.к.}$) were established. Similar studies, which provide establishing analytical dependences between quantitative indicators of production manufacturability and structural elements of process operations, were performed for other groups of process facilities.

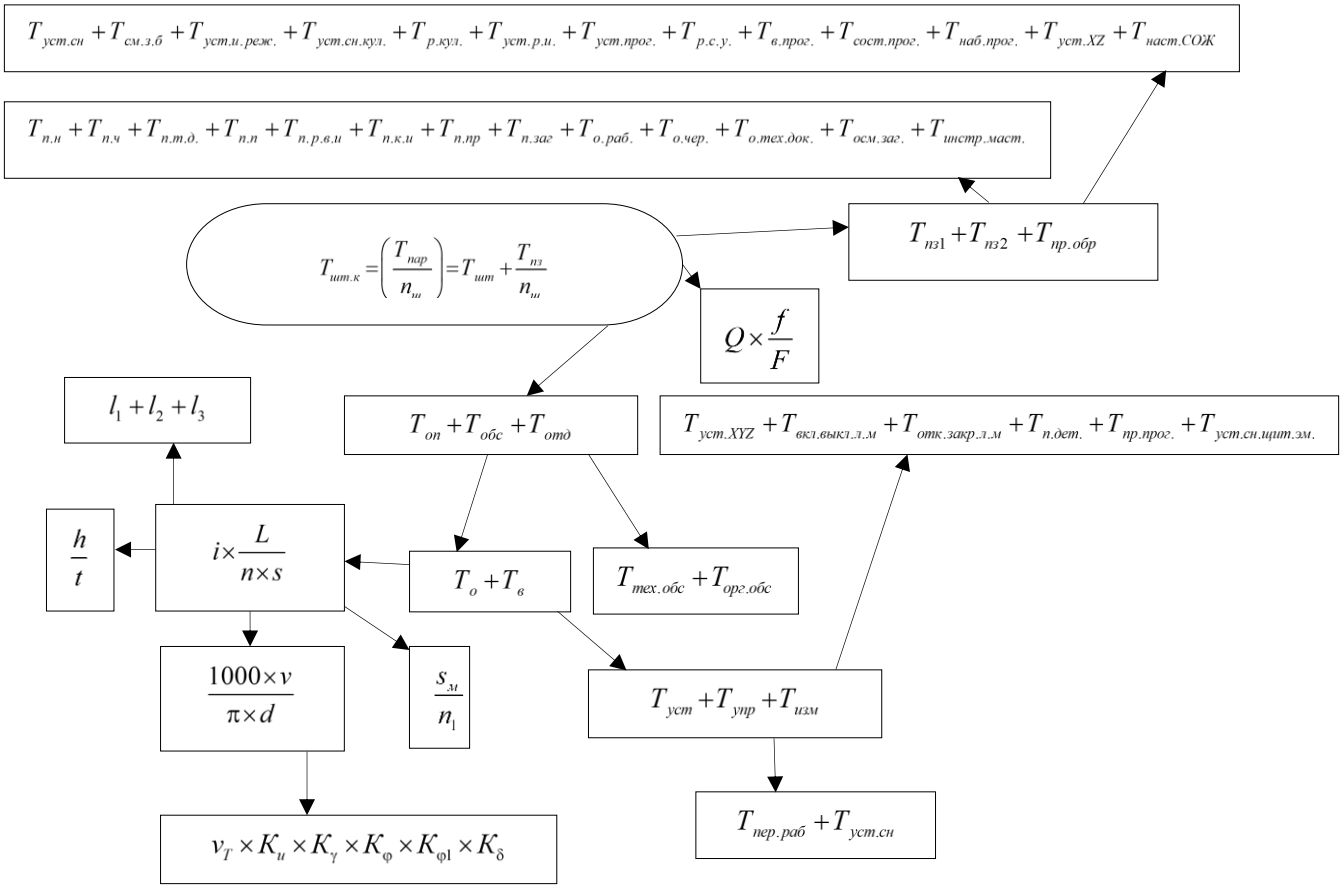


Fig. 1. Block diagram $T_{шт.к.}$ under turning

The interrelations between the manufacturability indicators and the expenses for all items of the production cost of machining workpiece $C_{ОП}$ (Fig. 2), as well as the expenses associated with organizational downtime of equipment, are established $C_{пр.о.i.}$(Fig. 3).

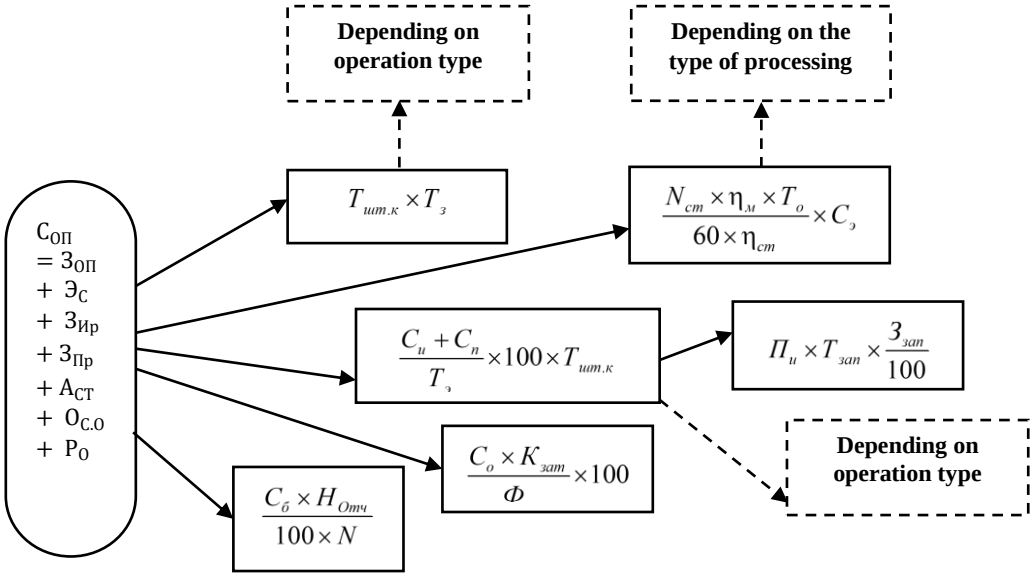


Fig. 2. Components $C_{ОП}$ — expenses for all items of the production cost

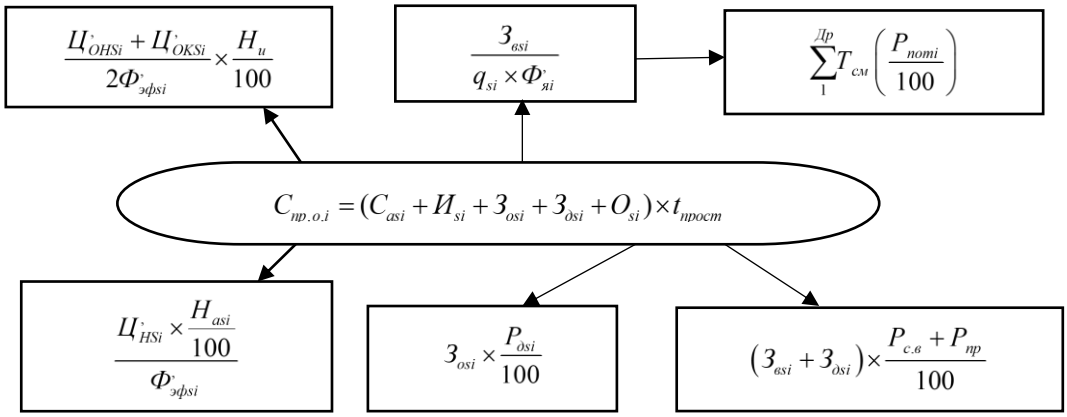


Fig. 3. Costs associated with organizational downtime of process facilities $C_{пр.о.i.}$

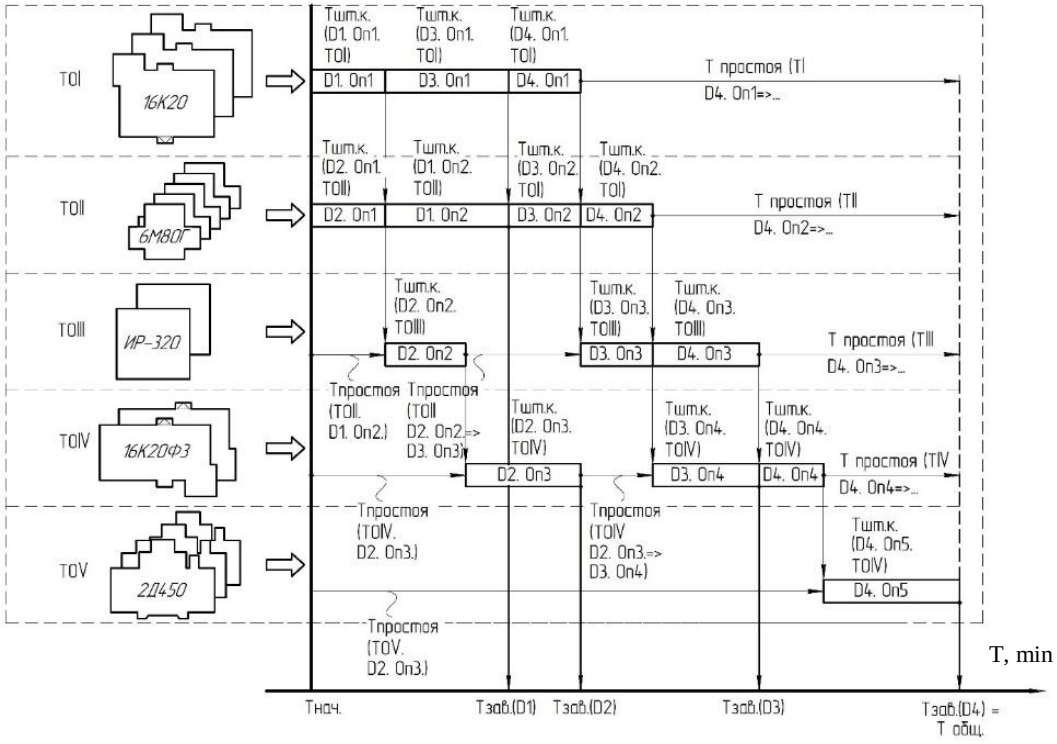


Fig. 4. Gantt chart in process implementation

To assess the impact of quantitative production indicators associated with equipment downtime, a model has been created for predicting event flows of delivery of batches of parts for manufacturing for a specific operation and flows of processed parts using the queuing theory apparatus. This technical approach is used in the process planning system [8, 9]. As an example, Figure 4 shows the results in the form of a Gantt chart. This approach enables to take into account the design-engineering characteristics of parts, the features of a specific production system and the emerging production situation.

Figure 5 shows an enlarged diagram of the structure for determining production costs under manufacture of parts, used to assess the specific weight of quantitative indicators for assessing manufacturability in the process planning system. The analysis of the possibility of the influence of each indicator on the efficiency of the entire production system in the manufacture of a batch of selected parts for specific production conditions was carried out.

The results of the presented analysis and the established relationships between the manufacturability indicators and the efficiency of machining systems allowed us to move on to solving the issue of establishing the significance of quantitative indicators of industrial manufacturability. The presented fragment (Fig. 6) contains information about the above links in relational form and is supplemented with information about the specific weight of cost elements (as a percentage). The data are obtained on the basis of statistical processing of the results of the real production system operation. In the absence of such information, it is possible to use general machine-building or industry-specific data.

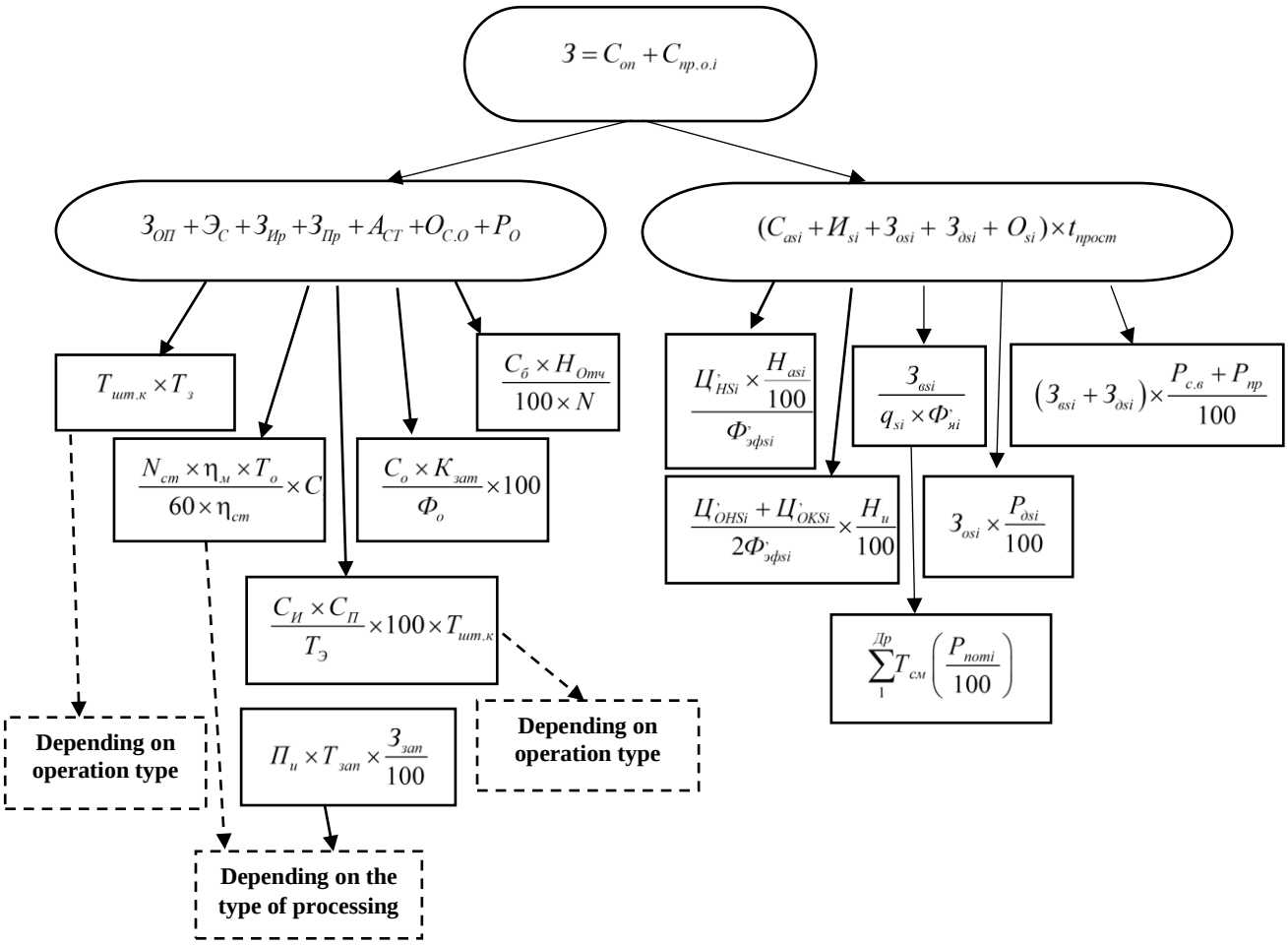


Fig. 5. Enlarged scheme for determining production costs

1	2	3	4	5	6	7	8	9	10	Processability indices						
										11	12	13	14	15	16	17
										Machinability index	Part design complexity indicator	Coefficient of accuracy and surface roughness of the part	Indicator of unification of structural elements	Material usage rate	Indicator of the possibility of manufacturing a given range of parts	Indicator of using capabilities of production systems
3	СО (78 %)	ЗОП (32 %)	$T_{шт.к.}$ (29 %)	$T_{шт}$ (25 %)	Топ (22 %)	То (14 %)	n	v	v_T	+	+	+	+	+	+	+
									K_u	+	+	+	+	+	+	+
									K_γ	+	+	+	+	+	+	+
									K_φ	+	+	+	+	+	+	+
									$K_{\varphi 1}$	+	+	+	+	+	+	+
								d	–	+	+	+	+	+	+	+
								π	–	+	+	+	+	+	+	+

Fig. 6. Fragment of the impact of quantitative indicators of production manufacturability on cost components

To establish the weighting significance of the indicators of production manufacturability, it remains only to solve the problem of determining their impact directly on each element in the presented structure of production costs. The solution to this problem was carried out using the apparatus of paired comparisons in decision-making (the iterative Berge's process [10]). This method provided determining the degree of impact of manufacturability indicators with other indicators in relation to all structural elements of production costs. Table 1 shows a pairwise comparison of technological performance indicators relative to the basic time as an example (*To*).

Table 1

Pairwise comparison of manufacturability indicators relative to basic time (*To*)

To (14 %)														
Measured indicators	11	12	13	14	15	16	17	19	20	21	22	23	Total:	Weight indicators
11	–	2	2	1	0	0	0	1	2	1	2	2	13	0.098485
12	0	–	1	0	0	2	0	0	1	1	2	2	9	0.068182
13	0	1	–	0	1	1	0	0	2	2	0	0	7	0.05303
14	1	2	2	–	1	1	1	1	1	1	1	1	13	0.098485
15	2	2	2	1	–	1	1	1	1	1	2	2	16	0.121212
16	2	2	1	1	0	–	0	0	1	1	0	0	8	0.060606
17	2	0	1	1	1	1	–	1	1	1	0	0	9	0.068182
19	0	2	2	1	1	2	2	–	1	1	0	0	12	0.090909
20	0	2	0	1	1	1	1	1	–	1	1	1	10	0.075758
21	1	1	0	1	1	1	1	1	1	–	1	1	10	0.075758
22	0	0	2	1	2	2	2	1	1	1	–	1	13	0.098485
23	0	0	2	1	0	2	2	2	1	1	1	–	12	0.090909
												Sum	132	1

Industrial testing and approbation of the developed models was carried out under the conditions of "GAZPROMMASH" LLC, specializing in the batch production of direct-acting gas heaters with an intermediate coolant and a modified series of stations, regulators, filter blocks and valves, high-pressure valves. The initial data for the experiments were: a generated and completed database containing information on the process capabilities of the

equipment (production unit No. 1), information on the actual condition and technical and economic characteristics of the site performance (Tables 2–6), the program of manufactured parts (drawings of individual parts are shown in Fig 7).

Table 2

Ratio of the components of production costs
(site no. 1, LLC “GAZPROMMASH” LLC)

$З_{Общ.}$ = 5,682 (100 %)										
$C_{ОП}$ = 4,662 Rub. (78 %)					$C_{пр.о.i}$ = 1,020 Rub. (22 %)					
$З_{ОП}$ 1,872 Rub. (32 %)	$Э_С$ 710 Rub. (14 %)	$З_{И.Р}$ 534 Rub. (6 %)	$З_{Пр}$ 4 Rub. (0,1 %)	$А_{СТ}$ 1,545 Rub. (26 %)	$И_{si}$ 4 Rub. (0.1 %)	$С_{asi}$ 3,806 Rub. (2 %)	$З_{osi}$ 12,500 Rub. (5 %)	$З_{osi}$ 3,125 Rub. (1.9 %)	$О_{si}$ 21,250 Rub. (11 %)	$t_{пр.i}$ (2 %)

Table 3

Costs for the worker's salary for performing the operation

$З_{ОП}$ = 1,872 Rub. (32 %)							
$Т_{шт.к.}$ (29 %)							$Т_{з.}$ (3 %)
$Т_{шт.}$ (25 %)				$Т_{п.з.}$ (4 %)			–
$Т_{оп.}$ (22 %)	$Т_{обс.}$ (2 %)		$Т_{отд.}$(1%)	$Т_{п.з.1}$ (1.5 %)	$Т_{п.з.2}$ (1 %)	$Т_{пр.}$ (0.5 %)	–
See Table 9	$Т_{тех.обс}$ *(1.0 %)*	$Т_{орг.обс}$ *(1.0 %)*	–	*See Table 8*	*See Table 9*	–	–

Table 4

Costs for manufacturing the product

$Т_{оп.}$ (22 %)												
$Т_{о.}$ (14 %)	$Т_{в.}$ (8 %)										$Т_{тех.обс.}$ (1 %)	$Т_{орг.об.}$ (2 %)
–	$Т_{уст.}$ (4 %)		$Т_{упр.}$ (3 %)						$Т_{изм.}$ (1 %)		–	–
–	$Т_{пер.раб.}$ (2 %)	$Т_{уст.сн.}$ (2 %)	$Т_{уст.XYZ}$ (0.5 %)	$Т_{вк.\ Выкл\ л.м.}$ (0.5 %)	$Т_{откр.\ закр.\ л.м.}$ (0.5 %)	$Т_{п.\ дет.}$ (0.5 %)	$Т_{пр.про.}$ (0.5 %)	$Т_{уст.сн.цц\ т.эм.}$ (0.5 %)	–		–	–

Table 5

Costs for organizational preparation

$Т_{п.з.1}$ (1.5 %)												
$Т_{п.н}$ (0.2 %)	$Т_{п.ч}$ (0.2 %)	$Т_{п.т.о}$ (0.1 %)	$Т_{п.п}$ (0.1 %)	$Т_{п.р.в.и}$ (0.1 %)	$Т_{п.к.и}$ (0.1 %)	$Т_{п.пр}$ (0.1 %)	$Т_{п.заг}$ (0.1 %)	$Т_{о.раб.}$ (0.1 %)	$Т_{о.чер.}$ (0.1 %)	$Т_{о.тех.док}$ (0.1 %)	$Т_{осм.заг}$ (0.1 %)	$Т_{инст.маст.}$ (0.1 %)

Table 6

Costs for setting up the machine

$Т_{п.з.2}$ (1.0 %)												
$Т_{ус.сн}$ (0.1 %)	$Т_{см.з.б}$ (0.1 %)	$Т_{уст.и.реж}$ (0.1 %)	$Т_{уст.сн.кул}$ (0.1 %)	$Т_{р.кул}$ (0.1 %)	$Т_{уст.р.и}$ (0.1 %)	$Т_{уст.прог}$ (0.05 %)	$Т_{р.с.у}$ (0.05 %)	$Т_{в.прог}$ (0.05 %)	$Т_{сост.прог}$ (0.1 %)	$Т_{наб.прог}$ (0.05 %)	$Т_{уст.XZ}$ (0.05 %)	$Т_{наст.СОЖ}$ (0.05 %)

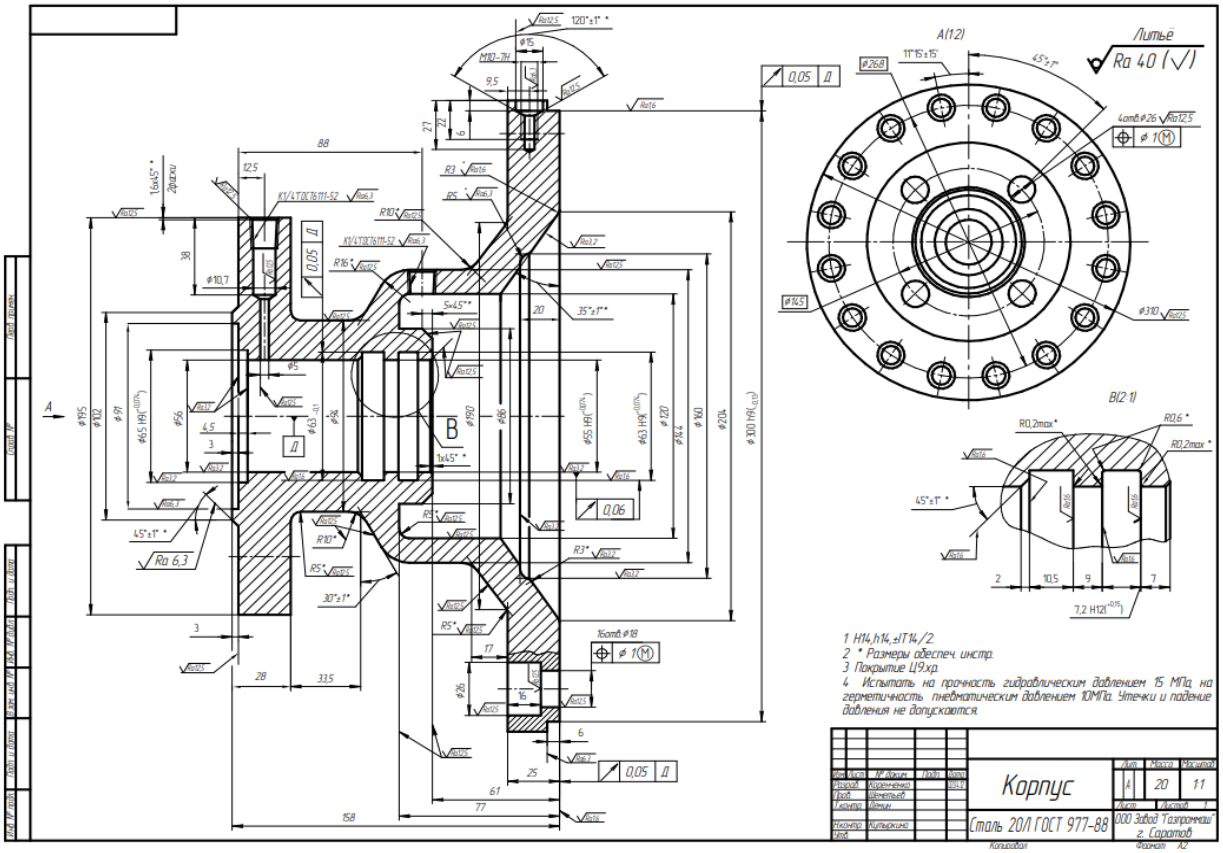


Fig. 7. Examples of design drawings of machined parts ("GAZPROMMASH" LLC) (part 1)

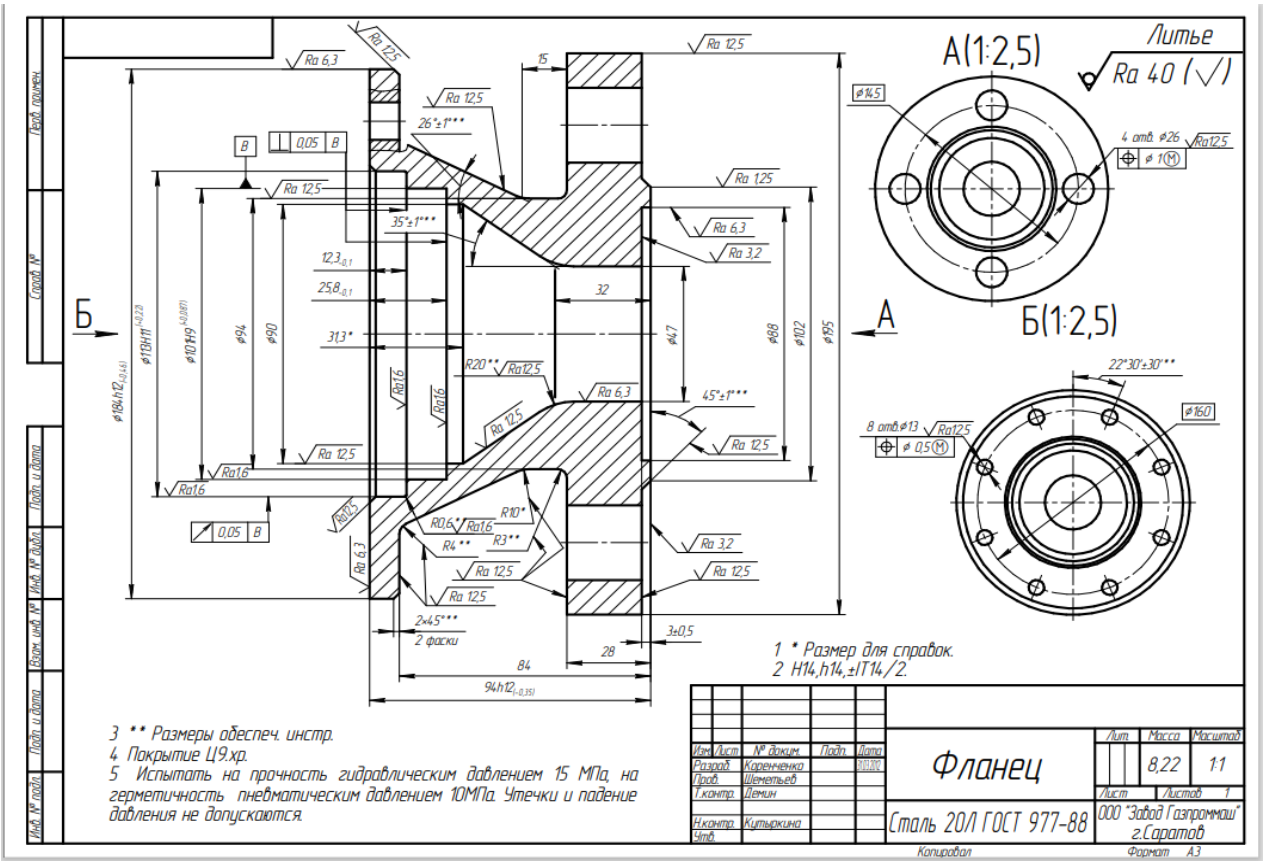


Fig. 8. Examples of design drawings of machined parts ("GAZPROMMASH" LLC) (part 2)

The results of calculations of the degree of relative influence of quantitative indicators of the manufacturability of parts on individual structural elements of production costs (site No. 1, "GAZPROMMASH" LLC) are presented in Tables 7–10.

