VESTNIK
of
DON STATE
TECHNICAL
UNIVERSITY

Vol. 19, № 1

Theoretical
and scientific-practical journal

Published since 1999

4 issues a year
January-March 2019

ISSN 1992-5980
eISSN 1992-6006
DOI: 10.23947/1992-5980

Founder and publisher — Don State Technical University

Included in the list of peer-reviewed scientific editions where the basic research results of doctoral, candidate's theses should be published (State Commission for Academic Degrees and Titles List) in the following research areas:

01.02.01 – Analytical Mechanics (Engineering Sciences)
01.02.04 – Deformable Solid Mechanics (Engineering Sciences)
01.02.04 – Deformable Solid Mechanics (Physicomathematical Sciences)
01.02.06 – Dynamics, Strength of Machines, Gear, and Equipment (Engineering Sciences)
05.02.02 – Engineering Science, Drive Systems and Machine Parts (Engineering Sciences)
05.02.04 – Machine Friction and Wear (Engineering Sciences)
05.02.07 – Technology and Equipment of Mechanical and Physicotechnical Processing (Engineering Sciences)
05.02.08 – Engineering Technology (Engineering Sciences)
05.02.10 – Welding, Allied Processes and Technologies (Engineering Sciences)
05.02.11 – Testing Methods and Diagnosis in Machine Building (Engineering Sciences)
05.13.11 – Software and Mathematical Support of Machines, Complexes and Computer Networks (Engineering Sciences)
05.13.17 – Foundations of Information Science (Engineering Sciences)
05.13.18 – Mathematical Simulation, Numerical Methods and Program Systems (Engineering Sciences)

The journal is indexed and archived in the Russian Science Citation Index (RSCI), and in EBSCO International Database
The journal is a member of Directory of Open Access Journals (DOAJ), Association of Science Editors and Publishers (ASEP) and Cross Ref

Certificate of mass media registration ПИ № ФС 77-66004 of 06.06.2016 is issued by the Federal Service for Supervision of Communications, Information Technology, and Mass Media

The subscription index in Rospechat catalogue is 35578

The issue is prepared by:
Inna V. Boyko, Marina P. Smirnova (English version)
Passed for printing 28.03.2019,
imprint date 30.03.2019.
Format 60×84/8. Font «Times New Roman».
C.p.sh. 22.6. Circulation 1000 cop. Order no. 28/03 Free price.

Founder's, Publisher's and Printery Address:
Gagarin Sq. 1, Rostov-on-Don, 344000, Russia. Phone: +7 (863) 2-738-372
E-mail: vestnik@donstu.ru http://vestnik.donstu.ru/

CONTENT

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 4–12. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 4–12. ISSN 1992-5980 eISSN 1992-6006

MECHANICS
МЕХАНИКА

UDC 519.6 https://doi.org/10.23947/1992-5980-2019-19-1-4-12

Study on stationary solutions to the problem of phytoplankton dynamics considering transformation of phosphorus, nitrogen and silicon compounds*

Yu.V. Belova[1], A.M. Atayan[2], A.E. Chistyakov[3], A.V. Strazhko[4]**
[1,2,3,4] Don State Technical University, Rostov-on-Don, Russian Federation

Исследование стационарных решений задачи динамики фитопланктона с учетом трансформации соединений фосфора, азота и кремния ***

Ю. В. Белова[1], А. М. Атаян[2], А. Е. Чистяков[3], А. В. Стражко[4]**
[1,2,3,4] Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация

Introduction. The solution to the problem of transformation of phosphorus, nitrogen and silicon forms is studied. This problem arises under modeling phytoplankton dynamics in shallow-water bodies including the Azov Sea. The phytoplankton dynamics model is formulated as a boundary value problem for the system of diffusion-convection-response equations and takes into account the absorption and release of nutrients by phytoplankton, as well as the transition of nutrients from one compound to another. To calculate the initial conditions and parameters of the equations under which the steady-state regime occurs, the software is developed, which is based on the model describing changes in phytoplankton concentrations without considering current effects. This model is represented by a system of inhomogeneous differential equations. Based on the developed software, the initial conditions and parameters of the phytoplankton dynamics model in the Azov Sea are calculated experimentally.

Materials and Methods. A 3D model of phytoplankton dynamics is considered taking into account the transformation of phosphorus, nitrogen and silicon compounds based on the system of nutrient transport equations. The case of a spatially uniform distribution of substances is considered to specify the parameters of the model at which the stationary modes occur. Because of simplification, a system of ordinary differential equations solved through the Runge-Kutta method is obtained.

Research Results. The software is developed to specify the initial conditions and parameters of the phytoplankton dynamics model considering the transformation of phosphorus, nitrogen and silicon compounds.

Several numerical experiments are performed under the assumption that the development of phytoplankton is limited by

Введение. Работа посвящена исследованию решения задачи трансформации форм фосфора, азота и кремния. Данная проблема возникает при моделировании динамики фитопланктона в мелководных водоемах, в том числе в Азовском море. Модель динамики фитопланктона сформулирована как краевая задача для системы уравнений диффузии-конвекции-реакции и учитывает поглощение и выделение питательных веществ фитопланктоном, а также переход питательных веществ из одного соединения в другое. Для расчета начальных условий и параметров уравнений, при которых наступает стационарный режим, разработано программное обеспечение, основой которого послужила модель, описывающая изменения концентраций фитопланктона без учета влияния течений. Данная модель представлена системой неоднородных обыкновенных дифференциальных уравнений. На основе разработанного программного обеспечения экспериментальным образом рассчитаны начальные условия и параметры модели динамики фитопланктона в Азовском море.

Материалы и методы. Рассматривается трехмерная модель динамики фитопланктона с учетом трансформации соединений фосфора, азота и кремния, основанная на системе уравнений транспорта биогенных веществ. Для уточнения параметров модели, при которых наступают стационарные режимы, рассматривается случай пространственно-однородного распределения субстанций. В результате упрощения получена система обыкновенных дифференциальных уравнений, которая решена методом Рунге-Кутты.

Результаты исследования. Разработано программное обеспечение для уточнения начальных условий и параметров модели динамики фитопланктона с учетом трансформации соединений фосфора, азота и кремния. Проведены несколько численных экспериментов в предположении, что развитие фитопланктона лимитируется единственным биогенным веществом. В результате вычисли-

* The research is supported by the RSF (project no. 17-11-01286).
** E-mail: yvbelova@yandex.ru, atayan24@yandex.ru, cheese_05@mail.ru, strajcko2@gmail.com
*** Работа выполнена при поддержке РНФ (проект № 17-11-01286).

a single biogenic substance. As a result of the computational experiment, it can be seen that with the obtained values of the initial concentrations and parameters of the equations, stationary modes occur for the system of ordinary differential equations describing the case of the spatially uniform distribution of substances.

Discussion and Conclusions. The mathematical model of the transformation of phosphorus, nitrogen and silicon forms in the problem of phytoplankton dynamics is studied. Stationary modes for the system of ordinary differential equations are obtained, for which the values of the system parameters and initial conditions are determined. The results obtained can be used in further simulation of the phytoplankton dynamics considering the transformation of phosphorus, nitrogen and silicon compounds with account for convection-diffusion, salinity, and temperature.

Keywords: phytoplankton, phosphorus, nitrogen, silicon, biogen, chemical-biological source, convection-diffusion-response equation, Cauchy problem for system of ordinary differential equations, stationary mode.

For citation: Yu.V. Belova, et al. Study of stationary solutions to the problem of phytoplankton dynamics considering transformation of phosphorus, nitrogen and silicon compounds. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 4–12. https://doi.org/10.23947/1992-5980-2019-19-1-4-12

тельного эксперимента видно, что при полученных значениях начальных концентраций и параметров уравнений наступают стационарные режимы для системы обыкновенных дифференциальных уравнений, описывающей случай пространственно-равномерного распределения субстанций.

Обсуждение и заключения. В работе исследована математическая модель трансформации форм фосфора, азота и кремния в задаче динамики фитопланктона.

Получены стационарные режимы для системы обыкновенных дифференциальных уравнений, для которых определены значения параметров системы и начальные условия. Полученные результаты могут быть использованы в процессе дальнейшего моделирования динамики фитопланктона с учетом трансформации соединений фосфора, азота и кремния с учетом конвекции-диффузии, солености, температуры.

Ключевые слова: фитопланктон, фосфор, азот, кремний, биоген, химико-биологический источник, уравнение конвекции-диффузии-реакции, задача Коши для системы обыкновенных дифференциальных уравнений, стационарный режим.

***Образец для цитирования*.** Белова, Ю. В. Исследование стационарных решений задачи динамики фитопланктона с учетом трансформации соединений фосфора, азота и кремния / Ю. В. Белова [и др.] // Вестник Донского гос. техн. ун-та. — 2019. — Т.19, №1. — С. 4–12. https://doi.org/10.23947/1992-5980-2019-19-1-4-12

Introduction. Because of the development of major cities on the coast of shallow water bodies and river systems that flow into these water bodies, eutrophication has become more frequent. The growth of algae in reservoirs is caused by an increase in the flow of nitrogen and phosphorus compounds from the adjacent land areas. Each water body is unique and requires a thorough study. Field investigations [1] and mathematical modeling are used to explore water bodies. Without downplaying the role of field experiments, we can say that mathematical modeling is less costly, and it allows us to predict the behavior of the ecosystem.

To study the Sea of Azov, a three-dimensional model of hydrodynamics [2, 3] including the equations of motion in three spatial directions was developed. In [4], this model was made for the case of dynamic rebuilding of the computational domain geometry due to the tidal effects. The investigation of this model accuracy is given in [5]. In [6–8], the reconstruction of an ecological catastrophe that occurred in 2001 caused by an excessive concentration of algae in the eastern part of the Sea of Azov is given. In [9], methods of controlling the suffocation phenomena arising in the Sea of Azov were proposed. The [10–12] papers are devoted to studying the dynamics of phyto- and zooplankton.

The water condition in shallow water bodies is changing rapidly, and mathematical models need to be refined. The parameters determination of the three-dimensional model of the phyto- and zooplankton dynamics is laborious; therefore, it is proposed to use a simplified model to calculate these parameters.

The work objective is to improve the parameters of the model of the phytoplankton dynamics considering the transformation of phosphorus, nitrogen and silicon compounds, under which stationary regimes occur with the assumption of a spatially uniform distribution of substances.

Materials and Methods. The model is based on the system of equations for the transport of nutrients [15, 16], the form of which for each F_i model block is

$$\frac{\partial q_i}{\partial t}+u\frac{\partial q_i}{\partial x}+v\frac{\partial q_i}{\partial y}+w\frac{\partial q_i}{\partial z}=div\left(k\,\mathrm{grad}\,q_i\right)+R_{q_i}, \quad (1)$$

where q_i is concentration of the *i*-th component, [mg/l]; $i \in M$, $M=\{F_1, F_2, F_3, PO_4, POP, DOP, NO_3, NO_2, NH_4, Si\}$; $\{u,v,w\}$ are components of the velocity vector of the water flow, [m/s]; k is turbulent exchange coefficient, [m^2/s]; R_{q_i} is function-source of nutrients, [mg/l · s].

In equation (1), i index indicates the type of substance (Table 1).

Table 1

Biogenic substances in model of phytoplankton dynamics

No.	Notation	Name
1	F_1	Chlorella vulgaris green algae
2	F_2	Aphanizomenon flos-aquae green-blue algae
3	F_3	Sceletonema costatum diatom
4	PO_4	phosphates
5	POP	suspended organic phosphor
6	DOP	soluble organic phosphor
7	NO_3	nitrates
8	NO_2	nitrites
9	NH_4	ammonium
10	Si	soluble inorganic silica (silicic acids)

Chemical and biological reactions are described by the following equations $R_{F_i}=C_{F_i}(1-K_{F_iR})q_{F_i}-K_{F_iD}q_{F_i}-K_{F_iE}q_{F_i}$,

$$i=\overline{1,3},$$

$$R_{POP}=\sum_{i=1}^{3}s_P K_{F_iD}q_{F_i}-K_{PD}q_{POP}-K_{PN}q_{POP},$$

$$R_{DOP}=\sum_{i=1}^{3}s_P K_{F_iE}q_{F_i}+K_{PD}q_{POP}-K_{DN}q_{DOP},$$

$$R_{PO_4}=\sum_{i=1}^{3}s_P C_{F_i}\left(K_{F_iR}-1\right)q_{F_i}+K_{PN}q_{POP}+K_{DN}q_{DOP},$$

$$R_{NH_4}=\sum_{i=1}^{3}s_N C_{F_i}\left(K_{F_iR}-1\right)\frac{f_N^{(2)}\left(q_{NH_4}\right)}{f_N\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)}q_{F_i}+\sum_{i=1}^{3}s_N\left(K_{F_iD}+K_{F_iE}\right)q_{F_i}-K_{42}q_{NH_4},$$

$$R_{NO_2}=\sum_{i=1}^{3}s_N C_{F_i}(K_{F_iR}-1)\frac{f_N^{(1)}\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)}{f_N\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)}\cdot\frac{q_{NO_2}}{q_{NO_2}+q_{NO_3}}q_{F_i}+K_{42}q_{NH_4}-K_{23}q_{NO_2},$$

$$R_{NO_3}=\sum_{i=1}^{3}s_N C_{F_i}\left(K_{F_iR}-1\right)\frac{f_N^{(1)}\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)}{f_N\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)}\cdot\frac{q_{NO_3}}{q_{NO_2}+q_{NO_3}}q_{F_i}+K_{23}q_{NO_2},$$

$$R_{Si}=s_{Si}C_{F_3}\left(K_{F_3R}-1\right)q_{F_3}+s_{Si}K_{F_3D}q_{F_3}.$$

Here, K_{F_iR} is specific breathing rate of phytoplankton; K_{F_iD} is specific die-off rate of phytoplankton; K_{F_iE} is specific excretion rate of phytoplankton; K_{PD} is specific rate of *POP* autolysis; K_{PN} is coefficient of *POP* phosphatification; K_{DN} is coefficient of *DOP* phosphatification; K_{42} is specific rate of ammonium oxidation to nitrites under nitrification; K_{23} is specific rate of nitrite oxidation to nitrates under nitrification; s_P, s_N, s_{Si} are normalization coefficients between the content of *N, P, Si* in organic matter.

Phytoplankton growth rate is determined by the following expressions:

$$C_{F_{1,2}}=K_{NF_{1,2}}\,min\left\{f_P\left(q_{PO_4}\right),f_N\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)\right\},$$

$$C_{F_3}=K_{NF_3}\,min\left\{f_P\left(q_{PO_4}\right),f_N\left(q_{NO_3},q_{NO_2},q_{NH_4}\right),f_{Si}\left(q_{Si}\right)\right\},$$

where K_{NF} is maximum specific phytoplankton growth rate.

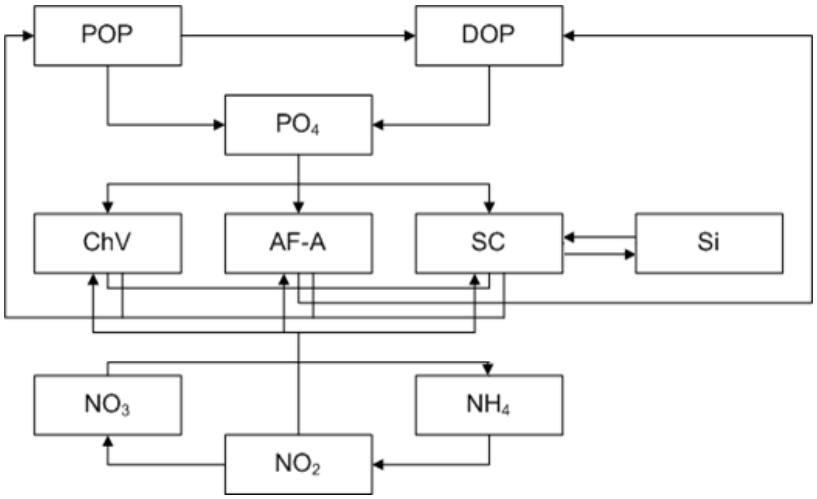


Fig. 1. Model scheme of biogeochemical transformation of phosphorus, nitrogen and silica forms

Functions describing nutrient content

– for phosphorus: $f_P\left(q_{PO_4}\right)=\dfrac{q_{PO_4}}{q_{PO_4}+K_{PO_4}}$,

where K_{PO_4} is phosphates half saturation constant;

– for silica: $f_{Si}\left(q_{Si}\right)=\dfrac{q_{Si}}{q_{Si}+K_{Si}}$,

where K_{Si} is silica half saturation constant;

– for nitrogen: $f_N\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)=f_N^{(1)}\left(q_{NO_3},q_{NO_2},q_{NH_4}\right)+f_N^{(2)}\left(q_{NH_4}\right)$,

$$f_N^{(1)}(q_{NO_3},q_{NO_2},q_{NH_4})=\frac{(q_{NO_3}+q_{NO_2})\exp(-K_{psi}q_{NH_4})}{K_{NO_3}+(q_{NO_3}+q_{NO_2})},\ f_N^{(2)}(q_{NH_4})=\frac{q_{NH_4}}{K_{NH_4}+q_{NH_4}},$$

where K_{NO_3} is nitrates half saturation constant, K_{NH_4} is ammonium half saturation constant, K_{psi} is ammonium inhibition ratio.

For the system (1), it is necessary to specify the vector field of the water flow velocities, as well as q_i initial values of the concentration functions

$$q_i\left(x,y,z,0\right)=q_i^0\left(x,y,z\right),\left(x,y,z\right)\in\overline{G}.,\ t=0,\ i\in M. \tag{2}$$

Assume Σ boundary of G cylindrical region is piecewise smooth and $\Sigma=\Sigma_H\cup\Sigma_o\cup\sigma$, where Σ_H is surface of the reservoir bottom, Σ_o is still water surface, σ is lateral (cylindrical) surface. Suppose u_n is vector component of the water flow velocity normal to Σ, and n is outward normal vector to Σ. For q_i concentrations, we assume:

– on side boundary:

$$q_i=0,\ \text{on}\ \sigma,\ \text{if}\ u_n<0,\ i\in M; \tag{3}$$

$$\frac{\partial q_i}{\partial n}=0,\ \frac{\partial q_i}{\partial n}=0,\ \text{on}\ \sigma,\ \text{if}\ u_n\geq 0,\ i\in M;$$

– on Σ_o there is reservoir surface:

$$\frac{\partial q_i}{\partial z}=0,\ i\in M; \tag{5}$$

– on the bottom Σ_H:

$$k\frac{\partial q_i}{\partial z}=\varepsilon_{1,i}q_i,\ i\in\{\mathrm{F}_1,\mathrm{F}_2,\mathrm{F}_3\},$$

$$k\frac{\partial q_i}{\partial z} = \varepsilon_{2,i} q_i,\ i \in \{\mathrm{PO_4},\ \mathrm{POP},\ \mathrm{DOP},\ \mathrm{NO_3},\ \mathrm{NO_2},\ \mathrm{NH_4},\ \mathrm{Si}\}, \tag{6}$$

where $\varepsilon_{1,i}$, $\varepsilon_{2,i}$ are sedimentation rates of algae and nutrients to the bottom.

Stationary Mode. Consider the case of a spatially uniform distribution of substances (phytoplankton, forms of phosphorus, nitrogen and silica); then each of the equations (1) is simplified; and as a result, we get the following system of ordinary differential equations (ODE):

$$\frac{dq_{F_i}}{dt} = C_{F_i}(1-K_{F_iR})q_{F_i} - K_{F_iD}q_{F_i} - K_{F_iE}q_{F_i},\ i=\overline{1,3}, \tag{7}$$

$$\frac{dq_{POP}}{dt} = \sum_{i=1}^{3} s_P K_{F_iD} q_{F_i} - K_{PD}q_{POP} - K_{PN}q_{POP},$$

$$\frac{dq_{DOP}}{dt} = \sum_{i=1}^{3} s_P K_{F_iE} q_{F_i} + K_{PD}q_{POP} - K_{DN}q_{DOP},$$

$$\frac{dq_{PO_4}}{dt} = \sum_{i=1}^{3} s_P C_{F_i}\left(K_{F_iR}-1\right)q_{F_i} + K_{PN}q_{POP} + K_{DN}q_{DOP},$$

$$\frac{dq_{NH_4}}{dt} = \sum_{i=1}^{3} s_N C_{F_i}\left(K_{F_iR}-1\right)\frac{f_N^{(2)}(NH_4)}{f_N(NO_3,NO_2,NH_4)}q_{F_i} + \sum_{i=1}^{3} s_N\left(K_{F_iD}+K_{F_iE}\right)q_{F_i} - K_{42}q_{NH_4},$$

$$\frac{dq_{NO_2}}{dt} = \sum_{i=1}^{3} s_N C_{F_i}(K_{F_iR}-1)\frac{f_N^{(1)}(NO_3,NO_2,NH_4)}{f_N(NO_3,NO_2,NH_4)}\cdot\frac{q_{NO_2}}{q_{NO_2}+q_{NO_3}}q_{F_i} + K_{42}q_{NH_4} - K_{23}q_{NO_2},$$

$$\frac{dq_{NO_3}}{dt} = \sum_{i=1}^{3} s_N C_{F_i}\left(K_{F_iR}-1\right)\frac{f_N^{(1)}\left(NO_3,NO_2,NH_4\right)}{f_N\left(NO_3,NO_2,NH_4\right)}\cdot\frac{q_{NO_3}}{q_{NO_2}+q_{NO_3}}q_{F_i} + K_{23}q_{NO_2},$$

$$\frac{dq_{Si}}{dt} = s_{Si}C_{F_3}\left(K_{F_3R}-1\right)q_{F_3} + s_{Si}K_{F_3D}q_{F_3}.$$

We solve the system of ordinary differential equations by the Runge – Kutta method [15–17]. We will conduct several numerical experiments, assuming that the development of phytoplankton depends on a single limiting substance.

Research Results. For the ODE system (7), we calculate the initial conditions and parameters of the equations at which the stationary regimes occur. Let us take the initial concentration values: $q_{F_1}(0) = 2.5$ *mg/l*, $q_{F_2}(0) = 2.6$ *mg/l*, $q_{F_3}(0) = 0.91$ *mg/l*, $q_{POP}(0) = 0.07$ *mg/l*, $q_{DOP}(0) = 0.07$ *mg/l*, $q_{PO_4}(0) = 0.005$ *mg/l*, $q_{NH_4}(0) = 0.11$ *mg/l*, $q_{NO_2}(0) = 0.0178$ *mg/l*, $q_{NO_3}(0) = 0.304$ *mg/l*, $q_{Si}(0) = 0.4$ *mg/l*; coefficients: $K_{NF_i} = 2.8$ *day*$^{-1}$, $K_{F_iR} = 0.15$ *day*$^{-1}$, $K_{F_iD} = 0.05$ *day*$^{-1}$, $K_{F_iE} = 0.15$ *day*$^{-1}$, $K_{PD} = 0.015$ *day*$^{-1}$, $K_{PN} = 0.02$ *day*$^{-1}$, $K_{DN} = 0.1$ *day*$^{-1}$, $K_{42} = 0.9$ *day*$^{-1}$, $K_{23} = 2.5$ *day*$^{-1}$, $K_{psi} = 1.46$ *day*$^{-1}$, $s_P = 0.01$, $s_N = 0.016$, $s_{Si} = 0.023$, $K_{PO_4} = 0.024$, $K_{NO_3} = 3.0$, $K_{NH_4} = 2.0$, $K_{Si} = 3.0$.

The obtained stationary modes of the ODE system (7) on the assumption that the development of phytoplankton is limited by a single nutrient (phosphorus, nitrogen or silica) are shown in Fig. 2–4, respectively. Fig. 2 describes the effect of phosphorus on the development of various phytoplankton species; Fig. 3 describes the effect of nitrogen on the development of various phytoplankton species; Fig. 4 describes the effect of nitrogen on diatom development.