Table 7

Impact of manufacturability quantitative indicators *To*
(basic (process) time for the manufacture or processing of a unit of product)

Estimated indicators	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.9	1.10	1.11	1.12	1.13	Total	Weight indicators
1.1	–	2	2	1	0	0	0	1	2	1	2	2	13	0.098485
1.2	0	–	1	0	0	2	0	0	1	1	2	2	9	0.068182
1.3	0	1	–	0	1	1	0	0	2	2	0	0	7	0.05303
1.4	1	2	2	–	1	1	1	1	1	1	1	1	13	0.098485
1.5	2	2	2	1	–	1	1	1	1	1	2	2	16	0.121212
1.6	2	2	1	1	0	–	0	0	1	1	0	0	8	0.060606
1.7	2	0	1	1	1	1	–	1	1	1	0	0	9	0.068182
1.9	0	2	2	1	1	2	2	–	1	1	0	0	12	0.090909
1.10	0	2	0	1	1	1	1	1	–	1	1	1	10	0.075758
1.11	1	1	0	1	1	1	1	1	1	–	1	1	10	0.075758
1.12	0	0	2	1	2	2	2	1	1	1	–	1	13	0.098485
1.13	0	0	2	1	0	2	2	2	1	1	1	–	12	0.090909

Table 8

Impact of quantitative indicators of manufacturability *Tynp*
(time to supply tool to workpiece)

Estimated indicators	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.9	1.10	1.11	1.12	1.13	Total	Weight indicators
1.1	–	2	2	0	1	1	2	2	2	2	2	2	18	0.136364
1.2	0	–	0	0	0	1	1	1	1	1	1	1	7	0.05303
1.3	0	2	–	2	1	1	1	1	1	2	2	2	15	0.113636
1.4	2	2	0	–	0	1	1	1	1	1	0	0	9	0.068182
1.5	1	2	1	2	–	2	2	1	2	2	2	2	19	0.143939
1.6	1	1	1	1	0	–	1	1	1	1	1	1	10	0.075758
1.7	0	1	1	1	0	1	–	1	1	1	1	1	9	0.068182
1.9	0	1	1	1	1	1	1	–	2	2	2	2	14	0.106061
1.10	0	1	1	1	0	1	1	0	–	1	2	2	10	0.075758
1.11	0	1	0	1	0	1	1	0	1	–	2	2	9	0.068182
1.12	0	1	0	2	0	1	1	0	0	0	–	1	6	0.045455
1.13	0	1	0	2	0	1	1	0	0	0	1	–	6	0.045455

Table 9

Impact of quantitative indicators of manufacturability *Tn.з.1*
(time limit for organizational preparation)

Estimated indicators	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	Total	Weight indicators
1.1	–	2	2	1	1	1	2	2	2	2	2	2	2	21	0.135484
1.2	0	–	0	1	0	0	0	0	0	0	0	0	0	1	0.006452
1.3	0	2	–	1	1	1	1	1	1	1	1	0	0	10	0.064516

Estimated indicators	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	Total	Weight indicators
1.4	1	1	1	–	0	0	0	0	0	0	0	0	0	3	0.019355
1.5	1	2	1	0	–	2	2	2	2	2	2	2	2	20	0.129032
1.6	1	2	1	2	0	–	2	2	2	2	2	1	1	18	0.116129
1.7	0	2	1	2	0	0	–	1	1	1	1	1	1	11	0.070968
1.8	0	2	1	2	0	0	1	–	1	1	1	1	1	11	0.070968
1.9	0	2	1	2	0	0	1	1	–	1	1	1	1	11	0.070968
1.10	0	2	1	2	0	0	1	1	1	–	1	1	1	11	0.070968
1.11	0	2	1	2	0	0	1	1	1	1	–	1	1	11	0.070968
1.12	0	2	2	2	0	1	1	1	1	1	1	–	1	13	0.083871
1.13	0	2	2	2	1	1	1	1	1	1	1	1	–	14	0.090323

Table 10

Impact of quantitative indicators of manufacturability *Acm*
(costs for the use of process facilities)

Estimated indicators	1.8	1.9	1.10	1.11	1.12	1.13	Total	Weight indicators
1.8	–	1	1	2	1	1	6	0.2
1.9	1	–	2	1	1	1	6	0.2
1.10	1	0	–	0	0	0	1	0.033333
1.11	0	1	2	–	1	1	5	0.166667
1.12	1	1	2	1	–	1	6	0.2
1.13	1	1	2	1	1	–	6	0.2

1.1 — material machinability index;

1.2 — part design complexity index;

1.3 — coefficient of accuracy and surface roughness of the part;

1.4 — indicator of unification of structural elements;

1.5 — material usage rate;

1.6 — indicator of the possibility of manufacturing a given range of parts;

1.7 — indicator of the use of production system capabilities;

1.8 — indicator of the manufacturability of the part by the uniformity of process facilities;

1.9 — indicator of predicting the level of loading of process facilities when processing a given range of parts;

1.10 — indicator of multivariate decision-making when designing a process;

1.11 — indicator of multivariate decision-making in the implementation of processes;

1.12 — indicator of the manufacturability of the part, reflecting the possibility of observing the principle of unity of bases under the process development in terms of the surface of the part that is the main design base;

1.13 — indicator of the manufacturability of the part, reflecting the possibility of observing the principle of unity of bases in the development of the technological process in terms of the surfaces of parts that are auxiliary design bases.

Based on the presented models and previously known dependences of the calculation of quantitative indicators, an assessment of the manufacturability of parts was carried out. At the same time, the software developed and registered by the authors was used. Thus, taking into account the information about the real state of the production system, the adjustment of design documentation, range, sequence of implementation of the manufacture of individual groups of parts and production planning was performed. A comparative analysis of the calculation results is presented in Table 11.

Table 11

Comparison results

No.		1 option	2 option	Performance
1	Total time for the manufacture of products of 20 items with an annual production program of 16,600 pcs.	61193.42 h	53073.35 h	15 %
2	Estimated number of equipment participating in the process	23	17	35 %
3	Operation factor	0.67	0.72	7 %

Discussion and Conclusion. The results of the presented theoretical studies and their approbation under real production conditions allowed us to propose a method for assessing the manufacturability of parts. It provides for a comprehensive assessment based on the developed analytical dependences for determining the weighting coefficients that characterize the significance of each indicator of manufacturability from the standpoint of the efficiency of the machining system. A distinctive feature and scientific novelty of the work is the consideration of the actual emerging production situation when assessing manufacturability. This makes it possible to use this design procedure not only traditionally at the initial stages of production planning, but also at the stages of implementation of processes for the purpose of rational organization of the production.

The developed formalized models create the basis for complete sequential automation of design actions when evaluating the manufacturability of products, and provide prerequisites for constructing a promising intelligent system of predicting the efficiency of manufacturing parts in a particular production and making informed organizational and engineering decisions.

References

1. Bazrov BM. *Mechanical Engineering Technology Fundamentals*. Moscow: Mashinostroenie; 2005. 736 p. URL: https://studizba.com/files/show/djvu/1875-1-bazrov-b-m--osnovy-tehnologii.html (accessed: 12.04.2023) (In Russ.).

2. Suslov AG. *Mechanical Engineering Technology*. Moscow: Knorus; 2013. 336 p. (In Russ.)

3. Bochkarev PYu. Systemic Presentation of the Mechanical Processing Technology Planning. *Tekhnologiya Mashinostroeniya*. 2002;1:10–14. (In Russ.).

4. Bazrov BM. Ensuring the Processability of the Product Design. *Science Intensive Technologies in Mechanical Engineering*. 2020;8(110):18–22. https://doi.org/10.30987/2223-4608-2020-8-18-22 (In Russ.).

5. Vartanov MV, Chushenkov II. Methodology for Evaluating the Manufacturability of Engineering Products. *STANKOINSTRUMENT*. 2019;2(015):14–23. https://doi.org/10.22184/2499-9407.2019.15.02.14.22 (In Russ.).

6. Bazrov BM, Troitskii AA. The System of the Product Design Manufacturability Coefficients. *STIN*. 2020;3:22–26. (In Russ.).

7. Bokova LG, Bochkarev PYu. Development of Indicators for Assessment of Parts Operability in the System of Planning of Engineering Mechanical Processing. *Frontier Materials and Technologies*. 2015;1(3):29–35. https://doi.org/10.18323/2073-5073-2015-3-29-35

8. Mitin SG, Bochkarev PYu, Bokova LG. Automation of the Product Manufacturability Assessment within the Multi-Nomenclature Production Systems. *Science Intensive Technologies in Mechanical Engineering*. 2014;9(39):45–48. (In Russ.).

9. Ivanov AA, Bochkarev PYu. Formalizing the Description and Methods for Optimization of Mechanical Treatment Technologies within the System of Planning Technological Processes. *Vestnik Saratov State Technical University*. 2015;3(80):76–85. URL: http://lib.sstu.ru/doc/ibo/vestniki/2015/3(80)_2015.pdf (accessed: 25.03.2023).

10. Rastegayev EV. Technological Effectiveness Assessment Indicators of a Product Design at Small-Lot and Medium-Batch Manufacturing. *Vestnik of P.A. Solovyov Rybinsk State Aviation Technical University*. 2021;1:32–35.

Received 09.04.2023
Revised 05.05.2023
Accepted 12.05.2023

About the Authors:

Peter Yu. Bochkarev, Dr.Sci. (Eng.), Professor of the Mechanical Engineering Technology and Applied Mechanics Department, Kamyshin Technological Institute, VSTU branch (6 a, Lenina St., Kamyshin, Volgograd Region, 403805, RF), Professor of the AIC Engineering Support Department, Saratov State Vavilov Agrarian University (1, Teatralnaya sq., Saratov, 410012, RF), ORCID, AuthorID, bpy@mail.ru

Richard D. Korolev, Design Engineer, OOO PCF "Ex-Form" (13, Shkolnaya St., Berezina Rechka village, Saratov, 410512, RF), ORCID, AuthorID, rihardkorolev@mail.ru

Larisa G. Bokova, Cand.Sci. (Eng.), Associate Professor of the Mechanical Engineering Department, Yuri Gagarin State Technical University of Saratov (77, Politechnicheskaya St., Saratov, 410054, RF), ResearcherID, ORCID, AuthorID, bokovalg@mail.ru

Claimed contributorship:

PYu Bochkarev: academic advising; development of a methodological approach to assessing the impact of quantitative indicators of manufacturability; analysis of the research results; revision of the text.

RD Korolev: development of models of the relationship between indicators of manufacturability and production time costs, industrial approbation and processing of experimental results; text preparation.

LG Bokova: determination of composition and calculation of indicators of production manufacturability.

Conflict of interest statement: the authors do not have any conflict of interest.

All authors have read and approved the final manuscript.

Поступила в редакцию 09.04.2023
Поступила после рецензирования 05.05.2023
Принята к публикации 12.05.2023

Об авторах:

Петр Юрьевич Бочкарёв, доктор технических наук, профессор кафедры технологии машиностроения и технической механики Камышинского технологического института (филиал) Волгоградского технического университета (403874, РФ., Волгоградская обл., г. Камышин, ул. Ленина, д. 6а), профессор кафедры технического обеспечения АПК Вавиловского университета (РФ, г. Саратов, Театральная пл., 1), ORCID, AuthorID, bpy@mail.ru

Рихард Джахангилович Королёв, инженер конструктор завода ООО ЭКС-ФОРМА (410512, РФ, г. Саратов, с. Березина речка, ул. Школьная, д. 13), ORCID, AuthorID, rihardkorolev@mail.ru

Лариса Геннадьевна Бокова, кандидат технических наук, доцент кафедры технологии машиностроения Саратовского государственного технического университета им. Ю. А. Гагарина (410054, РФ, г. Саратов, ул. Политехническая, д. 77), ResearcherID, AuthorID, ORCID, bokovalg@mail.ru

Заявленный вклад:

П.Ю. Бочкарев — научное руководство, разработка методического подхода оценки влияния количественных показателей технологичности, анализ результатов исследований, доработка текста.

Р.Д. Королев — разработка моделей взаимосвязи показателей технологичности и затрат производственного времени, промышленная апробация и обработка результатов экспериментов, подготовка текста.

Л.Г. Бокова — определение состава и оценка показателей производственной технологичности.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Все авторы прочитали и одобрили окончательный вариант рукописи.

INFORMATION TECHNOLOGY, COMPUTER SCIENCE AND MANAGEMENT
ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

UDC 004.896

Original article

https://doi.org/10.23947/2687-1653-2023-23-2-169-179

Predicting the Behavior of Road Users in Rural Areas for Self-Driving Cars

Sergey A. Ivanov ✉, **Bader Rasheed**

Laboratory of Unmanned Technology, Innopolis University, Innopolis, Russian Federation

✉ se.ivanov@innopolis.ru

Abstract

Introduction. The prediction module generates possible future trajectories of dynamic objects that enables a self-driving vehicle to move safely on public roads. However, all modern prediction methods evaluate their performance only under urban conditions and do not consider their applicability to the domain of rural roads. This work examined the adaptability of existing methods to work under rural unstructured conditions and suggested a new, improved approach.

Materials and Methods. As a solution, we propose to use a neural network that includes the following submodules: a graph-based scene encoder, a multimodal trajectory decoder, and a trajectory filtering module. Another proposed feature is to use an adapted loss function that penalizes the network for generating trajectories that go beyond the drivable area. These elements use standard practices for solving the prediction problem and adapting it to the domain of rural roads.

Results. The presented analysis described the basic features of the prediction module in the rural road domain, showed a comparison of popular models, and discussed its applicability to new conditions. The paper describes the new approach that is more adaptive to the considered domain of study. A simulation of the new domain was performed by modifying existing public datasets.

Discussion and Conclusion. Comparison to other popular methods has shown that the proposed approach provides more accurate results. The disadvantages of the proposed approach were also identified and possible solutions were described.

Keywords: trajectory prediction, behavior prediction, neural networks, self-driving cars, artificial intelligence, autonomous cars

Acknowledgements: the authors appreciate the “Center for Autonomous Technologies”, Innopolis University, for their assistance in conducting the research.

For citation. Ivanov SA, Rasheed B. Predicting the Behavior of Road Users in Rural Areas for Self-Driving Cars. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):169–179. https://doi.org/10.23947/2687-1653-2023-23-2-169-179

Научная статья

Прогнозирование поведения участников дорожного движения в условиях проселочных дорог для беспилотных автомобилей

С.А. Иванов ✉, **Б. Рашид**
Центр беспилотных технологий университета Иннополис, г. Иннополис, Российская Федерация
✉ se.ivanov@innopolis.ru

Аннотация

Введение. Благодаря модулю прогнозирования траекторий движения динамических объектов беспилотный автомобиль способен безопасно двигаться по дорогам общего пользования. Однако все современные методы прогнозирования оценивают производительность только в городских условиях и не рассматривают свою применимость к домену проселочных дорог. Цель данного исследования заключается в анализе адаптивности существующих методов прогнозирования и разработке подхода, который будет демонстрировать лучшую производительность при работе в новых условиях.

Материалы и методы. В качестве решения предлагается использовать нейронную сеть, включающую в себя следующие подмодули: графовый кодировщик сцены, мультимодальный декодировщик траекторий, модуль фильтрации траекторий. Также предлагается применить адаптированную функцию потерь, которая штрафует сеть за генерацию траекторий, выходящих за границы дорожного полотна. Данные элементы задействуют распространённые практики решения задачи прогнозирования, а также адаптируют её для домена проселочных дорог.

***Результаты исследования*.** Проанализированы основные отличия и условия работы модуля прогнозирования в условиях проселочных дорог. Выполнена симуляция нового домена путем модификации существующих наборов данных. Проведено сравнение популярных методов прогнозирования и оценена их применимость к новым условиям. Представлен новый, более адаптивный к новому домену, подход.

Обсуждение и заключение. Проведенное сравнение с другими популярными методами показывает, что предложенное авторами решение обеспечивает более точные результаты прогнозирования. Также были выявлены недостатки предложенного подхода и описаны возможные пути их устранения.

Ключевые слова: прогнозирование траекторий, прогнозирование поведения, нейронные сети, беспилотные автомобили, искусственный интеллект, автономные автомобили

Благодарности: авторы выражают благодарность центру беспилотных технологий университета Иннополис за помощь в проведении исследования.

Для цитирования. Иванов С.А., Рашид Б. Прогнозирование поведения участников дорожного движения в условиях проселочных дорог для беспилотных автомобилей. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):169–179. https://doi.org/10.23947/2687-1653-2023-23-2-169-179

Introduction. The latest achievements in the field of artificial intelligence (AI) are being actively implemented in various areas of activity. One of such achievements is autonomous vehicles (AV). The current research was aimed at creating algorithms that allow AV to move safely on public roads. This will significantly reduce the number of road accidents [1].

The scientific community has already identified the basic modules of an autonomous vehicle. One of them is a system for predicting the future behavior of road users (agents) [2]. A clear understanding of how the environment will develop and in which direction dynamic objects (pedestrians, cars, cyclists) will move is urgently needed for AV to search and use a safe and effective trajectory of movement.

Numerous scientific papers are devoted to the problem of predicting such trajectories [3–12]. However, there is currently no active research on the application of existing methods outside of urban conditions. And this is extremely important, since autonomous cars will be used on country roads, too [13]. Urban conditions are highly structured: cars

mostly follow traffic lanes, and pedestrians move through special zones. In this sense, the area of country roads is the complete opposite, which means that it will have additional difficulties during development. In the given paper, attention is focused on these difficulties: the existing predicting methods and their applicability to new, less structured conditions are considered.

The objective of the study involved:

– analysis of the major differences and working conditions of the prediction module under the conditions of country roads;

– simulation of a less structured country road domain by modifying the existing datasets;

– comparison of modern prediction methods, including their applicability to new conditions;

– description of the new approach and proof of its higher accuracy in comparison to other prediction methods.

Materials and Methods. At first glance, it would seem that the domain presented is a simpler version of urban conditions due to the fact that country roads are characterized by less traffic. However, the absence of complex multi-level junctions, special traffic-free zones, a large number of signs, markings, etc., makes the domain of country roads less structured, i.e., fewer rules and specific traffic patterns increases the randomness and reduces the predictability of the behavior of cars and pedestrians.

The following features of the country road domain influence strongly on the selection of the architecture of the prediction module:

– crossroads. Undoubtedly, they are more simple on country roads in comparison to urban ones, but at the same time this fact of simplicity means that the model must take into account multimodality and assess the probability of choosing each possible direction of movement at the crossroad when the agent approaches it;

– country roads do not have lane markings, pedestrian crossings, bike paths, etc. Instead, the HD map will contain only information about the boundaries of the roadway. Therefore, the stage of encoding the scene should take into account this feature to describe the surrounding context more effectively;

– pedestrians and cyclists will move along the same road with conventional and autonomous vehicles. Therefore, the model should be adaptive for predicting the future trajectories of both cars and pedestrians/cyclists.

The prediction module implies the presence of AV recognition, tracking and localization systems and their accurate operation. The authors of the article use the Argoverse dataset, which stores the required records of the operation of all systems in a convenient form [14].

The dataset consists of recordings of road scenes observed on the streets of Miami and Pittsburgh, USA. Each of the entries contains a local part of the terrain map (lane boundaries, roads, pedestrian crossings) and a list of all recognized agents, including the current position and movement history of each of them. Each of the records is divided into two parts: two seconds of the observation history and three subsequent seconds for which prediction is made (prediction horizon). Data on the future movement of objects is also available and used to calculate the accuracy of prediction methods and model training.

Information about agents is presented in a discrete format. The time interval between measurements is fixed, in this work it is equal to 0.1 seconds (10 Hz).

For each moment of time *t*, the module receives the observation history S_i^p for each detected agent *i*. The observation history consists of the agent's current and past states, where each of the states s_t is a 2D position in the global coordinate system. The authors make the assumption that the height information is redundant.

The dataset also provides access to an HD map that contains information about road borders and roadway, pedestrian crossings. To simulate the domain of country roads, the dataset was modified in such a way as to exclude all information from road maps, except for the boundaries of the roadway *D*. This reduces the amount of information about the road context and complicates the task of prediction.

Hence, the context of the scene is represented as

$$c = \left\{S_1^p, S_2^p, \ldots, S_k^p, D\right\}, \quad (1)$$

where *k* — the total number of tracked agents on the scene.

This approach implies predicting the trajectory for only one agent per execution, therefore further S_i^p is treated as S^p for simplification. To generalize the model for all recognized agents, it is required to repeat the proposed approach for all *k* agents on the scene. The agent for which the prediction is currently being made is considered a target agent.

To assess the accuracy of prediction methods, the dataset contains recorded future trajectories S^f for each target agent:

$$S^f = \{s_1, s_2, \ldots, s_H\}, \quad (2)$$

where *H* indicates the number of next time steps. In this case, parameter *H* will be equal to 30, since the planning horizon is three seconds with a sampling frequency of 10 Hz.

The domain of the prediction module is multimodal, i.e., the future behavior of agents may differ significantly in absolutely identical traffic situations. Let us say, a car approaching a crossroad may continue straight ahead or make a turn. To take this into account, it is required to generate *M* possible future trajectories and *M* probabilities of each of them at the model output.

Therefore, the purpose of the prediction module is to create function f, that takes the context of the scene *c* as input and generates *M* pairs of possible future trajectories and their probabilities:

$$f(c) = \left\{S_1^f, S_2^f, \ldots, S_M^f, p_1, p_2, \ldots, p_M\right\}. \quad (3)$$

Here, at least one generated trajectory S^f should be as close as possible to the real trajectory S^f, and the probability of its execution *p* should be close to unity.

Model architecture. The proposed approach involves the use of a neural network consisting of submodules of scene encoding, decoding and filtering trajectories. The architecture of the system is shown in Figure 1.

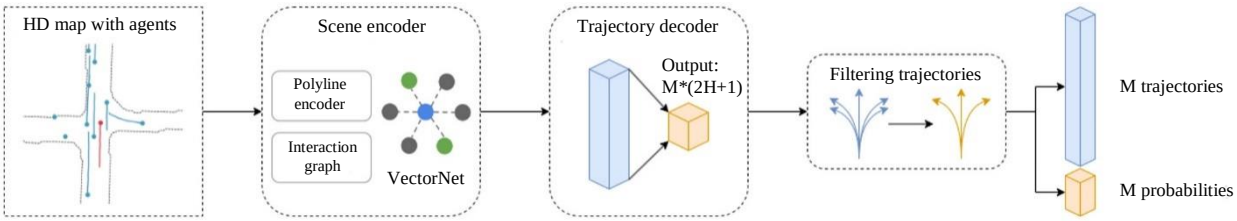


Fig. 1. System architecture

A neural network adapted to the new conditions, based on a vector representation, is responsible for encoding the scene. This selection is due to the fact that on country roads, the HD map will contain a limited amount of information (only the roadway boundaries and the history of observations of dynamic objects). Popular methods represent the context of a road scene *c* in an image format and process it using convolutional neural networks. However, vector coding avoids the overhead associated with image generation [4–5].

The presented encoder is based on the VectorNet model, but its input data format has been modified to receive information only about the boundaries of the roadway and the state of the agents. [3]. This encoder represents the boundaries of the road and the state of agents using polylines, which are further processed by a graph neural network. This provides encoding the interaction between polylines. Details of the implementation are described in paper [3].

A trajectory decoder is a task of regressing several possible trajectories and generating a set of probabilities. To solve this problem, a multilayer perceptron model is used. The decoder implementation is inspired by the MTP model [4], however, the authors of the article propose a different formula for calculating the best trajectory *m** from the set of *M* trajectories. It is also proposed to use an additional mechanism that penalizes the model for predictions that go beyond the area of movement.

The authors of the original MTP model propose to train a multilayer perceptron using the loss function that represents the sum L_{reg} and L_{class}, where:

$$L_{reg} = L\left(S^f, S_{m^*}^f\right) \quad (1)$$

$$\mathrm{L}_{class} = -\sum_{m=1}^{M} I_{m=m^*} \log p_m \tag{5}$$

In this case, L_{reg} — mean-square error between the real trajectory S^f and the best trajectory m^* of M generated.

$$\mathrm{L}\left(S^f, S^f_{m^*}\right) = \frac{1}{H}\sum_{i=1}^{H}\left(s_i - s^*_i\right)^2, \tag{6}$$

where s_i — the agent's actual future position at time i, and $\tilde{s}^*_i$ — the predicted future state of the best trajectory m^*.

L_{class} — loss function based on cross entropy, which increases the probability of executing the best of the predicted trajectories m^* to 1 and reduces the probability of other trajectories to 0.

I_c is a binary indicator equal to 1, if condition c is true, and 0 — otherwise.

In the original article, the best of the predicted trajectories m^* is defined as the one that has the minimum value of the root-mean-square error in comparison to the real trajectory:

$$m = \underset{m}{argmin}\ \mathrm{L}\left(S^f, S^f_m\right). \tag{7}$$

The authors of the article suggest using the following modification:

$$m = \underset{m\in\Delta}{argmin}\ \mathrm{L}\left(S^f, S^f_m\right), \tag{8}$$

where Δ — subset of generated trajectories that has a similar final direction to the real trajectory S^f.

The idea is to remove from consideration trajectories in which the final direction of the agent differs significantly from the direction in the real trajectory when calculating the best trajectory m^*. If the difference in directions features less than certain threshold γ, then the generated trajectory is considered correct, i.e., $m \subset \Delta$. In the case under consideration $\gamma = 30^o$. Therefore, the best trajectory m^* should have a similar final direction and the lowest value of the loss function.

This work also involves prior knowledge of the domain to achieve greater convergence of the model [15]. Since only information about the roadway boundaries is available from the HD map when driving in the domain of country roads, an additional variable is introduced — L_{da} into the loss function. Thanks to it, the model will penalize the predicted trajectories that go beyond the road in cases where at least one state is $s_i \notin D$. The model penalizes only the best trajectory, since only in this case, it is possible to determine the direction of error reduction by approximating the best of the generated trajectories m^* to the real trajectory S^f.

Thus, L_{da} is defined as:

$$\mathrm{L}_{da} = \frac{1}{H}\sum_{i=1}^{H} I_c \cdot \left(s_i - s^*_i\right)^2, \tag{9}$$

where I_c is equal to 1, if $s_i \notin D$, and 0 — otherwise.

The final loss function is defined as

$$\mathrm{L} = \mathrm{L}_{class} + \alpha \cdot \mathrm{L}_{reg} + \beta \cdot \mathrm{L}_{da}, \tag{10}$$

where α and β — neural network hyperparameters used for training. In this case, both of these parameters are equal to 0.5.

To filter similar and duplicate trajectories, the proposed approach uses the filtering of a finite set of trajectories M at the final stage. This module is required because in some cases, the number of possible agent trajectories may be less than M, e.g., when a car is moving along a straight road at a constant speed, the model can generate only one trajectory: the car continues to move straight. However, the need to generate exactly M trajectories will result in the situation when all predictions are similar to each other.

The proposed filtering is based on the final direction and positions of states s_i: if the direction and the sum of the deviations between states s_i of the real and generated trajectories are less than the threshold value σ, then the trajectories are considered similar. The authors average each state of the trajectories and sum up the probabilities of the trajectories p_i.

This approach was implemented in the Python programming language on the PyTorch deep learning framework. The model was trained on GeForce RTX 2080 Ti graphics card for 40 epochs, the training took four hours.

Research Results. To assess the accuracy of prediction models, this section applies widely used metrics for the trajectory prediction problem: average displacement error, ADE, final displacement error, FDE [6], MissRate (MR), and Offroad rate (OR).

For multimodal cases with the generation of several trajectories, ADE and FDE are taken as the minimum ADE and FDE among *M* trajectories (the trajectory with the lowest metric value) [5].

The prediction is considered "missed" if ADE metric of the generated trajectory is more than two meters. OR metric is calculated as the percentage of trajectories in which at least one state s_i goes beyond the range of motion *D*.

To visualize the context of scene *c*, as well as the real and predicted future trajectories S^f and S^f a script in the Python programming language was implemented using the Matplotlib library.

This section compared the operation of several different methods in the case of an unstructured domain. The following methods were used in comparison:

– Kalman filter;

– Single trajectory output – the proposed scene encoder with the generation of a single trajectory;

– Fixed set classification – the proposed scene encoder with the reduction of the task to classification among predefined trajectories: by sets of 64 and 415 predefined trajectories;

– Proposed approach.

Table 1 presents comparison of the accuracy of the methods when working under unstructured conditions. Several methods are compared, including the proposed approach.

Table 1

Comparison of models in the unstructured domain of work

Method	Modes	ADE_1	FDE_1	ADE_6	FDE_6	$MR2_1$	$MR2_6$	OR
Kalman filter	1	3.78	8.05	3.78	8.05	0.89	0.89	5.89
Single trajectory output	1	3.12	6.75	3.12	6.75	0.89	0.89	3.26
Fixed set classification	415	3.27	7.00	1.74	3.57	0.84	0.52	3.61
Fixed set classification	64	2.6	5.63	1.52	2.91	0.82	0.49	2.58
Proposed approach	6	2.36	5.29	1.32	2.55	0.78	0.38	1.84

Kalman filter. The simplest way to predict behavior is to obtain the current state of the object (current lane, speed, direction, etc.) and extend this state to future steps based on some assumptions, e.g., that the car will continue to follow its lane or will have a constant speed and/or acceleration. Another popular method for such tasks is to use the Kalman filter [12].

According to Table 1, the Kalman filter works worse than all the presented methods based on neural networks.

Figure 2 shows two cases. In the first case, the Kalman filter successfully performs prediction because the vehicle is moving straight, without any turns or speed variation. In the second case, the Kalman filter mispredicts due to lack of knowledge about the context of the traffic situation.

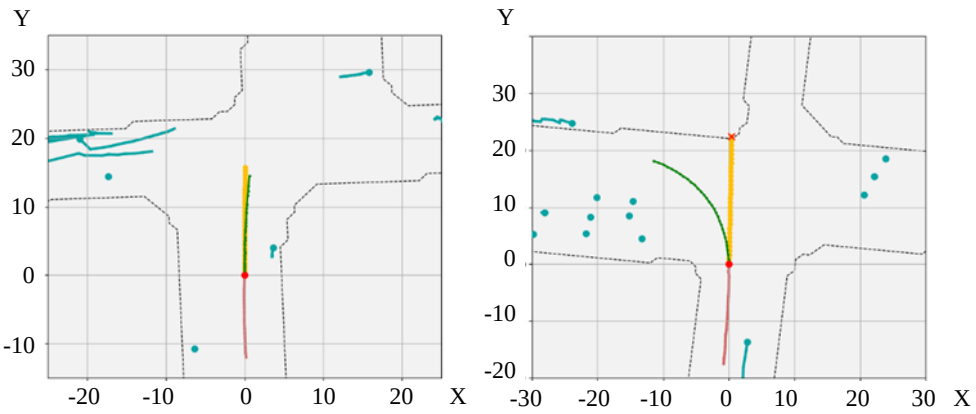


Fig. 2. Example of predictions using the Kalman filter. Dotted lines — roadway boundaries, red lines — target agent with history of observations, blue — other agents, green — real trajectory, yellow — predicted trajectory, red crosses indicate predicted states outside the roadway

Single trajectory output. This method involves the use of a graph scene encoder, which is identical to the one used in the proposed approach. The output of the network implies the generation of only one trajectory. This model is trained using the root-mean-square loss function.

As shown in Table 1, the neural network, even with the generation of a single trajectory, demonstrates better results in comparison to the Kalman filter.

Figure 3 shows the visualization of this prediction method operation. The image on the left shows that the model can successfully predict the agent's turn. The image on the right shows that generating one trajectory is not enough. The neural network tries to imagine both possible outcomes: going straight and turning right. As a result, the model outputs the average of the two outcomes.

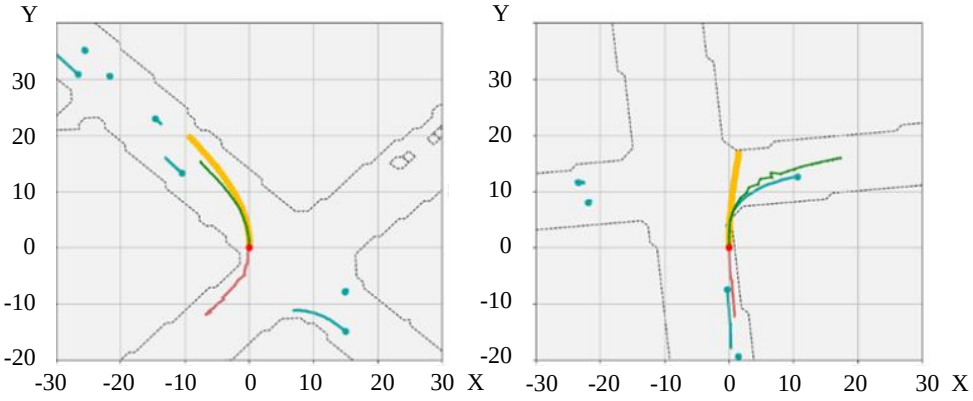


Fig. 3. Example of generating a single trajectory. Red line shows the target agent with history of observations, green — real trajectory of movement, yellow — predicted trajectory

Fixed set classification. The implementation was inspired by the CoverNet prediction method [5]. This model consists of a proposed vector scene encoder, followed by a different trajectory decoder. The decoder is a classification task based on a predefined set of trajectories consisting of physically realizable vehicle trajectories with sufficient coverage. Two sets were created for experiments: of 415 and 64 possible trajectories. The second set has the same coverage as the first, but provides a lower density of trajectories. Detailed information about the sets of trajectories is contained in paper [5].

The visualization of the work is shown in Figure 4. The classification model successfully copes with multimodality at crossroads, but in some cases, the lack of sufficient coverage by a set of trajectories negatively affects the results.

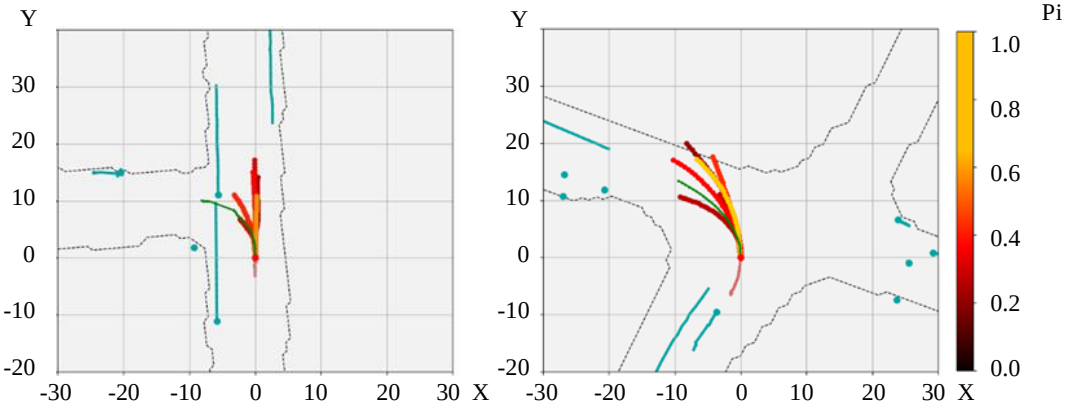


Fig. 4. Example of prediction using a classification model. Red lines represent the target agent with history of observations, green lines — real trajectory. *M* predicted trajectories with different probability of execution p_i are presented using red-yellow hues

As shown in Table 1, this method works more accurately than generating a single trajectory, but worse than the proposed approach. In addition, increasing the density of the set of trajectories by using a set of 415 trajectories did not improve the results. The authors attribute this to the presence of noise in the dataset, which comes from the tracking system used in the data collection.

Proposed approach. The proposed approach eliminates the disadvantages of all the methods described above. This is a multimodal forecasting method that does not suffer from the limitations of a predefined set of trajectories.

Moreover, according to Table 1, the proposed approach surpasses all other methods in all indicators. As shown in Figure 5, the method successfully captures two possible outcomes at the crossroad: driving straight or making a turn.

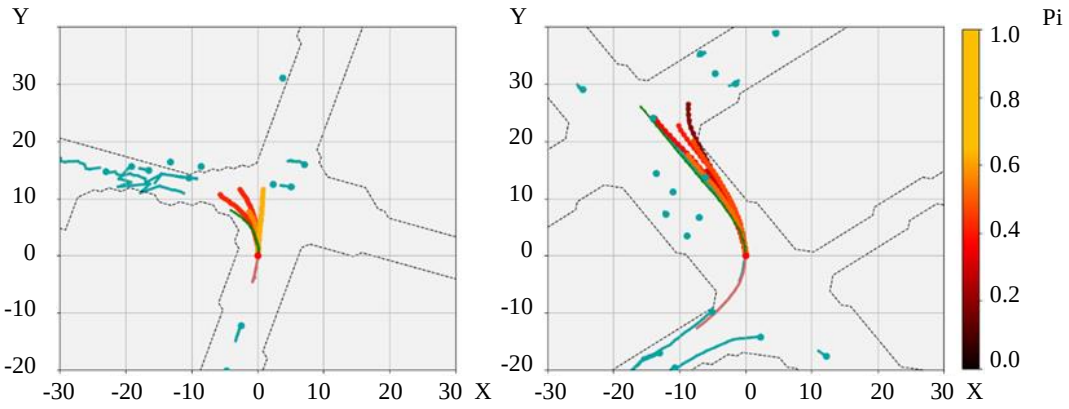


Fig. 5. Example of prediction using a classification model. Red lines represent the target agent with history of observations, green lines — real trajectory. *M* predicted trajectories with different probability of execution p_i are presented using red-yellow hues

Figure 6 shows an example of filtering similar trajectories in the case of a single possible outcome. The probability that the agent will complete the initiated turn is close to 1, since he is already in the process of turning. Therefore, in this case, the probability of other outcomes is close to 0. The proposed module successfully filters similar trajectories.

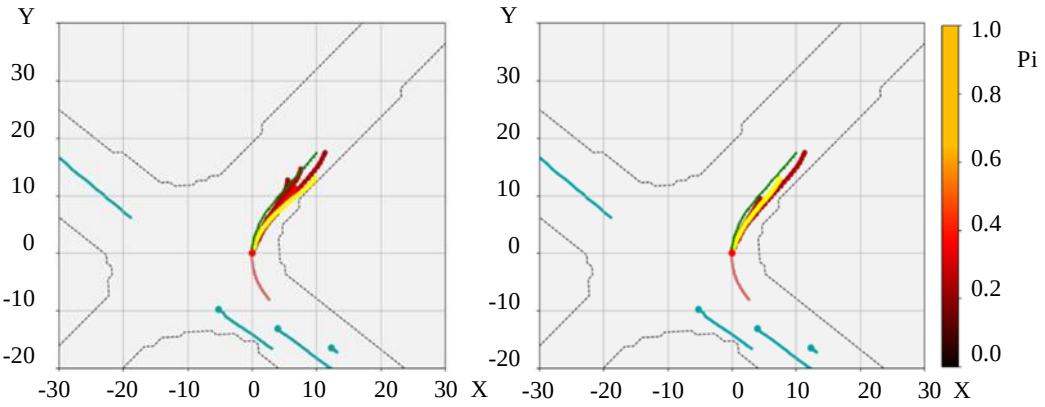


Fig. 6. Filtering effect. The entire set of predictions is shown on the left, and only filtered set — on the right

Limitations. Although the authors of the original article on the MTP model [4] indicate that their method solves the problem of mode collapse, the experiments conducted by the authors of this article do not confirm this. The problem still occurs in some cases. It is assumed to be due to the following features: the loss function does not penalize the neural network for generating all possible trajectories that the target agent can execute, as long as the best of them is as close as possible to the real trajectory. But also, the model does not encourage the network in any way to predict a variety of possible trajectories. Therefore, it is advantageous for the network to make several similar predictions in one direction, in which it is more confident than to make one prediction for each possible trajectory.

One of the possible solutions to this problem may be the use of a trajectory decoder presented in the TnT, DenseTnT models [10–11], which imply the generation of final goals at the first stages of work. In these models, all possible final goals for the agent are generated first, and then trajectories that describe the movement from the starting position to each of the goals, are generated. This provides filtering out similar final goals in the early stages, and preventing the mode collapse.

Discussion and Conclusion. In the work performed, modern methods of solving the trajectory prediction problem are investigated. The adaptability of the methods to unstructured road conditions — country roads, is considered. Insufficient accuracy of the methods is established, and a new approach to predicting is proposed.

The proposed approach is based on the VectorNet and MTP models, but has been adapted for the country road domain. In addition, a trajectory filtering module and an additional mechanism for the loss function, which penalizes trajectories for going beyond the movement zone, are proposed.

The presented comparison shows that the proposed approach is superior to other popular methods.

Limitations of the MTP approach have been identified: the output data still tends to mode collapse. The suggestion for further modifications is to use methods that generate the final goal at the early stages of prediction and thus are less susceptible to regime collapse.

References

1. Qing Rao, Jelena Frtunikj. *Deep Learning for Self-Driving Cars: Chances and Challenges*. In: Proc. 1st International Workshop on Software Engineering for AI in Autonomous Systems. New York, NY: Association for Computing Machinery; 2018. P. 35–38. https://doi.org/10.1145/3194085.3194087

2. Shaoshan Liu, Liyun Li, Jie Tang, et al. *Creating Autonomous Vehicle Systems*. San Rafael, CA: Morgan & Claypool; 2020. 216 p.

3. Jiyang Gao, Chen Sun, Hang Zhao, et al. *VectorNet: Encoding HD Maps and Agent Dynamics from Vectorized Representation.* In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA: IEEE; 2020. P. 11525–11533. https://doi.org/10.48550/arXiv.2005.04259

4. Henggang Cui, Vladan Radosavljevic, Fang-Chieh Chou, et al. *Multimodal Trajectory Predictions for Autonomous Driving Using Deep Convolutional Networks*. In: Proc. IEEE International Conference on Robotics and Automation (ICRA). Montreal, BC: IEEE; 2019. P. 2090–2096. https://doi.org/10.48550/arXiv.1809.10732

5. Tung Phan-Minh, Elena Corina Grigore, Freddy A. Boulton, et al. *CoverNet: Multimodal Behavior Prediction Using Trajectory Sets*. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA: IEEE; 2020. P. 14074-14083. https://doi.org/10.48550/arXiv.1911.10298

6. Abduallah Mohamed, Kun Qian, Mohamed Elhoseiny, et al. *Social-STGCNN: A Social Spatio-Temporal Graph Convolutional Neural Network for Human Trajectory Prediction*. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA: IEEE; 2020. P. 14424–14432. https://doi.org/10.48550/arXiv.2002.11927

7. Biktairov Yu, Stebelev M, Rudenko I, et al. *PRANK: Motion Prediction Based on RANKing*. In: Neural Information Processing Systems. Vancouver: Virtual Conference; 2020. P. 2553–2563. https://doi.org/10.48550/arXiv.2010.12007

8. Yuning Chai, Benjamin Sapp, Mayank Bansal, et al. MultiPath: Multiple Probabilistic Anchor Trajectory Hypotheses for Behavior Prediction. *Proceedings of the Conference on Robot Learning*. 2020;100:86–99. https://doi.org/10.48550/arXiv.1910.05449

9. Ajay Jain, Sergio Casas, Renjie Liao, et al. Discrete Residual Flow for Probabilistic Pedestrian Behavior Prediction. In: Proc. 3rd Conference on Robot Learning, Osaka, Japan, 2019. *Proceedings of Machine Learning Research*. 2019;100:407–419. https://doi.org/10.48550/arXiv.1910.08041

10. Hang Zhao, Jiyang Gao, Tian Lan, et al. *TNT: Target-driveN Trajectory Prediction*. In: Conference on Robot Learning. Cambridge, MA: Virtual Conference; 2020. P. 895–904. https://doi.org/10.48550/arXiv.2008.08294

11. Junru Gu, Chen Sun, Hang Zhao. *Dense TNT: End-to-end Trajectory Prediction from Dense Goal Sets*. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal, BC: IEEE; 2021. P. 15303–15312. https://doi.org/10.48550/arXiv.2108.09640

12. Prévost CG, Desbiens A, Gagnon E. *Extended Kalman Filter for State Estimation and Trajectory Prediction of a Moving Object Detected by an Unmanned Aerial Vehicle*. In: Proc. American Control Conference. New York, NY: IEEE; 2007. P. 1805–1810. https://doi.org/10.1109/ACC.2007.4282823

13. Zeyu Zhu, Nan Li, Ruoyu Sun, et al. *Off-road Autonomous Vehicles Traversability Analysis and Trajectory Planning Based on Deep Inverse Reinforcement Learning*. In: IEEE Intelligent Vehicles Symposium (IV). Las Vegas, NV: IEEE; 2020. P. 971–977. https://doi.org/10.1109/IV47402.2020.9304721

14. Mig-Fang Chang, John Lambert, Patsorn Sangkloy, et al. *Argoverse: 3D Tracking and Forecasting with Rich Maps*. In: Proc. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA: IEEE; 2019. P. 8748–8757. https://doi.org/10.1109/CVPR.2019.00895

15. Casas S, Gulino C, Suo S, et al. *The Importance of Prior Knowledge in Precise Multimodal Prediction*. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Las Vegas, NV: IEEE; 2020. P. 2295–2302. https://doi.org/10.48550/arXiv.2006.02636

Received 09.04.2023
Revised 28.04.2023
Accepted 10.05.2023

About the Authors:

Sergey A. Ivanov, Senior Engineer, Center for Autonomous Technologies, Innopolis University (1, Universitetskaya St., Innopolis, 420500, RF), ORCID, se.ivanov@innopolis.ru

Bader Rasheed, Head of the Recognition Systems Development Department, Center for Autonomous Technologies, Innopolis University (1, Universitetskaya St., Innopolis, 420500, RF), ResearcherID, ScopusID, ORCID, b.rasheed@innopolis.university

Claimed contributorship:

SA Ivanov: basic concept formulation; research objective and tasks; computational analysis; text preparation; formulation of conclusions.

B Rasheed: academic advising; analysis of the research results; hypotheses advancement; revision of the text.

Conflict of interest statement: the authors do not have any conflict of interest.

All authors have read and approved the final manuscript.

Поступила в редакцию 09.04.2023
Поступила после рецензирования 28.04.2023
Принята к публикации 10.05.2023

Об авторах:

Сергей Александрович Иванов, старший инженер центра беспилотных технологий университета Иннополис (420500, РФ, г. Иннополис, ул. Университетская, 1), ORCID, se.ivanov@innopolis.ru

Бадер Рашид, руководитель отдела разработки систем распознавания центра беспилотных технологий университета Иннополис (420500, РФ, г. Иннополис, ул. Университетская, 1), ResearcherID, ScopusID, ORCID, b.rasheed@innopolis.university

Заявленный вклад соавторов:

С.А. Иванов — формирование основной концепции, цели и задачи исследования, проведение расчетов, подготовка текста, формирование выводов.

Б. Рашид — научное руководство, анализ результатов исследований, выдвижение возможных гипотез, доработка текста.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Все авторы прочитали и одобрили окончательный вариант рукописи.

INFORMATION TECHNOLOGY, COMPUTER SCIENCE AND MANAGEMENT
ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

УДК 004.946
https://doi.org/10.23947/2687-1653-2023-23-2-180-190

Original article

Visual Coherence for Augmented Reality

Andrey L. Gorbunov [1,2]
[1] Moscow State Technical University of Civil Aviation, Moscow, Russian Federation
[2] "Aviareal" LLC, Moscow, Russian Federation
✉ a-gorbunov@mail.ru

Abstract

Introduction. The 2020s were marked by the emergence of a new generation of computer simulators using augmented reality. One of the promising advantages of augmented reality technology is the ability to safely simulate hazardous situations real-world. A prerequisite for realizing this advantage is to provide the visual coherence of augmented reality scenes: virtual objects must be indistinguishable from real ones. All IT leaders consider augmented reality as a next "big wave"; thus, the visual coherence is becoming a key issue for IT in general. However, it is in aerospace applications that the visual coherence has already acquired practical significance. An example is Boeing's development of an augmented reality flight simulator, which began in 2022. Visual coherence is a complex problem, one of the aspects of which is to provide the correct overall coloration of virtual objects in an augmented reality scene. The objective of the research was to develop a new method of such tinting.

Materials and Methods. The developed method (called spectral transplantation) uses two-dimensional spectral image transformations.

Results. A spectral transplantation technology is proposed that provides direct transfer of color, brightness, and contrast characteristics from the real background to virtual objects. An algorithm for automatic selection of the optimal type of spectral transformation has been developed.

Discussion and Conclusion. Being a fully automatic process without recording lighting conditions, spectral transplantation solves a number of complex problems of visual coherence. Spectral transplantation can be a valuable addition to other methods of providing visual coherence.

Keywords: computer simulators, augmented reality, visual coherence

Acknowledgements: the author would like to thank Alessandro Terenzi, Inglobe Technologies Srl, Ceccano, Italy, for his support in software development.

For citation. Gorbunov AL. Visual Coherence for Augmented Reality. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2):180–190. https://doi.org/10.23947/2687-1653-2023-23-2-180-190

Научная статья

Визуальная когерентность в дополненной реальности

А.Л. Горбунов [1,2] ✉

[1]Московский государственный технический университет гражданской авиации, г. Москва, Российская Федерация

[2]ООО «Авиареал», г. Москва, Российская Федерация

✉ a-gorbunov@mail.ru

Аннотация

Введение. 2020-е годы ознаменовались появлением нового поколения компьютерных тренажеров с применением технологии дополненной реальности. Одно из преимуществ данной технологии — возможность безопасного моделирования опасных ситуаций в реальном мире. Необходимым условием использования этого преимущества является обеспечение визуальной когерентности сцен дополненной реальности: виртуальные объекты должны быть неотличимы от реальных. Все мировые IT-лидеры рассматривают дополненную реальность как следующую волну радикальных изменений в цифровой среде, поэтому визуальная когерентность становится ключевым вопросом для будущего IT, а в аэрокосмических приложениях визуальная когерентность уже приобрела практическое значение. Примером может служить разработка корпорацией Боинг пилотского тренажера с дополненной реальностью (2022). Визуальная когерентность — сложная комплексная проблема, одним из аспектов которой является обеспечение корректной колористической тонировки виртуальных объектов в сцене дополненной реальности. Цель работы — разработка нового метода такой тонировки.

Материалы и методы. В разработанном методе (названном спектральной трансплантацией) используются двумерные спектральные преобразования изображений.

Результаты исследования. Предложена технология спектральной трансплантации, обеспечивающая прямую передачу характеристик цвета, яркости и контраста от реального фона к виртуальным объектам. Разработан алгоритм автоматического выбора оптимального вида спектрального преобразования.

Обсуждение и заключение. Будучи полностью автоматическим процессом без регистрации условий освещенности, спектральная трансплантация решает ряд сложных проблем визуальной когерентности. Спектральная трансплантация может стать ценным дополнением к другим методам обеспечения визуальной когерентности.

Ключевые слова: компьютерные тренажеры, дополненная реальность, визуальная когерентность

Благодарности: автор выражает благодарность А. Теренци (Inglobe Technologies Srl, Чеккано, Италия) за поддержку в разработке программного обеспечения.

Для цитирования. Горбунов А.Л. Визуальная когерентность в дополненной реальности. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2):180–190. https://doi.org/10.23947/2687-1653-2023-23-2-180-190

Introduction. Modern simulators actually by default imply the use of virtual reality (VR). The advantages of this approach are well known; therefore, we will not dwell on them, but we will note a number of significant and, more importantly, insurmountable disadvantages due to the very nature of virtual reality technology. VR is a digital, discrete technology, while the real world is continuous. Therefore, modeling the real world in VR is inevitably associated with errors, which reduces the efficiency of training. However, for training systems, an even more serious negative aspect is that human decisions are largely based on subconscious consideration of numerous details of the real picture of the world. This process is fundamentally impossible to reproduce using purely computer technologies (e.g., VR) for two reasons: we still do not know (and are unlikely to ever know) what the mechanism of the human brain is. The latest speculations on the topic of artificial intelligence only confirm this. The details of the real world taken into account when making decisions are almost infinite in number, they arise randomly and are of quite a different nature (visual, acoustic, tactile ...).

The emergence of augmented reality (AR) training systems in the 2020s reduced the severity of this problematic situation. Examples are the development by Boeing of an augmented reality pilot simulator based on the well-known R6 ATARS project, which began in the fall of 2022, as well as a similar project launched by British BAE Systems or an air traffic control training simulator from this article. All the information wealth of the world around us in AR is presented explicitly and does not require modeling. But it is needed to solve the problem of visual coherence (VC) to realize the advantages of AR associated with the parallel presence of real and virtual objects in scenes: virtual objects must be indistinguishable from real ones. This article proposes a method for solving the problem of visual coherence in the framework of a project on the development of a training system for air traffic controllers.

AR is a derivative form of VR. AR retains all the features of VR, but, in addition, as a hybrid technology, it has significant advantages arising from the parallel coexistence of virtual and real objects, which attracts the attention of developers to VC. Moreover, studies [1] show that among the negative psychophysiological consequences of using augmented reality devices, optical discomfort dominates, which occurs due to the difference in perception of real and virtual objects in the same scene due to the absence of VC. IT industry leaders see AR as the next “big wave” of revolutionary changes in digital electronics. Therefore, the VC problem is becoming a key one for IT as a whole, and these leaders show a growing interest in methods of solving it [2]. However, the problem of visual coherence has already acquired practical significance in aerospace applications. The authors encountered a VC problem when developing a training system for air traffic controllers: the rapid increase in the intensity of air traffic at airports caused an increase in the frequency of collisions of aircraft with other aircraft and airfield transport during ground maneuvering (>50 cases worldwide in 2018 before the outbreak of the pandemic). Air traffic controllers working on airport towers are not always ready to respond adequately to such emergency situations, which requires additional training. The most effective form of such training involves presenting the dispatcher with a situation of hazardous proximity of objects on the airfield, which is impossible with real objects, but can be absolutely safely implemented in augmented reality scenes. In our application, emergency situations were safely simulated using AR at a real airfield, while the virtual aircraft used should be indistinguishable from real ones.

An exhaustive overview of the known VC methods can be found in [3]. According to the author's classification, all VC methods can be divided into two main classes: with the measurement of lighting parameters, and with the assessment of lighting conditions. In the first case, a mandatory procedure is a preliminary measurement of illumination conditions, carried out with the help of special equipment. This procedure is a long and labor-intensive process. It seems to be impossible if a pre-obtained image or video of the real world is used. In the second case, the complexity of reconstructing the lighting pattern from images causes assumptions and limitations, which makes the results ambiguous. Therefore, despite the impressive results obtained by researchers using the methods mentioned in the review [3], the VC level is still often insufficient, specifically, in AR scenes with real natural landscapes under ambient lighting conditions, which are typical for aviation applications. As the review of publications below shows, there is a shortage of research of this kind.

This work was aimed at developing a universal and automatic method to provide direct transfer of color, brightness and contrast characteristics from a real background to virtual objects without digital 3D modeling, which was required in existing VC approaches. The method is based on the mathematical apparatus of two-dimensional spectral transformations, we called it “spectral transplantation”.

The key results of this study are:

– basic scheme for the spectral transplantation method, which provides a direct transfer of color, brightness and contrast characteristics from the real background to virtual objects. The method involves replacing a part of the spectrum of the image of the virtual world with the same part of the spectrum of the image of the real world, followed by an inverse transformation of the spectrum with the transplanted part;

– algorithm for automatic selection of the optimal type of spectral transformation for use in spectral transplantation.

It is important to note that VC depends on many factors: lighting, shadows, color tone, mutual reflections, surface texture, optical aberrations, convergence, accommodation, etc. Accordingly, various AR visualization techniques were used. In our case, VC is provided only for the factors of general illumination and coloring of virtual objects in AR. This is one of the VC challenges, especially for outdoor scenes. Therefore, spectral transplantation should be used in combination with other VC methods to achieve full VC.

The list of sources in [3] includes 175 positions; this review includes almost all approaches to achievements in VC (with the exception of the latter, based on neural networks, discussed below). Therefore, here we will briefly describe some characteristic examples that correspond to the mentioned basic classes.

Measurement of lighting conditions

Using a light probe with diffuse bands between mirror spherical quadrants, P. Debevec and others [4] demonstrated how the full dynamic color range of a scene could be reconstructed from a single exposure. Based on the image obtained with the probe, the intensity of several light sources could be estimated by solving a simple linear system of equations. The results were used to render a virtual diffuse sphere.

A. Alhakamy and M. Tuceryan [5] estimated the direction of incident light (direct illumination) of a real scene using computer vision techniques with a 360° camera attached to an AR device. The system simulated the light reflected from surfaces when rendering virtual objects. Then, the shadow parameters for each virtual object were determined.

Assessment of lighting conditions

S.B. Knorr and D. Kurtz [6] proposed a scheme for assessing lighting conditions in the real world based on a photo of a human face. The method was based on training a model of the type of face based on a database of faces with known lighting. The authors then reconstructed the most plausible lighting conditions in the real world in the basis of spherical harmonics for the captured face.

We should mention work [7], which described a combination of measurement and evaluation of illumination. The authors measured the reflective properties of real objects using depth maps and color images of a rotating object on a turntable using an RGB-D camera. The shape of the object was reconstructed through integrating images of the depth of the object obtained from different viewpoints. The reflectivity of an object was determined by evaluating the parameters of the reflection model from reconstructed images of shape and color.

The closest analogues of the proposed method are approaches that, like spectral transplantation, do not involve preliminary measurements of lighting and simulation of lighting conditions, scene geometry, surface reflection, and also provide for automatic processing.

Among such analogues, there are methods of color transfer from image to image. Paper [8] presented a method for automatic transferring color statistics (averages and standard deviations) from the reference image to the target image. Additional parameters were used to avoid manual processing, which was required to determine the features of color transmission in cases where images had a strong difference in the color palette. These additional parameters combined the variances of the reference and target images. The authors of the article claimed that, although manual modification of these parameters was extremely rare, it was nevertheless sometimes necessary. In addition, the statistical nature of the method raised questions about the type and scope of statistics. Also, the ability of the method to process certain types of images (containing shiny objects, shadows) was not obvious.

Xuezhong Xiao and Lizhuang Ma [9] presented an algorithm to solve the problem of color transmission reliability in terms of scene details and colors. The authors considered the preservation of the color gradient as a necessary condition for the authenticity of the scene. They formulated the problem of color transfer as an optimization problem and solved it in two stages — histogram matching and gradient-preserving optimization. A metric was proposed for an objective assessment of the efficiency of color transfer algorithms based on examples.