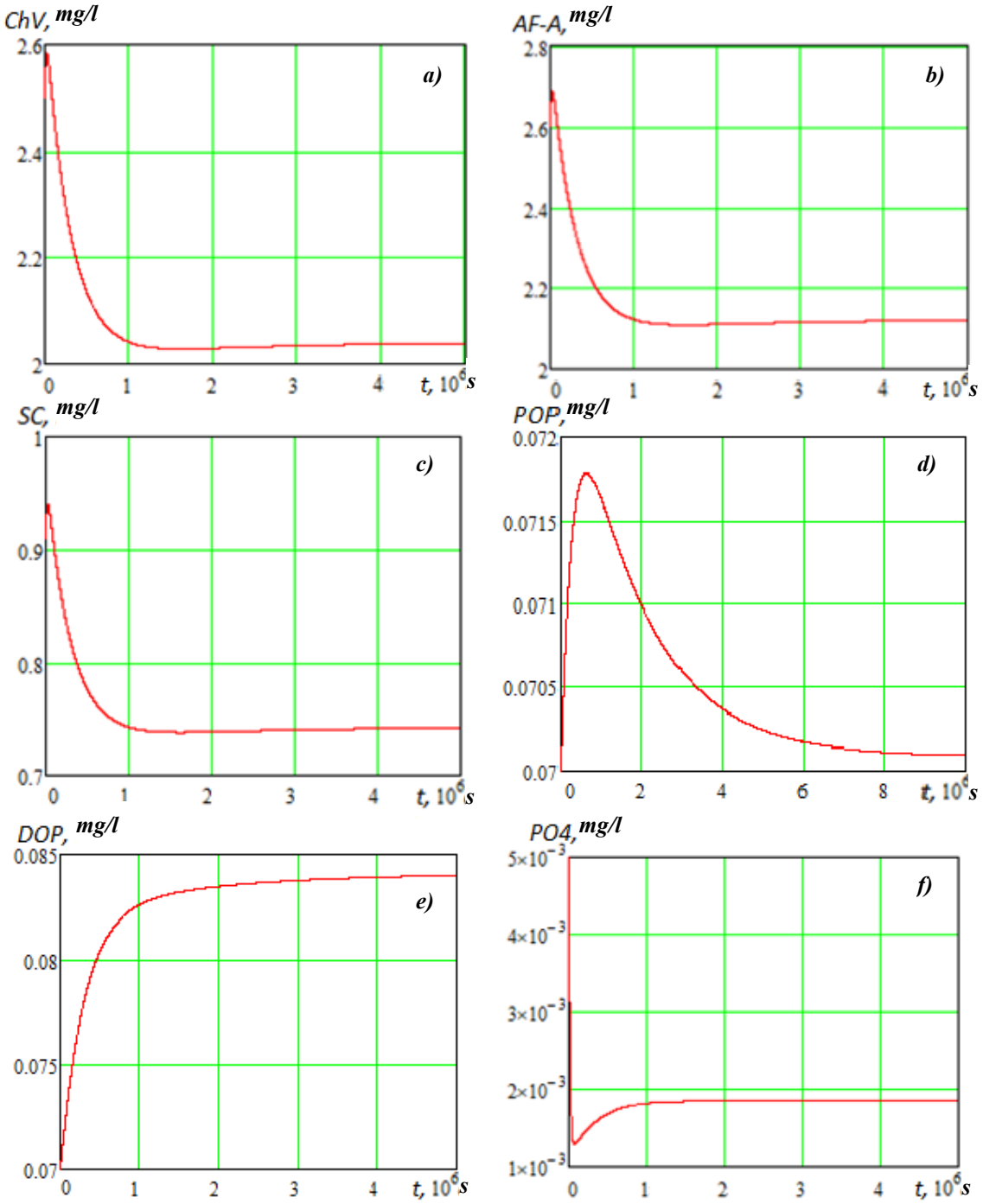


Fig. 2. Stationary mode of ODE system under assumption that phytoplankton development is limited by phosphorus: a) green algae (*ChV*), b) green-blue algae (*AF-A*), c) diatom (*SC*), d) suspended organic phosphorus (*POP*), e) soluble organic phosphorus (*DOP*), f) phosphates (PO_4)

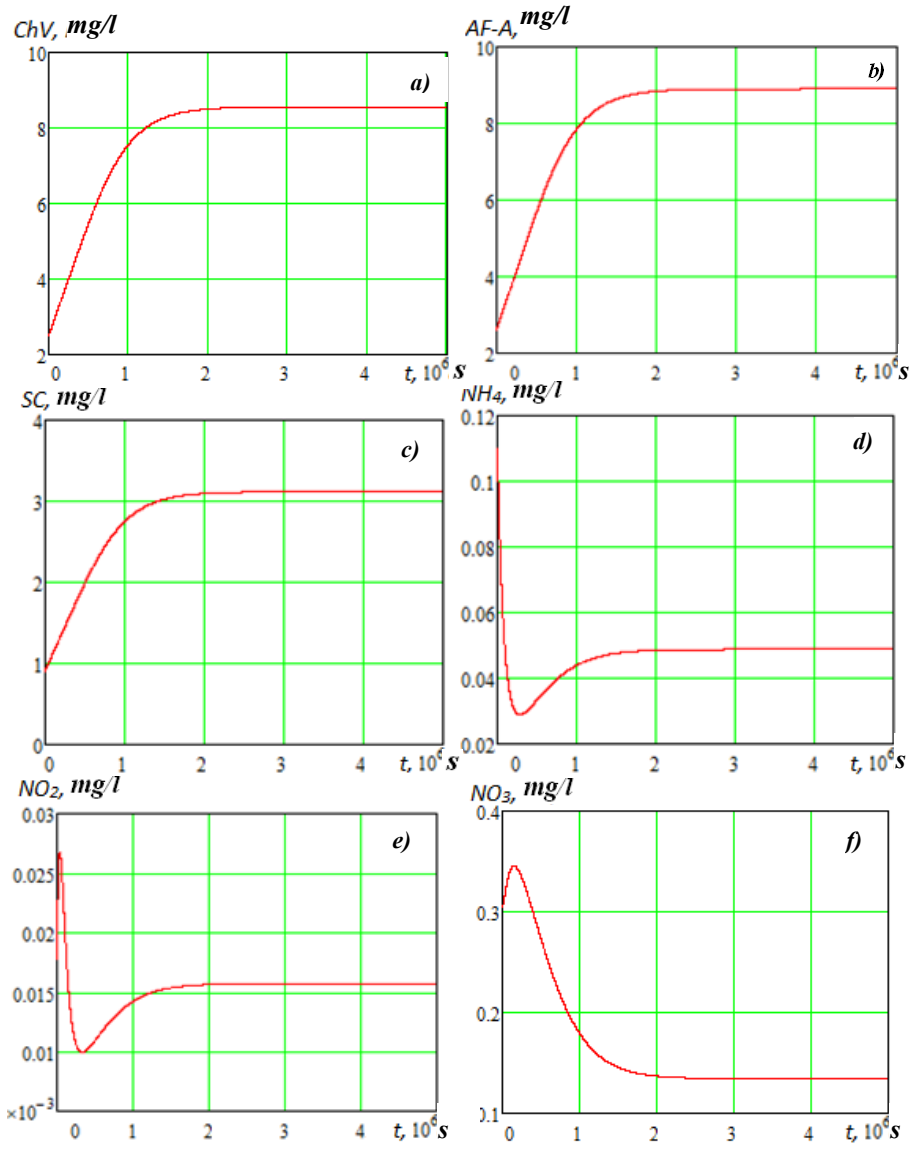


Fig. 3. Stationary mode of ODE system under assumption that phytoplankton development is limited by nitrogen:
a) green algae (*ChV*), b) green-blue algae (*AF-A*), c) diatom (*SC*), d) ammonium (NH_4), e) nitrites (NO_2), f) nitrates (NO_3).

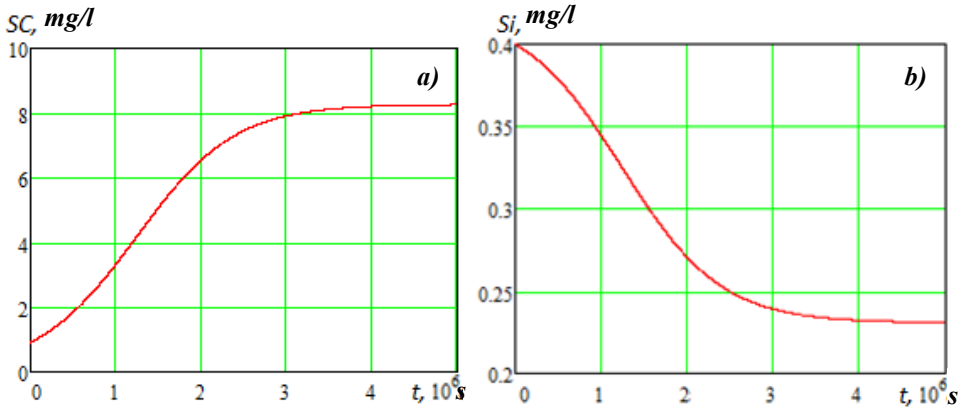


Fig.4. Stationary mode of ODE system under assumption that phytoplankton (diatoms) development is limited by silica:
a) diatom (*SC*), b) silica (*Si*).

The result of the computational experiment shows that with the above values of the initial concentrations and parameters of the equations, stationary modes occur for the ODE system (7), which describes the case of a spatially uniform distribution of substances. The obtained values will be used in further simulation of the spatially inhomogeneous distribution of substances, saltiness and temperature considering the movement of the aquatic environment [18].

Conclusion. A mathematical model of the transformation of forms of phosphorus, nitrogen and silica in the problem of phytoplankton dynamics is studied in the paper. The case of spatially uniform distribution of substances (phytoplankton, forms of phosphorus, nitrogen and silica) is considered. The system is divided into three systems of ordinary differential equations, each of which simulates the dependence of phytoplankton growth on a single nutrient. These systems are solved by the Runge-Kutta method (Fig. 2–4); stationary modes are obtained, for which the values of the system parameters and initial conditions are determined.

The results obtained will be used for the further simulation of the phytoplankton dynamics considering the transformation of phosphorus, nitrogen and silica compounds, taking into account diffusion-convection, saltiness, and temperature.

References

1. Yakushev, E.V., Sukhinov, A.I., et al. Kompleksnye okeanologicheskie issledovaniya Azovskogo morya v 28-m reyse nauchno-issledovatel'skogo sudna «Akvanavt». [Comprehensive oceanological studies of the Azov Sea on the 28th voyage of the “Aquanaut” research vessel.] Oceanology, 2003, vol. 43 no. 1, pp. 44–53 (in Russian).

2. Sukhinov, A.I., et al. Chislennoe modelirovanie ekologicheskogo sostoyaniya Azovskogo morya s primeneniem skhem povyshennogo poryadka tochnosti na mnogoprotsessornoy vychislitel'noy sisteme. [Numerical modeling of ecologic situation of the Azov Sea with using schemes of increased order of accuracy on multiprocessor computer system.] Computer Research and Modeling, 2016, vol. 8, no. 1, pp. 151–168 (in Russian).

3. Sukhinov, A.I., Sukhinov A.A. 3D model of diffusion-advection-aggregation suspensions in water basins and its parallel realization. Parallel Computational Fluid Dynamics 2004: Multidisciplinary Applications — 2005, pp. 223–230. DOI: 10.1016/B978-044452024-1/50029-4.

4. Sukhinov, A.I., Chistyakov, A.E., Shishenya, A.V., Timofeeva, E.F. Mathematical model for calculating coastal wave processes. Mathematical Models and Computer Simulations, 2013, vol. 5, no. 2, pp. 122–129. DOI: 10.1134/S2070048213020087.

5. Sukhinov, A.I., et al. Predskazatel'noe modelirovanie pribrezhnykh gidrofizicheskikh protsessov na mnogoprotsessornoy sisteme s ispol'zovaniem yavnykh skhem. [Predictive modeling of coastal hydrophysical processes in a multiprocession system based on explicit schemes.] Mathematical Models and Computer Simulations, 2018, vol. 30, no. 3, pp. 83–100 (in Russian).

6. Sukhinov, A.I., et al. Matematicheskoe modelirovanie usloviy formirovaniya zamorov v melkovodnykh vodoemakh na mnogoprotsessornoy vychislitel'noy sisteme. [Mathematical modeling of the formation of suffocation conditions in shallow basins using multiprocessor computing systems.] Numerical Methods and Programming, 2013, vol. 14, no. 1, pp. 103–112 (in Russian).

7. Sukhinov, A.I., Sukhinov A.A. Reconstruction of 2001 ecological disaster in the Azov Sea on the basis of precise hydrophysics models. Parallel Computational Fluid Dynamics 2004: Multidisciplinary Applications — 2005, pp. 231–238. DOI: 10.1016/B978-044452024-1/50030-0.

8. Debolskaya, E.I., Yakushev, E.V., Sukhinov, A.I. Formation of fish kills and anaerobic conditions in the Sea of Azov. Water Resources, 2005, vol. 32, no. 2, pp. 151-162. DOI: 10.1007/s11268-005-0020-5.

9. Nikitina, A.V., et al. Optimal'noe upravlenie ustoychivym razvitiem pri biologicheskoy reabilitatsii Azovskogo moray. [Optimal control of sustainable development in biological rehabilitation of the Azov Sea.] Mathematical Models and Computer Simulations, 2016, vol. 28, no. 7, pp. 96–106 (in Russian).

10. Sukhinov, A.I., Nikitina, A.V., Chistyakov, A.E. И. Matematicheskoe modelirovanie protsessov evtrofikatsii v melkovodnykh vodoemakh na mnogoprotsessornoy vychislitel'noy sisteme. [Mathematical modeling of eutrophication processes in shallow waters on multiprocessor computing system.] Bulletin of the South Ural State University: Series “Computer Technologies, Automatic Control & Radioelectronics”, 2016, vol. 5, no. 3, pp. 36 –53 (in Russian).

11. Nikitina, A.V., Puchkin, M.V., Semenov, I.S. Differentsial'no-igrovaya model' predotvrashcheniya zamorov v melkovodnykh vodoemakh. [Differential game of fish kill prevention in shallow water bodies.] Large-scale Systems Control, 2015, iss. 55, pp. 343–361 (in Russian).

12. Sukhinov, A.I., Belova, Yu.V. Matematicheskaya model' transformatsii form fosfora, azota i kremniya v dvizhushcheysya turbulentnoy vodnoy srede v zadachakh dinamiki planktonnykh populyatsiy. [Mathematical transform model of phosphorus, nitrogen and silicon forms in moving turbulent aquatic environment in the problems of plankton populations dynamics.] Engineering Journal of Don, 2015, vol. 37, no. 3, pp. 50 (in Russian).

13. Degtyareva, E.E., Protsenko, E.A., Chistyakov, A.E. Programmnaya realizatsiya trekhmernoy matematicheskoy modeli transporta vzvesi v melkovodnykh akvatoriyakh. [Software implementation of a 3D mathematical model of suspension transport in shallow water areas.] Engineering Journal of Don, 2012, vol. 2, no. 4 – 2, 30 p. Available at: ivdon.ru/ru/magazine/archive/n4p2y2012/1283 (accessed: 12.12.2018) (in Russian).

14. Samarskiy, A.A. Teoriya raznostnykh skhem. [Difference scheme theory.] Moscow: Nauka, 1989, 616 p. (in Russian).

15. Sukhinov, A.I., Sidoryakina, V.V., Sukhinov, A.A. Dostatochnye usloviya skhodimosti polozhitel'nykh resheniy linearizovannoy dvumernoy zadachi transporta nanosov. [Sufficient conditions for convergence of positive solutions to linearized two-dimensional sediment transport problem.] Vestnik of DSTU, 2017, vol. 17, no. 1 (88), pp. 5–17 (in Russian).

16. Samarskiy, A.A., Nikolaev, E.S. Metody resheniya setochnykh uravneniy. [Methods of finite-difference equation solution.] Moscow: Nauka, 1978, 532 p. (in Russian).

17. Marchuk, G.I. Matematicheskoe modelirovanie v probleme okruzhayushchey sredy. [Mathematical modeling in environmental problem.] Moscow: Nauka, 1982, 319 p. (in Russian).

18. Belova, Yu.V., Chistyakov, A.E., Protsenko, U.A. O chetyrekhsloynoy iteratsionnoy skheme. [On four-layer iterative scheme.] Vestnik of DSTU, 2016, vol.16, no. 4 (87), pp. 146–149 (in Russian).

Received 20.11 .2018
Submitted 21.12.2018
Scheduled in the issue 11.01.2019

Authors:

Belova, Yuliya V.,
Junior research scholar, Research Institute for Mathematical Modeling and Forecasting of Complex Systems, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: https://orcid.org/0000-0002-2639-7451
yvbelova@yandex.ru

Atayan, Asya M.,
teaching assistant of the Computer and Automated Systems Software Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: https://orcid.org/0000-0003-4629-1002
atayan24@yandex.ru

Chistyakov, Aleksandr E.,
professor of the Computer and Automated Systems Software Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Phys.-Math.), professor,
ORCID: https://orcid.org/0000-0002-8323-6005
cheese_05@mail.ru

Strazhko, Aleksandr V.,
student of the Computer and Automated Systems Software Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: https://orcid.org/0000-0002-2449-8531
strajcko2@gmail.com

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 13–23. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 13–23. ISSN 1992-5980 eISSN 1992-6006

МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ
MACHINE BUILDING AND MACHINE SCIENCE

UDC 621.382 https://doi.org/10.23947/1992-5980-2019-19-1-13-23

Study on oil pilot circuit of adaptive hydraulic drive of tool advance in mobile drilling machine *

V. S. Sidorenko[1], V. I. Grishchenko[2], S. V. Rakulenko [3], M.S. Poleshkin[4], D.D. Dymochkin [5] **
[1,2,3,4,5] Don State Technical University, Rostov-on-Don, Russian Federation

Исследование гидравлического контура управления адаптивного гидропривода подачи инструмента мобильной буровой машины***

В. С. Сидоренко[1], В. И. Грищенко[2], С. В. Ракуленко[3], М. С. Полешкин[4], Д. Д. Дымочкин[5] **
[1,2,3,4,5]Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация

Introduction. An adaptive hydraulic drive of the tool advance in a mobile drilling machine is studied on the example of the URB-2.5 installation. A typical technological cycle of the mobile drilling machine is considered; the performance criteria are defined. An original design of the adaptive hydraulic drive is proposed on the basis of the analysis. Adaptation of the hydraulic drive of the tool advance is carried out using an adjustable volumetric hydraulic motor with a hydraulic control circuit under discontinuous loads on the tool during the drilling process.

Materials and Methods. Through a preliminary computational experiment in the Matlab Simulink program, the following parameters of the control loop devices were determined: a hydromechanical sensor and a hydraulically controlled valve, on the basis of which the experimental setup was implemented. The performed multifactor experiment allowed identifying the processes in the original hydraulic control circuit of the hydraulic motor under various modes of tool loading.

Research Results. The kinematic and power characteristics of the hydromechanical system of a mobile drilling rig, the hydraulic control effect on the settings of the hydraulic control circuit devices were obtained and determined. The results enabled to specify the rational ranges of the hydromechanical system operation for a typical work cycle.

Discussion and Conclusions. The results obtained can be used to create hydraulic systems of new drilling machines with various characteristics. The application of the developed techniques of research and processing of their results will reduce the time and costs involved in designing adaptive hydraulic systems for mobile technological machines, creating prototypes and conducting commissioning procedures.

Введение. Статья посвящена исследованию адаптивного гидропривода подачи инструмента мобильной буровой машины на примере установки УРБ-2,5. Рассмотрен типовой технологический цикл мобильной буровой машины, определены критерии функционирования. По результатам анализа предложено оригинальное схемотехническое решение адаптивного гидропривода. Адаптации гидропривода подачи инструмента осуществляется при помощи регулируемого объемного гидродвигателя с контуром гидравлического управления при изменяющейся нагрузке на инструменте в процессе бурения.

Материалы и методы. Предварительным вычислительным экспериментом в программе Matlab Simulink определены параметры устройств контура управления: гидромеханического датчика и гидроуправлямого клапана, на основе которых реализована экспериментальная установка. Выполненный многофакторный эксперимент позволил идентифицировать процессы в оригинальном гидравлическом контуре управления гидромотором при различных режимах нагружения инструмента.

*Результаты исследования***.** Получены и определены кинематические и силовые характеристики гидромеханической системы мобильной буровой установки, влияние гидравлического управления на параметры настройки устройств контура. Результаты позволили определить рациональные диапазоны функционирования гидромеханической системы для типового рабочего цикла.

Обсуждение и заключения. Полученные результаты могут быть использованы при создании гидросистем новых буровых машин с различными характеристиками. Использование разработанных методик исследования и обработки их результатов позволит сократить затраты времени и средств при проектировании адаптивных гидросистем мобильных технологических машин, создании

* The research is done within the frame of the independent R&D.
** E-mail: v.sidorenko1942@gmail.com, vig84@yandex.ru, rakulenko84@mail.com, poleshkin.maks@gmail.com., dydedmi_77_06_02@mail.ru
***Работа выполнена в рамках инициативной НИР.

Keywords: adaptive hydraulic drive, mobile drilling unit, drilling technological cycle, hydraulic control circuit, hydromechanical sensor, kinematic and power characteristics.

For citation: V.S. Sidorenko, et al. Study on oil pilot circuit of adaptive hydraulic drive of tool advance in mobile drilling machine. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 13–23. https://doi.org/10.23947/1992-5980-2019-19-1-13-23

опытных образцов и проведении пуско-наладочных работ.

Ключевые слова: адаптивный гидропривод, мобильная буровая установка, технологический цикл бурения, контур гидравлического управления, гидромеханический датчик, кинематические и силовые характеристики.

Образец для цитирования. Сидоренко, В. С. Исследование гидравлического контура управления адаптивного гидропривода подачи инструмента мобильной буровой машины / В. С. Сидоренко [и др.] // Вестник Донского гос. техн. ун-та. — 2019. — Т.19, №1. — С. 13–23. https://doi.org/10.23947/1992-5980-2019-19-1-13-23

Introduction. Dynamic development of natural resources requires the improvement of existing and the creation of new automated complexes of processing equipment for drilling production with improved mechanical and energy characteristics. Analysis of the known circuit design solutions of hydromechanical systems (HMS) of mobile drilling machines (MDM) identified the key feature of their construction — a multi-engine system, in which it is important to consider the effect of the tool feed drive and the main motion drive during the technological cycle [1–4]. At this, the quality, productivity and safety of operation depend much on how the kinematic and force parameters (V, M, etc.) are matched under changing the tool loads [4, 5]. In view of the above, the work objective was to increase the efficiency of the hydromechanical system of MDM working motions through developing and studying its hydraulic control circuit.

Performance Criteria. In the drilling process, an important criterion is the fulfillment of the basic production requirement — operation matching of the main motion and feed drive, which would ensure a stable tool advance revolution [6].

Using the basic laws of similarity of various process cycles, it is easy to apply the above reasoning to the technique of drilling various wells [7, 8].

The performance ($П_О$) is determined as follows:

$$П_О = 1/T_{ОЦ} \quad (1)$$

$$П_{СЦ} = П_О \cdot k_ц \; k_ц = L_{СК} / L_{ОЦ} \quad (2)$$

$$T_{ОЦ} = T_{МАШ} + T_{ВС} + T_{ПЗ} + T_{ОБСЛ} + T_{РЕМ}. \quad (3)$$

$$T_{МАШ} = L_{ОЦ} / V_n; \; V_n = S_o \cdot n_И \quad (4)$$

$$n_И = 1000 \cdot V_{БУР} / 3{,}14 \cdot D_И. \quad (5)$$

$T_{ОЦ}$ is reference cycle time; $T_{ПЗ}$ is time of adjustment of the next reference operation cycle (increasing the tool length); $T_{МАШ}$ is machine time spent on the cutting operation (drilling) when moving the tool by the value of $L_{ОЦ}$ with V_n speed.

When the condition (S_0 = const) is fulfilled, the tool life corresponds to $T_Н = [T_{ИН}]$ standard with $D_И$ diameter and $n_И$ frequency of tool rotation. In this case, the number of tool changes throughout $L_{СК} = L_{ОЦ} \cdot k_ц$ drilling depth is reduced.

$T_{ВС}$ is time spent on installation, commissioning operations, and tool change. It is regulated by the equipment type [7];

$T_{ОБСЛ}$ is scheduled time of change (replacement) of the tool after the development of its technological stability ($T_Н$);

$T_{РЕМ}$ is time for eliminating failures, it is reduced with the increased reliability; $V_{БУР}$, S_o are standards for drilling soils of various categories by σ_i or from the practice of drilling in each region [5].

The tool life and capacity depend directly on the stability of $S_o(t)$ value (working process flow per revolution) [7, 9]. This is achieved in case, when $S_м(t)$ minute feed rate performed by the drive of the MDM dependent feed, decreases synchronously with reducing the tool n(t) rotational speed. When the elastic deformations are not considered in the kinematic tool feed chain, then the feed amount is determined as follows:

$$S_0(t) = \frac{2\pi \cdot v(t)}{\omega(t)} \tag{6}$$

$$S_{\text{M}}(t) = \frac{v(t)}{60} \tag{7}$$

$$\omega(t) = 2\pi \cdot n(t) \tag{8}$$

where $v(t)$ is linear tool advance speed, m/s; $\omega(t) = 2\pi \cdot n(t)$ are angular velocities of tool rotation, rad/s; $n(t)$ is tool rotation frequency, rev/s.

As is well known, the possibilities of rigid stabilization of each of the working motions under the conditions of hydraulic drive application are limited. In the technological machines of this type, the task of adapting a volumetric hydraulic drive is not solved automatically, but is done manually by the operator [3, 10].

Circuit Solution. On the basis of the previously proposed structural scheme [11], as well as the analysis of typical MDM operation cycles, the MDM URB-2.5 schematic hybrid (Fig. 1) is proposed. It considers the behavior and composition of its mechanical subsystem.