The advantages of the developed method, in comparison to [8, 9] and their numerous analogues, are its versatility, fully automatic nature, and the ability to transfer not only color, but also all the main characteristics of the image using one simple procedure.

The proposed method uses two-dimensional spectral transformations. Various types of images are optimally described by different types of spectral transformations (“optimally” — in the sense of matching visual perception for real and virtual objects). Actively used in digital image processing since the advent of digital television are the Discrete Fourier Transform, Discrete Cosine Transform, Hadamar Transform, S-Transform, and Karhunen-Loeve Transform.

Materials and Methods. The scheme of the spectral transplantation method (the version using the Fourier transform [10]) is shown in Figure 1. Frames of the real world (world frame — WF) and virtual world (virtual frame — VF) are used as input data (Fig. 2). This is natural for AR “video” (when the real world is observed through a video

camera). For “transparent” AR, when the real world is observed through transparent glasses, real images are captured using cameras located on AR glasses.

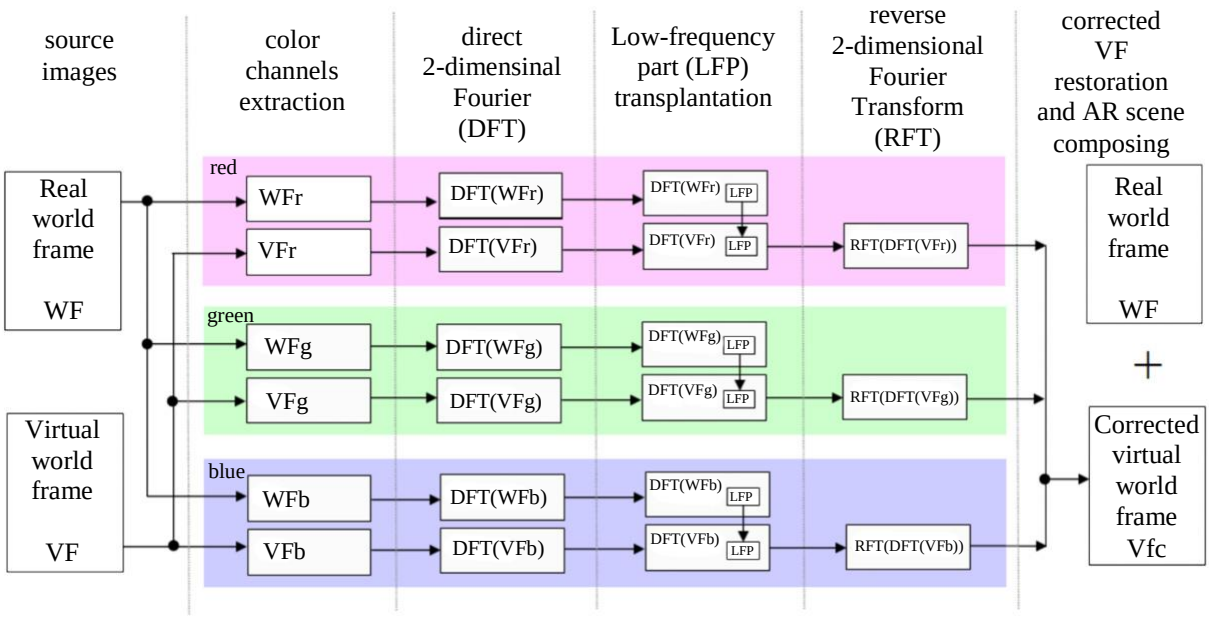


Fig. 1. Scheme of the spectral transplantation method.
A version using the Fourier transform

a) *b)* *c)*

Fig. 2. Real world (WF) and virtual world (VF) video frames:
a — WF, airport, cloudy weather; *b* — WF, airport, sunny weather; *c* — VF,
virtual airplane. WF are small fragments (<25%) of images published on websites sydneyairport.com.au and 6sqft.com.

The goal of this method is to transfer the main characteristics of the image from WF to VF. The scheme of the method is very simple, although the operations have a large computational volume. The method is implemented in five stages (Fig. 1):

1) Selection of color (RGB) channels for WF — WFr, WFg, WFb and for VF — VFr, VFg, VFb. The RGB model is used because of its generality and correlation between channels, which is specific to this model.

2) Calculation of two-dimensional direct Fourier transform (direct Fourier transform — DFT): DFT(WFr), DFT(WFg), DFT(WFb), DFT(VFr), DFT(VFg), DFT(VFb). The DFT formula is given below:

$$X_c(k,l) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x_c(m,n)\exp\left(-j2\pi\left(\frac{km}{M}+\frac{nl}{N}\right)\right) \quad (1)$$

where $c = R, G, B$ — index for red, green and blue color image channels; M, N — row and column numbers of the pixel matrix of the transformed image; k, l — spatial frequency arguments; $x_c(m,n)$ — pixel value with spatial coordinates (m,n) in channel c; $X_c(k,l)$ — complex numbers.

3) This is a key stage. Transplantation of low-frequency part (LFP) is carried out between pairs of WF and VF spectra for each of the red, green and blue channels. This means that VF LFP is replaced by the corresponding WF LFP. The idea of spectral transplantation is based on the following property of DFT: the general character of the image

(i.e., color hue, brightness, contrast) depends on the spatial frequencies contained in LFP (including the constant component) of its two-dimensional spectrum.

Thus, by transplanting WF LFP into VF spectrum, we transfer the main characteristics of the image from WF to VF. For this, it is more convenient to use a centered form of a two-dimensional spectrum, where the constant component is located in the center of the matrix of spectral coefficients, and the low-frequency components are symmetrically arranged around the constant component. In a centered spectrum, LFP is the central part of the DFT matrix, and LFP has the size Ml × Nl (Ml <M, Nl <N). If Ml =Nl, then the notation for the square matrix LFP can be LFP(012..F), where 0 — constant component, F — number of the largest spatial frequency in LFP matrix.

The size of the LFP for transplantation depends on the size of the transformed image (this size determines the spectral resolution) and on the volume of image characteristics that should be borrowed from WF. At this stage of research, the size of LFP is determined empirically. For example, the best visual results for 512×512-pixel images were obtained using LFP (012345).

4) Restoring RGB channels for VF using a two-dimensional reverse Fourier transform (RFT). While at this stage, the characteristics of WF and VF are mixed. As a result, RGB channels of the VF image are obtained with the main color, brightness and contrast characteristics of the WF, as well as with characteristics inherited from the original VF.

5) Restoring the corrected VF color through merging the RGB channels obtained at the previous stage, cutting out virtual objects and building an AR scene by superimposing the cut virtual objects on WF.

Obviously, if this method is used to process the WF video stream, then there is no need to calculate DFT, and, accordingly, LFP for each frame of the real world, since the main characteristics of the image are changed only with a radical change in the recorded scene. Such changes can be easily detected by jumps in the average pixel value. At these moments, it is needed to recalculate the spectral transformation for LFP.

Since various types of images are optimally described by different types of spectral transformations mentioned above, it is reasonable to develop an automatic algorithm for selecting the optimal type of transformation for use in spectral transplantation.

We propose to estimate the difference between the visual perception of VF and WF by the RMS distance Δ between the LFP power spectra of their images (for all color channels):

$$\Delta_c = \frac{1}{M_l N_l} \sum_{k=0}^{M_l} \sum_{l=0}^{N_l} \left| P_{Vc}(k,l) - P_{Wc}(k,l) \right|^2, \; c = R, G, B, \tag{2}$$

where P_V and P_W — two-dimensional power spectra of VF and WF, respectively. For example, in the case of the Fourier transform, the formula for P has the form:

$$P_c(k,l) = \left| \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_c(m,n) \exp\left(-j2\pi\left(\frac{km}{M} + \frac{nl}{N}\right)\right) \right|^2. \tag{3}$$

We propose to determine the optimal type of spectral transformation by the proximity of the vectors Δ and the mean vector calculated by the criterion of the minimum sum of squares of the distances between the mean vector and the vectors Δ for all transformations under consideration.

Let $\Delta_j(\Delta_j R, \Delta_j G, \Delta_j B)$ be the normalized vector of the distance between spectrum VF and WF LFPs for conversion j. Let $\Delta_a(\Delta_a R, \Delta_a G, \Delta_a B)$ be mean vector, and D_j — distance between Δ_j and Δ_a. Then, the sum of S squared distances from the vectors Δ_j of all transformations under consideration to the mean vector is equal to:

$$S = \sum_j D_j^{\,2} = \sum_j \sum_c (\Delta_{jc} - \Delta_{ac})^2, \; c = R, G, B. \tag{4}$$

Coordinates $\Delta_{aR}, \Delta_{aG}, \Delta_{aB}$ of the mean vector are calculated as the solution to a system of partial differential equations:

$$\frac{\partial S}{\partial \Delta_{ac}} = 0, \; c = R, G, B. \tag{5}$$

The selection of the optimal type of spectral transformation is determined by the proximity condition:

$$\min_j D_j. \tag{6}$$

Another obvious criterion for selecting the optimal type of transformation is the length of the vectors Δ. However, the extremes of such a criterion may be related to the ability or inability of certain transformations to correctly detect the difference between certain types of WF and VF. Therefore, we consider the use of the mean vector as a more reliable method of selection.

Similar to the DFT calculation for WF, the optimal type of transformation is selected only once at the beginning of spectral transplantation, unless WF is changed dramatically.

Research Results. The proposed method was tested using the Fourier transform without selecting the optimal type of transformation. WF (real airport scene) and VF (virtual airplane model) had a size of 512=12 pixels and 24-bit colors. Two different conditions were investigated:

1) WF — photo of the airport in cloudy weather (Fig. 2 *a*);

2) WF — photo of the airport in sunny weather (Fig. 2 *b*).

In both cases, VF contained a 3D model of the aircraft shown in Figure 2 *c*. LP(0), LP(01), LFP(012), LFP(0123), LFP(01234), LFP(012345) transplants were tested. Some of the test results are shown in Figure 3. The best visual results were obtained using LFP(012345). In Figure 3, the images after spectral transplantation are intentionally shown without other VC effects (shadows, lighting, etc.) to demonstrate the pure results of this method.

a) *b*) *c*)

Fig. 3. AR-scene: *a* — consisting of WF and VF without LFP transplantation; *b* — AP-scene after transplantation LFP(0123); *c* — AR-scene after transplantation LFP(012345)

The upper and lower rows in Figure 3 correspond to the opposite conditions for WF: light and dark WF with different shades. Experiments with any other WF will not add significantly new information since they will have conditions between those already presented in Figure 3.

Numerical simulation was carried out to demonstrate the mechanism of spectral transplantation. Figure 4 shows Fourier transplantation using a small (8×8) pixel matrix representing one of the color channels WF and VF. Such a small size of the matrix enables to clearly illustrate the transplantation procedure. In this example, the WF matrix can be associated with an image with a vertical gradient fill, and the VF matrix — with an image with a horizontal gradient fill. Another difference between WF and VF is the range of pixel values: 8-15 for WF (“lighter image”, 8 is a constant

component) and 0–7 for VF (“darker image”). LFP(01') transplantation is shown, where 1' means part of the first spatial frequency component (used because of the very low resolution of the 8×8 matrix). The 3D form of the VF matrix after transplantation indicates the transfer of properties from WF to VF: the edge of the surface has risen; the first pixel has received the value of the constant WF component. This example demonstrates how, as a result of spectral transplantation, VF starts to acquire a vertical gradient and a constant component.

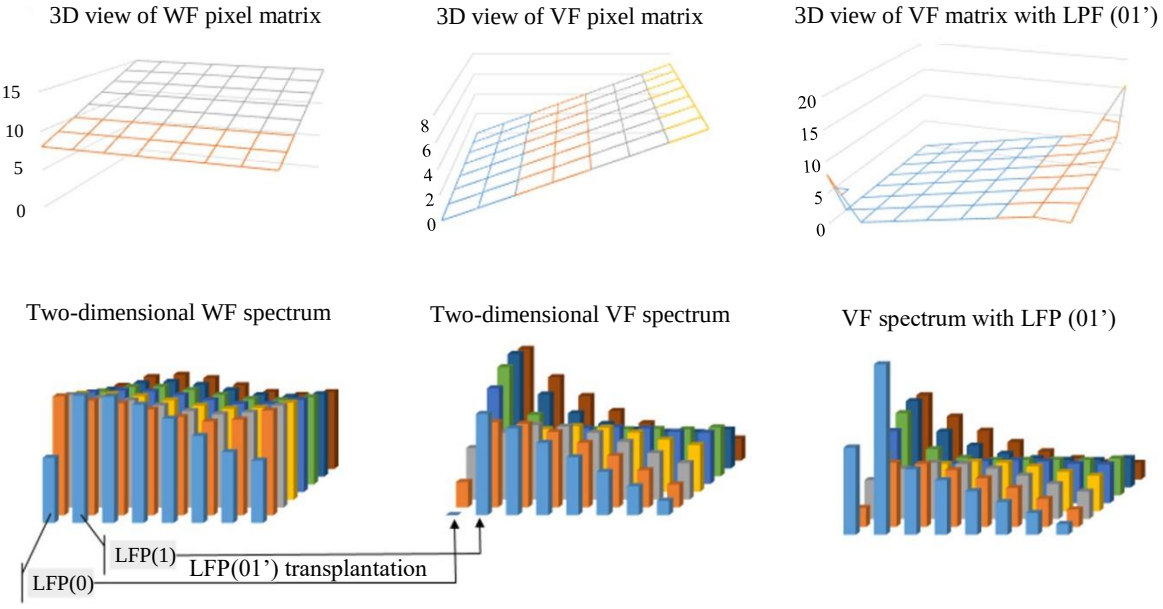


Fig. 4. Numerical simulation of spectral transplantation for 8×8-pixel matrices

Spectral transplantation provides several options for changing the parameters of this procedure: changing LFP size; selecting individual components of the spectrum for transplantation; using different transplant coefficients for various components to be transplanted.

Figure 5 shows the effect of transplantation with different parameters for various types of virtual objects — virtual aircraft models that differ in surface texture, markings, and gloss. Figures 5 *a* and *b* depict a virtual airplane with complex textures, text symbols and reflections of virtual light sources. Figures 5 *d*, *e* and *f* show a virtual plane with simple contrasting colors. Parts *a* and *d* contain virtual objects without transplantation; *b* and *e* contain virtual objects after LFP transplantation(0123); *c* and *f* contain virtual objects after LFP transplantation(012345). Virtual objects are intentionally shown without other VC effects (shadows, lighting, etc.) to demonstrate the pure results of the method.

a) *b)* *c)*

d) *e)* *f)*

Fig. 5. Scenes with cloudy WF: *a*, *d* — AR-scenes consisting of WF and VF without LFP transplantation; *b*, *d* — AR-scenes composed after LFP transplantation (0123); *c*, *f* — AR-scenes composed after LFP transplantation (012345)

It is important to emphasize that the presented figures illustrate the possibilities of tuning the proposed method, and not the final result, since it requires tuning to specific WF. The demonstration of well-executed, but incomplete results, as is often practiced in VC works, does not seem correct to us.

Discussion and Conclusions. The key complicating factor for the described method, presented in Figure 1, is the high computational costs. The most promising way to solve this problem is to directly convert WF LFP parameters into VF rendering parameters. This eliminates the cumbersome procedures of three DFT and three RFT calculations at the second and fourth processing steps, and requires only three WF DFT calculations, once for each section of the WF flow without significantly changing the basic characteristics of WF. This approach provides processing the VF flow in real time.

Another problem is selecting the optimal LFP size. As the volume of spatial frequencies used increases, they begin to hold information about the WF contents. Therefore, limiting the size of LFP is needed to eliminate the effect of a hybrid image [11]. The complexity of the optimal selection is conditioned by its association with both the LFP size and the nature of the image. Recent advances in deep learning suggest that a new approach related to visual coherence through spectral transplantation could be the use of generative adversarial networks (GAN) to transmit realistic lighting information from the source image to the target image in the same way that GAN do to transmit image style. In particular, it would be interesting to compare the performance of GAN in the case of data sets consisting of either RGB images or images represented in the frequency domain using DFT. We believe that the latter approach will help to select the optimal LFP. GAN are already widely used in VC study [2] as are neural networks in general [12].

In further research related to the topic of this paper, the following issues will be considered:

– automatic determination of the optimal LFP size for transplantation with a given volume of characteristics borrowed from WF;

– automatic detection of the exact moments when it is required to calculate new WP LFP for transplantation when processing WF video in real time (as mentioned above, this must be done if the basic characteristics of WF are radically changed);

– using the same approach in reverse (from virtual to real) to apply virtual lighting to real scenes (how virtual lighting affects the environment).

As a fully automatic process without measuring illumination, the proposed spectral transplantation method solves a number of complex VC problems. Let us say, how to best align the color, brightness, and contrast characteristics between real and virtual components in AR scenes. All these tasks are solved through one simple procedure without modeling lighting conditions, AR-scene geometry or BRDF, which eliminates the inevitable modeling errors. The proposed method can be a valuable addition to other VC tools.

References

1. Hughes CL, Fidopiastis C, Stanney KM, et al. The Psychometrics of Cybersickness in Augmented Reality. *Frontiers in Virtual Reality*. 2020;1:602954. https://doi.org/10.3389/frvir.2020.602954

2. Somanath G, Kurz D. *HDR Environment Map Estimation for Real-Time Augmented Reality*. Cupertino, CA: Apple Inc.; 2020. URL: https://arxiv.org/pdf/2011.10687.pdf (accessed: 17.06.2022).

3. Kronander J, Banterle F, Gardner A, et al. Photorealistic Rendering of Mixed Reality Scenes. *Computer Graphics Forum*. 2015;34(2):643–665. https://doi.org/10.1111/cgf.12591 (accessed: 17.06.2022).

4. Debevec P, Graham P, Busch J, et al. A Single-Shot Light Probe. *SIGGRAPH '12: ACM SIGGRAPH 2012 Talks*. 2012;10:1–19. URL: https://vgl.ict.usc.edu/Research/SSLP/A_Single_Shot_Light_Probe-SIGGRAPH2012.pdf (accessed: 17.06.2022).

5. Alhakamy A, Tuceryan M. *CubeMap360: Interactive Global Illumination for Augmented Reality in Dynamic Environment*. In: Proc. IEEE SoutheastCon. Huntsville, AL: IEEE; 2019. https://doi.org/10.1109/SoutheastCon42311.2019.9020588

6. Knorr SB, Kurz D. *Real-Time Illumination Estimation from Faces for Coherent Rendering*. In: Proc. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR). Munich: IEEE; 2014. P. 113–122. https://doi.org/10.1109/ISMAR.2014.6948483

7. Seiji Tsunezaki, Ryota Nomura, Takashi Komuro, et al. *Reproducing Material Appearance of Real Objects using Mobile Augmented Reality*. In: Proc. 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). Munich: IEEE; 2018. P. 196–197. https://doi.org/10.1109/ISMAR-Adjunct.2018.00065

8. Reinhard E, Akyuz AO, Colbert M, et al. *Real-Time Color Blending of Rendered and Captured Video*. In: Proc. Interservice/Industry Training, Simulation and Education Conference (I/ITSEC). Orlando, Florida: National Training and Simulation Association; 2004. P. 1–9. URL: https://user.ceng.metu.edu.tr/~akyuz/files/blend.pdf (accessed: 17.06.2022).

9. Xuezhong Xiao, Lizhuang Ma. Gradient-Preserving Color Transfer. *Computer Graphics Forum*. 2009;28(7):1879–1886. https://doi.org/10.1111/j.1467-8659.2009.01566.x

10. Gorbunov AL, et al. *Method of Forming an Augmented Reality Image that Ensures the Coincidence of Visual Characteristics of Real and Virtual Objects*. RF Patent No. 2667602. 2019. (In Russ.)

11. Oliva A, Torralba AJ, Schyns PhG. Hybrid Images. *ACM Transactions on Graphics*. 2006;25(3):527–532. https://doi.org/10.1145/1179352.1141919

12. Kán P, Kafumann H. DeepLight: Light Source Estimation for Augmented Reality Using Deep Learning. *The Visual Computer*. 2019;35:873–883. https://doi.org/10.1007/s00371-019-01666-x

Received 22.03.2023
Revised 18.04.2023
Accepted 20.04.2023

About the Author:

Andrey L. Gorbunov, Cand.Sci. (Eng.), Associate Professor, Air Traffic Control Department, Moscow State Technical University of Civil Aviation (20, Kronshtadtskii blvd, Moscow, 125993, RF), Director-General, "Aviareal" LLC (5, Zagoryevskaya St., Moscow, 115372, RF), ORCID, ResearcherID, ScopusID, AuthorID, a-gorbunov@mail.ru

Conflict of interest statement: the author does not have any conflict of interest.

The author has read and approved the final manuscript.

Поступила в редакцию 22.03.2023
Поступила после рецензирования 18.04.2023
Принята к публикации 20.04.2023

Об авторе:

Андрей Леонидович Горбунов, кандидат технических наук, доцент кафедры управления воздушным движением Московского государственного технического университета гражданской авиации (125993, РФ, г. Москва, Кронштадтский бул., 20), гендиректор ООО «Авиареал» (115372, РФ, г. Москва, ул. Загорьевская, 5), ORCID, ResearcherID, ScopusID, AuthorID, a-gorbunov@mail.ru

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Автор прочитал и одобрил окончательный вариант рукописи.

INFORMATION TECHNOLOGY, COMPUTER SCIENCE AND MANAGEMENT
ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

UDC 004.021:514.18
https://doi.org/10.23947/2687-1653-2023-23-2-191-202

Original article

Methods for Applying Matrices when Creating Models of Group Pursuit

Alexander A. Dubanov
Institute of Mathematics and Computer Science, BSU, Ulan-Ude, Republic of Buryatia, Russian Federation
✉ alandubanov@mail.ru

Abstract

Introduction. It is obvious that in the near future, the issues of equipping moving robotic systems with autonomous control elements will remain relevant. This requires the development of models of group pursuit. Note that optimization in pursuit tasks is reduced to the construction of optimal trajectories (shortest trajectories, trajectories with differential constraints, fuel consumption indicators). At the same time, the aspects of automated distribution by goals in group pursuit were not considered. To fill this gap, the presented piece of research has been carried out. Its result should be the construction of a model of automated distribution of pursuers by goals in group pursuit.

Materials and Methods. A matrix was formed to study the multiple goal group pursuit. The control parameters for the movement of the pursuers were modified according to the minimum curvature of the trajectory. The methods of pursuit and approach were considered in detail. The possibilities of modifying the method of parallel approach were shown. Matrix simulation was used to build a scheme of multiple goal group pursuit. The listed processes were illustrated by functions in the given coordinate systems and animation. Block diagrams of the phase coordinates of the pursuer at the next step, the time and distance of the pursuer reaching the goal were constructed as a base of functions. In some cases, the location of targets and pursuers was defined as points on the circle of Apollonius. The matrix was formed by samples corresponding to the distribution of pursuers by goals.

Results. Nine variants of the pursuit, parallel, proportional and three-point approach on the plane and in space were considered. The maximum value of the goal achievement time was calculated. There were cases when the speed vector of the pursuer was directed arbitrarily and to a point on the Apollonius circle. It was noted that the three-point approach method was convenient if the target was moving along a ballistic trajectory. To modify the method of parallel approach, a network of parallel lines was built on the plane. Here, the length of the arc of the line (which can be of any shape) and the array of reference points of the target trajectory were taken into account. An equation was compiled and solved with these elements. On an array of samples with corresponding time values, the minimum time was found, i.e., the optimal time for simultaneous group achievement of multiple goals was determined. For unified access to the library, the control vector was expressed through a one-parameter family of parallel planes. A library of calculations of control vectors was formed. An example of applying matrix simulation to group pursuit was shown. A scheme of group pursuit of multiple goals was presented. For two goals and three pursuers, six samples corresponding to the distribution of pursuers by goals were considered. The data was presented in the form of a matrix. Based on the research results, the computer program was created and registered — “Parallel Approach on Plane of Group of Pursuers with Simultaneous Achievement of the Goal”.

Discussions and Conclusion. The methods of using matrices in modeling group pursuit were investigated. The possibility of modifying the method of parallel approach was shown. Matrix simulation of group pursuit enabled to build its scheme for a set of purposes. The matrix of the distribution of pursuers by goals would be generated at each moment of time. Methods of forming matrices of the distribution of pursuers and targets are of interest in the design of virtual reality systems, for tasks with simulating the process of group pursuit, escape, evasion. The dynamic programming method opens up the possibility of automating the distribution with optimization according to the specified parameters under the formation of the matrix of the distribution of pursuers by goals.

Keywords: algorithm of group pursuit, optimization in pursuit tasks, automated distribution by goals, matrix of achievement of goals by pursuers, automated decision-making, autonomous control, parallel approach, proportional approach, three-point approach method, control vector calculation library

Acknowledgements: the author would like to thank Larisa V. Antonova, Director of the Institute of Mathematics and Computer Science, Banzarov Buryat State University, for her help provided in the work on the article.

For citation: Dubanov AA. Methods for Applying Matrices when Creating Models of Group Pursuit. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):191–202. https://doi.org/10.23947/2687-1653-2023-23-2-191-202

Научная статья

Методы применения матриц при создании моделей группового преследования

А.А. Дубанов

Институт математики и информатики Бурятского государственного университета им. Доржи Банзарова, г. Улан-Удэ, Российская Федерация

✉ alandubanov@mail.ru

Аннотация

Введение. Очевидно, что в ближайшее время сохранят актуальность вопросы оснащения движущихся робототехнических комплексов элементами автономного управления. Это требует развития моделей группового преследования. Отметим, что оптимизация в задачах преследования сводится к построению оптимальных траекторий (кратчайшие траектории, траектории с дифференциальными ограничениями, показатели расхода топлива). При этом не рассматриваются аспекты автоматизированного распределения по целям при групповом преследовании. Для восполнения этого пробела выполнена представленная научная работа. Ее результатом должно стать построение модели автоматизированного распределения преследователей по целям в групповом преследовании.

Материалы и методы. Для изучения группового преследования множества целей сформирована матрица. Управляющие параметры движения преследователей модифицированы по минимальной кривизне траектории. Детально рассмотрены методы погони и сближения. Показаны возможности модификации метода параллельного сближения. Матричное моделирование задействовали для построения схемы группового преследования множества целей. Перечисленные процессы проиллюстрированы функциями в заданных системах координат и анимацией. Как база функций построены блок-схемы фазовых координат преследователя на следующем шаге, времени и расстояния достижения преследователем цели. В ряде случаев расположение целей и преследователей определено как точки на окружности Аполлония. Матрица сформирована по выборкам, соответствующим распределению преследователей по целям.

Результаты исследования. Рассмотрены девять вариантов погони, параллельного, пропорционального и трехточечного сближения на плоскости и в пространстве. Рассчитано максимальное значение времен достижений целей. Отмечены случаи, когда вектор скорости преследователя направлен произвольно и в точку на окружности Аполлония. Отмечено, что трехточечный метод сближения удобен, если цель движется по баллистической траектории. Для модификации метода параллельного сближения на плоскости строится сеть параллельных линий. При этом учтены длина дуги линии (которая может быть произвольной формы) и массив опорных точек траектории цели. С данными элементами составлено и решено уравнение. На массиве выборок с соответственными значениями времен найдено минимальное время, то есть определено оптимальное время одновременного группового достижения множества целей. Для унифицированного обращения к библиотеке выражен управляющий вектор через однопараметрическое семейство параллельных плоскостей. Сформирована библиотека расчетов управляющих векторов. Показан пример применения матричного моделирования к групповому преследованию. Представлена схема группового преследования множества целей. Для двух целей и трех преследователей рассмотрены шесть выборок, соответствующих распределению преследователей по целям. Данные представлены в виде матрицы. По итогам научных изысканий создана и зарегистрирована программа для ЭВМ «Модель параллельного сближения на плоскости группы преследователей с одновременным достижением цели».

Обсуждение и заключение. Исследованы методы использования матриц при моделировании группового преследования. Показана возможность модификации метода параллельного сближения. Матричное

моделирование группового преследования позволяет выстроить его схему для множества целей. Матрица распределения преследователей по целям будет генерироваться в каждый момент времени. Методы формирования матриц распределения преследователей и целей представляют интерес при проектировании систем виртуальной реальности, для задач с моделированием процесса группового преследования, убегания, уклонения. Метод динамического программирования при формировании матрицы распределения преследователей по целям открывает возможность автоматизации распределения с оптимизацией по заданным параметрам.

Ключевые слова: алгоритм группового преследования, оптимизация в задачах преследования, автоматизированное распределение по целям, матрица достижения преследователями целей, автоматизированное принятие решений, автономное управление, параллельное сближение, пропорциональное сближение, трехточечный метод сближения, библиотека расчетов управляющих векторов

Благодарности: автор выражает признательность директору Института математики и информатики Бурятского государственного университета им. Д. Банзарова Антоновой Ларисе Васильевне за помощь, оказанную в работе над статьей.

Для цитирования. Дубанов А.А. Методы применения матриц при создании моделей группового преследования. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):191–202. https://doi.org/10.23947/2687-1653-2023-23-2-191-202

Introduction. The pursuit algorithms are studied from the point of view of their classical and optimal implementation. Their role in differential pursuit games is investigated. The applied sphere of ready-made solutions is very wide, because the results of such scientific research are applicable in various information technologies and systems, in particular, in search engines. Undoubtedly, the issues of equipping moving robotic complexes with autonomous control elements will be of current concern for a long time, which also requires high-quality implementation of the algorithms under consideration.

In [1–4], the coordinated behavior of a group of pursuers and targets was investigated. For general theoretical and practical issues in the problems of persecution, works [5–9] were considered. The guidance of the pursuer to the target was analyzed considering the information provided in [10–13].

With all the theoretical and practical interest in this topic, optimization in pursuit problems was limited to the construction of optimal trajectories. Specifically, the shortest trajectories, trajectories with differential constraints, fuel consumption indicators were proposed. But the aspects of automated distribution by goals in group pursuit were not considered. To fill this gap, this scientific work was carried out. Its basic result was the construction of a model of automated distribution of pursuers by goals in group pursuit. The formation of a matrix of achieving goals by pursuers was shown. When assigning goals to the pursuers, all possible combinations of achieving goals were sorted out, and a combination of the minimum value of the criterion from the generated set with the maximum value was selected.

Optimization of the multiple goal group pursuit is a promising direction for the development of such a discipline as optimal motion control in tasks related to automated decision-making and autonomous management.

Materials and Methods. In the model of group pursuit described in the paper, targets move along predefined trajectories. However, this predestination does not matter in principle. The pursuers are distributed by the targets automatically, based on the minimax solution of the goal function. Then the control parameters of the pursuers' movement are modified. In this paper, this is the parameter of the minimum curvature of the trajectory. This approach allows for simultaneous achievement of goals.

Consider a group pursuit of a set of goals: N pursuers catch up with M goals. We form a matrix of the distribution of pursuers by goals:

$$\Psi_{ij}, \text{ where } i = 1..N, j = 1..M.$$

Each cell Ψ_{ij} contains information about the phase coordinates of the i-th pursuer and the j-th target. Matrix Ψ_{ij} contains information about the method by which the i-th pursuer goes after the j-th goal.

The data stored in the cells of the matrix determines the access to the library of calculations of the control vectors of the pursuer.

In each cell of matrix Ψ_{ij}, the predicted time for the i-th pursuer to reach the j-th goal can be calculated: t_{ij}.

Research Results

In each received sample $A_k = \left\{\Psi_{i_{1k}j_{1k}} ... \Psi_{i_{nk}j_{nk}}\right\}$, it is required to find the maximum value of achievement times $t_k = Max\left\{t_{ij}\right\}$, e.g., from $\{t_{21}, t_{23}, t_{32}, t_{41}\}$ (Table 1).

Table 1

Samples corresponding to the distribution of pursuers by goals

Pursuers		Goals											
		1	2	1	2	1	2	1	2	1	2	1	2
	1	×		×		×			×		×		×
	2		×	×			×	×		×			×
	3		×		×	×		×			×	×	
Samples		A_1		A_2		A_3		A_4		A_5		A_6	

It is necessary to form matrices Ψ_{ij}, where i= 1…*3*, j=1…*2* according to the possible samples A_k,k = 1…6. Then, after the conversion, maximum value t_k=Max{t_{ij}} is found. The calculation made it possible to establish that pursuer P_1, demonstrated the greatest time of achievement, catching up with goal T_1 from sample A_2.

Thus, consider sample A_k. You can increase to the value of parameter t_k, all values t_{ij}, depending on the velocity vectors of the pursuers and goals, as well as their permissible angular velocities. This determines maximum value t_k.

Having received an array of samples {A_k} with corresponding time values {t_{ij}}, it is necessary to find minimum time t_{min}=*Min*{t_k}. This is how the optimal time for simultaneous group achievement of multiple goals is determined.

Algorithms for calculating the next step of the pursuer and estimating the time when the pursuer reaches the goal. Figure 1 shows the algorithm of the function for calculating the next step and the speed vector of the pursuer.

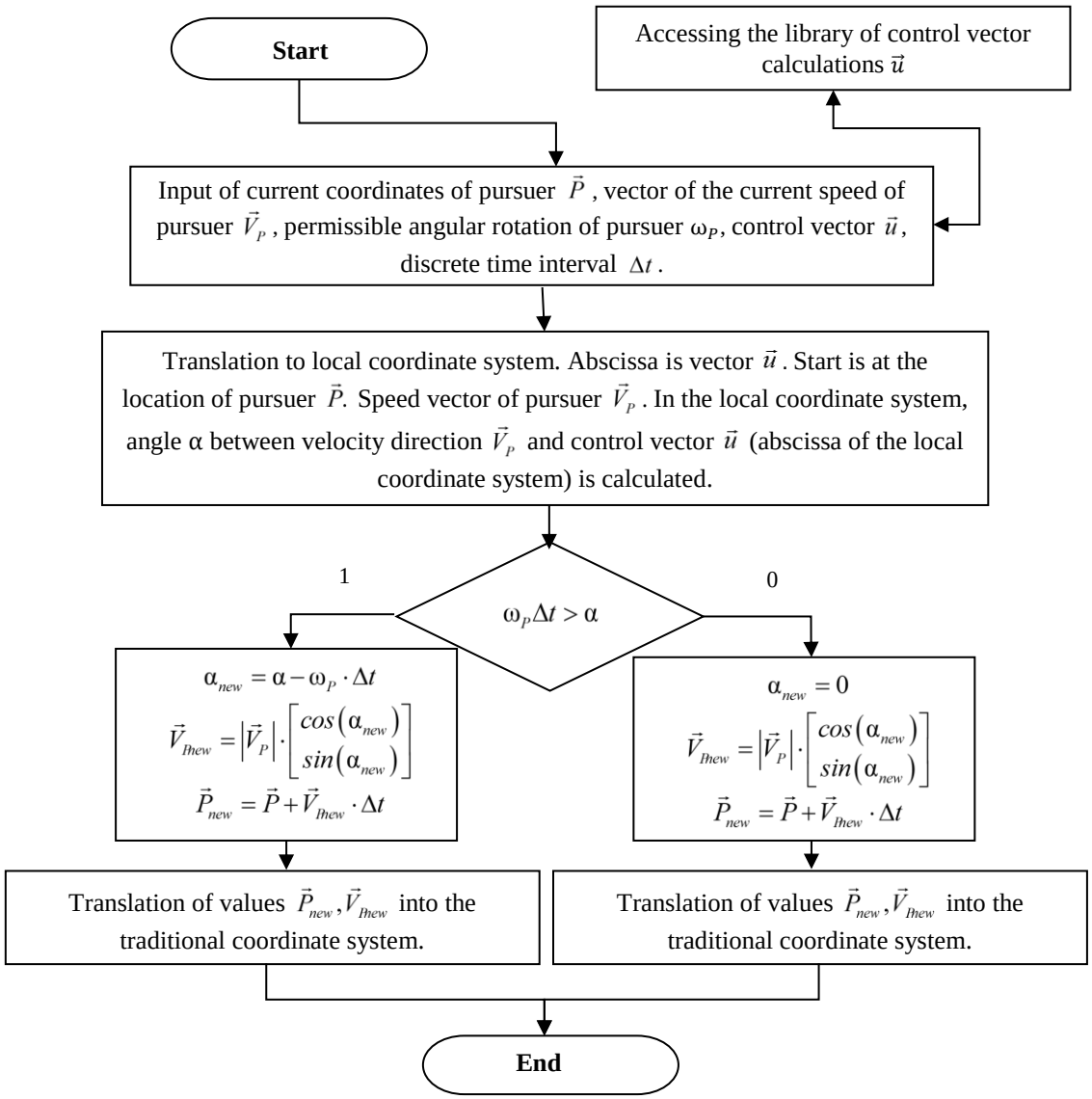


Fig. 1. Flowchart for calculating phase coordinates of the pursuer in the next step

Figure 2 shows an algorithm for calculating the time and distance of the pursuer reaching the goal. Variable ε is the threshold value of the distance from the pursuer to the target, at which the goal is considered to have been achieved.

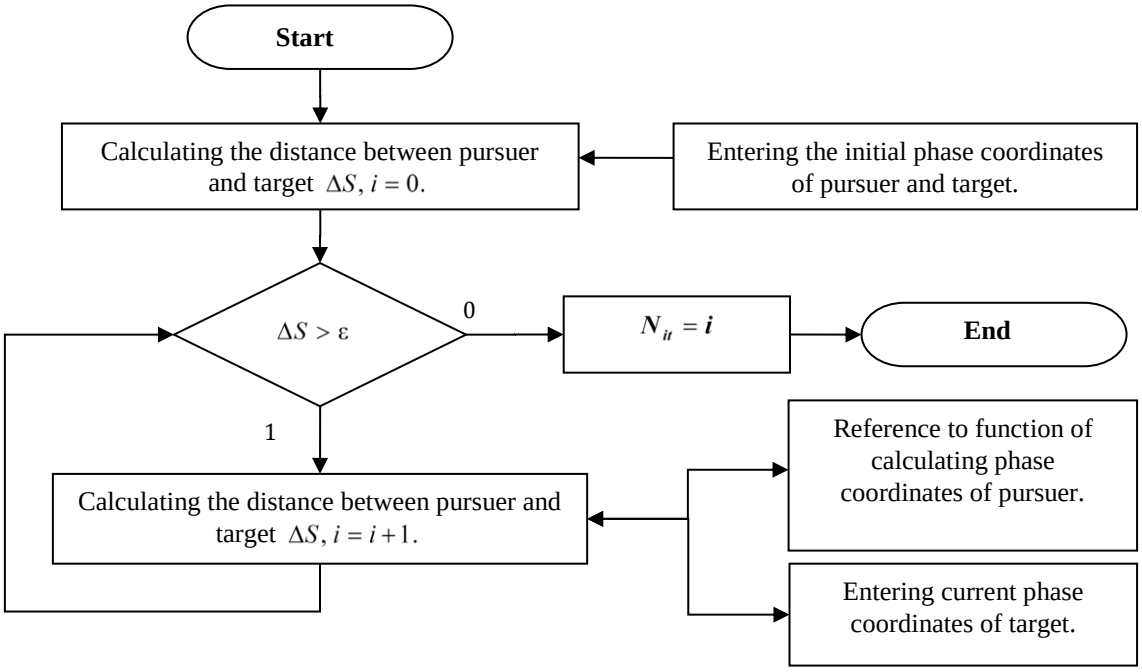


Fig. 2. Flowchart of the function for calculating the time and distance of reaching the goal by the pursuer

If the target moves along a predetermined trajectory, then the algorithm shown in Figure 2 can give an estimate of time t_{ij} for the i-th pursuer to reach the j-th goal. In this case, the output parameter of the function can be the number of iterations of the pursuit process N_{it}. Number of iterations N_{it} — output parameter of the function for calculating the time and distance of reaching the goal by the pursuer.

If the goal replies to avoid achievement, you should evaluate the time differently. It is necessary to build predicted trajectories as composite segments of straight lines, arcs of circles, square and cubic parabolas and other known lines. This will make it possible not to solve boundary value problems in the calculation cycle.

Formation of a library of control vector calculations. The distribution matrix Ψ_{ij}, where i= 1…N, j=1…M pursuers by goals is built on each discrete time interval. In each cell of matrix Ψ_{ij} , information about the method of persecution is stored. It is based on the reference to the library of functions for calculating control vectors $\vec{u}$ (Fig. 3–11).

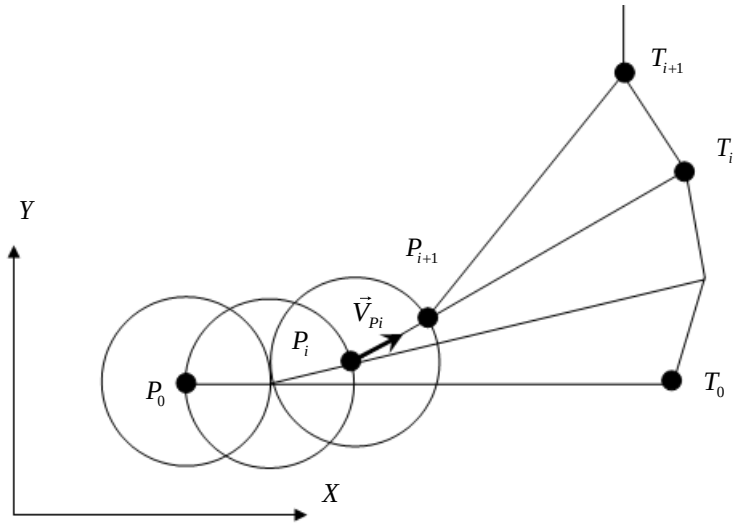


Fig. 3. Pursuit method on the plane and in space: $\vec{u}_i = \dfrac{\vec{T}_i - \vec{P}_i}{|\vec{T}_i - \vec{P}_i|}$. Here, $\vec{T}_i$ — target position point, $\vec{P}_i$ — point of the pursuer's position [14]

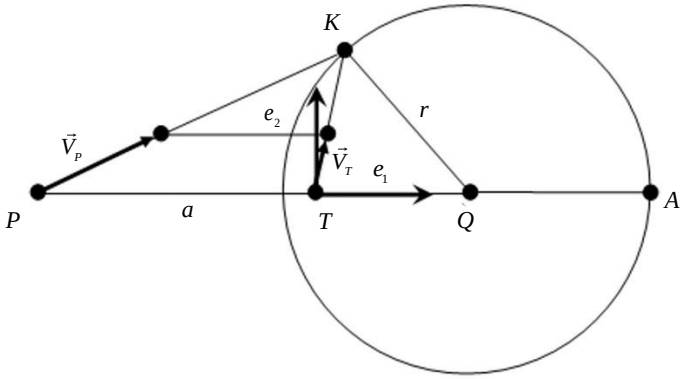


Fig. 4. Method of parallel approach on the plane: $\vec{u} = \dfrac{K-P}{|K-P|}$. T — target position point, P — point of the pursuer's position, K — point on the Apollonius circle. It is uniquely determined by points P, T and target velocity vector $\vec{V}_T$ [15]

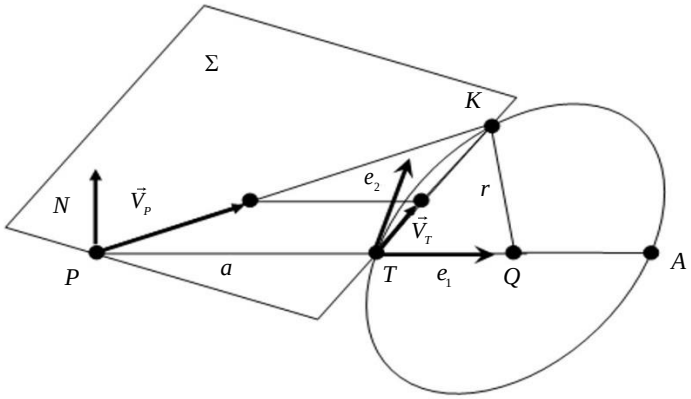


Fig. 5. Method of parallel approach in space: $\vec{u} = \dfrac{K-P}{|K-P|}$, T — target position point, P — point of the pursuer's position, K — point on the Apollonius circle. The Apollonius circle lies in plane Σ, formed by points P, T and target velocity vector $\vec{V}_T$[16]. The case is shown when the velocity vector of the pursuer is directed arbitrarily. As time passes, the velocity vector of the pursuer is directed to a point on the circle of Apollonius.

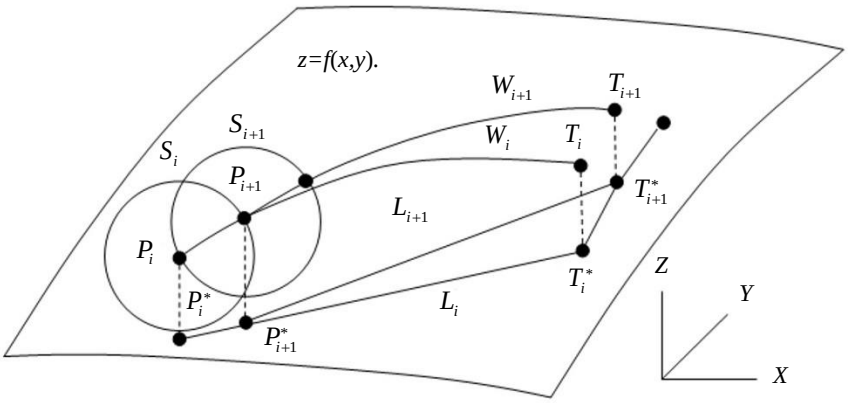


Fig. 6. Pursuit method on the plane: $\vec{u}_i = \dfrac{P_{i+1}-P_i}{|P_{i+1}-P_i|}$, where P_{i+1} — result of the intersection of surface $z=f(x,y)$, planes $P_iP_i^*T_i$ and sphere S_i centered at point P_i. Радиус $|\vec{V}_T| \cdot \Delta t$. Radius P_i^* — orthogonal projection of point P_i onto XY plane.

For unified access to the library, it is necessary to express the control vector [17].

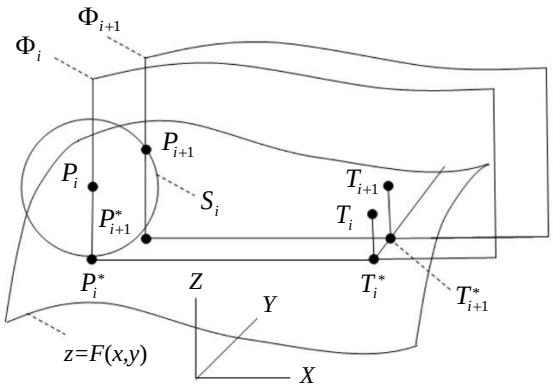


Fig. 7. Method of parallel approach on the plane: $\vec{u}_i = \dfrac{P_{i+1} - P_i}{|P_{i+1} - P_i|}$. Here, P_{i+1} — result of the intersection of surface $z=f(x,y)$, plane $P_{i+1}P^*_{i+1}T_{i+1}$ and sphere S_i centered at point P_i. Radius $|\vec{V}_T| \cdot \Delta t$.

Point P^*_{i+1} — orthogonal projection of point P_{i+1} onto XY plane. For unified access to the library, it is necessary to express the control vector. Φ_i — one-parameter family of parallel planes [18]

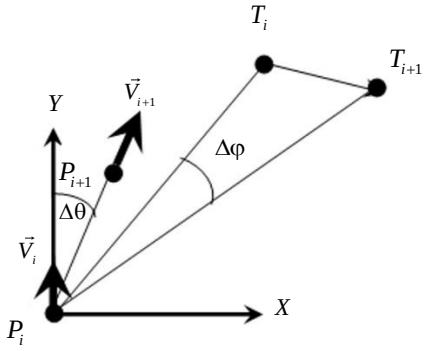


Fig. 8. Proportional approach method: $\dfrac{d\theta}{dt} = k \cdot \dfrac{d\varphi}{dt}$, $\Delta\varphi = arccos\left(\dfrac{|T_i|^2 + |T_{i+1}|^2 - |T_i - T_{i+1}|^2}{2 \cdot |T_i| \cdot |T_{i+1}|}\right)$,

$\Delta\theta = k \cdot arccos\left(\dfrac{|T_i|^2 + |T_{i+1}|^2 - |T_i - T_{i+1}|^2}{2 \cdot |T_i| \cdot |T_{i+1}|}\right)$, $P_{i+1} = \begin{bmatrix} V_P \cdot \Delta t \cdot cos\left(k \cdot arccos\left(\dfrac{|T_i|^2 + |T_{i+1}|^2 - |T_i - T_{i+1}|^2}{2 \cdot |T_i| \cdot |T_{i+1}|}\right)\right) \\ V_P \cdot \Delta t \cdot sin\left(k \cdot arccos\left(\dfrac{|T_i|^2 + |T_{i+1}|^2 - |T_i - T_{i+1}|^2}{2 \cdot |T_i| \cdot |T_{i+1}|}\right)\right) \end{bmatrix}$, $\vec{u}_i = \dfrac{P_{i+1} - P_i}{|P_{i+1} - P_i|}$

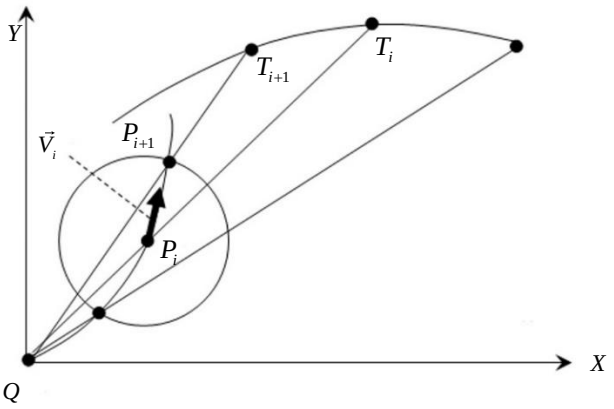


Fig. 9. Three-point approach method: $\begin{matrix} (P_{i+1} - P_i)^2 = (\vec{V}_i \cdot \Delta t)^2 \\ P_{i+1} = (1-\tau) \cdot Q + \tau \cdot T_{i+1} \end{matrix}$, $\vec{u}_i = \dfrac{P_{i+1} - P_i}{|P_{i+1} - P_i|}$.

The method is convenient if the target is moving along a ballistic trajectory.

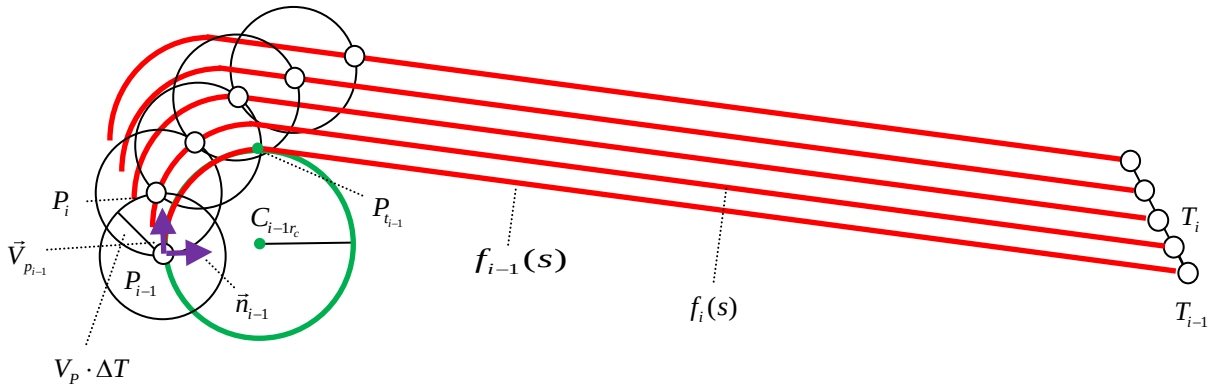


Fig. 10. Modification of the parallel approach method on the plane. A network of parallel lines is being built $f_{i+1}(s) = f_i(s) + T_{i+1} - T_i$, where s — arc length of the line, T_i — array of reference points of the target trajectory. Solving the equation $\left(f_{i+1}(s) - P_i\right)^2 = \left(V_P \cdot \Delta t\right)^2$ with respect to parameter s provides finding value s^*, which corresponds $P_{i+1} = f_{i+1}(s^*)$.

$\vec{u}_i = \dfrac{P_{i+1} - P_i}{|P_{i+1} - P_i|}$ Family $f_i(s)$ can have lines of any configuration [19].

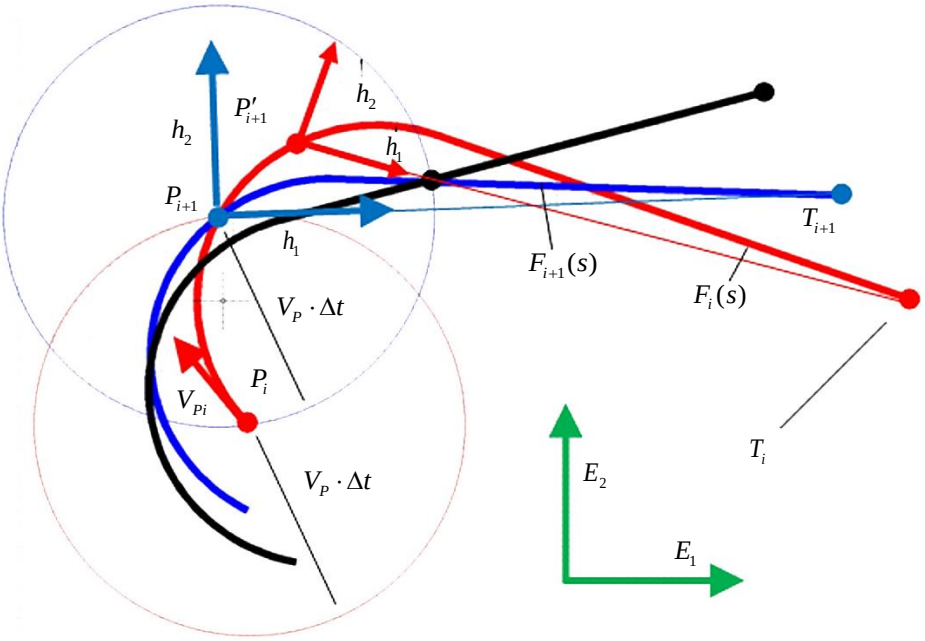


Fig. 11. Modification of the parallel approach method on the plane. Network $f_i(s)$ is built, where s — arc length of the line, T_i — array of reference points of the target trajectory. The condition is fulfilled that the end of line $f_i(s)$ passes through point T_i and point P_i is an incident on line $f_i(s)$, i.e., it is used as a pattern. Solving the equation $\left(f_{i+1}(s) - P_i\right)^2 = \left(V_P \cdot \Delta t\right)^2$ with respect to parameter s provides finding value s^*, which corresponds $P_{i+1} = f_{i+1}(s^*) \,.\, \vec{u}_i = \dfrac{P_{i+1} - P_i}{|P_{i+1} - P_i|}$.

Family $f_i(s)$ can have lines of any configuration [20].

Thus, the library of calculations of control vectors contains methods of pursuit on the plane, in space, and on the surface. Parallel approach methods are calculated on the plane, in space, and on the surface. Proportional approximation methods are calculated on the plane and in space. Three-point methods are calculated on the plane and in space. Modified pursuit methods are calculated on the plane and in space, when the permissible curvature of the trajectories is used to control the pursuer. Modified methods of parallel approach are calculated on the plane and in space.

Modification of the methods of parallel approach and pursuit provide building a network of predicted trajectories that allow for various boundary conditions. This is illustrated in Figures 3-11. But not all methods of calculating control vectors are presented in them. It is assumed that this is an open, replenished library of functions.

Case of applying matrix modeling to group pursuit. Consider a case of group pursuit (Fig. 12).

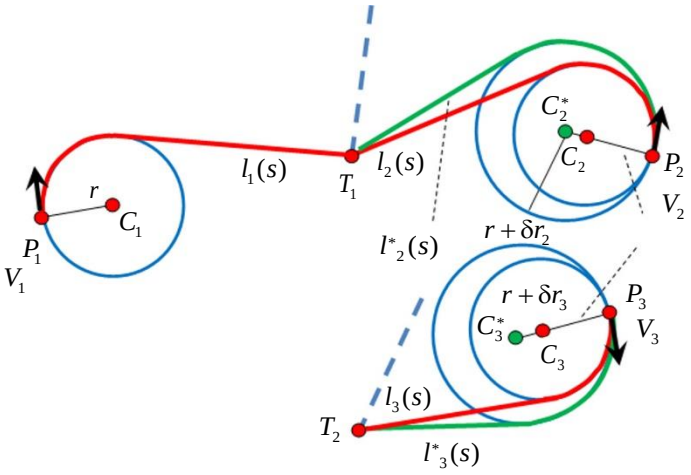


Fig. 12. Scheme of multiple goal group pursuit

Here, all pursuers achieve the goal using a modified method of parallel approach, which corresponds to Figure 10. In the pursuit model in Figure 10, the curvature of the trajectory should not be greater than a certain value. Therefore, the initial radius of curvature of the trajectory increases for pursuers P_2 and P_3 as shown in Figure 12.