The machine power system consists of hydraulic fixed axial-piston pumps (H1, H4), installed on the chassis transfer gear through power take-offs (PTO). The PTO control is electro-pneumatic and it is performed through switches installed in the cab of the auto chassis [10, 12].

The hydraulic control circuit (HCC) of the installation receives hydraulic energy from H2 two-stage pump driven by an electric motor. The electric motor receives power from the auto chassis generator. The first section of H2 directs the hydraulic energy to GMD1 sensor, the second one – to the hydraulically operated valve (HOV).

The operation of both HCC circuits occurs off-load, so overheating of H2 two-stage pump motor is excluded. Each two-stage pump has a safety relief valve with electrical control (KP2, KP3). Pressure is controlled by MN3 and MN5 gauges respectively. HCC has its own closed hydraulic tank (B2) [13]. The main hydraulic tank is B1 tank equipped with F1, F2, F3 drain filters, TO1 and TO2 heat exchangers.

The MDM hydraulic system (HS) is divided into two large circuits. The auxiliary circuit includes an outrigger drive (ГЦ1, ГЦ2), a mast lift-lower drive (ГЦ 3), and a winch drive (ГМ 3). The main contour includes a drive of the tool main motion (rotary table) and a feed drive (Fig. 1).

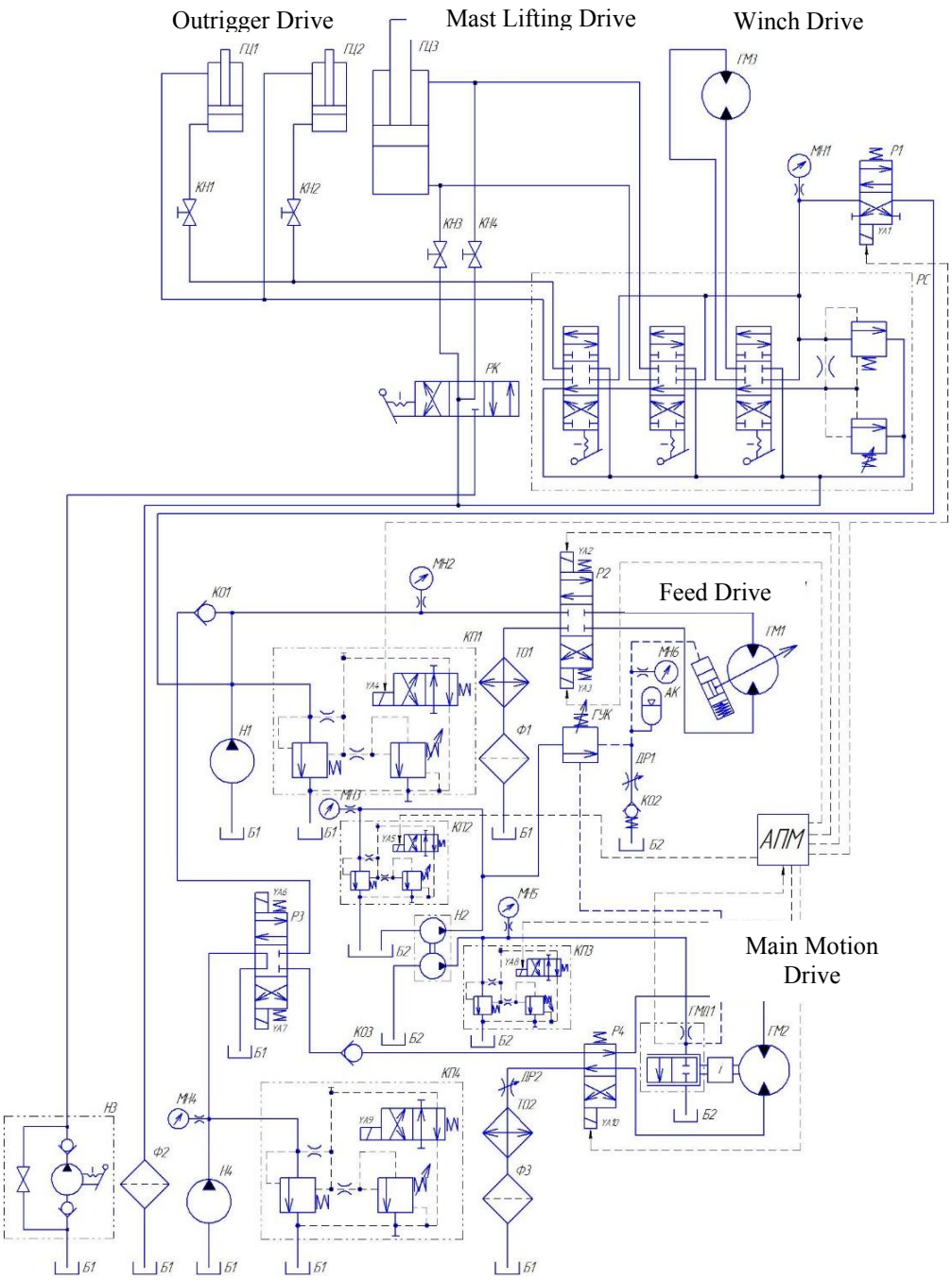


Fig. 1. Hydraulic schematic diagram of MDM URB-2.5

The tool feed drive entering the main circuit receives hydraulic energy from the hydraulic pump (H1). The pump is equipped with an electrically-controlled relief valve (КП1). Operating pressure is monitored by the pressure gauge (MH2). Reverse movement, stop are provided by a three-position electrically-controlled valve (P2) with a closed center. The hydromotor (ГМ1) of the feed drive, adjustable through mechanical transmission, provides a longitudinal feed of the drilling tool. The heat exchanger (ТО1) and filter (Ф1) provide filtration and conditioning of the working fluid.

The pump (H4) is rotated in the same way as the pump (H1) (Fig. 1). The valve (P3) initiates start, stop of the main drive or feed rate increase under the rapid traverse of the feed drive. In this case, the upper position of the valve is turned on, and the flow of the pump (H4) enters the drive supply circuit through the check valve (КО1), and is summed with the flow of the pump (H1) providing a quick lowering or lifting of the tool from the well. When the pump (H4) operates on the main drive (P3 in the lower position), the check valve (КО1) prevents the working fluid from being drained out of the supply circuit into the tank (B1) [13].

A similar function is performed by the check valve (КО3) when the pump is operated on the feed drive. The valve (Р4) provides reverse of the drive main motion. Using the choke (ДР2), the tool rotation speed is adjusted (at idle).

The heat exchanger (ТО2) and the drain filter (Ф3) provide filtration and conditioning of the working fluid of the rotational drive circuit.

The HCC circuit includes two original devices – a hydraulically controlled valve (HOV) (4) and a hydraulic multiparameter sensor (HMS) (5) (Fig. 2) [14]. Under the rapid traverse of the feed drive described above, the pump (Н2) (Fig. 1) is automatically shut off by the control system (АПМ).

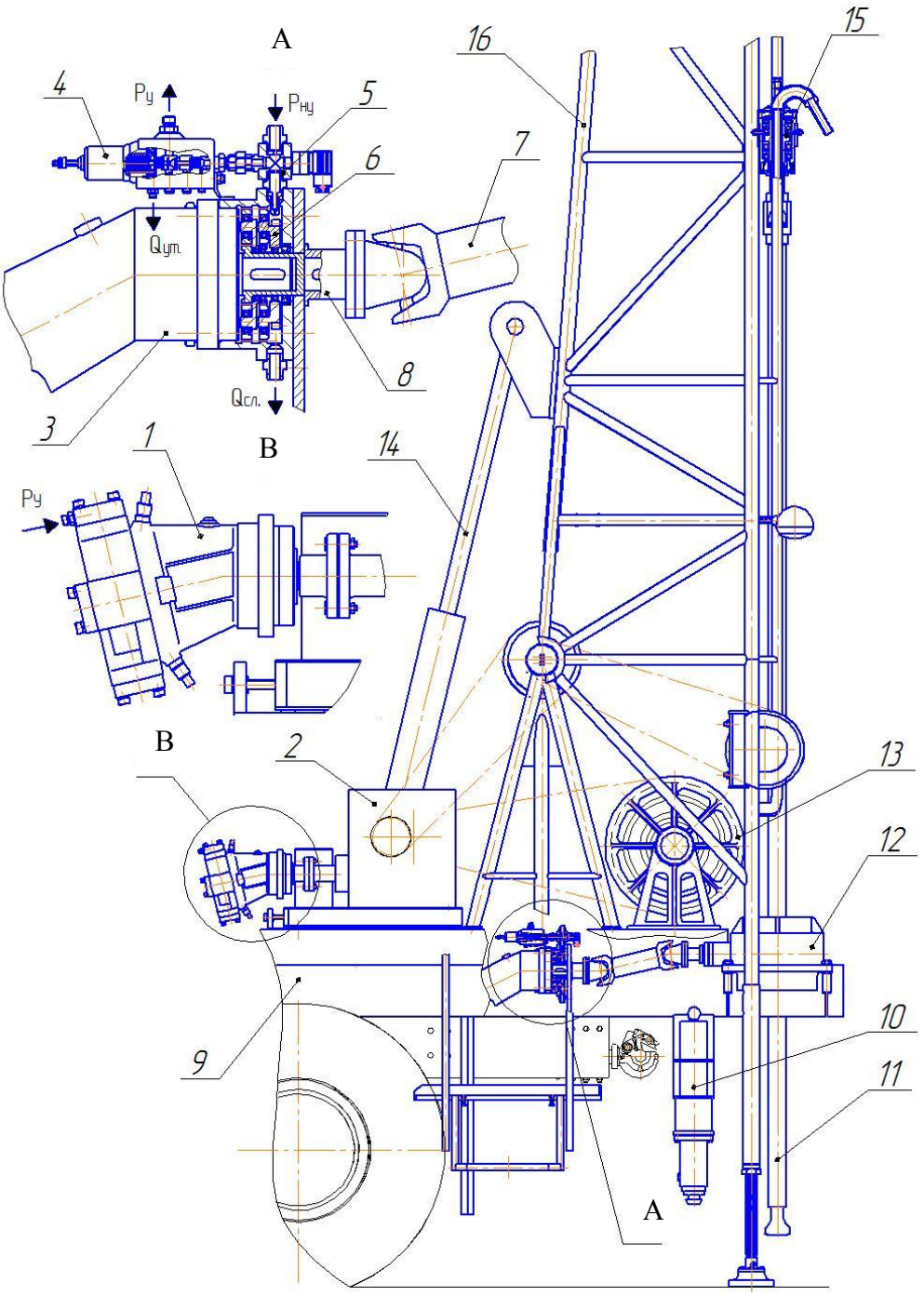


Fig. 2. MDM general view: 1 is adjustable hydraulic motor; 2 is tool feed drive; 3 is uncontrolled hydraulic motor; 4 is HOV valve; 5 is HMS sensor; 6 is gearbox; 7 is driveshaft; 8 is coupling; 9 is auto chassis; 10 is system of outriggers; 11 is tool; 12 is main motion drive; 13 is double drum winch; 14 is hydraulic cylinder of mast lifting (lowering); 15 is swivel; 16 is mast.

The HCC principle of operation (Fig. 2) as part of the MDM hydraulic system is as follows: the shaft (HMS) (5) is mechanically connected to the shaft (ГМ2) through the gearbox (6). Pressure fluctuation at the HMS input (5) is

transmitted to the control input of the HOV (4) which is adjusted – by selecting springs – to the mode of operation whereby, at its output, the average operating pressure proportional to the oscillation amplitude and, accordingly, to the rotational speed of the hydromotor shaft (ГМ1) is formed.

With resistance moment increment on the hydromotor shaft (ГМ2), the shaft rotation speed and the oscillation frequency at the HMD input decrease, and the pressure amplitude increases. As it increases, the average value of opening the HOV valve, which supplies a greater amount of fluid from the H2 to the gearbox (ГМ1) of the feed drive, increases. The control pressure (Py) at the HOV input increases, which provides an increase in the hydraulic motor capacity and reduces the feed speed. The mechanism of the HCC contour as a component of the HMS with a dependent tool advance, as well as the operation principle of the (АПМ) unit, is considered in [15].

The accumulator (АК) smoothes the pressure control pulsation (Py) at the hydraulic motor (ГМ1) input control. Using the setting of the chokes (ДР1, ДР2) and the selection of the pressure valve (КО3), the limiting values of the tool feed rate and rotation speed are adjusted, and the required functional relationship between the working motion speeds is formed.

The original HOV differs in function of continuous regulation of the valve flow section. Therefore, it is possible to control the performance of the hydromotor at all speeds.

Experimental Study. To identify the HMS parameters operating under the specified conditions, special bench equipment and accessories are developed [13]. The hydromotor flow rate was determined using a turbine sensor-flow meter connected through the converter board (ЦАП-АЦП). The adjustment ranges of the parameters under the study are shown in Table 1.

Table 1

Parameter ranges under HMS study

No.	Parameter	Designation	Range	Unit of measurement	Control device
1	Nozzle diameter	$d_{сп}$	0.5–1.2	mm	Plug gauge
2	Choke flange	$d_{др}$	0.8–2	mm	Plug gauge
3	Gap between nozzle and modulator	$y_з$	0.2–1	mm	Dial gauge
4	RPM	$n_{гмд}$	5–60	rpm	Rotation velocity sensor
5	Flow through HMS	$Q_{гмд}$	0.5–20	l/min	Flow gage, gage tank
6	Pressure in HMS	$p_{гмд}$	0.5–5	MPa	Pressure gauge
7	Characteristic design factor of HMS flow section	$x_д$	2–8	mm	Slide gage

The traverse speed and acceleration of the shaft (ГМ) (feed drive) was determined by sequential differentiation of motion over time using the following formulas implemented by post-processing of data in the *PowerGraph* program [8, 11, 16]:

$$\omega(t) = \frac{d\varphi}{dt}, \tag{9}$$

$$\varepsilon(t) = \frac{d\omega}{dt}, \tag{10}$$

where ω is angular velocity, rad/s; ε is angular acceleration of (ГМ), rad/s^2.

The results of the obtained experimental data processing through the known methods [17, 18] are presented in Fig. 3–9.

The graph in Fig. 3 explains the pressure fluctuation amplitude response of the HMS under a change in the hydromotor (ГМ1) speed variation within the range of 45–125 rad/s. The data were obtained when testing nozzles with the diameters: $d_{сп}$ = 2, 4, 6 mm.

As a result, the dependence of the pressure change was almost linear in nature, and it fell as the speed increased, which was associated with the non-stationary mode of the working fluid outflow through the flapper-nozzle unit.

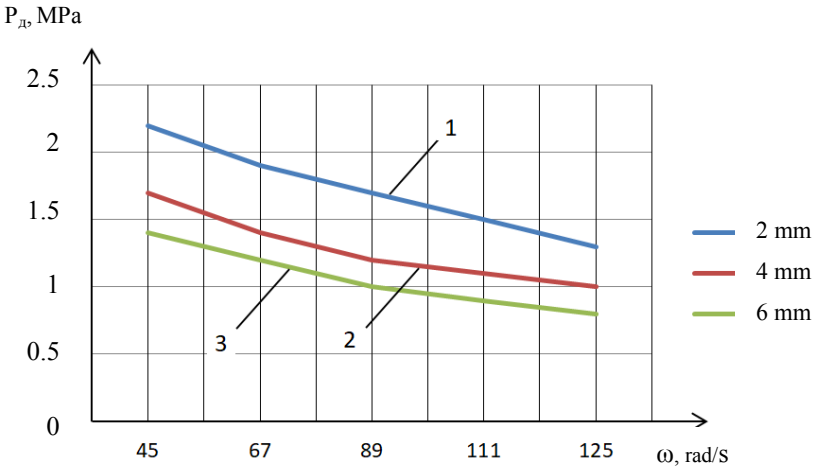


Fig.3. Dependence of pressure fluctuation value ($P_Д$) on ω speed of rotation of hydromotor (ГМ1), approximation: 1 is $P_Д=0.14\omega^2-3\omega+24.8$; 2 is $P_Д=0.36\omega^2-3.84\omega+20.4$; 3 is $P_Д=0.21\omega^2-2.8\omega+16.6$.

The results presented in Fig. 4 show that with an increase in the characteristic structural parameter of the flow section (ГМД $X_Д$) from 2 to 8 mm, the control pressure level changes from 1.9 to 1.4 MPa. This enables to make recommendations for further optimization of the HMS flow path [19], in particular, for some increase in (ХД). However, its further increase is impractical because it increases the geometric dimensions of the modulator disk.

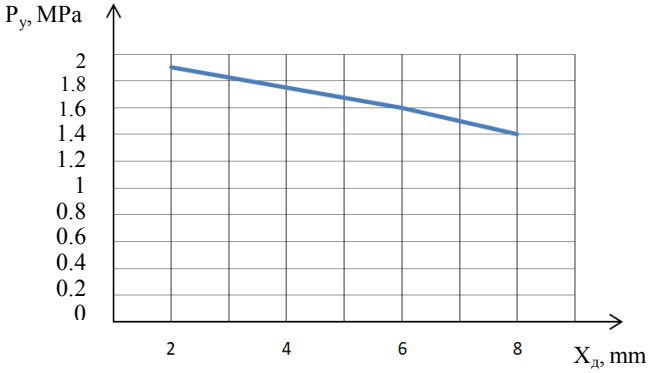


Fig. 4. Dependence control pressure ($P_у$) variation on effect of characteristic structural parameter of cross section of flow part (ГМД $X_Д$), approximation: $P_у=1.65X_Д+20.75$

Fig. 5 shows the effect of the design features of ($X_Д$) on the maximum pressure amplitude ($P_{дmax}$) used for the hydraulic control circuit during advance to the HOV.

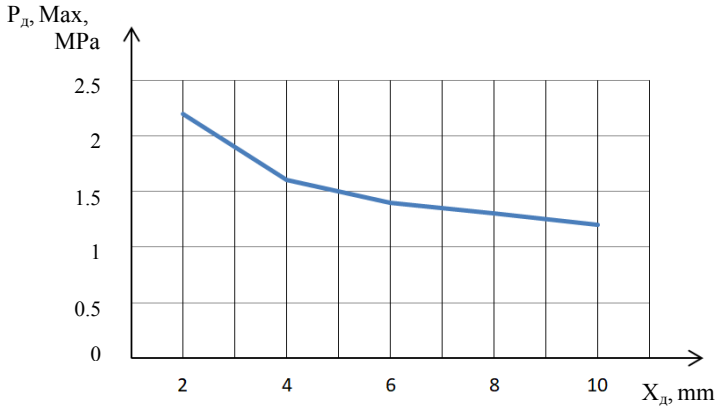


Fig.5. Dependence of change in maximum amplitude of control pressure ($P_{Д.MAX}$) on effect of characteristic design parameter of flow section of (ГМД $X_Д$), approximation: $P_{ДMAX}=21.46X_Д^{-0,372}$

During identification, an important step was the determination in the HCC circuit: HMS – HOV, the degree of impact of each control component setting. Thus, the HOV tuning property, under the control loop performance, is its spring constant ($C_{пр}$), the selection of which significantly affects the sensitivity of the control subsystem [20].

The characteristic built in Fig. 6 and approximated by the results of the experiment, is linear, and explains the magnitude of the maximum control pressure drop from 1.7 to 1.4 MPa under changing the HOV spring force.

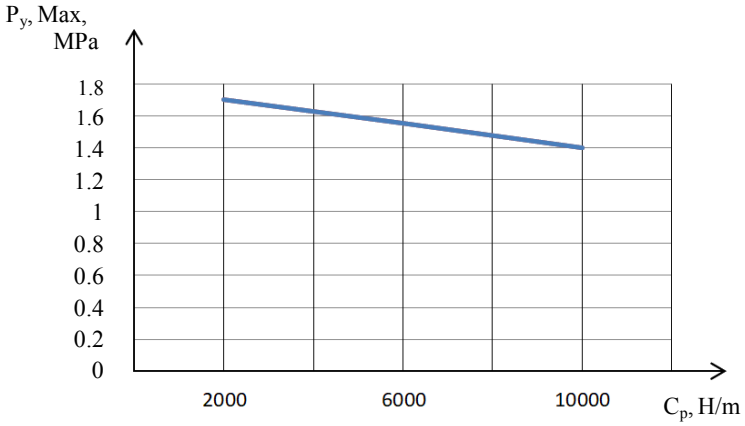


Fig. 6. Dependence of maximum control pressure ($P_{y.max}$) variation on magnitude of HOV spring force, approximation: $P_{УМАХ}=1.5C_Р+18.5$

The second tuning element in the circuit with the HOV is an adjustable choke in the shunt line (Fig. 1). Its setting parameter is the flow area regulated within the range from $2.5 \cdot 10^{-7}$ to $7.5 \cdot 10^{-7}$ m^2. The results of the experiment are shown in Fig. 7.

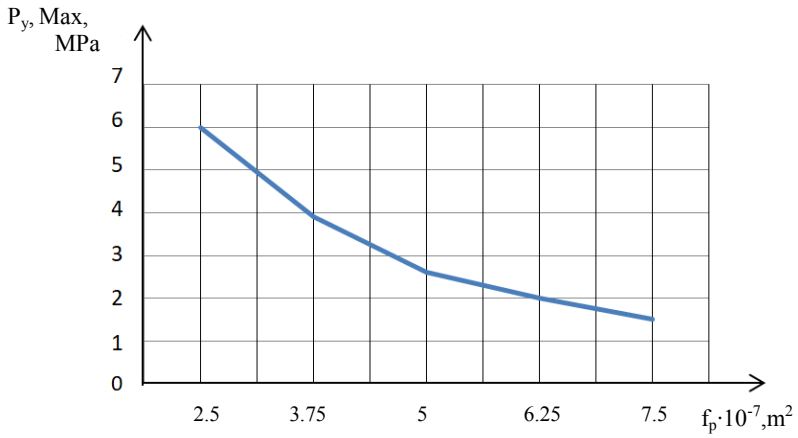


Fig. 7. Dependence of maximum control pressure ($P_{y.max}$) variation on flow area of choke (ДР1), approximation: $P_{УМАХ}=1.5f_{ДР}^2-27.6\cdot f_{ДР}+84.2$

Four operation modes of the tool-drive motor (hydraulic motor (ГМ 1)) under relay loading, and, accordingly, under changing ω angular velocity by: 22 rad/s; 49 rad/s; 71 rad/s and 85 rad/s, are studied. The dependence obtained from the experiment is shown in Fig. 8.

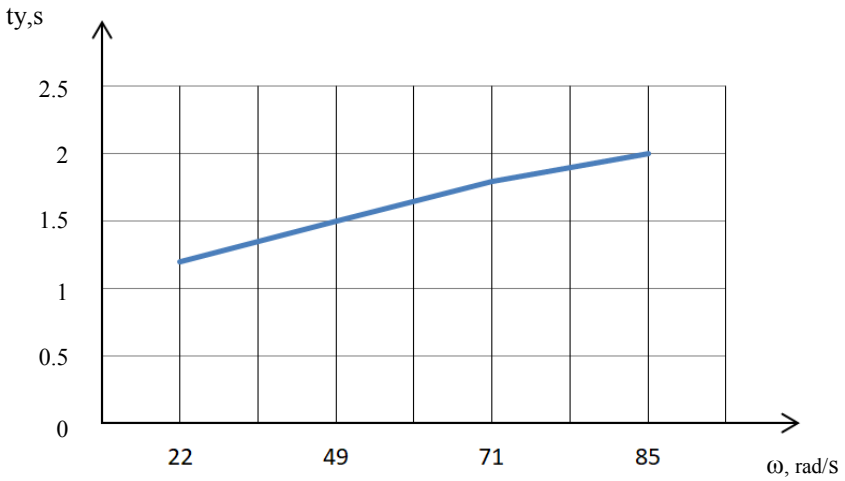


Fig. 8. t_y dependence of HCC response time on ω angular rotation velocity of hydromotor (ГМ1) shaft, approximation: $t_y=-0.002\omega^2-0.4\omega+0.83$

Since HCC is a part of the MDM hydraulic system, its length has a special impact on the quality and time of transient processes [16]. In order to determine the degree of impact on the above parameters, the pressure variation time in the control line with varying of its volume was studied (Fig. 9).