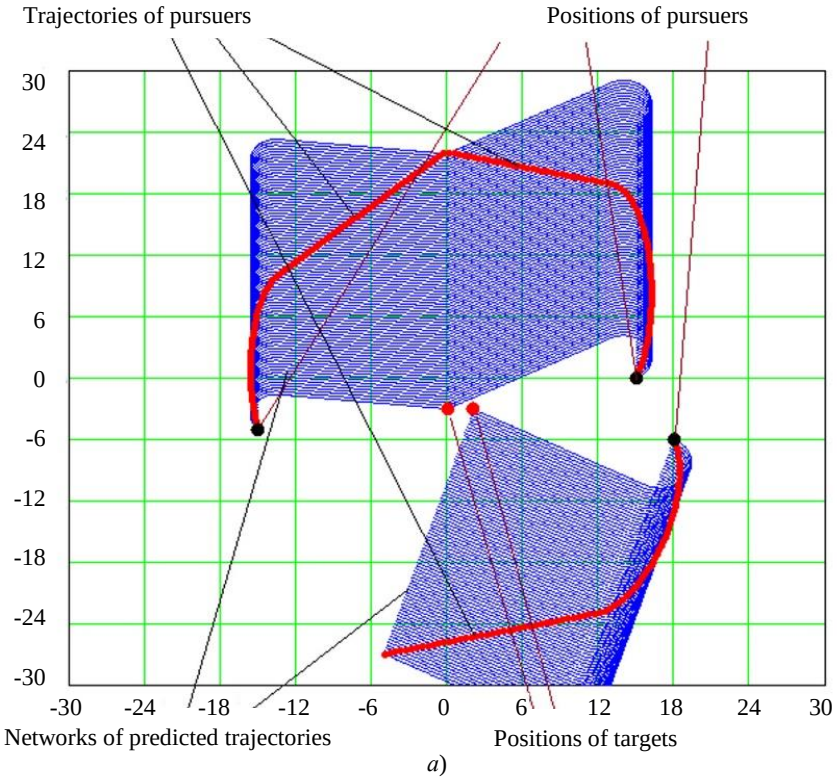
Sample A_k, has been formed, in which pursuer P_i catches up with T_j. Then there is a primary evaluation of the time of reaching t_{ij}. To estimate time t_{ij}, the following are calculated:

– length of the rectilinear section to the target,

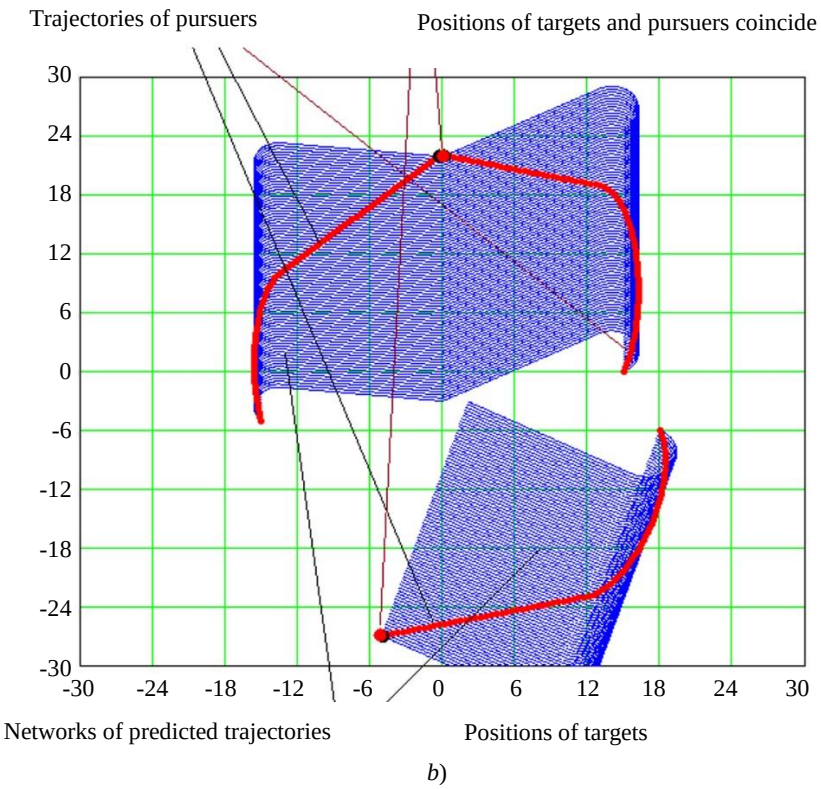
– length of the arc of the mating circle of the permissible radius.

Then maximum value $t_k=Max\{t_{ij}\}$ is selected. An increase in time t_{ij} to t_k occurs in this model due to an increase in radius of the mating circle from value r_i to $r_i+\delta r_i$ in pursuer P_i.

Figure 13 is supplemented with an animated image showing the process of multiple goal group pursuit [21].



a)



b)

Fig. 13. Schemes of group pursuit phases: *a* — initial phase; *b* — final phase

Based on the results of the study, a computer program, which implements an algorithm for group pursuit of several goals was created and registered [22]. This software solution is called “Parallel Approach on Plane of Group of Pursuers with Simultaneous Achievement of the Goal”.

Discussion and Conclusions. The methods of pursuit, parallel, proportional and three-point approach were described and visualized as functions on the plane, on the surface, and in space. In addition, the possibilities of modifying the method of parallel approach on the plane were shown. With the application of matrix modeling to group pursuit, a scheme of multiple goal group pursuit was built. The initial and final phases of this process were shown separately. The calculation of the achievement time made it possible to identify the pursuer who needed the most time to reach the goal from the sample under consideration.

Thus, it is assumed that the matrix of the distribution of pursuers by goals is generated at each moment of time. Goals and pursuers may disappear, new ones may appear. This matrix can also be used by the party representing the targets who evades prosecution. The results of the scientific research described in the article allow us to form the principles of automated distribution of pursuers by goals based on the selected target function. Algorithms for modifying the trajectories of pursuers to achieve goals simultaneously or according to a set schedule were proposed. The issues of forming a library of pursuit methods were also considered. The method of forming a matrix of the distribution of pursuers by goals can be in demand when designing virtual reality systems for game tasks in which the process of group pursuit, escape and evasion is simulated.

References

1. Rappoport IS. Strategies of Group Convergence in the Method of Resolving Functions for Quasi-Linear Conflict-Controlled Processes. *Cybernetics and Systems Analysis*. 2019;55(1):149–163. (In Russ.)

2. Bannikov AS. Some Non-Stationary Problems of Group Pursuit. *Proceedings of the Institute of Mathematics and Computer Science of UdSU*. 2013;1(41):3–46.

3. Khachumov MV. The Solution of the Problem of the Target Following by the Autonomous Aircraft. *Artificial Intelligence and Decision Making*. 2015;2:45–52.

4. Khachumov MV. Problems of Group Pursuit of a Target in a Perturbed Environment. *Artificial Intelligence and Decision Making*. 2016;2:46–54.

5. Abramyants TG, Maslov EP, Yahno VP. Evasion of Multiple Target in Three-Dimensional Space. *Automation and Remote Control*. 2008;5:3–14.

6. Samatov BT. About the Group Pursuit Problems under the Integral Control Constraints. *Cybernetics and Systems Analysis*. 2013;49(5):132–145. (In Russ.)

7. Chikrii AA. Game Dynamic Problems for Systems with Fractional Derivatives. In book: Altannar Chinchuluun, et al. (eds.) *Pareto Optimality, Game Theory and Equilibria*. New York, NY: Springer; 2008. Vol. 17. P. 349–387. https://doi.org/10.1007/978-0-387-77247-9_13

8. Borie RB, Tovey CA, Koenig S. *Algorithms and Complexity Results for Pursuit-Evasion Problems.* In: Proc. 21st Int. Joint Conf. on Artificial Intelligence (IJCAI). Pasadena, CA: Morgan Kaufmann Publishers Inc.; 2009. P. 59–66.

9. Sozinov PA, Gorevich BN. Kinematic Analysis of Proportional Navigation Methods as Applicable to Surface-to-Air Missile Guidance to a Ballistic Target. *Vestnik Koncerna VKO "Almaz – Antey"*. 2022;2:74–92.

10. Zarchan P. *Tactical and Strategic Missile Guidance*, 5th ed. Reston: American Institute of Aeronautics and Astronautics; 2006. 888 p.

11. Chikrii AA. *Conflict-Controlled Processes*. Dordrecht, Boston, London: Springer Science and Business Media; 2013. 424 p.

12. Chikrii AA, Chikrii GTs. Matrix Resolving Functions in Game Problems of Dynamics. *Proceedings of the Steklov Institute of Mathematics.* 2015;291(1):56–65. https://doi.org/10.1134/S0081543815090047

13. Chern F Chung, Tomonari Furukawa. A Reachability-Based Strategy for the Time-Optimal Control of Autonomous Pursuers. *Engineering Optimization*. 2008;40(1):67–93.

14. Dubanov AA. A Model of the Method of Pursuit in a Plane and in Space. URL: https://youtu.be/PAu9Qg1dySM (accessed: 16.01.2023). (In Russ.)

15. Dubanov AA. A Model of the Method of Parallel Convergence in a Plane. URL: https://youtu.be/hGieKXNiuz8 (accessed: 16.01.2023). (In Russ.)

16. Dubanov AA. A Model of the Parallel Convergence in Space. URL: https://youtu.be/8nDUSi3ENB4 (accessed: 16.01.2023). (In Russ.)

17. Dubanov AA. A Model of the Method of Pursuit in a Surface. URL: https://youtu.be/sU724Db_VMk (accessed: 16.01.2023). (In Russ.)

18. Dubanov AA. A Model of the Method of Parallel Convergence in a Surface. URL: https://youtu.be/06qgINE4j8U (accessed: 16.01.2023). (In Russ.)

19. Dubanov AA. Modification of the Parallel Convergence Method. URL: https://www.youtube.com/watch?v=qNXdykK21Z8 (accessed: 16.01.2023). (In Russ.)

20. Dubanov AA. Modification of the Pursuit Method. URL: https://www.youtube.com/watch?v=UQ5bVKjVqZ4 (accessed: 16.01.2023). (In Russ.)

21. Dubanov AA. The Results of the Problem Modeling. URL: https://www.youtube.com/watch?v=NNJDJOJT34I (accessed: 9.07.2022). (In Russ.)

22. Dubanov AA, et al. *A Model of the Parallel Convergence in a Plane of a Group of Pursuers for Simultaneous Achievement of a Goal*. RF Certificate of State Registration of Computer Program No. 2021618920, 2021. (In Russ.)

Received 11.04.2023
Revised 27.04.2023
Accepted 11.05.2023

About the Author:

Alexander A. Dubanov, Cand.Sci. (Eng), Associate Professor of the Department of Geometry and Teaching Methodology, Institute of Mathematics and Computer Science, BSU (5, Ranzhurova St., Ulan-Ude, Republic of Buryatia, 670000, RF), ResearcherID, ScopusID, ORCID, AuthorID, alandubanov@mail.ru

*Conflict of interest statement: t*he author does not have any conflict of interest.

The author has read and approved the final manuscript.

Поступила в редакцию 11.04.2023
Поступила после рецензирования 27.04.2023
Принята к публикации 11.05.2023

Об авторе:
Александр Анатольевич Дубанов, кандидат технических наук**,** доцент кафедры геометрии и методики преподавания математики института математики и информатики Бурятского государственного университета им. Д. Банзарова (670000, РФ, г. Улан-Удэ, ул. Ранжурова, 5) ResearcherID, ScopusID, ORCID, AuthorID, alandubanov@mail.ru

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Авторы прочитали и одобрили окончательный вариант рукописи.

INFORMATION TECHNOLOGY, COMPUTER SCIENCE AND MANAGEMENT
ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

UDC 004.58
https://doi.org/10.23947/2687-1653-2023-23-2-203-211

Check for updates

Original article

Placement of Multiple Virtual Objects in Physical Space in Augmented Reality Applications

Marianna V. Alpatova, Yuri V. Rudyak
Moscow Polytechnic University, Moscow, Russian Federation
✉ m.v.alpatova@yandex.ru

Abstract

Introduction. The challenges of placing virtual objects in a real-world environment limit the potential of augmented reality (AR) technology. This situation identifies a gap in scientific knowledge that requires additional research. Therefore, the main task of this study was to develop a method for optimal placement of virtual objects, in which the objective function of comfort was minimized. This approach is aimed at improving AR systems and developing the corresponding theory.

Materials and Methods. The conducted research was based on the analysis of the placement of virtual objects in AR/VR applications with particular emphasis on optimization. The concept of comfort of placement was proposed, taking into account the size of the object and the distance to the boundaries of free space in X, Y, Z coordinates.

Results. As part of the study, formulas were obtained for the optimal placement of objects with an arbitrary comfort function. The basic criterion was to minimize the difference between comfort levels from different sides of the object. It was found that a successful placement of objects required taking into account their size and comfort zones, as well as solving a system of *n* linear equations.

Discussion and Conclusion. The results obtained make an important contribution to the study of the problem of placing virtual objects in AR/VR/MR. They open up new opportunities for improving user interaction and conducting further research in the field of spatial computing. Possible directions for further development are dynamic adjustments and integration of the results into various XR scenarios.

Keywords: augmented reality, virtual objects, physical space, optimal placement, mathematical model, equations

Acknowledgements: this research was carried out thanks to the financial support of the RFBR in the framework of scientific project no. 21–510–07004. We appreciate our colleagues participating in this grant for their valuable contribution to the work. In addition, we would like to thank the editorial team of the journal and the reviewer for their competent expertise and valuable recommendations for improving the article.

For citation. Alpatova MV, Rudyak YuV. Placement of Multiple Virtual Objects in Physical Space in Augmented Reality Applications. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2): 203–211. https://doi.org/10.23947/2687-1653-2023-23-2-203-211

Научная статья

Размещение нескольких виртуальных объектов в физическом пространстве в приложениях дополненной реальности

М.В. Алпатова✉, **Ю.В. Рудяк**
Московский политехнический университет, г. Москва, Российская Федерация
✉ m.v.alpatova@yandex.ru

Аннотация

Введение. Проблемы, связанные с размещением виртуальных объектов в реальной среде, существенно ограничивают возможности технологии дополненной реальности (AR). Такая ситуация выявляет пробел в научных знаниях, требующий дополнительного исследования. Поэтому основной задачей данного исследования явилась разработка метода оптимального размещения виртуальных объектов, при котором происходит минимизация целевой функции комфортности. Такой подход направлен на усовершенствование систем AR и развитие соответствующей теории.

Материалы и методы. Проведенное исследование основывается на анализе размещения виртуальных объектов в AR/VR приложениях с особым акцентом на оптимизацию. Было предложено понятие комфортности размещения, учитывающее размеры объекта и расстояния до границ свободного пространства по координатам X, Y, Z.

Результаты исследования. В рамках исследования были получены формулы для оптимального размещения объектов с произвольной функцией комфортности. Основным критерием является минимизация разницы между уровнями комфортности с разных сторон объекта. Было выявлено, что успешное размещение объектов требует учета их размеров и зон комфортности, а также решения системы из n линейных уравнений.

Обсуждение и заключение. Полученные результаты представляют собой важный вклад в исследование проблемы размещения виртуальных объектов в AR/VR/MR. Они открывают новые возможности для улучшения взаимодействия с пользователями и проведения дальнейших исследований в области пространственных вычислений. Возможными направлениями для дальнейшего развития являются динамические корректировки и интеграция полученных результатов в различные XR-сценарии.

Ключевые слова: дополненная реальность, виртуальные объекты, физическое пространство, рациональное размещение, математическая модель, уравнения

Благодарности: данное исследование осуществлено благодаря финансовой поддержке РФФИ в рамках научного проекта № 21–510–07004. Выражаем признательность коллегам, участвующим в данном гранте, за их ценный вклад в работу. Кроме того, хотим поблагодарить редакционную команду журнала и рецензента за компетентную экспертизу и ценные рекомендации по улучшению статьи.

Для цитирования. Алпатова М.В., Рудяк Ю.В. Размещение нескольких виртуальных объектов в физическом пространстве в приложениях дополненной реальности. *Advanced Engineering Research (Rostov-on-Don).* 2023;23(2):203–211. https://doi.org/10.23947/2687-1653-2023-23-2-203-211

Introduction. Rapid development of augmented reality (AR) technology opens up new opportunities in various fields — from entertainment to education and industrial applications [1, 2]. However, despite considerable achievements, there are numerous problems that need to be solved, specifically, in the context of placing several virtual objects in a real-world environment. One of such problems is the optimal placement of virtual objects in augmented reality applications to provide for optimal and comfortable user experience [3–5].

This problem arose in connection with the need for the device to understand physical space. For effective placement of virtual objects in the real world, the application should be able to correctly interpret the material environment in which the user is located applying sensors and cameras of mobile devices [6].

This article presents a new approach to determining the optimal placement of virtual objects in physical space. This problem has some similarities with another close topic of generative contextual scene augmentation (CSA), where the key goal is to create a harmonious and convenient interaction between virtual and physical objects [7, 8]. However, the approach proposed by the authors differs from the one mentioned, since it focuses on determining the optimal distance between objects using a monotone comfort function, while CSA takes into account the semantics of the scene, the context and the meaning of virtual objects.

Existing approaches to solving the problem of rational location of virtual objects are usually limited by assumptions about the form of the comfort function, and they do not always guarantee the optimality of the solution. This article proposes a new approach that is significantly different from those currently used. This makes it more flexible to find a

rational arrangement of virtual objects and provides a universal solution for various scenarios and conditions. The term "optimal placement" is used for the following reasons: firstly, the final choice in the proposed recommendations is still at the discretion of the user; secondly, despite the fact that in the context of this work, the task of minimizing the objective function is solved, it contains elements of fuzzy sets.

Within the framework of this work, the concept of optimal placement of a set of virtual objects in a one-dimensional physical space is introduced. A model has been developed that allows solving the problem of optimal placement of a set of objects and results, regardless of the choice of the type of comfort function, in a system of linear equations for determining the optimal distances between objects. As an example, the solution to the problem for the case with two virtual objects is given.

The objective of this article is to propose and demonstrate a new approach to determining the optimal placement of virtual objects in physical space, to establish its applicability and efficiency. Mathematical formulations and methods for solving the optimal placement problem are presented, as well as examples of practical application of the results obtained. This will show new opportunities for improving the interaction between virtual and physical objects, as well as contribute to the development of the theory and practice of augmented reality.

Thus, this article is aimed at deepening the understanding of the problem of optimal placement of virtual objects in physical space and offering a new approach to its solution. The results of the study can be used to create more efficient and user-friendly virtual reality systems, as well as for further development of theory in this area.

Materials and Methods. Placing virtual objects in a real physical space is a task that arises in almost every AR/VR application. For all its simplicity, it can create challenges in case of insufficient attention to the issue of optimizing the placement of such objects, up to the complete refusal of a large number of potential users to work with the mentioned applications. The optimization problem is most acute when it is required to place several virtual objects at once in a given physical space. At the same time, even in the case of placing only one object, only recently the concept of "comfort" of its placement was formulated and a corresponding model was proposed [9], consisting of the following.

An object embedded in a three-dimensional physical area is presented as a rectangular parallelepiped with characteristic dimensions: *l* — length; *d* — width; *h* – height. At the same time, for each of *X, Z, Y* coordinates, the following concept of placement comfort is introduced. It is clear that the size of the free space should not be less than the size of the object, but, in addition, for each coordinate, the concept of comfortable distances from the object to the boundary of free space is introduced. For example, for *X* coordinate, we introduce the concept of comfortable distance on the left — D_- and right — D_+ and, respectively, left and right comfort — K_- and K_+. Denote the distances to the left and right of the object to the boundary of free space X_- and X_+. We assume that $K_- = 1$, if $X_- \geq D_-$ and decreases to 0 when approaching zero. For example, for simplicity, let us take linear dependences $K_-\left(X_-/D_-\right)$ and $K_+\left(X_+/D_+\right)$:

$$K_- = \begin{cases} \dfrac{X_-}{D_-},\ X_- < D_-, \\ 1,\ X_- \geq D_- \end{cases}. \tag{1}$$

Dependence $K_+\left(X_+/D_+\right)$ is similar (1). In exactly the same way, we introduce the concept of comfort on one side and on the other for *Z* and *Y* coordinates.

If the size of the free space horizontally is $L \geq D_- + l + D_+$, then the problem of comfortable placement $(K_- = K_+ = 1)$ does not arise; and all problems appear when $l \leq L \leq D_- + l + D_+$. In this case, the concept of comfort of the object placement is introduced, when comfort on the one hand is not obtained at the expense of comfort on the other hand. We introduce the target function of comfort:

$$K_2 = \left(K_- - K_+\right)^2. \tag{2}$$

By optimal placement, we will understand such placement, in which minimum K_2 is achieved. Obviously, this happens if $K_- = K_+$.

As it was shown in [9], if dependences $K_-\left(X_-/D_-\right)$ and $K_+\left(X_+/D_+\right)$ are linear, the minimum of the target function (2) corresponding to the optimal placement of the object is attained at the following values X_- and X_+:

$$X_- = \left(L - l\right)\frac{D_-}{D_- + D_+},$$

$$X_+ = L - X_- - l = (L - l)\frac{D_+}{D_- + D_+}. \quad (3)$$

Formulas (3) are simple and convenient for optimal placement of a single virtual object. As noted in [9], their disadvantage is that they were derived for the case of linear (1) comfort functions $K_-(X_-/D_-)$ and $K_+(X_+/D_+)$. We show that they are always valid if functions $K_-(X_-/D_-)$ and $K_+(X_+/D_+)$ are the same function *k(x)*, satisfying the condition that it increases monotonically for $0 \le x \le 1$, and is equal to 1 at $x > 1$. The type of such function $K_-(X_-/D_-) = k(x)$ is shown in Figure 1.

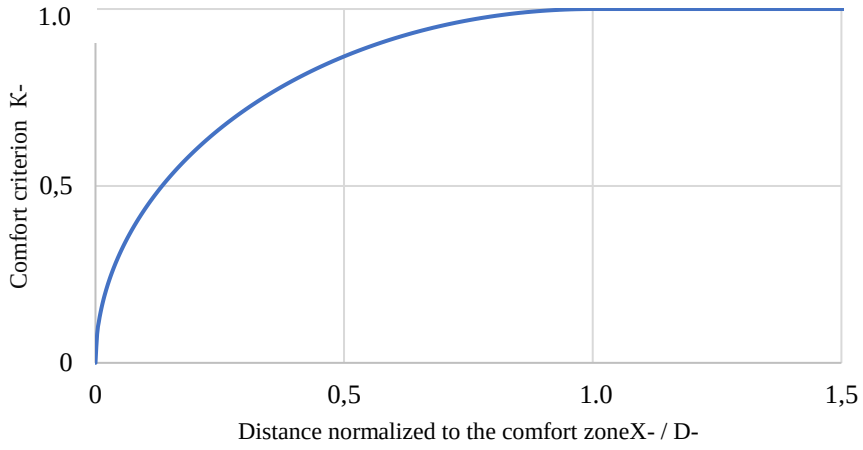


Fig. 1. Comfort function dependence $k(x) = \begin{cases} \sqrt{1-x^2},\ 0 \le x \le 1 \\ 1,\ x > 1 \end{cases}$

Figures 2 and 3 present additionally two more functions — cubic and linear, demonstrating similar described behavior.

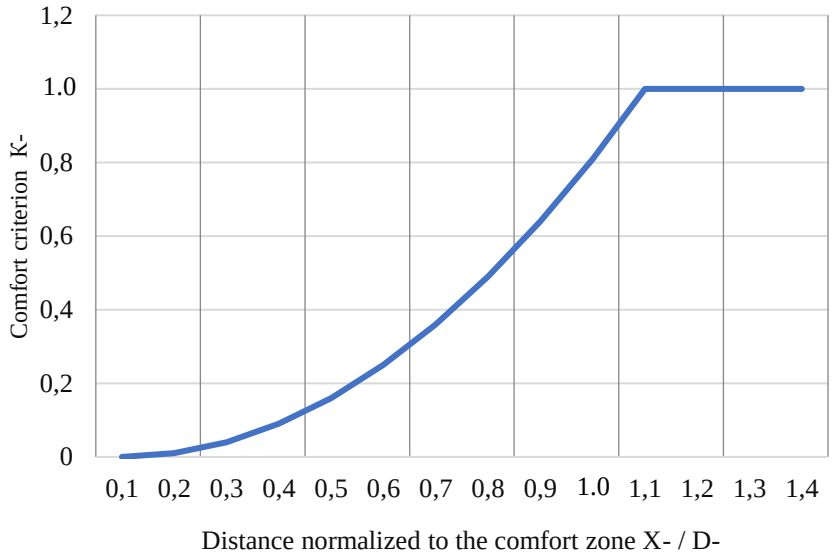


Fig. 2. Comfort function dependence $k(x) = \begin{cases} x^2,\ 0 \le x \le 1 \\ 1,\ x > 1 \end{cases}$

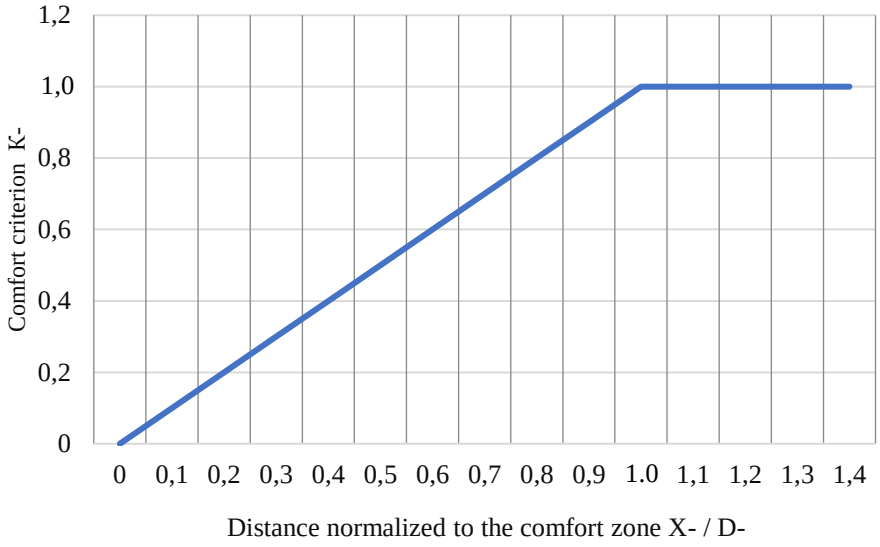


Fig. 3. Comfort function dependence $k(x)=\begin{cases} x,\ 0\le x\le 1 \\ 1,\ x>1 \end{cases}$

Indeed, suppose we need to embed an object of size *l* with comfortable distances on the left D_- and right D_+ into the space bounded by size *L*, and conditions $l \le L \le D_- + l + D_+$ are met. Then, if one-sided comforts are presentable in the form $K_-\left(X_-/D_-\right)=k\left(x_-\right)$, $K_+\left(X_+/D_+\right)=k\left(x_+\right)$, where $x_- = X_-/D_-$, $x_+ = X_+/D_+$, and dependence *k(x)* satisfies the above conditions, then, from the minimum condition of the objective function (2) we obtain:

$$K_- = K_+ \Rightarrow k\left(X_-/D_-\right)=k\left(X_+/D_+\right).$$

For the given nonlinear comfort function *k(x)*, the resulting equation can be solved numerically by one of the known methods. However, since one-sided comforts are described by self-similar function *k(x)*, from equation $k\left(X_-/D_-\right)=k\left(X_+/D_+\right)$, we obtain relation $X_-/D_- = X_+/D_+$, from which, taking into account equality $L = X_- + l + X_+$, equations follow (3).

Thus, it is shown that simple and rather convenient equalities (3), which provide embedding an object with optimal comfort, are valid for any one-sided comfort function *k(x)*.

Research Results. In a real situation, there is a need to place several objects at once. In this case, it makes no sense to solve sequentially the problems of optimal placement of the first object, then the second, third, etc., since when placing the next object, a need arises to shift previously installed objects so that the placement is optimal for the totality of all objects. If condition $L < \sum_{i=1}^{n} l^{(i)}$ is met, the objects in principle do not fit in the free space of length L. If $L \ge \sum_{i=1}^{n}\left(D_-^{(i)} + l^{(i)} + D_+^{(i)}\right)$, then the objects can be placed so that they do not interfere with each other. In reality, the problem of optimal placement arises if the following conditions are met:

$$\sum_{i=1}^{n} l^{(i)} < L < \sum_{i=1}^{n}\left(D_-^{(i)} + l^{(i)} + D_+^{(i)}\right). \quad (4)$$

In this case, we introduce the target function:

$$K_\Sigma = K_2^{(1)} + K_2^{(2)} + ... + K_2^{(n)}, \quad (5)$$

where $K_2^{(i)}$ — comfort of the *i*-th object determined by formula (2). Here, when there are two consecutive objects with numbers *i* and *i*+1, then their neighborhood will be comfortable if the distance between them is not less than $D_+^{(i)} + D_-^{(i+1)}$, which corresponds to the new comfortable distances $\tilde{D}_+^{(i)} = \tilde{D}_-^{(i+1)} = D_+^{(i)} + D_-^{(i+1)}$, *i*=1,2,…,(*n*-1) since a comfortable distance to the wall is one thing, and to another object, from where something can be pushed, is quite another. The rational arrangement of embedded objects is determined by the minimum of the target function (5).

The minimum of the target function (5), taking into account (2), gives us a system of *n* equations:

$$K_2^{(1)} = 0;\ K_2^{(2)} = 0;\ ...,\ K_2^{(n)} = 0\ . \tag{6}$$

From (6), it follows:

$$K_{-}^{(i)} = K_{+}^{(i)};\ i = 1,2,3,...,n\ . \tag{7}$$

The system of equations (7) can be written in the following form:

$$\begin{gathered} k\left(\frac{X_{-}^{(1)}}{D_{-}^{(1)}}\right) = k\left(\frac{X_{-}^{(2)}}{\tilde{D}_{+}^{(1)}}\right), \\ k\left(\frac{X_{-}^{(i)}}{\tilde{D}_{-}^{(i)}}\right) = k\left(\frac{X_{-}^{(i+1)}}{\tilde{D}_{+}^{(i)}}\right),\ \ i = 2,3,...(n-1), \\ k\left(\frac{X_{-}^{(n)}}{\tilde{D}_{-}^{(n)}}\right) = k\left(\frac{L - \sum_{i=1}^{n}\left(X_{-}^{(i)} + l^{(i)}\right)}{\tilde{D}_{+}^{(n)}}\right). \end{gathered} \tag{8}$$

That is, we have obtained *n* equations with respect to *n* unknowns $X_{-}^{(1)}, X_{-}^{(2)}, ..., X_{-}^{(n)}$, where $X_{-}^{(1)}$ — distance of the first object from the left edge of the embedding area, $X_{-}^{(i)},\ i = 2,3,...,n$ — distance between objects with numbers *i* and (*i*-1).