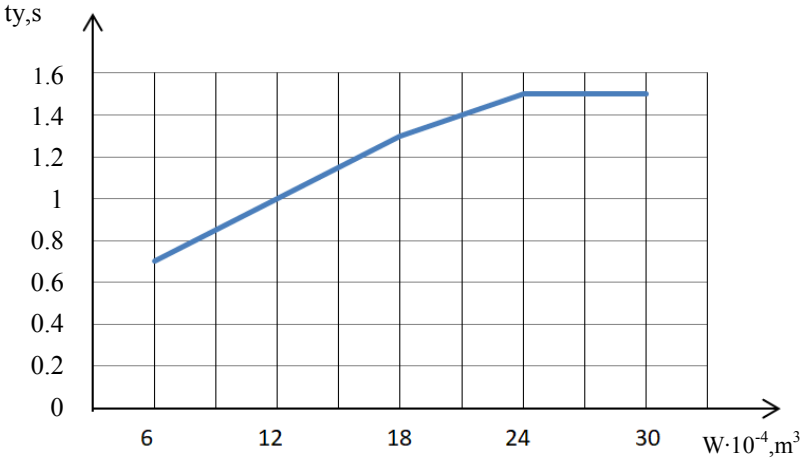


Fig. 9. t_y dependence of HCC response time on W volume in control flow line, approximation: $t_y=-0.02\omega^3+0.1\omega^2+0.12\omega+0.5$

The results enable to determine the critical value of the flow line volume: $W = 24\cdot10^{-4}$ м^3, over which the HCC response time does not exceed 1.5 s; that is due to wave processes taking place in the pipeline, as well as to the parameters of the lines themselves [20].

Conclusions. As a result of the research, the requirements were developed, and the HMS generalized structure of the MDM working motions was proposed, which made it possible to increase the efficiency of drilling production through introducing an internal kinematic hydromechanical connection between the main motion and the tool feed movement.

A hydraulic control circuit is developed and implemented on the basis of a multi-parameter sensor and a hydraulically controlled valve, which ensures the coordination of working motions under the discontinuous process duty. Considers the characteristic features of the drilling process, the configuration of the reference operating cycle, which allowed for the development of an automated system for driving work motions, is proposed.

References

1. Ozerskiy, A.I., Babenkov, Yu.I., Shoshiashvili, M.E. Perspektivnye napravleniya razvitiya silovogo gidravlicheskogo privoda. [Perspective ways of development of power hydraulic gear.] University News. North-Caucasian region. Technical Sciences Series, 2008, no. 6, pp. 55–61 (in Russian).

2. Sysoev, N.I., Mirny, S.G., Grinko, D.A. Obosnovanie struktury i ratsional'nykh rezhimnykh parametrov mekhatronnoy buril'noy mashiny. [Substantiation of the structure and rational drilling regimes of mechatronic drilling machine.] Mining Equipment and Electromechanics, 2011, no. 9, pp. 24–28 (in Russian).

3. Muravenko, V.A., Muravenko, A.D. Burovye mashiny i mekhanizmy. T.1. [Drilling machines and mechanisms. Vol. 1.] Moscow-Izhevsk: ICS Publ. House, 2002, 520 p. (in Russian).

4. Mirny, S.G. Obosnovanie i vybor ratsional'noy chastoty vrashcheniya shtangi mashin dlya sverleniya shpurov v porodakh povyshennoy kreposti i abrazivnosti: dis… kand. tekhn. nauk. [Validation and selection of rational frequency of rotation of the rod of machines for drilling holes in rocks of increased strength and abrasivity: Cand.Sci. (Eng.) diss.] Novocherkassk: NPI, 2005, 142 p. (in Russian).

5. Basarygin, Yu.M. Tekhnologiya burenie neftyanykh i gazovykh skvazhin. Ucheb. dlya vuzov. [Technology of drilling oil and gas wells.] Moscow: Nedra-Businesscentre, 2001, 679 p. (in Russian).

6. Kudinov, V.A. Dinamika stankov. [Machine dynamics.] Moscow: Mashinostroenie, 1967, 359 p. (in Russian).

7. Yashcheritsyn, P.I., Feldstein, E.E., Kornievich, M.A. Teoriya rezaniya. [Cutting theory.] Moscow: Novoe znanie, 2006, 512 p. (in Russian).

8. Le Chung Kien. Povyshenie effektivnosti gidroprivoda podachi tekhnologicheskogo oborudovaniya: dis…kand. tekhn. nauk. [Improving the efficiency of hydraulic supply of technological equipment: Cand.Sci. (Eng.) diss.] Rostov-on-Don, 2013, 165 p. (in Russian).

9. Grinko, D.A. Metod rascheta i podderzhaniya ratsional'nykh rezhimnykh parametrov buril'noy mashiny mekhatronnogo klassa: dis… kand. tekhn. nauk. [Method of calculating and maintaining rational operational parameters of a mechatronic class boring machine: Cand.Sci. (Eng.) diss.] Novocherkassk, 2015, 158 p. (in Russian).

10. Ustanovka razvedyvatel'nogo bureniya URB-2,5S-KAMAZ. Rukovodstvo po ekspluatatsii URB-2,5S.00.00.000 RE. [Installation of reconnaissance drilling – URB-2.5S-KAMAZ. Operation manual for URB-2.5S.00.00.000 RE.] Rostov-on-Don, StroyNefteMash LLC, 2012 (in Russian).

11. Sidorenko, V.S., et al. Adaptivnyy gidroprivod s ob"emnym regulirovaniem podachi instrumenta tekhnologicheskoy mashiny. [Adaptive hydraulic drive with delivery tool-feed control of production machine.] Vestnik of DSTU, 2017, vol. 17, no. 2(89), pp. 88–99 (in Russian).

12. Rakulenko, S.V. Adaptivnyy gidravlicheskiy privod podachi instrumenta burovoy ustanovki. [Adaptive hydraulic drive of tool feed for drilling rig.] Yubileynaya konf. stud. i molodykh uchenykh, posv. 85-letiyu DGTU: sb. dokl. nauch.-tekhn. konf. [Jubilee conference of students and young scientists dedicated to the 85th anniversary of DSTU: Proc. Sci.-Tech. Conf.] Rostov-on-Don, 2015, pp. 306–318 (in Russian).

13. Sidorenko, V.S., et al. Modelirovanie gidravlicheskoy sistemy s zavisimoy podachey instrumenta mobil'noy burovoy ustanovki. [Hydraulic system simulation with dependent tool feed of mobile drilling rig.] Gidravlicheskie mashiny, gidroprivody i gidropnevmoavtomatika. Sovremennoe sostoyanie i perspektivy razvitiya — 2016: sb. nauch. tr. IX mezhdunar. nauch.-tekhn. konf. [Hydraulic machines, hydraulic drives and hydropneumatic automation. Current state and development potential - 2016: Proc. IX Int. Sci.-Tech. Conf.] St.Petersburg: Polytechnic University Publ. House, 2016, pp. 365–375 (in Russian).

14. Sidorenko, V.S., Rakulenko, S.V., Le Chung Kien. Gidravlicheskiy datchik: patent 2538071 Ros. Federatsiya, MPK G01P3/32. [Hydraulic sensor.] RF patent no. 2538071, 2015 (in Russian).

15. Sidorenko, V.S., Rakulenko, S.V., Poleshkin, M.S. Dinamika gidromekhanicheskoy sistemy tekhnologicheskoy mashiny s adaptivnym privodom podachi instrumenta. [Dynamics of the hydromechanical system of a production machine with an adaptive tool-feeding drive.] VESTNIK of Samara University. Aerospace and Mechanical Engineering, 2017, vol. 16, no. 1, pp. 162–175 (in Russian).

16. Popov, D.N. Dinamika i regulirovanie gidro- i pnevmosistem: ucheb. dlya vuzov. 2-e izd., pererab. i dop. [Dynamics and regulation of hydraulic and pneumatic systems.] Moscow: Mashinostroenie, 1987, 464 p. (in Russian).

17. Mosteller, F., Tukey, J.W. Analiz dannykh i regressiya. [Data analysis and regression.]; Yu.N. Blagoveshchenskiy, trans. from English; Yu.P. Adler, ed. Moscow: Finansy i statistika, 1982, 258 p. (in Russian).

18. Stepnov, M.N. Statisticheskie metody obrabotki rezul'tatov mekhanicheskikh ispytaniy. [Statistical methods for processing mechanical test results.] Moscow: Mashinostroenie, 1985, 220 p. (in Russian).

19. Ivanov, G.M., et al. Proektirovanie gidravlicheskikh sistem mashin. [Design of hydraulic systems of machines.] Moscow: Mashinostroenie, 1992, 224 p. (in Russian).

20. Kazmirenko, V.F. Elektrogidravlicheskie mekhatronnye moduli dvizheniya. Osnovy teorii i sistemnoe proektirovanie. Uchebn. Posobie. [Electro-hydraulic mechatronic motion modules. Fundamentals and system design]. Moscow: Radio i svyaz', 2001, 432 p. (in Russian).

Received 21.12 .2018
Submitted 23.12.2018
Scheduled in the issue 11.01.2019

Authors:

Sidorenko, Valentin S.,
professor of the Hydraulics, Hydraulic and Pneumatic Control Systems and Heat Processes Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Eng.), professor,
ORCID: https://orcid.org/0000-0001-5124-6324
v.sidorenko1942@gmail.com

Grishchenko, Vyacheslav I.,
associate professor of the Hydraulics, Hydraulic and Pneumatic Control Systems and Heat Processes Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Eng.),
ORCID:https://orcid.org/0000-0003-1422-2811
vig84@yandex.ru

Rakulenko, Stanislav V.,
associate professor of the Hydraulics, Hydraulic and Pneumatic Control Systems and Heat Processes Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: https://orcid.org/0000-0001-8293-0305
rakulenko84@mail.com

Poleshkin, Maxim S.,
associate professor of the Hydraulics, Hydraulic and Pneumatic Control Systems and Heat Processes Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Eng.),
ORCID: https://orcid.org/0000-0002-5364-1106
poleshkin.maks@gmail.com

Dymochkin, Denis D.,
associate professor of the Hydraulics, Hydraulic and Pneumatic Control Systems and Heat Processes Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Eng.),
ORCID: https://orcid.org/0000-0003-0717-4348
dydedmi_77_06_02@mail.ru

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 24–30. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 24–30. ISSN 1992-5980 eISSN 1992-6006

МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ
MACHINE BUILDING AND MACHINE SCIENCE

UDC 621.893 https://doi.org/10.23947/1992-5980-2019-19-1-24-30

Effect of organic acid concentration in lubricant on tribological characteristics of friction couple*

V. E. Burlakova[1], E. G. Drogan[2]**
[1,2] Don State Technical University, Rostov-on-Don, Russian Federation

Влияние концентрации органической кислоты в составе смазки на трибологические характеристики пары трения***

В. Э. Бурлакова[1], Е. Г. Дроган[2]**
[1,2] Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация

Introduction. The possibility of using monocarboxylic acids as a lubricant composition additive, and the effect of their concentration in lubricant on the evolution of the friction factor of a brass-steel couple, as well as the morphology of the film surface under friction is considered. The work objective is to study the effect of the concentration of carboxylic acids in the lubricant composition on the evolution of the friction factor of copper – steel alloy.

Materials and Methods. Tribological studies of a brass-steel friction couple in aqueous solutions of monocarboxylic acids with the concentrations of 0.025; 0.05; 0.1; 0.2; 0.5 mol/l are carried out. Using scanning electron microscopy, we have studied the morphology of the servovite film surface that is formed on a steel disk after frictional interaction of a brass-steel couple in aqueous solutions of acids with the concentration of 0.1 mol/l.

Research Results. Tribological characteristics of the brass-steel tribocoupling in aqueous solutions of carboxylic acids of various concentrations are studied. The optimum acid concentration in the lubricant composition is specified. Herewith, a selective transfer and a wearless friction regime are implemented under friction of the brass 59–steel 40X couple. A decrease in the friction ratio to 0.009 and 0.007 is found out under friction in aqueous solutions of valeric and caproic acids, respectively. The formation of an anti-friction film on the steel surface is identified through the scanning electron microscopy. It is established that the film formed in an aqueous solution of caproic acid has a denser structure in

Введение. Статья посвящена исследованию возможности использования в качестве присадки к смазочной композиции одноосновных карбоновых кислот и изучению их влияния на эволюцию коэффициента трения пары латунь-сталь, а также изучению морфологии формирующихся при трении поверхности пленок.

Целью работы являлось изучение влияния концентрации карбоновых кислот в составе смазочной композиции на эволюцию коэффициента трения пары сплав медь-сталь.

Материалы и методы. Проведены трибологические исследования пары трения латунь-сталь в водных растворах одноосновных карбоновых кислот с концентрациями 0,025; 0,05; 0,1; 0,2; 0,5 моль/л. С помощью растровой электронной микроскопии изучена морфология поверхности сервовитной пленки, формирующейся на стальном диске после фрикционного взаимодействия пары трения латунь-сталь в водных растворах кислот с концентрацией 0,1 моль/л.

Результаты исследования. Изучены трибологические характеристики трибосопряжения латунь-сталь в водных растворах карбоновых кислот различной концентрации. Установлена оптимальная концентрация кислоты в составе смазки, при которой в результате трения пары латунь 59-сталь 40Х реализуется избирательный перенос и достигается режим безызносного трения. Обнаружено снижение коэффициента трения до 0,009 и 0,007 при трении в водных растворах валериановой и капроновой кислот соответственно. С помощью растровой электронной микроскопии выявлено формирование на стальной поверхности антифрикционной пленки. Установлено, что пленка, формирующаяся в водном растворе капроновой кислоты, имеет более плотную структуру, в сравнении с пленкой, формирующейся при трении в водных растворах масляной и капроновой кислот.

* The research is done within the frame of the independent R&D.
** E-mail: vburlakova@donstu.ru, ekaterina.drogan@gmail.com
*** Работа выполнена в рамках инициативной НИР.

comparison with the film formed under friction in aqueous solutions of butyric and caproic acids.

Discussion and Conclusions. Thus, the tribological studies of a brass-steel friction couple in aqueous acid solutions show that the optimum molar acid concentration in the lubricant composition is 0.1 mol/l. At this acid concentration, the values of the friction factor characteristic of the wearless mode are attained.

Обсуждение и заключения. В результате трибологических исследований пары трения латунь-сталь в водных растворах кислот выявлено, что оптимальной молярной концентрацией кислоты в составе смазки является концентрация 0,1 моль/л. При этой концентрации кислоты достигаются значения коэффициента трения, характерные для режима безызносности. Выявлено, что изменение концентрации кислоты приводит либо к увеличению значений коэффициента трению, либо к незначительному его снижению. При этом увеличение концентрации кислоты сопровождается коррозионными процессами на поверхности пары трения.

Keywords: friction factor, wear, selective transfer, servovite film, carboxylic acid, friction surface topography.

Ключевые слова: коэффициент трения, износ, избирательный перенос, сервовитная пленка, карбоновая кислота, топография поверхности трения.

For citation: V.E. Burlakova, E.G. Drogan. Effect of organic acid concentration in lubricant on tribological characteristics of friction couple. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 24-30. https://doi.org/10.23947/1992-5980-2019-19-1-24-30

Образец для цитирования. Бурлакова, В. Э. Влияние концентрации органической кислоты в составе смазки на трибологические характеристики пары трения / В. Э. Бурлакова, Е. Г. Дроган // Вестник Донского гос. техн. ун-та. — 2019. — Т.19, №1. — С. 24-30. https://doi.org/10.23947/1992-5980-2019-19-1-24-30

Introduction. Modern high-developing machine-building industry advances new demands on lubricants. Most conventional methods of the friction and wear control are based on the use of solid and liquid lubricants [1–4]. The principal function of lubrication in friction units is to keep two contacting surfaces of machine parts from wear. To perform the necessary functions, basic fluids need modification by functional additives that change the antiwear properties of base oils through improving their lubricity among other. For this reason, additives are an integral part of the design of modern lubricants. The studies [2–7] show that to reduce friction and wear, both nanoparticles of metals and various organic components are used as additives to lubricant compositions. Some of them contribute to the formation of protective antifriction films on the tribocontact surfaces due to the presence of metal powders with particle sizes in the micro-range and nanoscale in the lubricant composition [7–9]; others – as a result of selective dissolution of tribo-conjugated surfaces (in the case of copper-steel alloy friction pair) under friction [10–17]. As is known, the lubricating medium plays a critical role in the selective transfer of copper to the steel surface under frictional interaction. For example, during the copper alloy – steel friction, various polar compounds, including carboxylic acids, are formed in the aqueous-glycerin medium. It was interesting to study the possibility of using them as additives to the lubricant composition for implementing a selective transfer and wearless friction.

Materials and Methods. The evolution of the friction factor of the "brass 59-aqueous solution of carboxylic acid – 40X steel" system was investigated on the AE-5 type face friction machine. Aqueous solutions of saturated monobasic carboxylic acids with the general formula: $R–COOH$ ($R = C_nH_{2n+1}$), with the concentrations of 0.025–0.5 mol/l were used as a lubricant composition. Before the tribological studies, samples of steel 40X and brass 59 were cleaned with abrasive paper, washed with distilled water, degreased with hexane, and air-dried. The friction assembly consisted of a steel rigidly fixed disk and three movable brass fingers arranged in a circle at the angle of 120° relative to each other. A friction couple of a ring steel specimen and brass fingers was placed in the lubricant composition into the working part of the friction machine made in the form of a textolite bath. A PHYWE Cobra force sensor was attached to the front face of the working part of the friction machine for continuous recording of the friction variation. Tribological studies were carried out under the following modes: sliding speed of the specimen was 0.45 m/s; axial load was 1.7 MPa; test time was 10 hours; temperature of the working environment was 37° C; slide path was 15,260 m.

The morphology of the servo film was studied using the scanning electron microscopy (SEM) on a Carl ZEISS microscope in the Electron and Optical Microscopy Laboratory and the Center for Collective Use of Scientific Equipment, REC "Materials" (http: //nano.donstu.ru). The studies were conducted under high vacuum conditions. Accelerating voltage under the scanning mode was 1–3 kV.

Research Results. The dependence of the friction factor on the concentration of carboxylic acid for various volume acid densities under the brass-steel couple friction is given in Fig. 1 and in Table 1. From the analysis of the

data obtained, it follows that the dependence in Fig. 1 is nonmonotonic, with a minimum concentration of acid in the solution of 0.1 mol/l.

Table 1

Dependence of friction factor on acid concentration in lubricant

Lubricant composition, acid solution	Acid chemical formula: *R-COOH,* where *R*	Acid concentration, mol/l				
		0.025	0.05	0.1	0.2	0.5
		friction factor, μ				
Formic	*-H*	0.307	0.348	0.274	0.391	0.545
Acetic	$-CH_3$	0.285	0.312	0.258	0.367	0.462
Propionic	$-CH_2CH_3$	0.195	0.129	0.087	0.239	0.335
Butyric	$-CH_2CH_2CH_3$	0.112	0.083	0.037	0.184	0.307
Valeric	$-CH_2CH_2CH_2CH_3$	0.071	0.031	0.009	0.045	0.154
Caproic	$-CH_2CH_2CH_2CH_2CH_3$	0.054	0.016	0.007	0.028	0.126

Studying the effect of acid concentration on the friction factor variation in the solution shows that strengthening of formic and acetic acids in the lubricant composition from 0.025 mol/l to 0.05 mol/l leads to the increase in the acidity of the medium and, consequently, to an increase of the values of the friction factor from 0.3 to 0.35.

As a result of frictional interaction, tribocorrosion occurs on the metal surface in the tribosystem. A further increase in the acid concentration to 0.1 mol/l causes a decrease in the friction factor, and then it is accompanied by its sharp increase and strong wear of friction couple materials as a result of the adhesion-mechanical interaction of the protrusions of the surface microrelief (Fig. 1, 2).

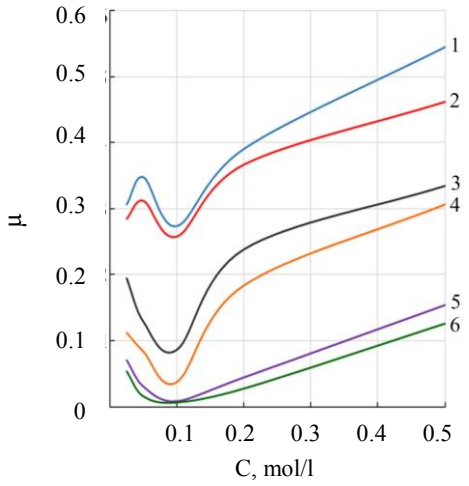


Fig. 1. Dependence of friction factor (μ) on acid concentration (C) in lubricant composition under friction of brass-steel couple: 1 is formic acid, 2 is acetic acid, 3 is propionic acid, 4 is butyric acid, 5 is valeric acid, 6 is caproic acid

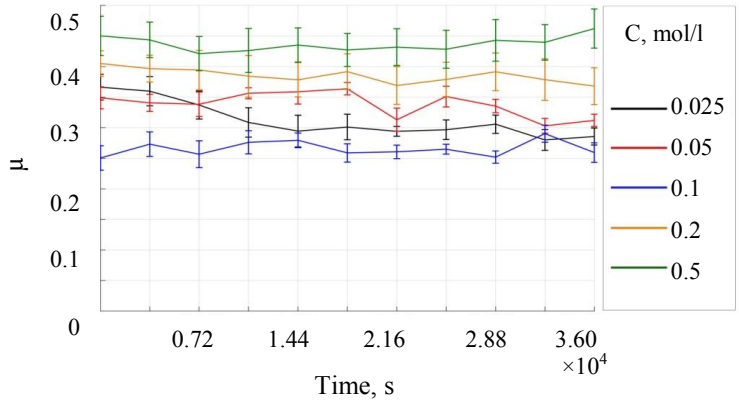


Fig. 2. Evolution of friction factor (μ) on concentration (C) in “brass-aqueous solution of acetic acid-steel” system

A similar dependence of the friction factor on the acid concentration in the lubricant composition is also observed under friction of a brass-steel couple in aqueous solutions of propionic and butyric acids (Fig. 3).

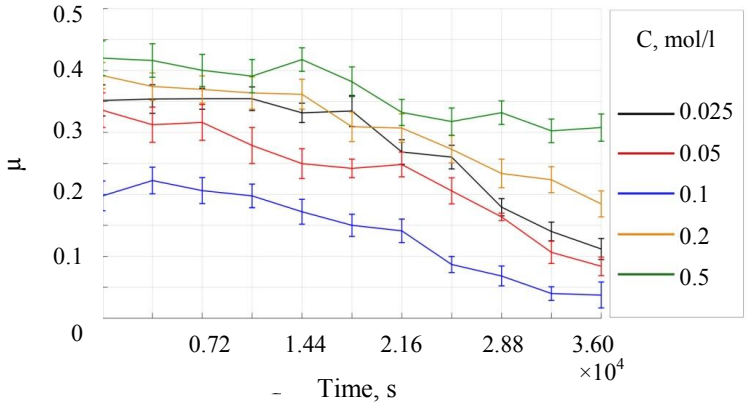


Fig. 3. Evolution of friction factor (μ) on concentration (C) in "brass-aqueous solution of butyric acid-steel" system

At this, the friction factor decreases significantly at the acid concentration of 0.1 mol/l and does not exceed 0.1. A further increase in the acid concentration in the lubricant composition, as well as in the case of formic and acetic acids, is accompanied by its sharp increase (Fig. 1). As a result, wear products are formed in the lubricating fluid volume. In this case, the friction surface is subjected to corrosion-mechanical wear, and the friction factor has rather high values from 0.15 to 0.35 (Fig. 3).

Analysis of the friction factor variation of a brass-steel couple in aqueous solutions of valeric and caproic acids with the concentrations of 0.025 and 0.05 mol/l determines rather low value to 0.07 (Fig. 1, 4).

Application of valeric and caproic acids in a lubricant composition with the concentration of 0.1 mol/l enables to obtain the lowest values of the friction factor to 0.007; optimal conditions for self-organization on the steel surface; the formation of a visually detectable servovite film [13, 18]; and the establishment of a wear-free mode in the tribological system. At this, reduction of the friction factor in the copper-steel alloy couple is associated with the best damping of tribo-film stresses caused by friction.

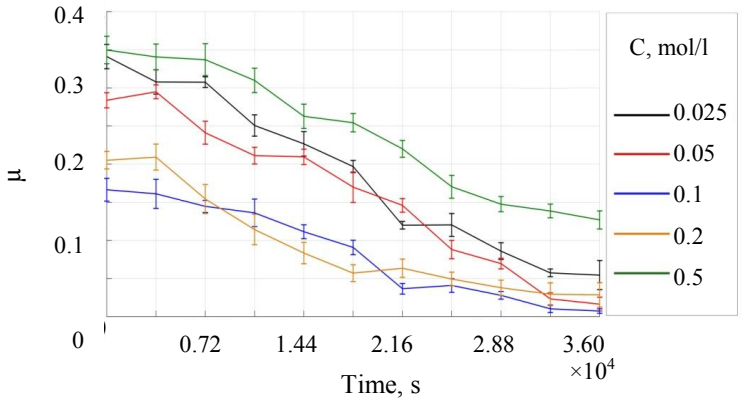


Fig. 4. Evolution of friction factor (μ) on concentration (C) in "brass-aqueous solution of caproic acid-steel" system

Triboelectrochemical reactions occurring in the frictional contact zone in aqueous solutions of valeric and caproic acids not only initiate the formation of a servovite film, which helps to reduce the friction factor, to heal surface microdefects, but intensify the chemosorption which enhances the ordering effect of the substrate on the orientation of molecules under the formation of adsorption acid molecule layer.

Active *–COOH* polar groups in a carboxylic acid molecule lead to its interaction with metal surfaces with the formation of chemically adsorbed compounds of the bidentate-ligand type [12]. Aliphatic monocarboxylic acids (*R – COOH*) form layers in which the hydrocarbon radicals of the molecules form a closely-packed structure. The adsorbed carboxylic acid molecules on the metal surface not only reduce the friction factor, as compared to a clean friction surface, as it follows from the results obtained, but also increase it. At very low acid concentrations in the solution – to 0.05 mol/l, the degree of filling the adsorption layer is very small; the molecules do not form a continuous film and move freely along the metal surface. Aside from the acid molecules, water molecules can adsorb on the metal surface; however, they do not have an effective shade action, therefore, the friction factor is almost the same as for clean surfaces. As the acid concentration in the solution rises to 0.1 mol/l, the occupancy of the adsorbed friction surface layer

increases, while the acid molecules run parallel to the surface reducing the friction factor as compared to the values for clean friction surfaces. The generated boundary layers at low acid concentrations reduce the friction factor. With further raise of the acid concentration, molecules are arranged perpendicular to the surface; the adsorbed layer increases the surface roughness; and the friction factor grows. The closedness of intermolecular bindings inside the adsorption layer leads to low surface energy and virtual absence of secondary adsorption of acids or other components [19–21].

Analysis of the results obtained through the scanning electron microscopy (SEM) indicates drastic structural changes in the friction surface in acid aqueous solutions when moving from formic to caproic acid (Fig. 4). As can be seen from the SEM visualization of the results obtained, with the relative movement of two surfaces in an aqueous solution of acetic and butyric acids, the surface of the tribocontact has a heterogeneous structure with a large number of pores and irregularities resulting from mechanochemical corrosion; as well as the occurrence of wear particles on the friction surface which causes abrasive wear of rubbing metals. At this, cracks and scratches appear on the friction surface, which significantly reduces wear resistance of the material. Micrographs of the sample surface after friction in aqueous solutions of acetic and butyric acids show major roughness of the friction track, which leads to an increase in the friction factor (Fig. 4a, b). The friction force growth also provokes an increase in temperature in the friction unit, and, as a result, thermal stress of the metal, which is also one of the causes of the crack formation, both on the material surface and in its volume.

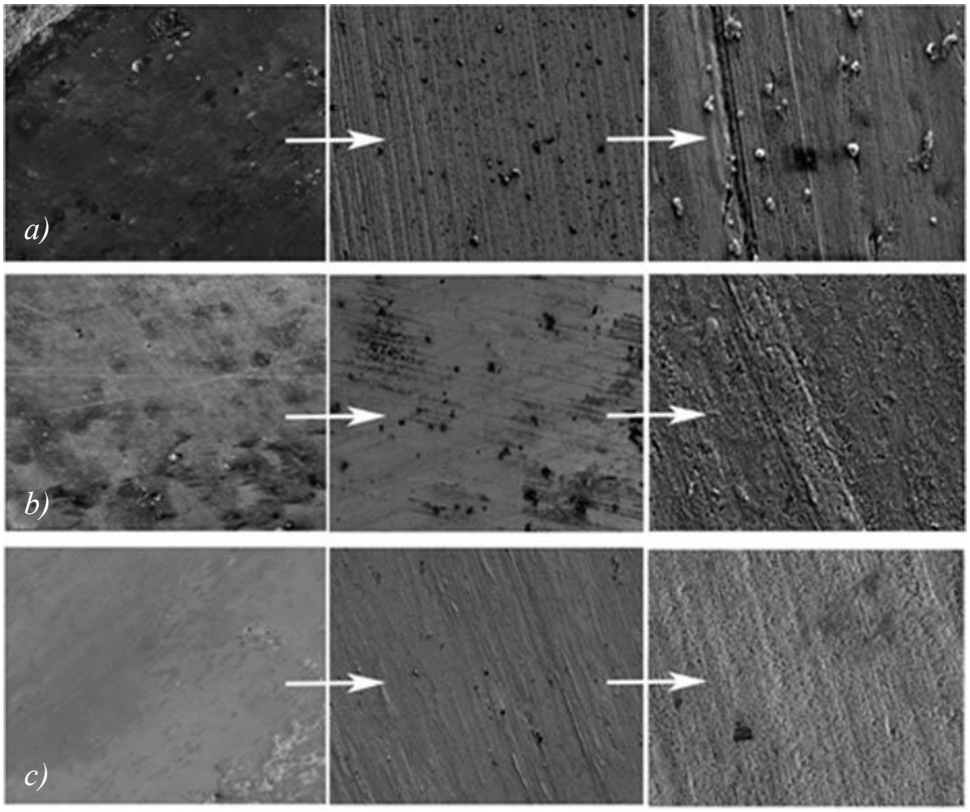


Fig. 5. SEM results of servovite film obtained under friction in “brass-aqueous solution of acid-steel” system:
(a) is acetic acid, (b) is butyric acid, (c) is caproic acid

When an aqueous solution of caproic acid is used as a lubricant composition under friction of a copper alloy over steel, the formation of a copper non-oxidizing layer on the surface of a steel disk is observed (Fig. 4c) [22]. This effect is explained by the catalytic effect of copper, which converts the lubricant monomers into polymers, ensuring that copper remains in an unoxidized state [12]. Copper film, which is formed through friction in an aqueous solution of caproic acid, has a dense structure with a minimum number of pores, ensuring wearless friction.

Conclusion. As a result of the tribological studies of a brass-steel friction couple in acid aqueous solutions, the optimum acid concentration in the lubricant composition was found. It is established that a change in the acid concentration causes a change in its adsorption on the friction surface, leading to a change in the morphology of the tribo-conjugated surfaces.

References

1. Calhoun, S.F. Antiwear and extreme pressure additives for greases. Tribology Transactions, 1960, vol. 3, pp. 208—214. DOI: 10.1080/05698196008972405

2. Ettefaghi, E., Ahmadi, H., Rashidi, A., Mohtasebi, S. S., Alaei, M. Experimental evaluation of engine oil properties containing copper oxide nanoparticles as a nanoadditive. International Journal of Industrial Chemistry, 2013, vol. 4, pp.1–6. DOI: 10.1186/2228-5547-4-28

3. Wu, Y., Tsuia, W., Liub, T. Experimental analysis of tribological properties of lubricating oils with nanoparticle additives. Wear, 2007, vol. 262, pp. 819–825. DOI: 10.1016/j.wear.2006.08.021

4. Su, F., Chen, G., Huang, P. Lubricating performances of graphene oxide and onion-like carbon as water-based lubricant additives for smooth and sand-blasted steel discs. Friction, 2018, pp. 1–11. DOI: 10.1007/s40544-018-0237-3

5. Kuzharov, A.S., et al. Nanotribologiya vodnykh rastvorov karbonovykh kislot pri trenii bronzy po stali. [Nanotribology of aqueous solutions of carboxylic acids under bronze friction on steel.] Innovatsii, ekologiya i resursosberegayushchie tekhnologii: materialy XI mezhdunar. nauch.-tekhn. foruma. [Innovation, ecology and resource-saving technologies: Proc. XI Int. Sci.-Tech. Forum.] Rostov-on-Don, 2014, pp. 712–717 (in Russian).

6. Burlakova, V.E., et al. Vliyanie prirody organicheskoy komponenty na tribotekhnicheskie svoystva sistemy «bronza-vodnyy rastvor karbonovoy kisloty-stal'». [Effect of organic component nature on tribological properties of “bronze-aqueous solution of carboxylic acid-steel” system.] Vestnik of DSTU, 2015, vol. 15, no. 4 (83), pp. 63–68. DOI: 10.12737/16067 (in Russian).

7. Yu, H. et al. Tribological properties and lubricating mechanisms of Cu nanoparticles in lubricant. Transactions of Nonferrous Metals Society of China, 2008, vol. 18, no. 3, pp. 636–641. DOI: 10.1016/S1003-6326(08)60111-9

8. Hu, Z.S, Lai, R., Lou, F., Wang, L., Chen, Z., Chen, G., et al. Preparation and tribological properties of nanometer magnesium borate as lubricating oil additive. Wear, 2002, vol. 252, pp. 370–374. DOI: 10.1016/S0043-1648(01)00862-6

9. Rastogi, R., Yadav, M., Bhattacharya, A. Application of molybdenum complexes of 1-aryl-2.5-dithiohydrazodicarbonamides as extreme pressure lubricant additives. Wear, 2002, vol. 252, no. 9–10, pp. 686–692. DOI: 10.1016/S0043-1648(01)00878-X

10. Kragelsky, I.V., Alisin, V.V. Friction wear lubrication: tribology handbook. Elsevier; 2016, 263 p.

11. Burlakova, V.E. Triboelektrokhimiya effekta bezyznosnosti. [Triboelectrochemistry of the wearlessness effect.] DSTU Publ. Centre, 2005, 211 p. (in Russian).

12. Burlakova, V.E., Milov, A. A., Drogan, E. G. Nanotribology of Aqueous Solutions of Monobasic Carboxylic Acids in a Copper Alloy–Steel Tribological Assembly. Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques, 2018, vol. 12, no. 6, pp. 1108–1116. DOI: 10.1134/S1027451018050427

13. Burlakova, V.E., et al. Vliyanie sostava smazochnoy sredy na strukturu poverkhnostnykh sloev formiruyushcheysya pri trenii servovitnoy plenki. [Effect of the lubricating medium composition on the structure of surface layers of a servovite film formed under friction.] Journal of Surface Investigation. X-Ray, Synchrotron and Neutron Techniques, 2019, no. 4. DOI: 10.1134/S0207352819040061 (in Russian).

14. Myshkin, N.K. Friction transfer film formation in boundary lubrication. Wear, 2000, vol. 245, iss. 1-2, pp. 116–124. DOI: 10.1016/S0043-1648(00)00472-5

15. Bulgarevich, S.B. Boiko, M.V., Feizova, V.A., Akimova, E.E. Effect of pressure on chemical reactions in the zone of direct friction contact of systems with selective transfer. Journal of Friction and Wear, 2011, vol. 32, iss. 3, pp. 145–149. DOI: 10.3103/S1068366611030020

16. Burlakova, V.E., et al. Mekhanicheskie svoystva servovitnykh plenok, formiruyushchikhsya pri trenii v vodnykh rastvorakh karbonovykh kislot. [Mechanical properties of servovite films formed in aqueous solutions of carboxylic acids under friction.] Vestnik of DSTU, 2018, vol. 18, no. 3, pp. 280–288. DOI: 10.23947/1992-5980-2018-18-3-280-288 (in Russian).

17. Burlakova, V.E., Kosogova, Yu.P., Drogan, E.G. Vliyanie nanorazmernykh klasterov medi na tribotekhnicheskie svoystva pary treniya stal'-stal' v vodnykh rastvorakh spirtov. [Effect of copper nanoclusters on the tribological properties of steel-steel friction pair in alcohol aqueous solutions.] Vestnik of DSTU, 2015, vol. 15, no. 2(81), pp. 41–47. DOI: 10.12737/11590 (in Russian).

18. Burlakova, V.E., Drogan, E.G., Gerashchenko, D.Yu. Tribologicheskie vozmozhnosti pary treniya latun'-stal' v vodnykh rastvorakh organicheskikh kislot. [Tribological capabilities of a of brass-steel friction pair in aqueous solutions of organic acids.] Tribologiya-mashinostroeniyu: trudy XII mezhdunar. nauch.-tekhn. konf., posvyashchennoy

80-letiyu IMASh RAN. [Tribology – to Mechanical Engineering: Proc. XII Int. Sci.-Tech. Conf. dedicated to the 80th anniversary of IMASH RAS.] Izhevsk, 2018, pp. 92–95 (in Russian).

19. Mukhortov, I.V. Polimolekulyarnaya adsorbtsiya smazochnykh materialov i ee uchet v teorii zhidkostnogo treniya. [Multimolecular adsorption lubricants and its integration in the fluid friction theory.] Bulletin of the South Ural State University. Series "Mechanical Engineering Industry", 2011, no. 31 (258), pp. 62–67 (in Russian).

20. Buyanovskiy, I.A., Ignatieva, Z.V., Levchenko, V.A., Matveenko, V.N. Orientatsionnaya uporyadochennost' granichnykh sloev i smazochnaya sposobnost' masel. [Orientation ordering of boundary layers and lubricity of oils.] Journal of Friction and Wear, 2008, vol. 29, no. 4, pp. 375–381 (in Russian).

21. Novoselova, M.V., Vilms, M.P. Tribologicheskie svoystva tonkikh plenok zhirnykh kislot. [Tribological properties of thin films of fatty acids.] Herald of TvSU. Series "Physics", 2011, no. 15, pp. 86–91 (in Russian).

22. Garkunov D.N. Scientific Discoveries in Tribotechnologies. No-wear effect under friction: Hydrogen wear of metals, 2007.

Received 12.12.2018
Submitted 20.01.2019
Scheduled in the issue 20.01.2019

Authors:

Burlakova, Victoria E.,
Head of the Chemistry Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Eng.), professor,
ORCID:http://orcid.org/0000-0003-3779-7079
vburlakova@donstu.ru

Drogan, Ekaterina G.,
junior research scholar, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID:https://orcid.org/0000-0002-4002-2082
ekaterina.drogan@gmail.com

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 31–37. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 31–37. ISSN 1992-5980 eISSN 1992-6006

МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ
MACHINE BUILDING AND MACHINE SCIENCE

UDC 621.7/9 https://doi.org/10.23947/1992-5980-2019-19-1-31-37

Formation features of composite electrochemical nickel and nanostructured zirconium boride coatings*

L. A. Degtyar[1], I. S. Ivanina[2], I. Yu. Zhukova[3]**
[1,2,3]Don State Technical University, Rostov-on-Don, Russian Federation
[1]Don State Agrarian University, Persianovsky village, Russian Federation

Особенности формирования композитных электрохимических покрытий на основе никеля и наноструктурного диборида циркония***

Л. А. Дегтярь[1], И. С. Иванина[2], И. Ю. Жукова[3]**
[1,2,3]Донской государственный технический университет, Ростов-на-Дону, Российская Федерация
[1]Донской государственный аграрный университет, п. Персиановский, Российская Федерация

Introduction. The electrodeposition of composite electrochemical coatings from electrolyte-colloid nickel plating containing ultradisperse zirconium boride powder is studied. The work objectives are as follows: to study mechanical-and-physical properties of the composites based on nickel and nanostructured zirconium boride, and to determine optimal conditions for the application of such electrochemical coatings.

Materials and Methods. Microhardness of composite electrochemical coatings was measured using PMT-3 microhardness tester on samples with the layer thickness of 30 μm under the indentation load of 100 g. A three-ball machine was used to determine wear resistance of the coatings. Sample tests were carried out under dry friction modes and with the use of 3% RV coolant. WSD values were measured by MIR-3 TU 3-3.1954-86 microscope. To determine the internal stresses in the coating, we used a flexible cathode method up to GOST 9.302-88.

Research Results. The electrolyte-colloid composition and modes of electrodeposition of composite nickel - nanostructured zirconium boride coatings are developed. Mechanical-and-physical properties (microhardness, wear resistance and internal stresses) of the obtained composite electrochemical coatings are analyzed. Recommendations for use of the developed electrolyte and the application of a composite coating on machine parts for their surface hardening are formulated.

Discussion and Conclusions. $Ni–ZrB_2$ CEC (composite electrochemical coating) has high microhardness (10–11 hPa at the indentation load of 100 g), which exceeds the microhardness of pure nickel by 1.5–2 times. As the microhardness increases, the internal stresses of $Ni–ZrB_2$ CEC decrease. The proposed coatings were compared to chromium ones deposited from the environmentally hazardous electrolytes. The wear resistance of

Введение. Исследован процесс электроосаждения композитных электрохимических покрытий из электролита-коллоида никелирования, содержащего ультрадисперсный порошок диборида циркония.

Цели работы: исследование физико-механических свойств композитов на основе никеля и наноструктурного диборида циркония, а также определение оптимальных условий нанесения такого электрохимического покрытия.

Материалы и методы. Микротвердость композитных электрохимических покрытий измеряли с помощью микротвердомера ПМТ-3 на образцах с толщиной слоя 30 мкм при нагрузке на индентор 100 г. Для определения износостойкости покрытий использовали трехшариковую машину трения. Испытания образцов проводили в режимах сухого трения и с применением 3 % смазочно-охлаждающей жидкости РВ. Значения диаметра пятна износа измерили под микроскопом МИР-3 ТУ 3-3.1954-86. Для определения внутренних напряжений в покрытии воспользовались методом гибкого катода в соответствии с ГОСТ 9.302-88.

Результаты исследования. Разработан состав электролита-коллоида и режимы электроосаждения композитных покрытий никель — наноструктурный диборид циркония. Проведен анализ физико-механических свойств (микротвердость, износостойкость и внутренние напряжения) полученных композитных электрохимических покрытий. Сформулированы рекомендации по использованию разработанного электролита и нанесению композитного покрытия на детали машин для их поверхностного упрочнения.

Обсуждение и заключение. КЭП $Ni–ZrB_2$ имеет высокую микротвердость (10–11 гПа при нагрузке на индентор 100 г), что превышает микротвердость чистого никеля в 1,5–2 раза. При возрастании микротвердости снижаются внутренние напряжения КЭП $Ni–ZrB_2$. Предлагаемые покрытия сравнивались с хромовыми, осаждаемыми из эко-

* The research is done within the frame of the independent R&D.
**E-mail: degtiar@yandex.ru, i.ivanina96@mail.ru, iyuzh@mail.ru
***Работа выполнена в рамках инициативной НИР.

Ni–ZrB_2 CEC is 2–5 times higher than that of chromium coatings. Thus, instead of chromic coatings, it is recommended to use the proposed composition for surface hardening of parts of the specialty machinery and industrial equipment.

логически опасных электролитов. Износостойкость КЭП *Ni–ZrB_2* в 2–5 раз больше, чем у хромовых покрытий. Таким образом, рекомендуется вместо хромовых покрытий использовать предлагаемый состав для поверхностного упрочнения деталей специальной техники и промышленного оборудования.

Keywords: electrolyte, composite electrochemical coating, nickel, zirconium boride, microhardness, wear resistance, internal stresses.

For citation: L.A. Degtyar, et al. Formation features of composite electrochemical nickel and nanostructured zirconium boride coatings. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 31–37. https://doi.org/10.23947/1992-5980-2019-19-1-31-37

Ключевые слова: электролит, композитное электрохимическое покрытие, никель, диборид циркония, микротвердость, износостойкость, внутренние напряжения.

Образец для цитирования: Особенности формирования композитных электрохимических покрытий на основе никеля и наноструктурного диборида циркония / Л. А. Дегтярь [и др.] // Вестник Дон. гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 31–37 https://doi.org/10.23947/1992-5980-2019-19-1-31-37

Introduction. Composite materials are widely used to strengthen the surface of typical nodes and parts of the petroleum-gas field, processing and other industrial equipment. As a result of a combination (composition) of dissimilar substances, a new material appears whose characteristics differ from the properties of its components [1].

To improve the performance properties of electroplating, particulate fillers are introduced into the solution. Thus, a composite electrochemical plating (CEP) with a matrix of the base metal is obtained. Precisely the creation of the CEP is one of the current areas of modern electroplating. To obtain the CEP, dispersed particles of various sizes and types from electrolytes codeposit simultaneously with metals. These particles are included in the coating; they improve drastically the operational properties (hardness, wear resistance, corrosion resistance) of the coatings, and give them new qualities (anti-friction, magnetic, catalytic).

Various technologies of producing nickel-based CEP [2–5] are known. The selection of technique is determined by the operating conditions of the composite coating (CC). At this, wealth of fillers for obtaining coatings with specific performance properties complicates the choice of dispersed particles [3, 5–9]. In [2, 3, 6–8, 10], the physicomechanical properties of various nickel-based CC with particulate fillers (for example, nanocarbon material taunite, aluminum oxide, silicon carbide, graphite, etc.) are considered. In the mentioned works, in particular, it was noted what determines the structural and morphological characteristics of the CEP [7–9], their resistance to corrosion [7, 11], the coefficient of sliding friction [8], heat resistance, microhardness, and other properties [11] . Conditions and modes of CC deposition, composition of electrolytes, nature and degree of fineness of powder or other additive materials are mentioned as determining factors. The use of micropowders with a grain size of more than 1 μm and ultrafine powders or suspensions can change the nature of the codeposition of particles with the excreted metal, and, therefore, the characteristics of the CEP [9].

The use of micropowders with a grain size of more than 1 μm and ultra-disperse powders or suspensions can change the nature of the codeposition of particles with the excreted metal, and, consequently, the characteristics of the CEP [9]. During the dispersion phase, the properties of electrolytes and coatings from them gain a number of advantages [12, 13, 14] critical from the point of view of the applicability.

Nickel CC obtained from electrolytes with additives based on zirconium dioxide to improve surface morphology, microhardness, heat resistance and other properties [5, 15] are of interest. Zirconium boride (ZrB_2) as an additive to electrolytes for obtaining CC was not studied. However, it is known that zirconium and boron compounds improve drastically the mechanical and processing properties of electroplating [16]. . This work objective is to study the physicomechanical properties of the CEP based on nickel and nanostructured ZrB_2, as well as to determine the optimal conditions for its deposition.

Materials and Methods. Reagents of *Aldrich* Company and distilled water were used for the fabrication of electrolytes. The content of ZrB_2 in the coating was determined by the weight (gravimetric) method. The CEP microhardness was measured with PMT-3 microhardness tester. For this purpose, samples with the layer thickness of 30 μm were used under the indentation load of 100 g.

The current output of the coatings was obtained using a copper coulometer. The wear resistance of the coatings was determined on a three-ball friction machine. Samples were tested under dry friction modes and with the use of 3% RV coolant [17]. The wear resistance of the CEP was investigated as follows. The balls of steel ShKh 15 with the area of 0.05 dm^2 were covered with 30 μm thickness coating. Washers of St 45 steel served as a counterbody. The wear scat

diameter was measured using MIR-3 TU3-3.1954-86 microscope. To determine the internal stresses in the coating, we used the flexible cathode method in accordance with GOST 9.302-88.

Research Results. The electroplating process and some features of the CEC with ultradisperse powder (UDP) ZrB_2 (0.04–0.06 μm) are studied. The creation of nickel-based CEC with certain physical and mechanical characteristics requires consideration of several factors. In this case, first and foremost, the challenge of the interaction of the dispersed phase, introduced earlier into the electrolyte, and colloid particles formed in the solution under the electrolysis or in the preparation of electrolyte is considered [12, 13].

For the studies, sulfate-chloride nickel plating electrolyte-colloid of the following composition was used, g/l: nickel sulfate – 250, nickel chloride – 60, α-aminoacetic acid – 20. ZrB_2 concentration varied from 1 to 60 g/l.

The optimal electrolysis conditions in the presence of ZrB_2 UDP are specified. For this purpose, ranges of the cathode current density and the electrolyte pH, in which the CEC of good quality is formed, are determined. In the electrolyte composition under study, α-aminoacetic acid was present as an additive agent. Its maximum buffer properties were shown at pH 2–2.5: in this case, light matte coatings of good quality were formed. The presence of an effective additive agent in the electrolyte allows for electrolysis at the high cathode current density.

If ZrB_2, is absent in the electrolyte, then the limiting value of the cathode current density is 2 A/dm^2. When this value is exceeded, a dark contour is formed on the test samples. This indicates alkalization in the cathode layer, the formation of nickel hydroxide in a coarsely dispersed form, and its inclusion in the coating [12]. In this case, the physicomechanical properties of the precipitation deteriorate, their fragility increases, and the coating cracks along the edges of the cathode.

The introduction of ZrB_2 UDP in the electrolyte affects drastically the limiting value of the current density. Good precipitates are formed above the permissible current density characteristic of the electrolyte under study without adding ZrB_2. In the experiments, coatings without a dark contour were obtained at 3; 4; 6 and 8 A/dm^2. At the current density above 4 A/dm^2, dendrites were formed at the corners of the cathode. Their appearance means that the delivery of nickel ions is the limiting stage of the process [12]. The concentration polarization is eliminated under increasing the concentration of nickel salts in the electrolyte. At the nickel sulfate concentration of 300 g/l and high current densities, dendrites were not formed at the cathode.

The change in the current efficiency of the coating depending on the cathode current density is due to the appearance of dendrites at the cathode. With increasing current density, the current efficiency increases, reaches the maximum value at 3 A/dm^2, and then decreases sharply (Fig. 1).