Since function *k(x)* is monotonic, system (8) is reduced to a linear system of equations, which does not depend on the type of the comfort function itself *k(x)*.

$$\begin{gathered} \frac{X_{-}^{(1)}}{D_{-}^{(1)}} = \frac{X_{-}^{(2)}}{\tilde{D}_{+}^{(1)}}, \\ \frac{X_{-}^{(i)}}{\tilde{D}_{-}^{(i)}} = \frac{X_{-}^{(i+1)}}{\tilde{D}_{+}^{(i)}}, \\ \frac{X_{-}^{(n)}}{\tilde{D}_{-}^{(n)}} = \frac{L - \sum_{i=1}^{n}\left(X_{-}^{(i)} + l^{(i)}\right)}{\tilde{D}_{+}^{(n)}}. \end{gathered} \tag{9}$$

System (9) can be solved by one of the known methods. However, due to the fact that the matrix of system (9) is highly sparse, the solution to the system can be found quite simply. For example, in the first (*n*-1) equations, it is possible to express $X_{-}^{(i+1)}$ by $X_{-}^{(i)}$ in each *i*-th equation, then, substituting this into the last equation, to obtain a linear equation with respect to $X_{-}^{(1)}$. After that, moving from the first equation to (*n*-1)-th, we can consistently find $X_{-}^{(2)}, X_{-}^{(3)}, ..., X_{-}^{(n)}$.

In the case of placing two virtual objects, *n*=2, system (9) takes the following form:

$$\begin{gathered} \frac{X_{-}^{(1)}}{D_{-}^{(1)}} = \frac{X_{-}^{(2)}}{\tilde{D}_{+}^{(1)}}, \\ \frac{X_{-}^{(2)}}{\tilde{D}_{-}^{(2)}} = \frac{L - X_{-}^{(1)} - X_{-}^{(2)} - l^{(1)} - l^{(2)}}{\tilde{D}_{+}^{(n)}}. \end{gathered} \tag{10}$$

From system (10), we find:

$$\begin{gathered} X_{-}^{(1)} = \frac{\left(L - l^{(1)} - l^{(2)}\right)\tilde{D}_{-}^{(2)} D_{-}^{(1)}}{\left(\tilde{D}_{-}^{(2)} + D_{+}^{2}\right)\tilde{D}_{+}^{(1)} + D_{-}^{(1)}\tilde{D}_{-}^{(2)}}, \\ X_{-}^{(2)} = \frac{\left(L - l^{(1)} - l^{(2)}\right)\tilde{D}_{-}^{(2)}\tilde{D}_{+}^{(1)}}{\left(\tilde{D}_{-}^{(2)} + D_{+}^{2}\right)\tilde{D}_{+}^{(1)} + D_{-}^{(1)}\tilde{D}_{-}^{(2)}}. \end{gathered} \tag{11}$$

where $X_{-}^{(1)}$ — distance between the first object and the left edge of space; $X_{-}^{(2)}$ — distance between the right edge of the first object and the left edge of the second object.

Consider example (12) when:

$$L = 100,\ l^{(1)} = 40,\ l^{(2)} = 30,\ D_{-}^{(1)} = 10, D_{+}^{(1)} = 5, D_{-}^{(2)} = 11, D_{+}^{(1)} = 13\ . \quad (12)$$

Since conditions (13) are met, then:

$$l^{(1)} + l^{(2)} < L < l^{(1)} + l^{(2)} + D_{-}^{(1)} + D_{+}^{(1)} + D_{-}^{(2)} + D_{+}^{(1)} \quad (13)$$

For optimal arrangement of two objects, we can use expressions (11), which give regardless of the form *k(x)* in this case $X_{-}^{(1)} \approx 7.692$; $X_{-}^{(2)} \approx 12.308$; $K_2^{(1)} = K_2^{(2)} = 0$. Here, the value of one-sided comfort depends on the type *k(x)*. For linear dependence *k(x)*, shown in Figure 3, $K_{-}^{(1)} = K_{+}^{(1)} = K_{-}^{(2)} = K_{+}^{(2)} \approx 0.769$. If dependence *k(x)* corresponds to Figure 1, we get the comfort value $K_{-}^{(1)} = K_{+}^{(1)} = K_{-}^{(2)} = K_{+}^{(2)} \approx 0.973$. Thus, the paper introduces the concept of optimal placement of a set of virtual objects in physical space. A model has been developed that provides solving the problem of optimal placement of a set of objects. It is shown that the solution to this problem does not depend on the type of monotone comfort function.

Discussion and Conclusion. The theoretical aspect of the important issue of optimal placement of a set of virtual objects in physical space (the problem often encountered in augmented reality applications) was considered. By proposing a new mathematical model and including fuzzy logic, we laid the foundation for an algorithm that could potentially help users find a rational and convenient location of virtual objects in their real environment.

The foundations laid in this study show that the proposed model effectively solves the problems associated with the placement of a set of virtual objects in a given physical space. By analyzing virtual planes and taking into account the distances between virtual objects and the edges of these virtual planes, our method provides optimal placement considering the linear dimensions of virtual objects and the comfort zone around them.

The results obtained contribute to the current development of augmented and mixed reality applications, providing a theoretical solution to the problem of optimal placement, which, in turn, can improve user interaction and overall satisfaction with the tools of the technology under discussion. Moreover, the possible applications of this research go beyond AR applications, they open up new avenues for research in the related fields, such as virtual reality, mixed reality, and spatial computing.

Considering the results of this study, future developments may be aimed at verifying the algorithm through empirical testing, enabling dynamic real-time adjustments based on user behavior, and exploring the integration of our approach into various XR application scenarios. As the area of augmented reality continues to evolve, we expect that our research will make a significant contribution to the development of the technology, inspiring its wide dissemination and further enriching the user experience.

References

1. Mekni M, Lemieux A. *Augmented Reality: Applications, Challenges and Future Trends*. In: British Library Conference Proceedings: Applied Computational Science. Athens: WSEAS; 2014. P. 205–214. URL: https://www.cs.ucf.edu/courses/cap6121/spr2020/readings/Mekni2014.pdf (accessed: 29.03.2023)

2. Rick Van Krevelen, Ronald Poelman. A Survey of Augmented Reality Technologies, Applications and Limitations. *International Journal of Virtual Reality*. 2019;9(2):1–20. https://doi.org/10.20870/IJVR.2010.9.2.2767

3. Rui Nóbrega, Diogo Cabral, Giulio Jacucci, et al. *NARI: Natural Augmented Reality Interface — Interaction Challenges for AR Applications*. In: Proc. Int. Conf. on Computer Graphics Theory and Applications (GRAPP 2015). Pasadena, CA: Morgan Kaufmann Publishers Inc.; 2015. P. 504–510. https://doi.org/10.13140/RG.2.1.2240.1440

4. Kurkovsky SA, Koshy R, Novak V, et al. *Current Issues in Handheld Augmented Reality*. In: Proc. 2012 Int. Conf. on Communications and Information Technology (ICCIT): 2012 International Conference on Communications and Information Technology (ICCIT). Hammamet: IEEE; 2012. P. 68–72. https://doi.org/10.1109/ICCITechnol.2012.6285844

5. Irshad Sh, Rambli DRA. Advances in Mobile Augmented Reality from User Experience Perspective: A Review of Studies. In book: HB Zaman, et al. (eds.). *Advances in Visual Informatics*. Cham: Springer International Publishing; 2017. P. 466–477. https://doi.org/10.1007/978-3-319-70010-6_43

6. Turk M, Fragoso V. Computer Vision for Mobile Augmented Reality. In book: Gang Hua, Xian-Sheng Hua (eds.). *Mobile Cloud Visual Media Computing*. Cham: Springer International Publishing; 2015. P.3–42. https://doi.org/10.1007/978-3-319-24702-1_1

7. Keshavarzi M, Yang A, Caldas L, et al. *Optimization and Manipulation of Contextual Mutual Spaces for Multi-User Virtual and Augmented Reality Interaction*. In: Proc. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Atlanta, GA: IEEE; 2020. P.353–362. https://doi.org/10.1109/VR46266.2020.00055

8. Z Sadeghipour Kermani, Liao Z, Tan P, et al. Learning 3D Scene Synthesis from Annotated RGB-D Images. *Computer Graphics Forum*. 2016;35(5):197–206. https://doi.org/10.1111/cgf.12976

9. Alpatova MV, Glazkov AV, Rudyak YuV. *Mathematical Model of Rational Location of Augmented Reality Objects in User's Environment*. In: Proc. Int. Sci. Conf. "Smart Nations: Global Trends In The Digital Economy". Cham: Springer International Publishing; 2022. P. 248–254. https://doi.org/10.1007/978-3-030-94873-3_30

Received 02.04.2023
Revised 18.04.2023
Accepted 02.05.2023

About the Authors:

Marianna V. Alpatova, Senior lecturer of the Computer Science and Information Technology Department, Moscow Polytechnic University (38, Bolshaya Semyonovskaya St., Moscow, 107023, RF), ScopusID, ORCID, AuthorID, m.v.alpatova@yandex.ru

Yuri V. Rudyak, Dr.Sci. (Phys.-Math.), Professor of the Computer Science and Information Technology Department, Moscow Polytechnic University (38, Bolshaya Semyonovskaya St., Moscow, 107023, RF), ORCID, AuthorID, rudyak@mail.ru

Claimed contributorship:

MV Alpatova: basic concept formulation; research objectives and tasks; text preparation; formulation of conclusions.

YuV Rudyak: academic advising; calculation analysis; revision of the text; correction of conclusions.

Conflict of interest statement: the authors do not have any conflict of interest.

All authors have read and approved the final manuscript.

Поступила в редакцию 02.04.2023
Поступила после рецензирования 18.04.2023
Принята к публикации 02.05.2023

Об авторах:

Марианна Валерьевна Алпатова, старший преподаватель кафедры информатики и информационных технологий Московского политехнического университета (107023 г. Москва, ул. Большая Семеновская, д. 38) ScopusID, ORCID, AuthorID, m.v.alpatova@yandex.ru

Юрий Владимирович Рудяк, доктор физико-математических наук, профессор кафедры информатики и информационных технологий Московского политехнического университета (107023 г. Москва, ул. Большая Семеновская, д. 38), ORCID, AuthorID, rudyak@mail.ru

Заявленный вклад соавторов:

М.В. Алпатова — формирование основной концепции, цели и задачи исследования, подготовка текста, формирование выводов.

Ю.В. Рудяк — научное руководство, проведение расчётов, доработка текста, корректировка выводов.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Все авторы прочитали и одобрили окончательный вариант рукописи.

INFORMATION TECHNOLOGY, COMPUTER SCIENCE AND MANAGEMENT
ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

UDC 004.942: 519.63
https://doi.org/10.23947/2687-1653-2023-23-2-212-224

Original article

Simulation of Vertical Movements of Seawater in Stratified Reservoirs

Nikita V. Kudinov ✉[1], **Alena A. Filina**[2], **Alla V. Nikitina**[1], **Denis V. Bondarenko**[1], **Irina F. Razveeva**[1]

[1]Don State Technical University, Rostov-on-Don, Russian Federation

[2] "Supercomputers and Neurocomputers Research Center" Co Ltd, Taganrog, Russian Federation

✉ kudinov_nikita@mail.ru

Abstract

Introduction. In the field of computational mathematics, there are many ways to approximate the model of fluid mechanics. Methods and estimates of approximation quality criteria, such as stability and convergence, are developed, while a combination of approaches to constructing economical difference schemes, such as splitting by physical processes, regularization by B. N. Chetverushkin, a linear combination of the Upwind and Standard Leapfrog difference schemes in aggregate has not been implemented and evaluated before. The authors were faced with the task of approximating each part of the hydrodynamic model split by physical processes with the most adequate scheme and further investigating the correctness of this approach.

Materials and Methods. The mathematical model of hydrophysical processes is closed by the empirical equation of the state of salt water. Significant properties were selected, a mathematical model was built. Difference operators approximated differential operators. An algorithm for layer-by-layer modeling of transients was constructed. The algorithm has been implemented in the form of the program, which mainly contains elementwise (massively-parallel) operations.

Results. Mathematical models of hydrodynamic processes in reservoirs were obtained, taking into account three equations of motion in the presence of a density gradient of the aqueous medium when hydrostatic approximation was abandoned. A new method of calculating the pressure field using B.N. Chetverushkin's regularizers in the continuity equation was tested. A software module for numerical simulation of hydrophysical processes of water movement with different salinity and density was developed. This is open-source software that provides not only the redefinition of empirical dependences (as algebraic functions), but also the connection of external simulating modules to display dependences algorithmically.

Discussion and Conclusion. The developed model of hydrophysics, taking into account the properties of salt water and the dynamic relationship of the mechanical movement of water with salinity, can be used to study the formation of a nonequilibrium distribution of parameters and identify the most stable parameters of the aquatic environment. The model explains the downward movement of oxygen. That will help in the future to estimate the values of the parameters of the aquatic environment, which are difficult to measure directly. It can be used in the procedure of parametric identification of hard-to-measure parameters of the aquatic environment.

Keywords: mathematical model, stratification, seawater dynamics model, quasi-hydrodynamic model, cell occupancy method, central difference scheme, sweep method, FTCS scheme, Upwind Leapfrog, Standard Leapfrog

Acknowledgements: the authors would like to thank A.I. Sukhinov, Corresponding Member, Russian Academy of Sciences, mentor of the scientific school, for organizing the research; and also, they appreciate the Editorial board of the journal "Advanced Engineering Research (Rostov-on-Don)" for the help provided in the preparation of the article.

Funding information. The research is done with the financial support from Russian Science Foundation (project no. 22–71–10102).

For citation. Kudinov NV, Filina AA, Nikitina AV, et al. Simulation of Vertical Movements of Seawater in Stratified Reservoirs. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2):212–224. https://doi.org/10.23947/2687-1653-2023-23-2-212-224

Научная статья

Моделирование вертикальных движений морской воды в стратифицированных водоемах

Н.В. Кудинов ✉[1], **А.А. Филина**[2], **А.В. Никитина**[1], **Д.В. Бондаренко**[1], **И.Ф. Развеева**[1]

[1]Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация

[2]Научно-исследовательский центр супер-ЭВМ и нейрокомпьютеров, г. Таганрог, Российская Федерация

✉ kudinov_nikita@mail.ru

Аннотация

Введение. В области вычислительной математики известно множество способов аппроксимации модели механики жидкости. Учеными выработаны методы и оценки критериев качества аппроксимации, таких как устойчивость и сходимость. Комбинация подходов построения экономичных разностных схем, таких как расщепление по физическим процессам, регуляризация по Б.Н. Четверушкину, линейная комбинация разностной схемы «кабаре» и «крест» в совокупности ранее не реализовывалась и не оценивалась. Перед авторами стояла задача аппроксимировать каждую часть расщеплённой по физическим процессам модели гидродинамики наиболее адекватной схемой и далее исследовать корректность данного подхода.

Материалы и методы. Математическая модель гидрофизических процессов замыкается эмпирическим уравнением состояния соленой воды. Выбираются значимые свойства, строится математическая модель. Разностные операторы аппроксимируют дифференциальные операторы. Строится алгоритм послойного моделирования переходных процессов. Алгоритм реализован в виде программы, которая, в основном, содержит поэлементные (массивно параллельные) операции.

Результаты исследования. Получены математические модели гидродинамических процессов в водоемах, учитывающие три уравнения движения при наличии градиента плотности водной среды при отказе от гидростатического приближения. Апробирован новый способ вычисления поля давления с применением регуляризаторов по Б. Н. Четверушкину в уравнении неразрывности. Разработан программный модуль численного моделирования гидрофизических процессов движения воды с различной солёностью и плотностью. Это открытое программное обеспечение, допускающее не только переопределение эмпирических зависимостей (как алгебраических функций), но и подключение внешних моделирующих модулей для отображения зависимостей алгоритмически.

Обсуждение и заключение. Разработанная модель гидрофизики, учитывающая свойства солёной воды и динамическую связь механического движения воды с солёностью, может применяться для изучения формирования неравновесного распределения параметров и идентификации наиболее стабильных параметров водной среды. Модель объясняет нисходящее движение кислорода, что позволит в будущем оценивать величины параметров водной среды, которые сложно измерить непосредственно. Она может быть использована в процедуре параметрической идентификации трудноизмеряемых параметров водной среды.

Ключевые слова: математическая модель, стратификация, модель динамики морской воды, квазигидродинамическая модель, метод заполненности ячеек, центрально-разностная схема, метод прогонки, ВВЦП, Кабаре, Крест

Благодарности: авторы благодарят наставников научной школы, члена-корреспондента РАН А.И. Сухинова за организацию научной работы, а также выражают признательность редакции журнала «Advanced Engineering Research» за помощь, оказанную в процессе подготовки статьи.

Финансирование. Работа выполнена при поддержке РНФ (проект № 22–71–10102).

Для цитирования. Кудинов Н.В., Филина А.А, Никитина А.В. и др. Моделирование вертикальных движений морской воды в стратифицированных водоемах. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2):212–224. https://doi.org/10.23947/2687-1653-2023-23-2-212-224

Introduction. One of the critical tasks related to the ecology and life safety of people living in coastal areas is forecasting and modeling the movement of water in the seas and large regional reservoirs. Moreover, it is important to study the process of transport of substances dissolved in an aqueous medium, taking into account stratification and the dependence of water density on many variable factors. Such predictive modeling is likely to make it possible not only to assess water quality, but also to manage it under conditions of climate change and industrial impacts. A more general and major goal is to link water quality to the number and species diversity of natural hydrobiocenoses living in the hydrosphere. When solving these problems, it is necessary to take into account the hydrodynamic characteristics of the aquatic environment, features of external influencing factors, such as the heterogeneity of temperature distribution, salinity, oxygen saturation of water, the amount of gases dissolved in water, acidity. These parameters are some key parameters of the biological activity of the aquatic ecosystem [1].

The water area of the reservoir can also be considered as a transport system that transfers oxygen from the atmosphere to bottom sediments. However, there are cases of the formation of chemical gradients of large magnitude at relatively small depth differences — in boundary layers with a relatively small value outside these zones. The reason for the appearance of such sites with the general stratification of water by density in universal gravity on the one hand, and the radiation effect of the sun on the water, causing its heating, on the other. These processes can cause a decrease in the rate of production, destruction and recycling of biogenic elements and bio-organisms until these processes stop, as well as predetermine the biodiversity of hydrobionts in general and species composition in particular [2]. Temperature stratification affects significantly the distribution of organisms in the water column, the transfer and deposition of impurities harmful to bio-organisms. An increase in the temperature of surface waters causes a violation of vertical water exchange and, accordingly, a decrease in aeration of the deep-water zone, a decrease in solubility and oxygen concentration in water. Stratification by density, temperature and chemical composition limits the convective rise of biogenic elements, carbon dioxide and products of incomplete oxidation of organic substances entering the hypolimnion (cold, salty, dense layers of water) as a result of sedimentation of seston into the surface layers of water. From the beginning of stratification until its end, the surface layer is depleted, and the hypolimnion, on the contrary, is enriched with these substances. As a result, physico-chemical stratification causes an uneven distribution of a number of biologically significant substances in depth and is the reason for the self-organization of a complex structure of ecological niches [3]. Mathematical modeling of mechanical, chemical and biological processes occurring in aquatic ecosystems is urgent, it is associated with problems of ecology and life safety of the population of coastal territories.

We should name the outstanding scientists who made a significant contribution to the study of hydrology and oceanology. V.P. Dymnikov was engaged in the study of climate and oceanology, modeling of the atmosphere and ocean. A.S. Monin and M.Y. Belevich investigated the processes of kinematics of the aquatic environment, turbulence and microstructure of the ocean. Soviet scientist A.S. Sudolsky studied the dynamics of waters and coastal processes in various reservoirs in relation to solving the problems of designing, building and operating specific hydraulic structures, and rational economic use of reservoirs in general [4]. According to V.I. Vernadsky, one of the most important manifestations of life is the gas exchange of organisms and the environment, mainly respiration processes based on oxygen consumption [5]. Some outstanding Russian scientists, including S.V. Bruevich, whose works were devoted to the development of analytical research methods, the formulation of the basics of hydro- and bio-hydrochemistry, were engaged in the study of hydrobiological processes of reservoirs. G.G. Matishov and V.G. Ilyichev actively study the conditions of optimal exploitation of water resources, develop models of transport of pollutants in water bodies, and investigate the assessment of their impact on the bioresources of the aquatic environment [6].

To study the influence of these processes of the aquatic environment, a complex of interrelated mathematical models is being developed. It is based on the use of accurate predictive models and software implementation of economical numerical methods, which provide detailed investigation of the kinematics of the process, cause-effect relationships, and the state of the modeling object. The existing methods and means of predictive modeling of the state of the aquatic environment, taking into account a number of biotic and abiotic factors, including the processes of

distribution of oxygen, carbon dioxide, salts, are based on general scientific approaches, simplified mathematical models with low adaptability, lack of modeling of nonlinear dynamic processes characteristic of most aquatic ecosystems, incorrect setting of the boundaries of the computational domain. In some cases, the studies are accompanied by a formal definition of boundary conditions, and give rather rough and approximate simulation results.

When modeling the substance transfer process based on the advection-diffusion equations, a good approximation of advective terms, which are gradients of pressure, mass density and total energy, momentum of motion, is required. The use of standard difference schemes with large estimates of similarity parameters causes loss of calculation accuracy due to an increase in the approximation error and increased restrictions on the time step due to the stability condition of the difference scheme. In the works of A.I. Suhinov, A.E. Chistyakov and others [7, 8], it is shown how to effectively use a linear combination of the Upwind and Standard Leapfrog difference schemes with the optimal values of weight coefficients for the approximation of the transport equation. The efficiency of these methods is achieved by optimization of the approximation error of discrete-continuum model, the exact solution of the mass transfer with constant speed. Studies have shown that this approach also extends to hydrodynamic models with a variable (sign-alternating) velocity without the effect of grid viscosity. Another positive quality of such approximations is that they can be used to simulate complex flow structure, e.g., vortex. Currently, numerous researchers use such a scheme for the simulation of turbulent flows. Members of leading foreign research organizations, such as Stanford University, Imperial College London, etc., as well as members of the Institute of Computational Mathematics of the Russian Academy of Sciences E.M. Volodin, A.V. Glazunov, A.S. Gritsun, N.G. Yakovlev and others [9] published works in which mathematical modeling of climatic changes, hydrodynamic and atmospheric processes and phenomena are carried out on the basis of eddy-resolving schemes. The quasi-hydrodynamic approximation of a continuous medium enables to further reduce the time step requirement and increase the spatial resolution of the model with limited computer memory. In practice, a small term proportional to the second derivative in time or density is added to the system of Navier-Stokes and continuity equations. This approach makes it possible to smooth out non-physical fluctuations in mass density and momentum, as well as total energy, brought along the spatial grid faster than the speed of sound.

Existing universal application software packages (e.g., “Mars3d” software package, Ecointegrator, CHARISMA software package, SALMO complex, CHTDM, CARDINAL software package, packages for modeling various processes of aerohydrodynamics, PHOENICS, FLUENT, GAS DYNAMICS TOOL software packages) do not take into account some properties of the simulated complex systems, thus reducing the accuracy and efficiency of modeling. These properties include spatial heterogeneity of the aquatic environment motion, vortex structures of currents. Mathematical models and algorithms for their numerical implementation do not take into account the probability of a significant change in the depth and density of the aquatic environment, which can cause instability of the numerical solutions obtained. For this reason, such specialized software packages can be used to simulate a limited variety of hydrophysical processes of water systems. Most of the well-known specialized software (ADAM, CAL3QHC, Chensi, TASCflow, ISC–3, PANACHE, REMSAD, UAM-IV, ECOLOGIST, PRISM, VITECON), designed to simulate the process of spreading pollutants, the interaction of hydrobionts, is mainly focused on uniprocessors, represented mainly by personal computers. In such systems, only single component modules of these systems (e.g., ECOSIM and MAQSIP) are scaled to parallel systems. In practice, a small term proportional to the second derivative in time or density is added to the system of Navier-Stokes equations and continuity. This approach makes it possible to smooth out non-physical fluctuations in mass density, momentum and total energy brought along the spatial grid faster than the speed of sound.

Materials and Methods. The success of the development of a mathematical model of hydrophysical and hydrobiological processes depends on the availability and elaboration of test cases and tasks for investigating steadily observed phenomena in the seas, such as vertical mixing and redistribution of salinity and oxygen, halocline and thermocline. To study these phenomena, the paper uses a model of hydrodynamics that takes into account the balance of mass forces and cross-border flows [1, 2]:

$$\frac{\partial \rho}{\partial t} + (\nabla, \rho v) = \sigma,$$

$$\frac{\partial v}{\partial t} + (v, \nabla) v = -\frac{1}{\rho} \nabla p + b,$$

$$\rho\frac{\partial}{\partial t}\left(\frac{k+\varepsilon}{\rho}\right)=div(Tv)+(v,\rho b)+(\nabla,h)+\rho q,$$

where ρv — flow density; $\partial k/\partial t$ — kinetic-energy density rate; $\partial\varepsilon/\partial t$ — — self-energy density rate; T — stress tensor $Tn=-pn$; b — mass force; $h=h(t,x)$ — heat flux density; q — specific heat inflow due to radiation. Due to the fact that these phenomena are usually described by a vertical distribution of parameters over depth, it is reasonable to obtain a simplified model that provides rapid identification of unobservable parameters. We select a cylindrical region V with bases on the bottom and surface of the water with a cross-sectional area S. We project the velocities, flows and forces to the vertical direction, assuming that the partial derivatives in *x*, *y* of the parameters to the horizontal direction are equal to zero outside the cylinder, i.e., assume horizontal uniformity of the parameters of the aqueous medium. Let us write system (1) in a conservative form so that it enables to determine the mass density (ρ), mechanical momentum (ρv) and total energy density ($\rho(k+\varepsilon)$), $k=v^2/2$. In the cylinder, we allocate infinitesimal volume and assume that each such volume, which is V, is affected by the reaction force of the support (bottom of the reservoir), equal to hydrostatic pressure, and a force similar to the friction force caused by the viscosity of the fluid and momentum transfer, not equal to zero with vertical movements of the fluid.

If we neglect horizontal fluid movements and assume that only vertical movements are essential, and the fact that the density of the aqueous medium significantly depends on salinity, accept the simplifications and conventions outlined earlier, then the equations of hydromechanics [1–3] in compact form are presented by a system of partial differential equations [2]:

$$\frac{\partial\rho}{\partial t}+\frac{\partial(\rho v)}{\partial x}=0,$$

$$\frac{\partial(\rho v)}{\partial t}+\frac{\partial(\rho v^2)}{\partial x}=-\frac{\partial p}{\partial x}+F/S,F=-\rho gS-\frac{\partial}{\partial x}\left(\mu\frac{\partial v}{\partial x}\right),p=f(\rho,T)$$

$$\frac{\partial}{\partial t}\left[\rho(\varepsilon+v^2/2)\right]+\frac{\partial}{\partial x}\left[v\left(\rho(\varepsilon+v^2/2)\right)\right]=-\frac{\partial(pv)}{\partial x}+\frac{Fv}{S}+\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)-k(T-T_{env}),T=f(\varepsilon)=\varepsilon/c_v \qquad (1)$$

$$\frac{\partial s}{\partial t}+\frac{\partial(vs)}{\partial x}=\frac{\partial}{\partial x}\left(\mu\frac{\partial s}{\partial x}\right),$$

where $p=f(\rho,T,s)$, $\varepsilon=f(\rho,v,T)$ — the empirical equation of the state of seawater and the equation, closing the system by internal energy, respectively; p — total hydrodynamic pressure; ρ — local density of the aquatic environment; v — projection of the velocity vector function on the vertical (z axis is directed upwards from the bottom to the surface); ε — volume density of internal energy; p — pressure of the gas enclosed in an elementary volume between adjacent layers; F — volume density of the generalized force (sum of forces) applied to elementary volumes of fluid, in addition to pressure; $s\equiv s(x,t)$ — salt concentration; S — cross-sectional area of the cylindrical selected area in which the most intensive upwelling and salt transport process is assumed; T — absolute water temperature; T_{env} — temperature of the water external to the allocated volume; k — thermal conductivity of water; $g=9{,}8m/s^2$ — acceleration of gravity; μ — coefficient characterizing the intensity of momentum transfer due to viscosity.

The boundary conditions characterizing the property of impermeability of fluid through the rock that makes up the bottom of a water body for system (1) can be written in the form of equalities:

$$(v,n)=0,\frac{\partial p}{\partial n}=0,\frac{\partial T}{\partial n}=0,\frac{\partial s}{\partial n}=0. \qquad (2)$$

Where n — the normal vector directed inside the computational domain.

The International Thermodynamic Equation of Seawater [10] defines density ρ of seawater as a function of salinity s, temperature T and hydrostatic pressure p, which has the form:

$$\rho(s,T,p)=\frac{\rho(s,T,0)}{1-p/K(s,T,p)}, \tag{3}$$

where K(s, T, p) — average modulus of elasticity; digit 0 corresponds to one standard atmosphere (101,325 Pa).