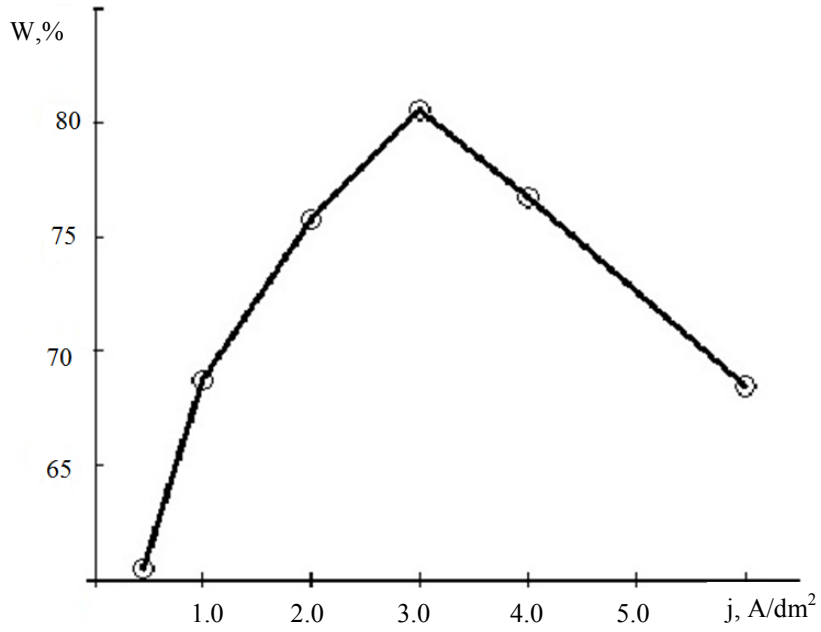


Fig. 1. Dependence of current efficiency of CEC $Ni–ZrB_2$ on cathode current density. Powder concentration in electrolyte is 40 g/l; electrolyte temperature is 50 °C; pH 2

When washing the samples, the dendrites are partially lost; as a result, the current output values of the coating decrease at the current density above 4 A/dm^2. The effect of increasing the upper value of the current density in the presence of ZrB_2 can be explained by the high degree of its dispersion. The particle size of the powder is comparable with the thickness of the electrical double layer, therefore, when particles move near the cathode surface, they can fall

into the double layer zone. In this case, colloidal nickel compounds are involved in the cathode layer. They are formed in the electrolyte under the preparation and electrodeposition, and they prevent sudden alkalization in the process of electrolysis.

The buffer capacity of the electrolyte remains almost unchanged with the addition of ZrB_2. pH in the cathode layer seems to decrease due to the addition of new portions of ZrB_2 from the electrolyte volume. The probability of this phenomenon is confirmed by the fact that the upper value of the operating current density does not increase in the presence of ZrB_2 micropowders (5 μm) in the electrolyte of similar composition.

The optimal concentration of ZrB_2 in the electrolyte was determined from its content in the coating and from the microhardness values of the CEC. It was established that in the CEC under study, the content of ZrB_2 reached maximum (1.1 wt.%) when its concentration in the electrolyte was 40–60 g/l. With an increase in the cathode current density at all the studied concentrations of ZrB_2 additive into the electrolyte, the content of ZrB2 in the CEC increases to the maximum limit (1.1 wt.%). CEC with stable microhardness characteristics is formed when ZrB_2 concentration in electrolyte is 30–40 g/l.

For production and cost reasons, the powder concentration should not exceed 40 kg/m^3. Otherwise:

- mixing efficiency of the electrolyte decreases;
- it is difficult to exclude the presence of particles in the workpiece surface areas which are hard-to-reach for the electrolyte circulation;
- quality of coatings deteriorates.

The dependence of the ZrB_2 content in the CEC on its concentration in the electrolyte is considered. It is shown that when the concentration of ZrB_2 particles is 10 g/l with the grain size of 2–9 μm, the distance between the ZrB_2 particles in the coating is approximately 1/3 of the distance between the particles in the electrolyte. As the ZrB_2 concentration increases by 10–20 times, the distance between the particles in the electrolyte decreases, and the coating remains almost unchanged. The concentration of ZrB_2 in the electrolyte can be significantly reduced, and the content of particles in the coating will not decrease. Thus, the use of electrolytes with a high concentration of ZrB_2 is impractical, since the maximum content of ZrB_2 in the coating is reached at its lower concentration in the electrolyte.

The microhardness of the CEC with ZrB_2 rises with increasing the powder concentration in the electrolyte. It reaches its highest value (11 GPa) at the ZrB_2 UDP concentration of 50–60 g/l and the current density of 5–6 A/dm^2. At low current densities (0.5 A/dm^2 and 1 A/dm^2), the microhardness gradually grows with increasing the concentration of the second phase. At the current densities above 2 A/dm^2, the microhardness of the CEC increases sharply, and already at the ZrB_2 concentration of 1 g/l it exceeds the microhardness of pure nickel by 1.5–2 times.

The increase in the CEC microhardness in the presence of the ultradisperse second phase is explained not only by the inclusion of solid and colloidal nickel particles in the coating, but also by a decrease in the crystalline grain at high current densities, as well as by the dispersion hardening of the nickel matrix. There is an opinion that a necessary condition for the dispersion hardening of the coating is the occurrence of ultrafine particles in it, which prevent recrystallization and the formation of coarse grains [18]. On the one hand, submicrometer-sized particles inhibit the growth of crystal grains; on the other hand, they form agglomerates in the electrolyte and cannot penetrate into in the coating as separate particles. Evidently, some part of the second phase is included in the sediment (CEC) in the form of individual particles and stimulates the formation of a finely-crystalline structure of a nickel matrix with high hardness. The probability of this phenomenon cannot be denied if for no reason than because particles of the same nature, differing in the dispersion degree, exhibit different ability to increase the microhardness of the CEC.

The possibility of obtaining CEC with a low content of the second phase and high hardness takes on particular importance for increasing the wear resistance of the surface, since, in this case, the matrix must have specific elasticity. A significant increase in the volumetric content of solid particles causes the fragility of materials. Thus, the use of UDP to obtain wear-resistant coatings is based on the practice. Equally important is the reduction of the consumption of the second phase, and, consequently, the process efficiency.

It is known that the hardness of coatings allows for indirect assertions about some physical and mechanical properties of the CEC. In some cases, it correlates adequately with wear resistance [4]. However, maximum microhardness does not always correspond to high wear resistance. The latter depends on the interaction of various factors of the process, as well as on the specific friction conditions. Hence, the wear resistance of the CEC with ZrB_2 was compared to chrome galvanic coatings traditionally used as wear resistant [19]. At this, both lubricated and unlubricated friction modes were considered. According to the results obtained, the service lifetime of the parts with $Ni – ZrB_2$ coating is 2–5 times longer than that of the parts with hard chrome plating.

Internal stresses of the $Ni–ZrB_2$ composite coatings decrease with an increase in the ZrB_2 concentration in the electrolyte and the coating thickness. Differences in the internal stress values of nickel CEC grow with increasing the sediment thickness (Fig. 2).

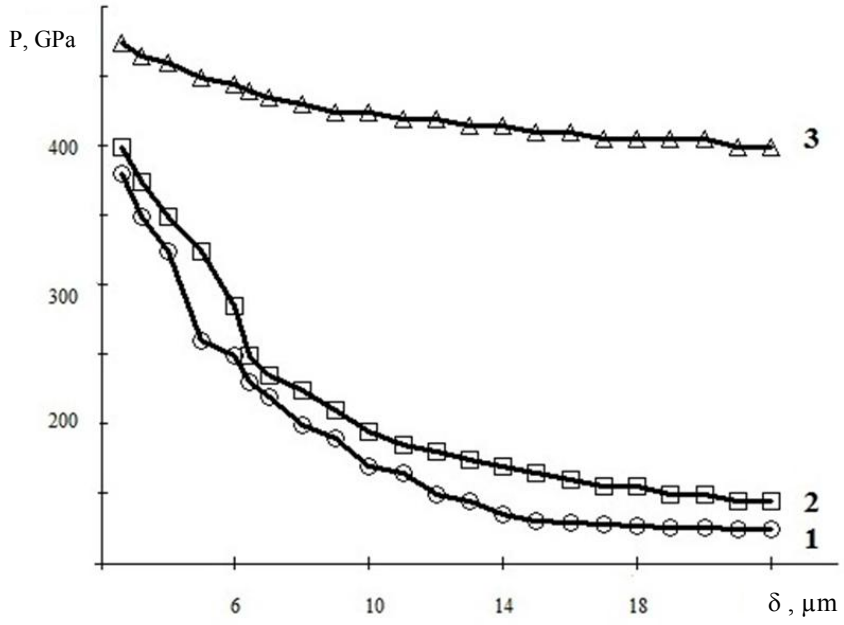


Fig. 2. Dependence of internal stresses on coating thickness and zirconium boride concentration in electrolyte, g/l: 1 — 30; 2 — 5; 3 — without second phase. Cathodic current density is 2 A/dm^2; electrolyte temperature is 50°C; pH 2

For galvanic coatings, the following regularity is peculiar: the finer the crystal grain and the higher the hardness, the greater the value of the internal stresses. In the presence of ZrB_2, the microhardness of the coatings increases sharply, and the internal stresses decrease. To explain this common pattern mismatch, it is necessary to consider the effect of the second-phase particles (particles of ZrB_2 UDP and colloidal particles formed in the electrolyte) on the precipitation hydrogenization ratio. The discharge of the hydrogen ions (and, consequently, a more complete removal of the molecular hydrogen) occurs faster on the second-phase particles than on the surface of nickel. So, the degree of precipitation hydrogenization decreases, which leads to a reduction of the internal voltage of the coating.

Discussion and Conclusions. In summary, the authors of the paper propose the following composition of nickel sulfate-chloride electrolyte-colloid for deposition of the nickel-based CEC and nanostructured ZrB_2, g/l: nickel sulfate - 250; nickel chloride - 60; α-aminoacetic acid - 20; ZrB_2 UDP - 30–40. The electrolysis mode is as follows: cathode current density is 2–4 A/dm^2; electrolyte pH is 2–2.6; temperature is 50° C; it is necessary to mix the electrolyte during the electrolysis process.

The electrolyte composition is developed; and the process condition for the application of the nickel-based CEC with the addition of ZrB_2 is established. The physical and mechanical properties of the obtained $Ni – ZrB_2$ CEC are studied: microhardness, wear resistance, and internal stresses.

CEC has high microhardness (10–11 GPa with the indentation load of 100 g), which is 1.5–2 times higher than the microhardness of pure nickel. The wear resistance is 2–5 times more than that of chrome coatings. As microhardness increases, the internal stresses of the $Ni – ZrB_2$ CEC decrease.

Considering the listed parameters, the proposed method can be used for surface hardening of parts for specialty machinery and industrial equipment.

References

1. Tseluikin, V.N. Kompozitsionnye elektrokhimicheskie pokrytiya: poluchenie, struktura, svoystva. [Composite electrochemical coatings: preparation, structure, properties.] Protection of Metals and Physical Chemistry of Surfaces, 2009, vol. 45, no. 3, pp. 287–301 (in Russian).

2. Desyatkova, G.I., et al. Kompozitsionnye elektrokhimicheskie pokrytiya na osnove nikelya. [Composite nickel-based electroplates.] Protection of Metals, 2002, vol. 38, no. 5, pp. 525–529 (in Russian).

3. Ovcharenko, O.A., Sakhnenko, N.D., Ved, M.V. Elektroosazhdenie i fiziko-mekhanicheskie svoystva kompozitsionnykh pokrytiy Ni–Al_2O_3. [Electrodeposition and physicomechanical properties of composite Ni–Al_2O_3 coatings.] Sovremennye metody v teoreticheskoy i eksperimental'noy elektrokhimii: tezisy dokl. VII Mezhdunar. nauch.

konf. [Modern methods in theoretical and experimental electrochemistry: Proc. VII Int. Sci. Conf.] Ivanovo: G.A. Krestov Institute of Solution Chemistry, 2015, p. 159 (in Russian).

4. Degtyar, L.A. Fiziko-mekhanicheskie kharakteristiki kompozitsionnykh elektrokhimicheskikh pokrytiy na osnove nikelya. [Physico-mechanical characteristics of nickel-based composite electrochemical coatings.] Innovatsionnye puti razvitiya APK: problemy i perspektivy: mat-ly mezhdunar. nauch.-prakt. konf. [Innovation-based development of the agro-industrial complex: problems and prospects: Proc. Int. Sci.-Pract. Conf.] Persianovsky: DSAU, 2013, vol. 1, pp. 13–14 (in Russian).

5. Wang, W., et al. Fabrication and characterization of Ni-ZrO_2 composite nano-coatings by pulse electrodeposition. Scripta Materialia, 2005, pp. 613–618 (in Russian).

6. Litovka, Yu.V., et al. Razrabotka tekhnologii polucheniya nanomodifitsirovannykh gal'vanicheskikh pokrytiy. [Development of technology for nano-modified electroplating.] Pokrytiya i obrabotka poverkhnosti: 7-ya mezhdunar. konf. [Coatings and surface treatment: 7th Int. Conf.] Moscow: Mendeleev University of Chemical Technology of Russia, 2010, p. 55 (in Russian).

7. Haghmoradi, N., Dehghanian, C., Khanlarkhani, H. The correlation among deposition parameters, structure and corrosion behaviour of ZnNi/nano-SiC composite coating deposited by pulsed and pulsed reverse current. Transactions of the IMF, 2018, vol. 96, pp. 155–162.

8. Tseluikin, V.N. Tribologicheskie svoystva kompozitsionnykh elektrokhimicheskikh pokrytiy na osnove nikelya. [Tribological properties of composite electrochemical nickel-based coatings.] Journal of Friction and Wear, 2010, vol. 31, no. 5, pp. 475–478 (in Russian).

9. Degtyar, L.A., Kudryavtseva, I.D., Sysoev, G.N. Kompozitsionnoe elektrokhimicheskoe pokrytie: patent 2048573 Ros. Federatsiya: MPK C22C 19/03A. [Composite electrochemical coating.] RF Patent, no. 2048573, 1995 (in Russian).

10. Novotortseva, I.G., Gaevskaya, T.V. O svoystvakh kompozitsionnykh pokrytiy na osnove nikelya. [On the properties of nickel-based composite coatings.] Journal of Applied Chemistry, 1999, vol. 72, no. 5, pp. 789–791 (in Russian).

11. Mingazova, G.G., et al. Kompozitsionnye pokrytiya s razlichnoy metallicheskoy matritsey. [Composite coatings with different metal matrix.] Herald of Kazan Technological University, 2012, vol. 12, no. 20, pp. 81–83 (in Russian).

12. Kudryavtseva, I.D., Kukoz, F.I., Balakay, V.I. Elektroosazhdenie metallov iz elektrolitov-kolloidov. [Electrodeposition of metals from electrolyte colloids.] Itogi nauki i tekhniki. Series “Electrochemistry”, 1990, vol. 33, pp. 50–85 (in Russian).

13. Degtyar, L.A., et al. Electrodeposition from Colloid-electrolyte bath: Some new features. 50 ISE Conference, Pavia (Italy), 5–10 September, 1999, Russian Journal of Electrochemistry. Available at: https://link.springer.com/journal/11175 (accessed 06.02.19).

14. Fomina, R.E., et al. Kompozitsionnye elektrokhimicheskie pokrytiya na osnove nikelya. [Nickel-based electrochemical composite coatings.] Herald of Kazan Technological University, 2018, vol. 21, no. 2, pp. 70–73 (in Russian).

15. Gorelov, S.M., Tsupak, T.E., Yarovaya, O.V. Poluchenie i svoystva kompozitsionnogo pokrytiya na osnove nikelya s nanorazmernym dioksidom tsirkoniya. [Electrodeposition of Ni-based composite coating with nano-size ZrO_2 particles.] Electroplating & Surface Treatment, 2014, vol. 22, no. 4, pp. 32–36 (in Russian).

16. Rogozhin, V.V. Osobennosti katodnogo osazhdeniya nikel'-bornykh pokrytiy iz sernokislogo elektrolita s dobavkami poliedricheskikh boratov. [Specific features of cathodic deposition of nickel-boron coatings from a sulfuric acid electrolyte with addition of polyhedral borates.] Journal of Applied Chemistry, 2008, vol. 84, no. 5, pp. 757–760 (in Russian).

17. Shulga, G.I., Bessarabov, N.I., Skrinikov, E.V. Metodika issledovaniya smazochnykh materialov na mashine treniya SMTs-2 pri trenii pary vrashchayushchiysya tsilindr - nepodvizhnyy tor. [Methods of studying lubricants on SMC-2 friction machine under friction of rotating cylinder - fixed torus couple.] University News. North-Caucasian region. Technical Sciences Series, 2005, no. 4, pp. 58–61 (in Russian).

18. Martin, J. Mikromekhanizmy dispersionnogo tverdeniya splavov. [Dispersion hardening micromechanisms.] Moscow: Metallurgiya, 1983, 167 p. (in Russian).

19. Stekolnikova, N.Yu., et al. Vosstanovlenie iznoshennykh izdeliy sel'skokhozyaystvennoy tekhniki gal'vanicheskim khromirovaniem. [Restoration of worn details of agricultural machinery using galvanic chrome plating.] Transactions of TSTU, 2016, vol. 22, no. 4, pp. 679–686 (in Russian).

Received 02.11 .2018
Submitted 02.11.2018
Scheduled in the issue 15.01.2019

Authors:

Degtyar, Ludmila A.,
associate professor of the Chemical Technologies of Oil and Gas Complex Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), associate professor of the Natural Sciences Department, Don State Agrarian University (24, Krivoshlykova St., Persianovaky vill., Rostov Region, 346495, RF), Cand.Sci. (Eng.),
ORCID: https://orcid.org/0000-0002-7018-6483
degtiar@yandex.ru

Ivanina, Inna S.,
graduate student of the Chemical Technologies of Oil and Gas Complex Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: https://orcid.org/0000-0002-4635-3464
i.ivanina96@mail.ru

Zhukova, Irina Yu.,
Head of the Chemical Technologies of Oil and Gas Complex Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Eng.), professor,
ORCID: https://orcid.org/0000-0002-9902-3821
iyuzh@mail.ru

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 38–44. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 38–44. ISSN 1992-5980 eISSN 1992-6006

МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ
MACHINE BUILDING AND MACHINE SCIENCE

UDC 62.752; 621.534; 629.4.015 https://doi.org/10.23947/1992-5980-2019-19-1-38-44

Dynamic damping under introduction of additional couplings and external actions*

S. V. Eliseev[1], A. S. Mironov[2], Quang Truc Vuong[3]**
[1,2,3] Irkutsk State Transport University, Irkutsk, Russian Federation

Динамическое гашение колебаний при введении дополнительных связей и внешних воздействий***

С. В. Елисеев[1], А. С. Миронов[2], К. Ч. Выонг[3]**
[1,2,3] Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация

Introduction. The dynamic interaction features in mechanical oscillating systems, whose structure includes additional couplings, are considered. In practice, such cases occur when using various optional mechanisms and motion translation devices under the formation of technical objects. The study objective is to develop a method for constructing mathematical models in the problems of dynamics of the mechanical oscillating systems with optional devices and features in the system of external disturbing factors.

Materials and Methods. The techniques used to study properties of the systems and the dynamic effects are based on the ideas of structural mathematical modeling. It is believed that the mechanical oscillating system, considered as a design model of a technical object, can be compared to the dynamically equivalent automatic control system. The mathematical apparatus of the automatic control theory is used.

Research Results. A method for constructing mathematical models is developed. The essential analytical relations for plotting oscillating systems are obtained, which enable to form a methodological basis for the integral estimation and comparative analysis of the initial system properties in various dynamic states. Dynamic properties of the two-degree-of-freedom systems within the framework of the computer simulation are investigated. The implementability of dynamic oscillation damping mode simultaneously in two coordinates with the joint action of two in-phase kinematic perturbations in the mechanical oscillating systems is shown.

Discussion and Conclusions. The possibilities of new dynamic effects, which are associated with the change in the system

Введение. Рассматриваются особенности динамических взаимодействий в механических колебательных системах, в структуре которых имеются дополнительные связи. Практически такие ситуации возникают при использовании в формировании технических объектов различных дополнительных механизмов и устройств для преобразования движения. Цель исследования заключается в разработке метода построения математических моделей в задачах динамики механических колебательных систем с дополнительными устройствами и особенностями в системе внешних возмущающих факторов.

Методы, используемые для исследования свойств систем и изучения динамических эффектов, основаны на идеях структурного математического моделирования. Полагается, что механической колебательной системе, рассматриваемой в качестве расчетной схемы технического объекта, можно сопоставить эквивалентную в динамическом отношении систему автоматического управления. Используется математический аппарат теории автоматического управления.

Результаты исследования. Разработан метод построения математических моделей. Получены необходимые аналитические соотношения для построения частотных диаграмм колебательных систем, позволяющие сформировать методологическую основу для интегральной оценки и сравнительного анализа свойств исходных систем в различных динамических состояниях. Проведены исследования динамических свойств систем с двумя степенями свободы в рамках вычислительного моделирования. Доказаны возможности реализации в механических колебательных системах режимов динамического гашения колебаний одновременно по двум координатам при совместном действии двух синфазных кинематических возмущений.

Обсуждение и заключения. Отмечены возможности проявления новых динамических эффектов, которые связаны с изменением структуры системы при определенных формах динамических взаимодействий.

* The research is done within the frame of independent R&D.
** E-mail: eliseev_s@inbox.ru, art.s.mironov@mail.ru, trucvq1990@gmail.com
*** Работа выполнена в рамках инициативной НИР.

structure under certain forms of dynamic interactions, are noted. The study is of interest to experts in machine dynamics, robotics, mechatronics, nano and mesomechanics.

Работа представляет интерес для специалистов в области динамики машин, робототехники, мехатроники, нано- и мезомеханики.

Keywords: structure diagrams, transfer functions,frequency plots, dynamic damping, additional couplings, joint interactions

Ключевые слова: структурные схемы,передаточные функции, частотные диаграммы, динамическое гашение колебаний, дополнительные связи, совместные взаимодействия.

For citation: S.V. Eliseev, et al. Dynamic damping under introduction of additional couplings and external actions. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 38–44. https://doi.org/10.23947/1992-5980-2019-19-1-38-44

Образец для цитирования: Елисеев, С. В. Динамическое гашение колебаний при введении дополнительных связей и внешних воздействий / С. В. Елисеев, А. С. Миронов, К. Ч. Выонг // Вестник гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 38–44. https://doi.org/10.23947/1992-5980-2019-19-1-38-44

Introduction. Dynamic oscillation damping (DOD) is widely used in practice to change the state of technical objects and the local action on the forms of component interaction of mechanical oscillatory systems. In the papers [1–3], the features of the approaches to the implementation of the dynamic oscillation damping modes were considered, options of the design and engineering solutions and calculation technique of the system parameters were proposed.

A variety of dynamic tasks predetermines a wide variability of the proposed solutions, within the framework of which the dynamic features of the protected objects, the conditions of external disturbances of the original system and the design and engineering modes of the dynamic oscillation damping are considered [4–8].

DOD is used in the tasks of protecting instrumentation systems and tooling [9, 10]. However, some aspects of evaluating the dynamic properties of vibration protection systems have not obtained a proper level and detail of representations in the formulation of research tasks and parameter analysis for dynamic absorbers. This may be due to the introduction and use of additional links, considering the characteristics of external influences, as well as the effect of the simultaneous joint action of several external disturbing factors.

In this paper, we develop a method of constructing mathematical models and the formation of dynamic oscillation damping effects in the chain mechanical oscillating two-degrees-of-freedom systems.

I. Background. The original system is shown in Fig. 1. It represents two inertia members (m_1 and m_2), which are interconnected by the elastic elements with rigidities (k_1, k_2, k_3) and additional links in the form of motion translation devices (MTD) with reduced masses (L_1, L_2, L_3). The system performs small oscillations under the action of external in-phase harmonic effects. Resistant forces are not considered.

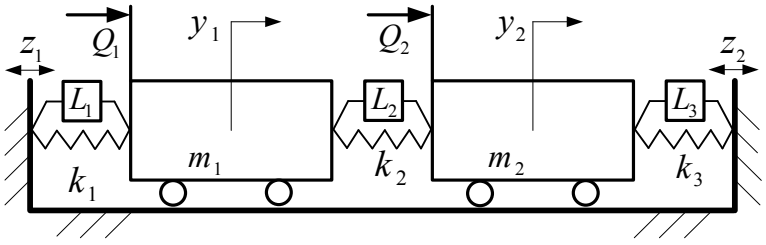


Fig. 1. Flowsheet of technical object in form of two-degrees-of-freedom lumped system

To describe the motion, the coordinate system (y_1, y_2) is used in a fixed basis. Suppose that the kinetic and potential energies of the system are determined by the following expressions:

$$\mathrm{T} = \frac{1}{2}m_1\dot{y}_1^2 + \frac{1}{2}m_2\dot{y}_2^2 + \frac{1}{2}L_1(\dot{y}_1 - \dot{z}_1)^2 + \frac{1}{2}L_2(\dot{y}_2 - \dot{y}_1)^2 + \frac{1}{2}L_3(\dot{y}_2 - \dot{z}_2)^2, \quad (1)$$

$$\Pi = \frac{1}{2}k_1(y_1 - z_1)^2 + \frac{1}{2}k_2(y_2 - y_1)^2 + \frac{1}{2}k_3(y_2 - z_2)^2. \quad (2)$$

The set of differential equations of the system motion in the time domain is obtained on the basis of the formalism of the Lagrange equations of second kind, and it has the form:

$$(m_1 + L_1 + L_2)\ddot{y}_1 + y_1(k_1 + k_2) - \ddot{y}_2 L_2 - k_2 y_2 = L_1\ddot{z}_1 + k_1 z_1 + Q_1, \quad (3)$$

$$(m_2 + L_2 + L_3)\ddot{y}_2 + y_2(k_2 + k_3) - \ddot{y}_1 L_2 - k_2 y_1 = L_3\ddot{z}_2 + k_3 z_2 + Q_2. \quad (4)$$

After the Laplace transformations under zero-initial conditions [11], the equation set (3)–(4) can be represented by a structural mathematical model in the form of a circuit that is dynamically equivalent to the automatic control system [12, 13] shown in Fig.2.