Under numerical modeling based on (1) – (3), also in demand as (3), we establish an algebraic relationship of pressure with density, temperature and salinity. The quadratic equation has two roots:

$$p_{1,2}=\frac{\pm\sqrt{D}+(1-A)\cdot\rho(s,T,p)+A\cdot p_0(s,T,p)}{D_N}, \tag{4}$$

$$D_N=2B\cdot(\rho(s,T,p)-\rho_0(s,T,p))$$

$$D=(-4BK_0(s,T,p)+A^2-2A+1)\cdot\rho(s,T,p)^2+(8B\cdot K_0(s,T,p)-2A^2+2A)\cdot p_0(s,T,p)\cdot \tag{5}$$

$$\cdot\rho(s,T,p)+(A^2-4B\cdot K_0(s,T,p))\cdot p_0(s,T,p)^2,$$

where A, B, $p_0(s,T,p)$, $K_0(s,T,p)$ — variable parameters whose relation to temperature, salinity and density is determined by the standard model of seawater; $\rho(s,T,p)=\rho$, where T — water Celsius temperature. The speed of sound in salt water can be expressed by the formula: $c(s,T,p)=\sqrt{\partial p/\partial\rho}$.

Suppose that the momentum, salinity and heat of a volume of water, bounded by a cylindrical surface, change only when the parameters at the boundary of the domain (Dirichlet boundary conditions) change, in particular, the salinity changes on the surface of the water. The behavior of mechanical parameters — density, momentum — is set by Neumann conditions and is inaccessible to external influence. Regardless of these assumptions, within the boundary of the aqueous medium, the system of equations (1) can be written in the vector form of a transfer-reaction model with a source additive [11, 5]:

$$\frac{\partial q}{\partial t}+\frac{\partial f}{\partial x}=T, \tag{6}$$

where $T=\left(0 \quad F/S+\tau\chi \quad Fv/S+k\partial^2T/\partial x^2-k(T-T_{env})\right)^\tau$ — vector characterizing the interaction of the flow with the surrounding fluid and the planet; $q=(q\ \rho v\ E)^\tau$ — vector of conservative state variables; $f=\left(\rho v\ p+\rho v^2\ v(E+p)\right)^\tau$ — vector of flows acting as feedback and closing the balance equation. From the components of vector f common multiplier can be derived $f=v\cdot q$.

In practice, to solve the transfer equation of form (6), a difference scheme has proven itself well, whose layered transition operator is obtained by a linear combination of similar transition operators of the Upwind and Standard Leapfrog schemes [12–15]. Taking into account the relationship between mass flows and density changes turned out to be more effective for a quasi-hydrodynamic system regularized according to B.N. Chetverushkin, approximated according to the FTCS scheme (forward-time central-space).

The discretization of the continuum model will be carried out by the integro-interpolation method on a uniform grid $S=S_1\times S_2$ $S_1=\{x_i=ih, i=0\div n, n\cdot h=L\}$, $S_2=\{t^j=j\tau,\ j=0\div m, m\cdot\tau=T\}$, where h — vertical grid step, i — index of the node (control/final volume in terms of the Godunov method) when numbering in space, τ — grid step in time, j — the number of the time layer. We describe a finite-difference approximation of the model. The difference schemes used to approximate the first-order balance equations give relatively acceptable results only with a very small grid step [3], which causes an active consumption of resources of computing devices and developers who create algorithms and programs that numerically implement them.

The problem of setting boundary conditions and their coordinated assignment at alternating speeds was solved by the method of filling control cells [7]. Taking into account the maximum occupancy and expressing its functional dependence on the node number makes it possible to increase the accuracy of approximation of boundary conditions.

Investigations on the theoretical and experimental selection of approximation methods and schemes led to the choice of splitting methods [13] and regularization. The continuity equations were regularized by B. N. Chetverushkin. The given equation was subjected to a difference approximation according to the FTCS scheme. The balance (transfer) equations of momentum, salinity and total energy were replaced by explicit equations obtained by linear combination of various approximations of the transfer operator (leapfrog and Upwind Leapfrog) [12, 15]. According to the splitting method, the transfer of momentum (and velocity), salinity, and total energy was approximated in a fractional step. The water velocity variation, which determined the intensity of transport and mass flows, was introduced at the second fractional step by solving the wave equation. Denote $\rho \equiv \rho_n$, $\tilde{\rho} = \rho^{n+\sigma}$, $\hat{\rho} = \rho^{n+1}$, $v \equiv v_n$, $\tilde{v} = v^{n+\sigma}$, $\hat{v} = v^{n+1}$, $0 < \sigma \le 1$. In the split form [16], after the introduction of a regularizer into the continuity equation and semi-discretization on grid S_2, approximation of the partial derivative in time by the FTCS scheme, the first two equations of system (1) can be written as:

$$\frac{\tilde{\rho}\tilde{v} - \rho v}{\tau} + v \cdot (\rho v)_x' = -\rho g S(x) - \frac{\partial}{\partial x}\left(\mu \frac{\partial v}{\partial x}\right), \tag{7}$$

$$\frac{1}{\tau}\left(\frac{\hat{\rho}_i - \rho_i}{\tau} + (\tilde{\rho}\tilde{v})_x'\right) = p_{xx}'' - \left(\frac{P_i^{n+1} - 2P_i^n + P_i^{n-1}}{c^2\tau^2}\right), p = f(\rho, T), c \equiv c(s, T, p), \tag{8}$$

$$\frac{\hat{\rho}\hat{v} - \tilde{\rho}\tilde{v}}{\tau} = -p_x'. \tag{9}$$

Equations (7) and (9) are a discrete analogue of the momentum transfer and change model — the second equation of system (1). Equation (8) predicts a change in the pressure field taking into account the continuity of the flow and information about internodal (intercellular) and boundary flows. Assuming time constancy at each step of the speed of sound, equation (8) can be approximated by an implicit time difference scheme and solved by sweeping. It was this direction in numerical modeling that was selected for the study.

Equation (7) and the fourth equation (1) are approximated on an equidistant grid by an explicit difference scheme [13]. Summands (8) determining the wave properties during the propagation of the pulse (the right side of the equation) are approximated by the central differences in space:

$$\frac{k_{0,i}}{c^2}\frac{p_i^{n+1} - 2p_i^n + p_i^{n-1}}{\tau^2} - \left(k_{1,i}\frac{p_{i+1}^{n+1} - p_i^{n+1}}{h^2} - k_{2,i}\frac{p_i^{n+1} - p_{i-1}^{n+1}}{h^2}\right), \tag{10}$$

where k_1 — occupancy of the cell to the right of the node with index i; k_2 — occupancy of the cell to the left of the node with index i; k_0 — degree of occupancy of the control area $x_{i-1/2} \le x \le x_{i+1/2}$, located in the neighborhood of the node with index i (k_0, k_1, k_2 — i index-dependent variables) [7].

The remaining two summands (8) determine the relationship of the rate of change in density with its flow. (8) includes the ratio of the rate of change in density to the time interval τ, expressed by the difference with a forward shift:

$$k_{0,i}\frac{\rho_i^{n+1} - \rho_i^n}{\tau^2},$$

the mass flow is transferred to the right side (8) and approximated by the central differences:

$$k_{0,i}\frac{1}{\tau}\frac{\partial(\tilde{\rho}\tilde{u})}{\partial x}\bigg|_{x=x_i} \approx k_{1,i}\frac{((\rho u)_{i+1/2} - (\rho u)_i)}{h\cdot\tau} + k_{2,i}\frac{((\rho u)_i - (\rho u)_{i-1/2})}{h\cdot\tau} \tag{11}$$

where k_0, k_1, k_2 — degree of occupancy of the areas located in the vicinity of the cells with number i [7]; k_0 characterizes the occupancy of the area $[x_{i-1}, x_{i+1}]$; k_1 — $[x_i, x_{i+1}]$, k_2 — $[x_{i-1}, x_i]$.

Balance equations of the mechanical impulse and pressure (9) are approximated by FTCS:

$$k_{0,i}\frac{\rho u_i^{n+1} - \rho u_i^{n+1/2}}{\tau} = -\left(k_{1,i}\frac{p_{i+1}^{n+1} - p_i^{n+1}}{2h} + k_{2,i}\frac{p_i^{n+1} - p_{i-1}^{n+1}}{2h}\right). \tag{12}$$

Let us present the problem of solving the difference approximation of equation (8) in the form of a matrix sweep problem with a vector variable in time on the right side $Ax = F$:

$$A_i x_{i-1} + C_i x_i + B_i x_{i+1} = F_i .$$

The approximation of equation (8), defined on a three-point difference template, in the form of a linear system of equations $Ap^{n+1} = f(p^n, p^{n-1}, (\tilde{\rho}\tilde{v})_x^{'})$, solvable with respect to pressure (p), has the form:

$$\begin{pmatrix} \frac{1}{2c^2\tau^2}+\frac{1}{h^2} & -\frac{1}{h^2} & 0 & \cdots & \cdots & \cdots & 0 \\ -\frac{1}{h^2} & \frac{1}{c^2\tau^2}+\frac{2}{h^2} & -\frac{1}{h^2} & & & & 0 \\ \vdots & \ddots & \ddots & \ddots & & & 0 \\ 0 & & -\frac{1}{h^2} & \frac{1}{c^2\tau^2}+\frac{2}{h^2} & -\frac{1}{h^2} & & \vdots \\ 0 & & & \ddots & \ddots & \ddots & \vdots \\ 0 & & & 0 & -\frac{1}{h^2} & \frac{1}{c^2\tau^2}+\frac{2}{h^2} & -\frac{1}{h^2} \\ 0 & & & \cdots & \cdots & -\frac{1}{h^2} & \frac{1}{2c^2\tau^2}+\frac{1}{h^2} \end{pmatrix} \cdot \begin{pmatrix} p_0^{n+1} \\ \vdots \\ p_{i-1}^{n+1} \\ p_i^{n+1} \\ p_{i+1}^{n+1} \\ \vdots \\ p_N^{n+1} \end{pmatrix} = \tag{13}$$

$$= \left(\frac{1}{2c^2\tau^2}\cdot(2p_0^n - p_0^{n-1}) \quad \cdots \quad \cdots \quad \frac{1}{c^2\tau^2}\cdot(2p_i^n - p_i^{n-1}) \quad \cdots \quad \cdots \quad \frac{1}{2c^2\tau^2}\cdot(2p_N^n - p_N^{n-1}) \right)^{\tau} +$$

$$+\frac{1}{\tau}\left((\hat{\rho}_0 - \rho_0)/2 \quad \cdots \quad \cdots \quad \hat{\rho}_i - \rho_i \quad \cdots \quad \cdots \quad (\hat{\rho}_N - \rho_N)/2 \right)^{\tau} +$$

$$+\frac{1}{2\tau h}\left((\tilde{\rho}\tilde{v})_1 \quad \cdots \quad \cdots \quad (\tilde{\rho}\tilde{v})_{i+1} - (\tilde{\rho}\tilde{v})_{i-1} \quad \cdots \quad \cdots \quad -(\tilde{\rho}\tilde{v})_{N-1} \right)^{\tau}.$$

A similar band matrix is obtained during a similar approximation of the other two terms (8).

The model can take into account dynamic changes in the flow at the boundaries and other parameters, such as temperature, salinity, oxygen content. Therefore, the simulation of water movement must be performed in a time cycle in layers with the retention of operational information about the parameters on at least two time-adjacent layers of the solution to the grid equation. The algorithm for calculating hydrodynamic parameters on a two-index grid in space and time includes:

– building of the projected pulse changing according to the first equation (3);

– approximate calculation of the function of the spatial distribution of pressure as a function of density and temperature from the previous time layer;

– assessment of the density changes according to the second equation (3);

– calculation of the pressure gradient from the new values of density and temperature, correction of the momentum distribution according to the third equation (3);

– finding a new distribution of the total energy and temperature according to the approximated third equation of the system (1).

When algorithmizing the method for solving a system of split equations, the following are introduced:

– binary-numeric masks that predetermine the switching of the difference scheme template with a change in the sign of the velocity;

– variable shifts of the indices of neighboring nodes, which make it possible to write the solution to equation (12) for boundary and internal nodes by a single system of computational operations.

Such variable index shifts are used in calculating gradient approximations both in the volume of the continuum model and at the boundaries with the second order of accuracy according to discrete analogues of equations (7) – (9) approximated by FTCS and by a linear combination of Leapfrog and Upwind Leapfrog schemes [12–14]:

$$\begin{aligned} &\frac{q_i^{n+1} - q_i^n}{\tau} + \frac{4}{3}\left(\frac{q_{i-1}^n - q_{i-1}^{n-1}}{2\tau} + v_i^n \frac{q_i^n - q_{i-1}^n}{h} \right) + \frac{q_i^n - q_i^{n-1}}{3\tau} + v_i^n \frac{q_{i+1}^n + q_{i-1}^n}{3h} = 0, v_i^n \geq 0, \\ &\frac{q_i^{n+1} - q_i^n}{\tau} + \frac{4}{3}\left(\frac{q_{i+1}^n - q_{i+1}^{n-1}}{2\tau} + v_i^n \frac{q_i^n - q_i^n}{h} \right) + \frac{q_i^n - q_i^{n-1}}{3\tau} + v_i^n \frac{q_{i+1}^n - q_{i-1}^n}{3h} = 0, v_i^n < 0, \end{aligned} \tag{14}$$

where is characteristic q for transfer of salt ($q \equiv s$) and momentum $q \equiv \rho u$.

Parameter *m* is a “switch” of the flow when the sign of the velocity changes in the difference approximations of the pressure gradients and the mass flow:

$$\bar{m} = 1 - m,\ m = \begin{cases} 0, & v < 0, \\ 1, & v \geq 0. \end{cases}$$

When using such a switch, the approximation of the gradient in solving the transfer problem for momentum and salinity can be written by the formula:

$$\left.\frac{\partial q}{\partial x}\right|_{x=x_i} \approx \frac{2}{3}\left(\mathrm{m}\frac{q_i - q_{i-1}}{h} + \bar{\mathrm{m}}\frac{q_{i+1} - q_i}{h}\right) + \frac{1}{3}\frac{q_{i+1} - q_{i-1}}{2h}. \quad (15)$$

The use of the cell occupancy method makes it possible to correctly approximate the boundary conditions when placing water state parameters in computer memory with an array of numerical values, and variable index shifts enable to reduce the volume of symbolic recording of the subroutine for performing layer-by-layer iterative changes of state variables assigned to the nodes of the difference scheme [17].

The matrix equation with a band tridiagonal matrix A (13) is solved by sweeping [18]. For the initial verification of the software implementation, a simplified model of the state of salt water in the form of P. S. Lineikin's equation was used.

Research Results. The simulation was performed on the basis of a software package written in MATLAB. Debugging was performed using the GNU Octave interpreter. Operation also presupposes the availability of libraries of this system. The software package consists of 16 functional modules. The selection of this interpreter and the corresponding language was due to the ability to write and test a program that operates with an array of state variables, which is a projection of the desired functions $\rho(x,t)$, $v(x,t)$, $\varepsilon(x,t)$ on a spatial grid. At this stage of the study, the authors abstracted from the specifics of performing element-by-element operations on arrays of real numbers.

The software system consists of the interconnected subprograms:

– execution of one step on the transfer equation according to formulas (10);

– calculation of seawater density (unesco_urs);

– calculation of the pressure of seawater located in the gravity field at a given temperature, salinity and density according to empirical equations of state (rhoTS2P);

– calculation of the speed of sound depending on temperature, salinity and pressure according to the standard and the updated UNESCO formula (speed_of_sound);

– estimation of the fluid viscosity when its volumes move under the influence of pressure forces at all points of the spatial grid (ForceOfFriction);

– cyclic variation of state and time variables during vertical movement of salt water (aqua_process);

– initial setting of constant values characterizing the fluid, initial and boundary conditions of its global hydrophysical equilibrium (start, set_parameters);

– formation of band matrix approximating the equation containing pressure (func5);

– solution to the matrix equation by the sweep method (run_sweep_shuttle);

– temperature ↔ total energy conversion (TFromE, EFromT);

– calculation of total energy balance (TotalPower);

– solution to the diffusion-convection-reaction problem, including the approximation of a combination of Leapfrog and Upwind Leapfrog difference schemes (ADR_solver);

– estimation of gradients and time derivatives with respect to central and directional differences, taking into account the change in the sign of the velocity and the pattern of the difference scheme (diff123).

The constructed model was used in a test run to estimate the density change under an increase in the salinity of the surface of the aquatic environment by 1% at the 30th second from the start of the simulation and a maximum depth of 1 km. Density transients in a set of cross-sections (function $\rho(t,x)\big|_{x\in S_1}$), spaced from each other, caused by an instantaneous change in salt concentration, are finite in time (Fig. 1), and the density pulse front is greatly weakened when moving in space.

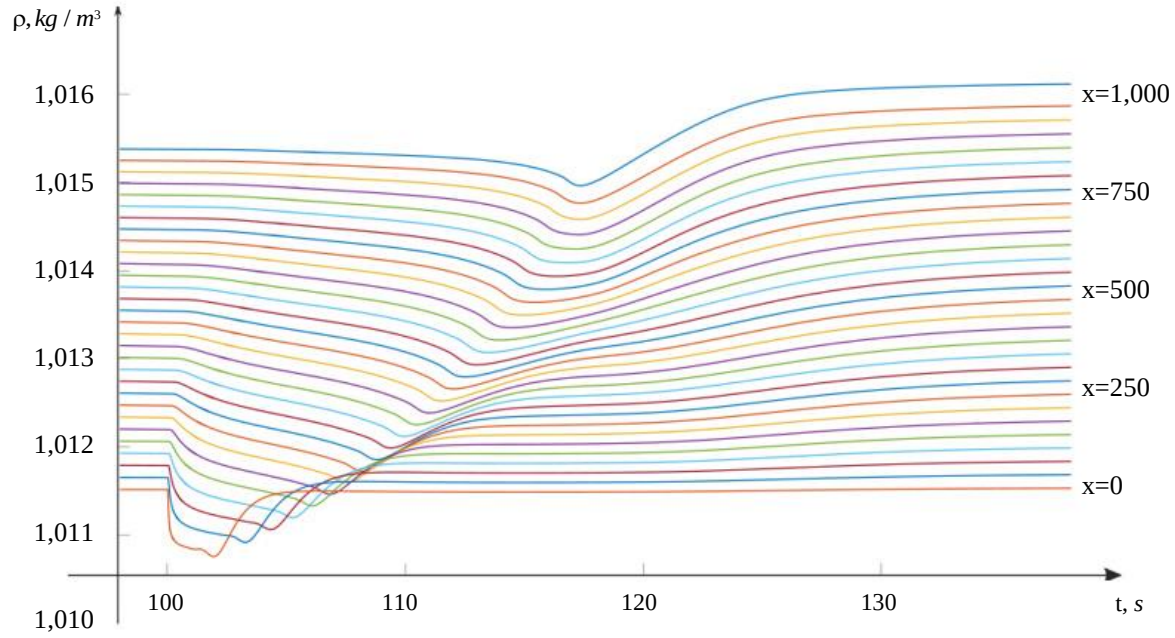


Fig. 1. Graphs of dependence of density on time with a sharp change in salinity on the water surface

Discussion and Conclusion. The presented spatially distributed model does not yet allow predicting a steady redistribution of density and a change in the salinity gradient in numerical calculations, because it does not take into account the buoyancy of less salty water and its change during salinization of the upper layers. Mathematics and software, which greatly simplify the modeling of processes leading to the observed effects of halocline, chemocline and pycnocline, has been tested and debugged on a variety of test tasks (momentum and salt transport, pressure wave propagation). For high-precision modeling of the observed physico-chemical phenomena in the seas, it is required to solve additional tasks of identifying model parameters, taking the data of observations and remote sensing of the Earth as initial information.

References

1. Mikhailov VN, Dobrovolskii AD, Dobrolyubov SA. *Hydrology*, 2nd ed. Moscow: Vysshaya shkola; 2007. 464 p. URL: https://www.geokniga.org/bookfiles/geokniga-mihaylov-vn-dobrovolskiy-ad-gidrologiya-2007.pdf (accessed: 01.04.2023) (In Russ.).

2. Il'ichev VG. *Sustainability, Adaptation and Management in Ecological Systems*. Monograph. Moscow: Fizmatlit; 2008. 231 p. URL: https://www.evolbiol.ru/docs/docs/large_files/ilyichev.pdf (accessed: 01.04.2023) (In Russ.).

3. Bogdanov NI. *Biological Restoration of Water Reservoirs.* Penza: RIO PGSKHA; 2008. 126 p. URL: https://microalgae.ru/f/bogdanov_na_biologicheskaya_reabilitaciya_vodoemov_2008.pdf (accessed: 01.04.2023) (In Russ.).

4. Sudolskii AS. *Dynamic Phenomena in Water Reservoirs*. Leningrad: Gidrometioizdat; 1991. 260 p. URL: http://elib.rshu.ru/files_books/pdf/img-217140610.pdf (accessed: 01.04.2023) (In Russ.).

5. Gevorkyan VKh. Lithological Aspects of V. I. Vernadsky's Doctrine about the Biosphere. *Geology and Mineral Resources of World Ocean.* 2010;21(3):37–56. URL: https://core.ac.uk/download/38371310.pdf (accessed: 01.04.2023) (In Russ.).

6. Il'ichev VG, Rokhlin DB. Internal Prices and Optimal Exploitation of Natural Resources. *Mathematics*. 2022;10(11):1860. https://doi.org/10.3390/math10111860

7. Sukhinov AI., Chistyakov AE., Belova YuV., et al. Supercomputer Modeling of Hydrochemical Condition of Shallow Waters in Summer Taking into Account the Influence of the Environment. *Communications in Computer and Information Science*. 2018;910:336–351. https://doi.org/10.1007/978-3-319-99673-8_24

8. Nikitina AV, Kravchenko L, Semenov IS, et al. Modeling of Production and Destruction Processes in Coastal Systems on a Supercomputer. *MATEC Web of Conference*. 2018;226(2):04025. https://doi.org/10.1051/matecconf/201822604025

9. Iakovlev NG, Volodin EM, Gritsun AS. Simulation of the Spatiotemporal Variability of the World Ocean Sea Surface Height by the INM Climate Models. *Atmospheric and Oceanic Physics*. 2016;52(4):376–385. https://doi.org/10.1134/S0001433816040125

10. Mcdougall TJ, Millero FJ, Feistel R, et al. *The International Thermodynamic Equation of Seawater - 2010: Calculation and Use of Thermodynamic Properties*. Paris: UNESCO; 2010. 196 p.

11. Kudinov NV, Neydorf RA, Zhuravlev LA, et al. Using Simulink Package for Transient Support-Parametric Simulation in Gas Pipeline Section. *Vestnik of DSTU*. 2012;12(1-2):60–66. URL: https://www.vestnik-donstu.ru/jour/article/view/498 (accessed: 02.02.2023).

12. Sukhinov AI, Chistyakov AE, Protsenko EA. Difference Scheme for Solving Problems of Hydrodynamics for Large Grid Peclet Numbers. *Computer Research and Modeling.* 2019;11(5):833–848. https://doi.org/10.20537/2076-7633-2019-11-5-833-848

13. Sukhinov AI, Chistyakov AE, Protsenko EA, et al. Linear Combination of Upwind and Standard Leapfrog Difference Schemes with Weight Coefficients Obtained by Minimizing the Approximation Error. *Chebyshevskii Sbornik*. 2020;21(4):243–256. https://doi.org/10.22405/2226-8383-2020-21-4-243-256

14. Sukhinov AI, Belova YuV, Chistyakov AE. Mathematical Modeling of Biogeochemical Cycles in Coastal Systems of the South Russia. *Mathematical Models and Computer Simulations*. 2021;13(6):930-942. https://doi.org/10.20948/mm-2021-03-02

15. Sukhinov AI, Chistyakov AE, Protsenko EA. Upwind and Standard Leapfrog Difference Schemes. *Numerical Methods and Programming.* 2019;20(2):170–181. https://doi.org/10.26089/NumMet.v20r216

16. Gushchin VA. Development and Application of the Method of Splitting by Physical Factors for the Study of the Incompressible Fluid Flows. *Computer Research and Modeling.* 2022;14(4):715–739. https://doi.org/10.20537/2076-7633-2022-14-4-715-739

17. Kudinov NV, Nikitina AV. *Computer Models of Axisymmetric Gas Flow through Channels for Solving Technical and Natural Science Problems*. In: Proc. Int. Conf. "Intelligent Information Technologies and Mathematical Modeling (IIT&MM-2022)". Don State Technical University; 2022. P. 93–100. (In Russ.).

18. Sukhinov AI, Chistyakov AE, Nikitina AV, et al. A Method of Solving Grid Equations for Hydrodynamic Problems in Flat Areas. *Mathematical Models and Computer Simulations*. 2023;35(3):35–58. https://doi.org/10.20948/mm-2023-03-03

Received 03.04.2023
Revised 24.04.2023
Accepted 04.05.2023

About the Authors:

Nikita V. Kudinov, Cand.Sci. (Eng.), Associated Professor of the Computer and Automated Systems Software Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), ScopusID, ResearcherID, ORCID, AuthorID, kudinov_nikita@mail.ru

Alena A. Filina, Cand.Sci. (Eng.), Researcher of the System Software Department, "Supercomputers and Neurocomputers Research Center" Co Ltd (106, Italyansky lane, Taganrog, Rostov Region, 347900, RF), ScopusID, ORCID, AuthorID, j.a.s.s.y@mail.ru

Alla V. Nikitina, Dr.Sci. (Eng.), Professor of the Computer and Automated Systems Software Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), Researcher, "Supercomputers and Neurocomputers Research Center" Co Ltd (106, Italyansky lane, Taganrog, Rostov Region, 347900, RF), ScopusID, ORCID, AuthorID, nikitina.vm@gmail.com

Irina F. Razveeva, Senior lecturer of the Mathematics and Informatics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), AuthorID, razveevai@mail.ru

Denis V. Bondarenko, Graduate student of the Computer and Automated Systems Software Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344003, RF), AuthorID, denis.bondarenko.dev@gmail.com

Claimed contributorship:

NV Kudinov: supplementing the control cell method with an algorithmic model of masks and indices; software implementation of the solution to hydrodynamic equations: an algorithm for updating values in virtual cells, data structures for storing hydrodynamic fields based on N-dimensional arrays; integration of software modules.

AA Filina: formalization of the relationship between bioecological and physicochemical processes in reservoirs of regional significance; testing the software implementation of the matrix sweep algorithm; selecting a method for spatiotemporal approximation of the continuity and Navier-Stokes equations.

AV Nikitina: preliminary scientific editing; scientific consulting of the working group.

IF Razveeva: analysis of hydrophysical properties of salt water according to field observations published on the Internet; correction of the text of the article.

DV Bondarenko: investigation, comparison, selection of equations of water states; implementation of an algorithm visualizing the dependence of the speed of sound on the density and salinity of the aquatic environment; testing of the software implementation of a computer model of salt water; correction of the text of the article.

Conflict of interest statement: the authors do not have any conflict of interest.

All authors have read and approved the final manuscript.

Поступила в редакцию 03.04.2023
Поступила после рецензирования 24.04.2023
Принята к публикации 04.05.2023

Об авторах:

Никита Валерьевич Кудинов, кандидат технических наук, доцент кафедры программного обеспечения вычислительной техники и автоматизированных систем Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), ScopusID, ResearcherID, ORCID, AuthorID, kudinov_nikita@mail.ru

Алёна Александровна Филина, кандидат технических наук, научный сотрудник отдела системного программного обеспечения, ООО «НИЦ супер-ЭВМ и нейрокомпьютеров» (347900, РФ, Таганрог, пер. Итальянский, 106), ScopusID, ORCID, AuthorID, j.a.s.s.y@mail.ru

Алла Валерьевна Никитина, доктор технических наук, профессор кафедры программного обеспечения вычислительной техники и автоматизированных систем Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), научный сотрудник, ООО «НИЦ супер-ЭВМ и нейрокомпьютеров» (347900, РФ, Ростовская область, Таганрог, пер. Итальянский, 106), ScopusID, ORCID, AuthorID, nikitina.vm@gmail.com

Ирина Федоровна Развеева, старший преподаватель кафедры математики и информатики Донского государственного технического университета (344003, г. Ростов-на-Дону, пл. Гагарина, 1), AuthorID, razveevai@mail.ru

Денис Вадимович Бондаренко, магистрант кафедры программного обеспечения вычислительной техники и автоматизированных систем Донского государственного технического университета (344003, г. Ростов-на-Дону, пл. Гагарина, 1), AuthorID, denis.bondarenko.dev@gmail.com

Заявленный вклад соавторов:

Н.В. Кудинов — дополнение метода контрольных ячеек алгоритмической моделью масок и индексов; программная реализация решения гидродинамических уравнений: алгоритм обновления значений в виртуальных ячейках, структуры данных для хранения гидродинамических полей на основе N-размерных массивов; комплексирование программных модулей.

А.А. Филина — формализация взаимосвязи биоэкологических и физико-химических процессов в водоёмах регионального значения; тестирование программной реализации алгоритма матричной прогонки, выбор способа пространственно-временной аппроксимации уравнений неразрывности и Навье-Стокса.

А.В. Никитина — предварительное научное редактирование, научное консультирование рабочей группы.

И.Ф. Развеева — анализ гидрофизических свойств солёной воды по данным натурных наблюдений, опубликованных в сети Интернет; корректировка текста статьи.

Д.В. Бондаренко — изучение, сравнение, выбор уравнений состояний воды; реализация алгоритма, визуализирующего зависимости скорости звука от плотности и солёности водной среды; тестирование программной реализации компьютерной модели солёной воды; корректировка текста статьи.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Все авторы прочитали и одобрили окончательный вариант рукописи.