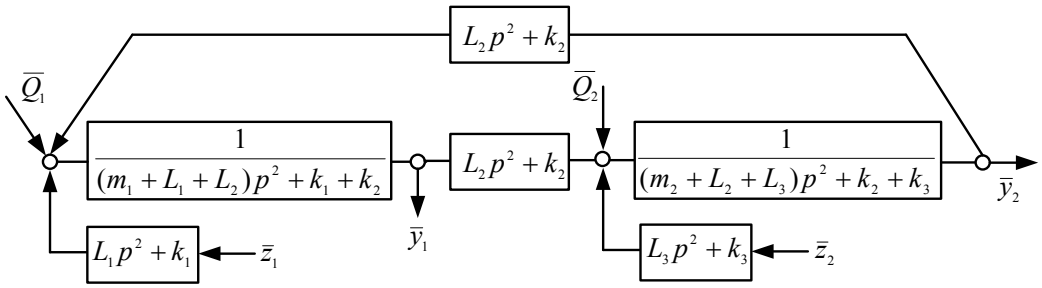


Fig. 2. Structural mathematical model of mechanical system shown in Fig. 1

II. Mathematical Models Development. The type of transfer functions depends on the nature of external disturbances, that is, on whether the disturbances are force ($\bar{Q}_1$ and $\bar{Q}_2$) or kinematic ($\bar{z}_1$ and $\bar{z}_2$). Further on, suppose that the force external actions ($\bar{Q}_1$ and $\bar{Q}_2$) have connectivity determined by the ratio

$$\bar{Q}_2 = \alpha \cdot \bar{Q}_1, \tag{5}$$

where α is coefficient of connectivity of external in-phase harmonic actions.

For kinematic effects ($\bar{z}_1$ and $\bar{z}_2$), it is assumed that

$$\bar{z}_2 = \beta \cdot \bar{z}_1, \tag{6}$$

where β is coefficient of connectivity of kinematic disturbances.

The connectivity coefficients (α and β) can have positive, negative, and zero values. Special cases of the effect of external disturbances can be considered through zeroing α and β properly.

1. The case of joint force perturbation at $\beta \neq 0$ ($\bar{Q}_1 = 0$ and $\bar{Q}_2 = 0$) is considered. The transfer functions of the system in this case take the form:

$$W_{\substack{1 \\ \bar{Q}_1=\bar{Q}_2=0}}(p) = \frac{\bar{y}_1}{\bar{z}_1} = \frac{(L_1p^2+k_1)[(m_2+L_2+L_3)p^2+k_2+k_3]+\beta(L_3p^2+k_3)(L_2p^2+k_2)}{A(p)}, \tag{7}$$

$$W_{\substack{2 \\ \bar{Q}_1=\bar{Q}_2=0}}(p) = \frac{\bar{y}_2}{\bar{z}_1} = \frac{\beta(L_3p^2+k_3)[(m_1+L_1+L_2)p^2+k_1+k_2]+(L_1p^2+k_1)(L_2p^2+k_2)}{A(p)}, \tag{8}$$

where

$$A(p) = [(m_1+L_1+L_2)p^2+k_1+k_2]\cdot[(m_2+L_2+L_3)p^2+k_2+k_3]-(L_2p^2+k_2)^2 \tag{9}$$

is a system frequency standard equation.

When considering the transfer functions (7), (8), it is assumed that the dynamic mode of oscillation damping is determined by the conditions for zeroing numerators (7), (8). Coordinate may cause two frequencies of dynamic oscillation damping. Along the coordinate ($\bar{y}_1$), the occurrence of two DOD frequencies is possible. Along the coordinate ($\bar{y}_2$), it is also possible to implement two DOD modes due to the virtual existence of roots of the biquadratic frequency equation.

Assuming that the variable factor is β (coefficient of connectivity), a frequency diagram can be developed considering the following frequencies:

1. partial frequencies:

$$n_1^2 = \frac{k_1+k_2}{m_1+L_1+L_2}, \tag{10}$$

$$n_2^2 = \frac{k_2+k_3}{m_2+L_2+L_3}; \tag{11}$$

2. critical frequency of interpartial communication:

$$n_{\text{пар}}^2 = \frac{k_2}{L_2}; \tag{12}$$

3. DOD frequencies determined from the solution of equations in $\bar{y}_1$ coordinate:

$$p^4[L_1(m_2+L_2+L_3)+\beta L_2L_3]+p^2[L_1(k_2+k_3)+k_1(m_2+L_2+L_3)+\beta(k_3L_2+k_2L_3)]+ \\ +k_1(k_2+k_3)+\beta k_2k_3=0; \quad (13)$$

and in $\bar{y}_2$ coordinate:

$$p^4[\beta L_3(m_1+L_1+L_2)+L_1L_2]+p^2[\beta L_3(k_1+k_2)+\beta k_3(m_1+L_1+L_2)+k_1L_2+k_2L_1]+ \\ +\beta k_3(k_1+k_2)+k_1k_2=0. \quad (14)$$

To construct a frequency diagram, the following parameters of the model problem are accepted: m_1 = 10 kg; m_2 = 10 kg; k_1 = 5000 N/m; k_2 = 10000 N/m; k_3 = 15000 N/m; L_1 = 5 kg; L_2 = 10 kg; L_3 = 10 kg.

The modes of dynamic oscillation damping in $\bar{y}_1$ and $\bar{y}_2$ coordinates are determined not only by the parameters of elastic inertia members, but also by the specificities of the external actions formation, in particular, by the connectivity value.

From the equations (13), (14), corresponding DOD frequencies can be found showing that two dynamic oscillation damping frequencies can be found in each of $\bar{y}_1$ and $\bar{y}_2$ coordinates. The frequencies values, as it follows from (13) - (14), depend on the connectivity coefficient of kinematic disturbances (β). Fig. 3 shows the system frequency diagram.

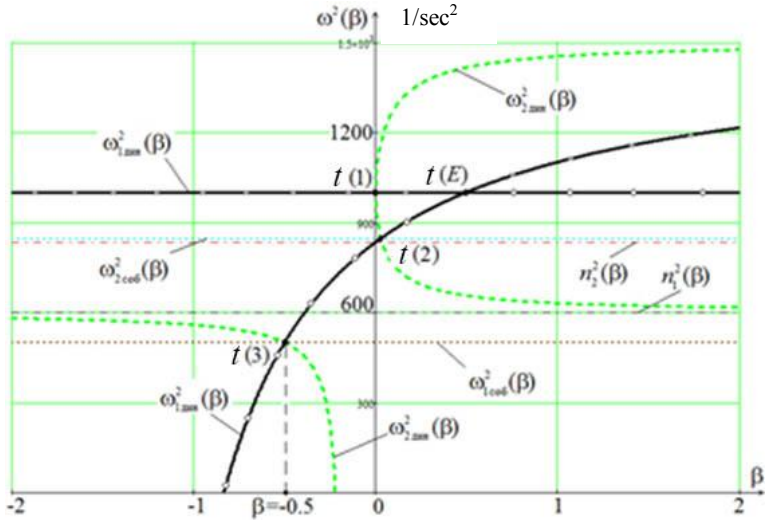


Fig.3. System frequency diagram shown in Fig. 1

In the diagram, the solid line (——) corresponds to $\omega^2_{1\,дин}(\beta)$ dependency graph. Since the frequency equations (13) - (14) are biquadratic, each of the equations has two roots. This is displayed in two graphs. For $\omega^2_{1\,дин}(\beta)$ graph of dependences, the solid line is marked with special symbols (×××), and for the second root, respectively (◦◦◦◦); $\omega^2_{1\,дин}(\beta)$ dependency graphs are touching in t. (*E*). Again, $\omega^2_{2\,дин}(\beta)$ dependency graphs are represented by dashed lines (— — —). $\omega^2_{2\,дин}(\beta)$ dependency graph consists of two non-contiguous blocks. Mutual intersections of $\omega^2_{1\,дин}(\beta)$ and $\omega^2_{2\,дин}(\beta)$ dependency diagrams occur in tt. (1), (2), (3). Each of the considered points determines the amplitude-frequency characteristics associated with the DOD modes features.

In the usual formulation of studying the dynamic oscillation damping, that is, under the action of a single perturbing factor correlated with a certain coordinate, one DOD frequency is determined in the two-degrees-of-freedom system. Such a frequency is determined by the partial frequency values of that system block, the movement of which demonstrates the dynamic oscillation damping (that is, "zeroing" the value of the corresponding coordinate).

Under the action of several simultaneous disturbances, it becomes possible to implement two DOD modes in each of the coordinates for the system as a whole. When additional links are introduced into the system, in particular, on the basis of the MTD, specific properties occur when the dynamic oscillation damping becomes possible simultaneously in two coordinates.

III. Comparative analysis of dynamic properties of the systems in the DOD modes.

1. Fig. 4 shows frequency-response characteristics (FRC) of the system, which manifest themselves under the conditions corresponding to the intersection of $\omega^2_{1\,дин}(\beta)$ and $\omega^2_{2\,дин}(\beta)$ graphs in t. (1) in Fig. 3. The intersection

corresponds to the case when β = 0. The solid line (——) in Fig. 4 corresponds to $\frac{\overline{y}_1}{\overline{z}_1}(\omega)$ dependency graph; the dotted line (·····) corresponds to $\frac{\overline{y}_2}{\overline{z}_1}(\omega)$ graph. The parameters of the system as a whole are also shown in Fig. 4.

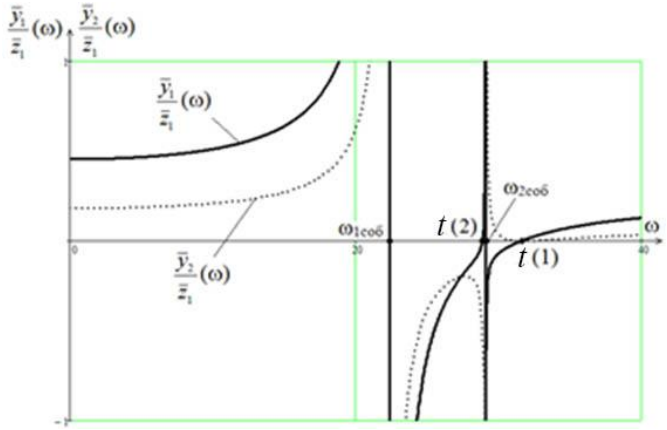


Fig. 4. Frequency-response characteristics of the system for parameters determined by t. in Fig.3

In t. (1) in Fig. 4, the frequency of the DOD mode is determined when $\overline{y}_1$ and $\overline{y}_2$ coordinate values are simultaneously "zeroed out".

It is possible to implement another DOD mode in $\overline{y}_1$ coordinate corresponding to $\frac{\overline{y}_1}{\overline{z}_1}(\omega)$ graph in t. (2) (Fig. 4). At this, the FRC reflect the properties of two-degrees-of-freedom systems. As follows from the FRC, it is possible to implement two DOD modes in tt. (1) and (2) in $\overline{y}_1$ coordinate. For the FRC corresponding to $\overline{y}_2$ coordinate, it is also possible to create two DOD modes at the double intersection of $\frac{\overline{y}_2}{\overline{z}_1}(\omega)$ graph by the abscissa line after $\omega^2_{2\,соб}$ eigenfrequency. Thus, with non-degenerate FRC for each of $\overline{y}_1$ and $\overline{y}_2$ coordinates, two DOD modes can be implemented; while at one of the frequencies, there is simultaneous DOD in two coordinates.

2. Fig. 5 shows the system FRC at $\beta = -\frac{1}{2}$, from which it follows that it becomes possible to restructure the system when one degree of freedom degrades at certain parameter ratios. In this case, in each of $\overline{y}_1$ and $\overline{y}_2$ coordinates (points (1) and (2) in Fig. 5), it is possible to implement the DOD modes, but this occurs when the system "degrades"; under the conditions when FRC ω → ∞ acquire limiting properties.

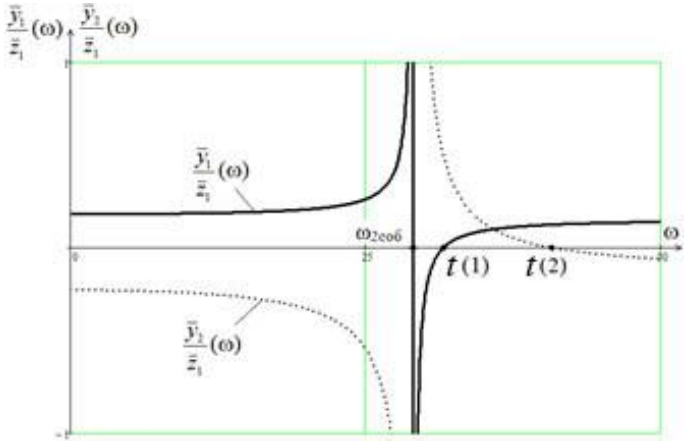


Fig. 5. Frequency-response characteristics of the system at $\beta = -\frac{1}{2}$ corresponding to t. (3) in Fig.3

3. Fig. 6 shows the system FRC at $\beta = \frac{1}{34}$ which corresponds to t. (2) in Fig. 3. For the given state of the system determined by the parameter values, a structural transformation of the system is also characteristic; the system

"degrades" to the status of a one-degree-of-freedom system. For the system on a whole, in each of the coordinates, the DOD implementation is possible, which corresponds to t. (1) and t. (2) in the graphs shown in Fig. 6.

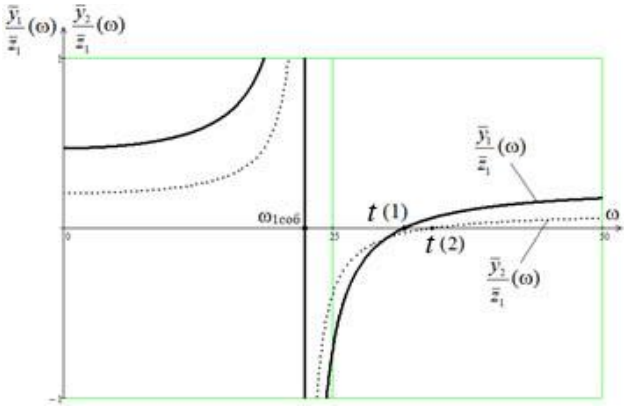


Fig. 6. Frequency-response characteristics of the system at $\beta = \frac{1}{34}$ corresponding to t. (2) in Fig. 3

In the high frequency area ($\omega \to \infty$), the system acquires limiting properties; at this, in comparison with the previous example, when $\beta = -\frac{1}{2}$, the oscillation amplitudes ratio will have a different sign at $\beta = \frac{1}{34}$. Consequently, changes in system parameters under the simultaneous action of two forces can change drastically the dynamic properties of the mechanical oscillatory systems.

Conclusion. The simultaneous action of external disturbances in the presence of additional links in the system, implemented by the MTD under the kinematic disturbance, can have a great impact on the change in the dynamic properties of the mechanical oscillatory systems with several degrees of freedom. So, the authors in this paper have obtained the following research results:

1. A technique of constructing mathematical models based on the use of methods of structural mathematical simulation, in which the mechanical oscillatory system is compared to the dynamically equivalent automatic control system, is proposed;

2. A technique for constructing frequency diagrams that allows for the integral estimation of the interdependence of frequency characteristics is proposed for the case when the system parameters and their perturbation conditions for various power factors change;

3. The analytical conditions for the implementation of the DOD modes simultaneously in two coordinates under the action of two interlinked disturbing factors are obtained;

4. The possibilities to control the structural states when the original mechanical oscillatory system can change the number of degrees of freedom and the system of its dynamic properties are proposed.

References

1. Eliseev, S.V., Nerubenko, G.P. Dinamicheskie gasiteli kolebaniy. [Dynamic Vibration Dampers.] Novosibirsk: Nauka, 1982, 142 p. (in Russian).

2. Korenev, B.G., Reznikov, L.M. Dinamicheskie gasiteli kolebaniy: teoriya i tekhnicheskie prilozheniya. [Dynamic vibration dampers: theory and engineering applications.] Moscow: Nauka, 1988, 304 p. (in Russian).

3. Eliseev, S.V., Khomenko, A.P. Dinamicheskoe gashenie kolebaniy: kontseptsiya obratnoy svyazi i strukturnye metody matematicheskogo modelirovaniya. [Dynamic vibration damping: feedback concept and structural methods of mathematical simulation.] Novosibirsk: Nauka, 2014. — 357 c. (in Russian).

4. Karamyshkin, V.V. Dinamicheskie gasiteli kolebaniy. [Dynamic vibration damping.] Leningrad: Mashinostroenie, 1988, 108 p. (in Russian).

5. Chelomey, V.N., ed. Vibratsii v tekhnike. [Vibrations in machinery.] Moscow: Mashinostroenie, 1981, 456 p. (in Russian).

6. Harris, C.M., Piersol, A.G. Shock and Vibration Handbook. New York: McGraw — Hill Book Co., 2002, 1457 p.

7. Eliseev, S.V., Belokobylskiy, S.V. Matematicheskie modeli mekhanicheskikh sistem s g-obraznymi dinamicheskimi gasitelyami. [Mathematical models of mechanical systems with L-shaped dynamic dampers.]

Matematika, ee prilozheniya i matematicheskoe obrazovanie: materialy IV mezhdunar. konf. [Mathematics, its applications and mathematical education: Proc. IV Int. Conf.] Ulan-Ude, 2011, pp. 150–155 (in Russian).

8. Eliseev, S.V., Artyunin, A.I., Kaimov, E.V. T-obraznye rychazhnye mekhanizmy v teorii transportnoy vibratsionnoy zashchity. [T-shaped lever mechanisms in the theory of transport vibration protection.] Kulaginskie chteniya: materialy XIV mezhdunar. nauch.-prakt. konf. [Kulagin readings: Proc. XIV Int. Sci.-Pract. Conf.] Chitaa, 2014, pp. 298–305 (in Russian).

9. Eliseev, S.V., Kashuba, V.B., Nikolaev, A.V., Quang Truc Vuong. Nekotorye vozmozhnosti dinamicheskogo gasheniya kolebaniy v sistemakh s neskol'kimi stepenyami svobody. [Some potentialities of dynamic oscillation damping in system with some degrees of freedom.] The Bryansk State University Herald, 2017, no. 1 (54), pp. 290–301 (in Russian).

10. Ilyinskiy, V.S. Zashchita apparatov ot dinamicheskikh vozdeystviy. [Protection of devices from dynamic effects.] Moscow: Energiya, 1970, 320 p. (in Russian).

11. Eliseev, S.V., Artyunin, A.I. Prikladnaya teoriya kolebaniy v zadachakh dinamiki lineynykh mekhanicheskikh system. [Applied theory of oscillations in problems of dynamics of linear mechanical systems.] Novosibirsk: Nauka, 2016, 459 p. (in Russian).

12. Eliseev, S.V., Reznik, Yu.N., Khomenko, A.P., Zasyadko, A.A. Dinamicheskiy sintez v obobshchennykh zadachakh vibrozashchity i vibroizolyatsii tekhnicheskikh ob"ektov. [Dynamic synthesis in generalized tasks of vibration protection and vibration isolation of technical objects.] Irkutsk: ISU Publ. House, 2008, 523 p. (in Russian).

13. Eliseev, S.V., Reznik, Yu.N., Khomenko, A.P. Mekhatronnye podkhody v dinamike mekhanicheskikh kolebatel'nykh system. [Mechatronic approaches in the dynamics of mechanical oscillatory systems.] Novosibirsk: Nauka, 2011, 384 p. (in Russian).

Received 01.11.2018
Submitted 01.11.2018
Scheduled in the issue 15.01.2019

Authors:

Eliseev, Sergey V.,
Director - Chief Research Scholar, REC of High Technologies, System Analysis and Modeling, Irkutsk State Railway Transport Engineering University (15, ul. Chernyshevskogo, Irkutsk, 664074, RF), Dr.Sci. (Eng.), professor,
ORCID: http://orcid.org/0000-0001-6876-8786
eliseev_s@inbox.ru

Mironov, Artem S.,
EdD Candidate, REC of High Technologies, System Analysis and Modeling, Irkutsk State Railway Transport Engineering University (15, ul. Chernyshevskogo, Irkutsk, 664074, RF),
ORCID : https://orcid.org/0000-0002-0921-0915
art.s.mironov@mail.ru

Quang Truc Vuong,
postgraduate student, REC of High Technologies, System Analysis and Modeling, Irkutsk State Railway Transport Engineering University (15, ul. Chernyshevskogo, Irkutsk, 664074, RF),
ORCID: https://orcid.org/0000-0003-3026-5301
trucvq1990@gmail.com

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 45–55. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 45–55. ISSN 1992-5980 eISSN 1992-6006

МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ
MACHINE BUILDING AND MACHINE SCIENCE

UDC 621.9 https://doi.org/10.23947/1992-5980-2019-19-1-45-55

Development of digital twin of CNC unit based on machine learning methods*

Yu. G. Kabaldin[1], D. A. Shatagin[2], M. S. Anosov[3], A. M. Kuzmishina[4]**
[1, 2, 3, 4] Nizhny Novgorod State Technical University, Nizhny Novgorod, Russian Federation

Разработка цифрового двойника станка с ЧПУ на основе методов машинного обучения***

Ю. Г. Кабалдин[1], Д. А. Шатагин[2], М. С. Аносов[3], А. М. Кузьмишина[4]**
[1, 2, 3, 4] Нижегородский государственный технический университет, г. Нижний Новгород, Российская Федерация

Introduction. It is shown that the digital twin (electronic passport) of a CNC machine is developed as a cyber-physical system. The work objective is to create neural network models to determine the operation of a CNC machine, its performance and dynamic stability under cutting.

Materials and Methods. The development of mathematical models of machining processes using a sensor system and the Industrial Internet of Things is considered. Machine learning methods valid for the implementation of the above tasks are evaluated. A neural network model of dynamic stability of the cutting process is proposed, which enables to optimize the machining process at the stage of work preparation. On the basis of nonlinear dynamics approaches, the attractors of the dynamic cutting system are reconstructed, and their fractal dimensions are determined. Optimal characteristics of the equipment are selected by input parameters and debugging of the planned process based on digital twins.

Research Results. Using machine learning methods allowed us to create and explore neural network models of technological systems for cutting, and the software for their implementation. The possibility of applying decision trees for the problem of diagnosing and classifying malfunctions of CNC machines is shown.

Discussion and Conclusions. In real production, the technology of digital twins enables to optimize processing conditions considering the technical and dynamic state of CNC machines. This provides a highly accurate assessment of the production capacity of the enterprise under the development of the production program. In addition, equipment failures can be identified in real time on the basis of the intelligent analysis of the distributed sensor system data.

Введение. В статье показано, что цифровой двойник (электронный паспорт) станка с ЧПУ разрабатывается как киберфизическая система.

Цель работы — создание нейросетевых моделей, определяющих функционирование станка с ЧПУ, его производительность и динамическую устойчивость при резании.

Материалы и методы. Рассматриваются вопросы создания математических моделей процессов механической обработки с использованием системы сенсоров и промышленного интернета вещей. Оценены методы машинного обучения, подходящие для реализации названных задач. Предложена нейросетевая модель динамической устойчивости процесса резания, позволяющая оптимизировать процесс механической обработки на этапе технологической подготовки производства. На основе подходов нелинейной динамики реконструированы аттракторы динамической системы резания и определены их фрактальные размерности. Выбраны оптимальные характеристики оборудования по входным параметрам и отладке планируемого технологического процесса на основе цифровых двойников.

Результаты исследований. Использование методов машинного обучения позволило создать и исследовать нейросетевые модели технологических систем обработки резанием и программное обеспечение для их реализации. Показана возможность применения деревьев решений для задачи диагностики и классификации неисправностей станков с ЧПУ.

Обсуждение и заключения. В реальном производстве технология цифровых двойников позволяет оптимизировать режимы обработки с учетом технического и динамического состояния станков с ЧПУ. Это обеспечивает высокоточную оценку производственных мощностей предприятия при составлении производственной программы. Кроме того, на основе интеллектуального анализа данных системы распределенных сенсоров можно выявить неисправности оборудования в режиме реального времени.

* The research is done within the frame of independent R&D.
** E-mail: Uru.40@mail.ru, dmitsanych@gmail.com, ansv-maksim@rambler.ru, Foxi-16@mail.ru
*** Работа выполнена в рамках инициативной НИР.

Keywords: cyber-physical system, neural network model, big data, Internet of Things, digital twin.

For citation: Yu.G. Kabaldin, et al. Development of digital twin of CNC unit based on machine learning methods. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 45–55 https://doi.org/10.23947/1992-5980-2019-19-1-45-55

Ключевые слова: киберфизическая система, нейросетевая модель, большие данные, интернет вещей, цифровой двойник.

Образец для цитирования: Разработка цифрового двойника станка с ЧПУ на основе методов машинного обучения / Ю. Г. Кабалдин [и др.] // Вестник Донского гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 45–55. https://doi.org/10.23947/1992-5980-2019-19-1-45-55

Introduction. The single core control platform proposed in [1] serves as the basis for the development of a new generation of the processing equipment control systems, and it also provides improvement of the numerical control systems (CNC) for digital productions. Open CNC systems of machines with long-lived computational resources and high speed processing of a large database (DB), embedded neural processor modules and communication modules with industrial Internet can be such a platform. In the latter case, we are talking about the applicability of cloud technologies for processing large amounts of data on enterprise servers (local networks) and providers. All this will create the basis of intelligent control for a wide range of the CNC process equipment.

Since the nineties of the 20th century [2], the authors of the presented research have been developing software products for intelligent control of the processing equipment. In particular, the criteria were proposed for evaluating the dynamic stability of the cutting process, based on the methods of nonlinear dynamics and fractal analysis of the vibroacoustic emission (VAE) signals. The authors have created a single platform for enhancing various technological equipment with CNC systems. It is implemented through embedding high-performance computing modules and deep learning of artificial neural networks into the CNC systems using CUDA nVidia technologies. In the framework of this work, the CNC process equipment accomplished with sensors is studied. It applies cloud technologies for collecting and processing information using the developed techniques. Such systems will be considered as cyber-physical ones [3].

It is assumed that the platform developed by the authors will become the basis for digitalization at all levels of the enterprise. To this end, it should not only analyze the data of equipment, systems, and devices, but also use the information obtained in this way to reduce time on market launch of new products, to increase production flexibility, product quality and efficiency of production processes.

Digital twin is a new word in the modeling of equipment, processes and production planning [3]. This is a set of mathematical models that describe reliably the processes and interrelationships on a single object and within an entire production enterprise using big data analysis and machine learning.

The leader in the application of digital twins is Siemens [3]. According to its definition, a digital twin is an ensemble of mathematical models. They characterize various equipment conditions, production and business processes over time, in accordance with current production conditions. Neural network circuits hold a specific place among such mathematical models; i.e. a neural network model of a process or product is its digital twin [2].

The base unit of the digital production is a cyber-physical system (CPS) [1]. Its high adaptive and intellectual capabilities are due to the following features:

- associated perception of information,
- ongoing training
- assessment of the current condition and forecasting of the future one.

CPS is able to analyze multidimensional data considering even hidden factors of the real production. Based on this data, it can autonomously solve optimization problems and make right decisions. Therefore, CPS is the key element in creating a digital twin. In this case, the following is meant by the digital twin:

- a set of mathematical models that characterize various equipment conditions, production processes over time, according to the current production conditions;
- detailed 3D assembly object models, reflecting connections and interactions between nodes.

From this point of view, a digital twin can be considered as a digital identity of the CPS, an electronic passport, which records all the data on the materials being processed, the performed technological operations, and testing.

Currently, digital twins are mainly created for commercial purposes. They operate successfully in the oil and gas industry. At the same time, there is no data on the twins of the equipment of machining industries based on CPS in the literature.

Digital twins based on CPS can be obtained through the following:

- traditional analytical approaches based on the mathematical description of physical processes;
- modern statistical methods including machine learning.

Materials and Methods. Machine learning methods used to build statistical models can be divided into three groups: regression analysis models, classification models, and outlier detection models [4–6] (Table 1).

Table 1

Basic Machine Learning Methods

Regression analysis	Classification	Outlier detection
Linear regression	Logistic regression	Support vector method
Bayesian regression	Decision Trees Forest	Principal component analysis
Decision Trees Forest	Decision Jungles	K-means
Decision Trees	Decision Trees	Neural networks
Quantile Fast Forest Regression	Support vector method	Kohonen self-organizing maps
Neural networks	Bayes point machine	
Poisson regression	One-vs-All	
Ordinal regression	Neural networks	

The selection of the machine learning method depends on the size, quality and nature of the data, as well as on the type of tasks to be solved. The existing methods require different computational performance, and they have varying degrees of accuracy. Mostly, they are evaluated by the possibility to achieve accurate approximation of the data and to identify the boundaries in the data space. The method of artificial neural networks (ANN) is the most universal and accurate one, it enables to operate a large amount of data and to build non-linear dependences. Neural networks use a large number of settings, which opens up possibilities for creating highly accurate models of processes operating in the regression analysis, classification, and outlier detection modes.

The basic method for analyzing model quality is cross-validation. It allows for evaluating the statistical quality of the source data through constructing and comparing several models obtained from different training and verification samples. When building models of complex objects and systems, it is required to reduce the data dimensionality and eliminate the effect of multicollinearity of variables. The solution to such problems is possible due to the application of the principal component method, which represents multidimensional data in the form of a limited number of components. Such a generalized approach can be applied to eliminate retraining of models. The basic methods of machine learning are presented in Table 1.

To improve the quality of the model, bagging and boosting algorithms are used. It is about building not one model, but a set of models solving the same task. The result of this work is a kind of integral assessment of the probability of some event. This assessment can be presented as a synergistic result of a set of models, each of which individually performs poorly. Thus, a digital twin can be a set of statistical models that use various combinations of the machine learning methods and have passed through various stages of verification and improvement.

So, in general, a digital twin based on the CPS is a multifactor model of the equipment [1], which includes an ensemble of electronic (i.e., neural network) models. In this case, the following neural networks will be decisive: those of dynamic stability of the cutting process, of cutting forces, and of the machined surface roughness. Cutting forces cause elastic pressing in the “tool - workpiece” system, which specifies errors in shape and size.

To solve this problem, a complex of statistical models was developed using machine learning methods. The resulting models are the basis for the digital twin of the CNC lathes. They provide solving the regression analysis problems to predict the dynamics of the cutting process under various machining conditions, roughness of the machined surface, and cutting forces. In addition, these models help to solve the classification problems for evaluating the machine current condition. A process diagram of creating a digital twin is shown in Fig. 1.

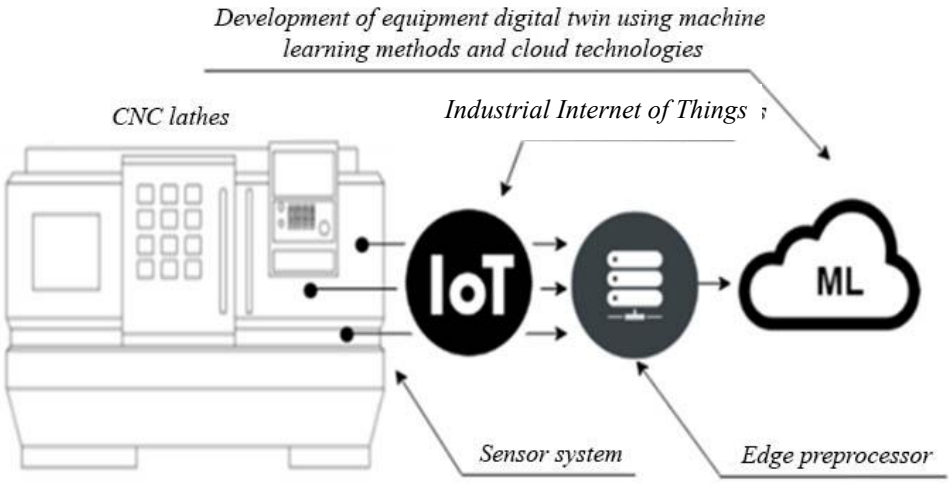


Fig.1. Process diagram of creating a digital twin of CNC machine

Consider in more detail the neural network model of the cutting process dynamics. The training set was obtained on the basis of telemetry data collected under machining through the distributed sensor system of the VAE signal, dynamometer, and Industrial Internet of Things technology (IIoT). The number of sensors, their type and spatial orientation were determined in accordance with the equipment layout. The presence of a large number of sensors with different spatial orientation in the system is explained by the heterogeneity of materials and structures, the signal propagation features, as well as the possibility of restructuring the oscillatory system under operation. Hence, the application of a heterogeneous sensor system could obtain the most complete dynamic picture of the processes in the n-dimensional "state – time" space. The standard TCP/IP protocol and JSON (JavaScript Object Notation) were used as the data transceiving protocol within the IIoT network.

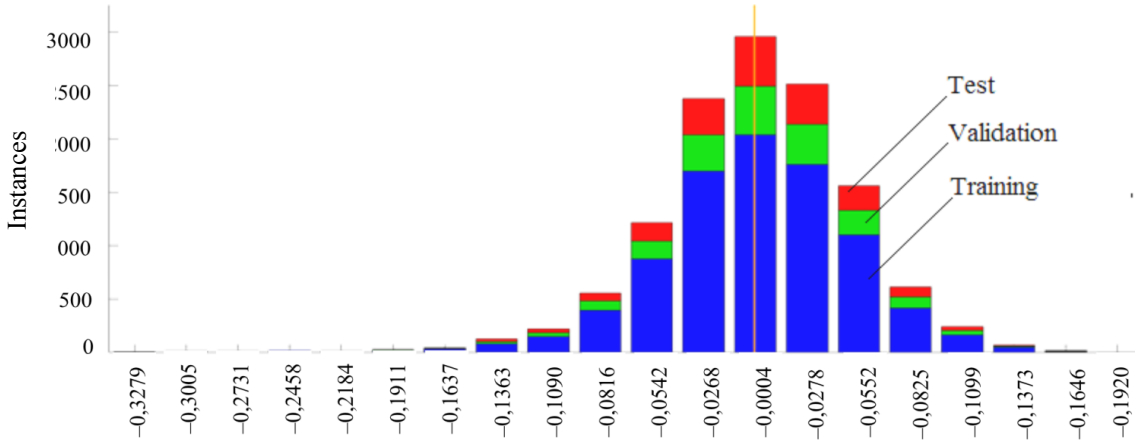
Analog and digital band pass filters, as well as wavelet filters, were used for the signal preprocessing. The application of wavelet filters provided:

- elimination of the noise term impact in the VAE signals;
- identification of periodic and chaotic components based on the entropy factors.

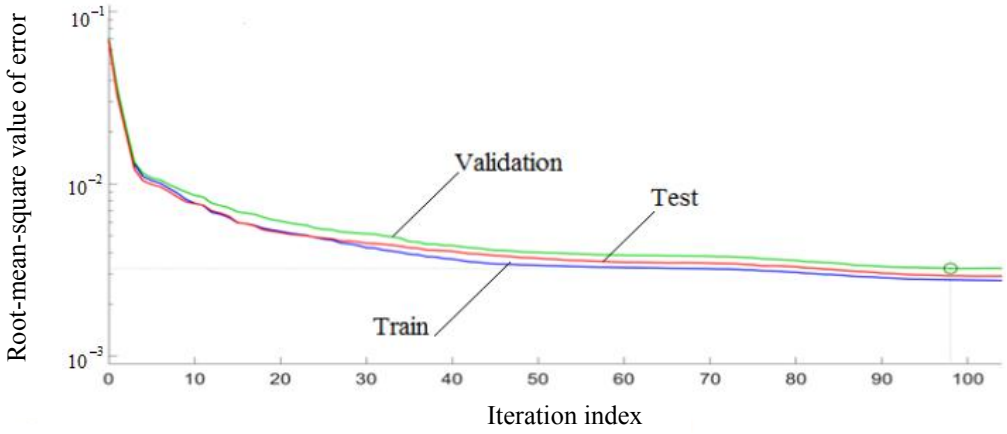
Data pre-processing has been performed on an "edge" device (Edge), which calculates the values of the VAE signal parameters and forms data packets for pushing them to the virtual cloud storage. The development of mathematical models based on the machine learning methods requires structuring and data separation. However, considering the IIoT technology aspects, the construction of relational DB is not always possible; therefore, an approach based on NoSQL technologies was applied. For this, data storage and processing were implemented on a virtual server that simulates the operation of a computing cluster. A special distributed, scalable file system was deployed on this server.

Freeware utilities, libraries and the Hadoop project frameworks (one of the most successful and common big data technologies) were used as a basis. In particular, the Hadoop Common software shell manages the HDFS distributed file system and the HBase database [4]. To perform the distributed computing and processing of large volumes of data, MapReduce was used, which provides automatic parallelization and distribution of tasks on a cluster. nVidia CUDA graphics processors were used to speed up the training of neural network models.

Research Results. A recurrent INS with the sigmoidal neuron activation function was selected to create a statistical model of the cutting dynamics. The data obtained under the industrial equipment operation was used as a training set. During training, the backpropagation algorithm was used. At the end of the learning process, the model obtained was verified on the basis of the mean square derivation values. Fig. 2 shows the monotonous decrease of error. Moreover, its distribution is normal, and it is near zero value, which implies good quality of the model obtained.



a)



b)

Fig. 2. Assessment of model quality obtained: diagram of training error distribution (*a*); dependence of root-mean-square error on learning iteration (*b*)

The ANN obtained (Fig. 3) consists of 17 input neurons that acquire information on processing conditions and on the previous dynamic condition.

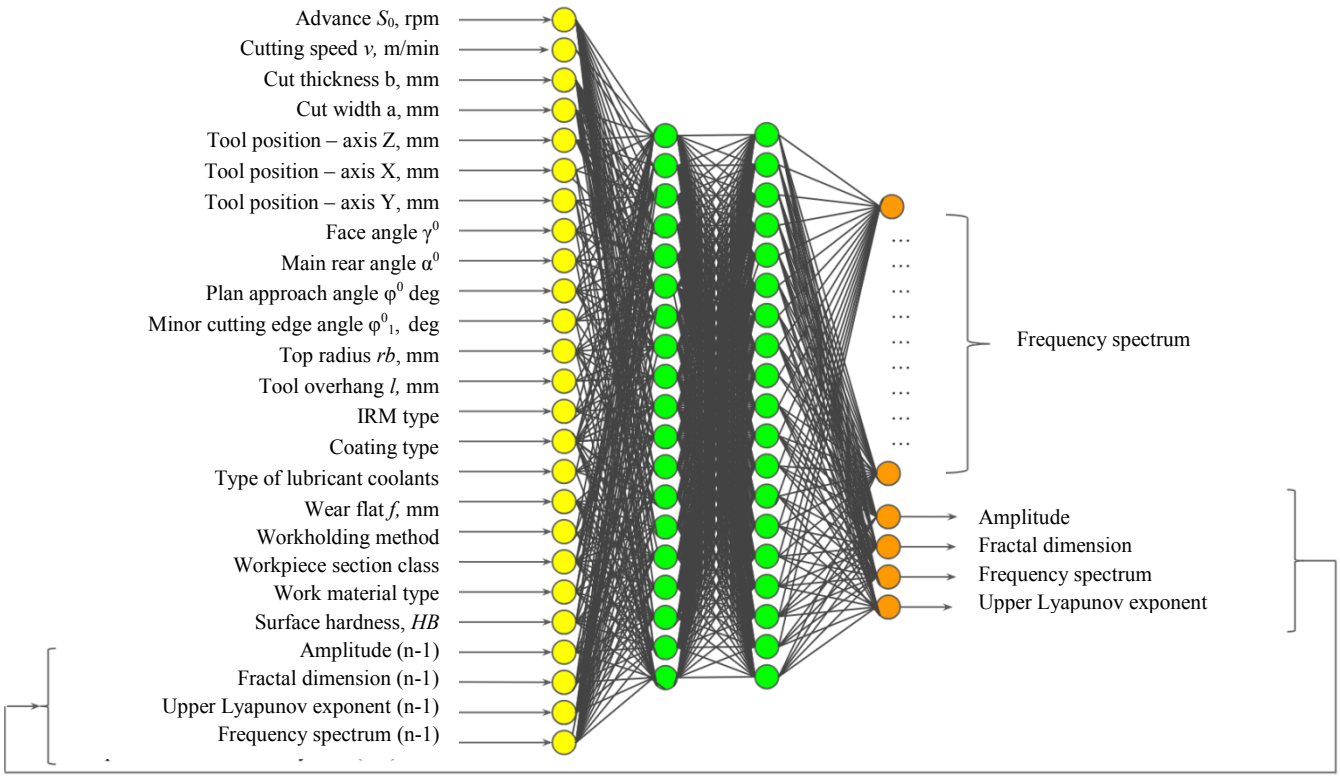
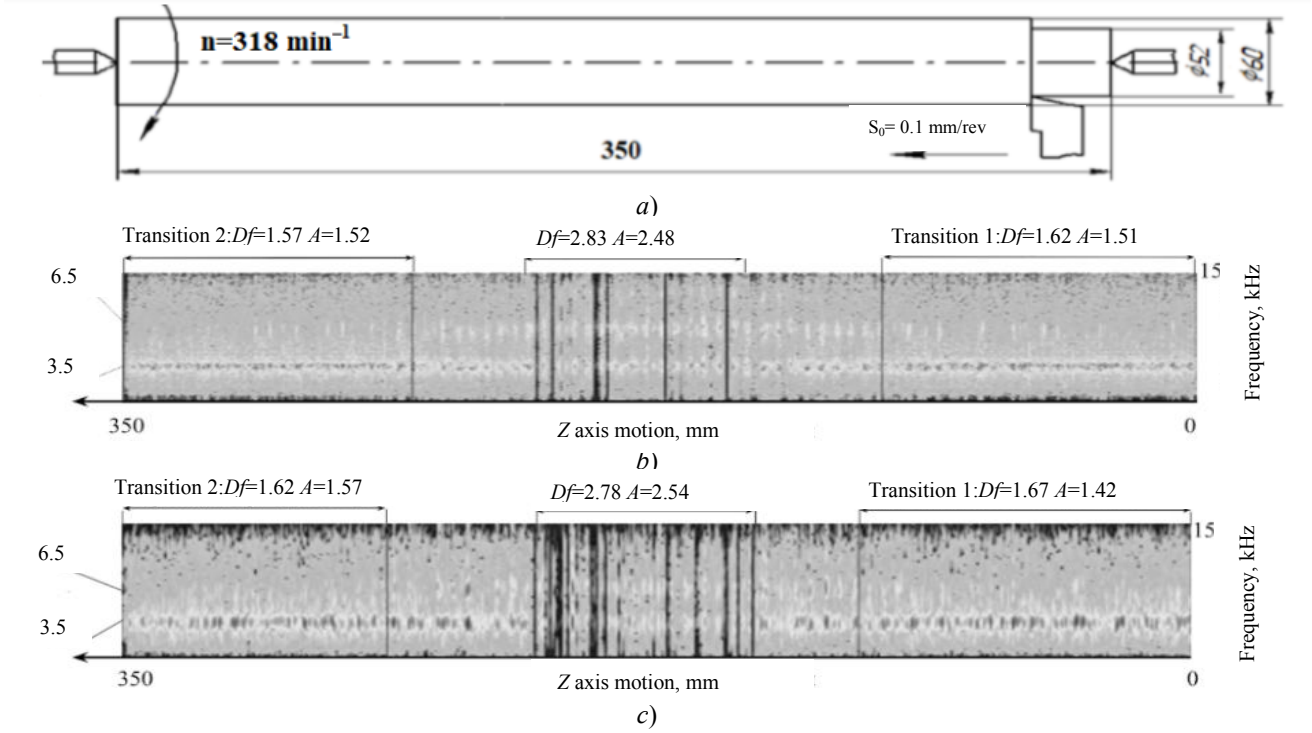


Fig. 3. Neural network model of dynamic stability of cutting process

The output layer of neurons gives information on the dynamic state of the cutting system. The hidden layer is formed by 38 neurons. This ANN is able to simulate the cutting system dynamics for various machining conditions at given instants.

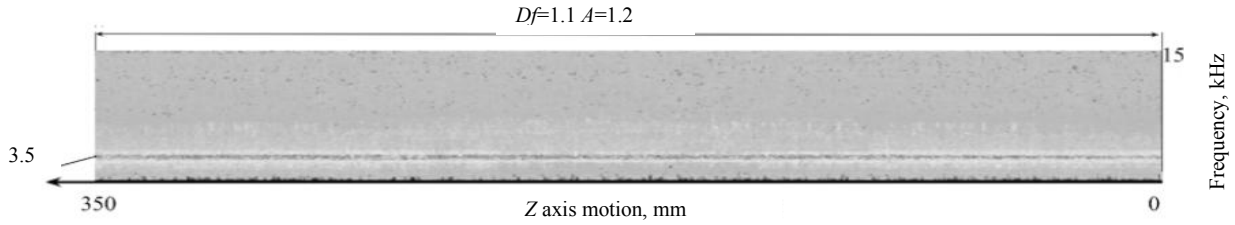
To assess the adequacy of the neural network model, a simulation of the machining process was carried out at a CNC turning center. A shaft of stainless 12X18H10T steel fixed on both sides in the center, as shown in Fig. 4, *a*, was used as the workpiece.



a)

b)

c)



d)

Fig. 4. Neural network simulation results:
treatment scheme (*a*), prototype cutting process (*b*), real-world cutting process (*c*), cutting process after optimization (*d*)

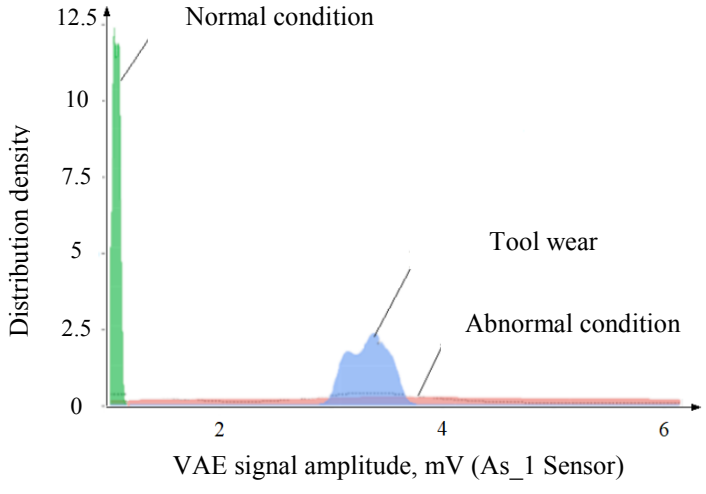
The modes were defined in accordance with the production standards for machining parts of this type on a CNC lathe. The simulation of the tool motion relative to the workpiece was provided by a variable on the ANN input layer responsible for the machine support position along *Z* axis. This variable spacing was 0.01 mm at the distance from 0 to 350 mm. Thus, the dynamic state of the cutting system was obtained for 35 thousand tool positions on *Z* axis. The recorded dynamic states are reflected in the corresponding spectrogram (see Fig. 4, *b*): there are two dominant self-oscillation frequencies (3.5 kHz and 6.5 kHz) and two phase transitions.

To confirm the simulation results, a full-scale experiment was carried out with similar processing conditions and recording of the VAE signal. According to the experimental data obtained, a spectrogram of the machining process was constructed (see Fig. 4, *c*). It also shows two dominant frequencies in the 3.5 kHz and 6.5 kHz areas and two phase transitions. As a result, the simulation error in different machining areas was from 3% to 7%. As can be seen from Fig. 4, the proposed production standards do not allow obtaining a product of a given quality. This may be the cause of defects, and, considering the workpiece runout, – of tool breakage.

The use of the above-mentioned norms leads to an inadequate assessment of the plant capacity and, accordingly, to a poor-quality formation of the production program. Removal of this disadvantage at the design-engineering stage is fraught with additional time and financial expenditures. Therefore, the modes factored in the norms of production are adjusted. For this, the optimal values of the technological conditions vector are calculated using the ANN and the gradient descent method. Objective optimization functions were the VAE signal amplitudes, fractal dimension and information entropy, which should tend to a minimum. The processing speed and the bevel value on the face surface were selected as optimization parameters. Upon completion of the optimization process, new treatment modes were obtained: $V = 90$ m/min; $S = 0.1$ mm/rev; $f\gamma = 0.2$ mm.

These modes were also evaluated according to the results of a full-scale experiment with recording and processing the VAE signal. As can be seen from Fig. 4, *d*, optimization of the processing conditions made it possible to increase the dynamic stability of the cutting process through reducing the amplitude of self-oscillations by half, and the fractal dimension ($Df = 1.1$). Hence, the dynamic machining quality was improved without loss of productivity.

In a similar way, models were obtained for predicting the roughness values of the machined surface and cutting forces. Cutting forces determine the processing accuracy, causing an elastic pressing of the part and tool, and, consequently, the part shape error. These neural network models are multilayered artificial neural feedforward networks (Fig. 5).



a)

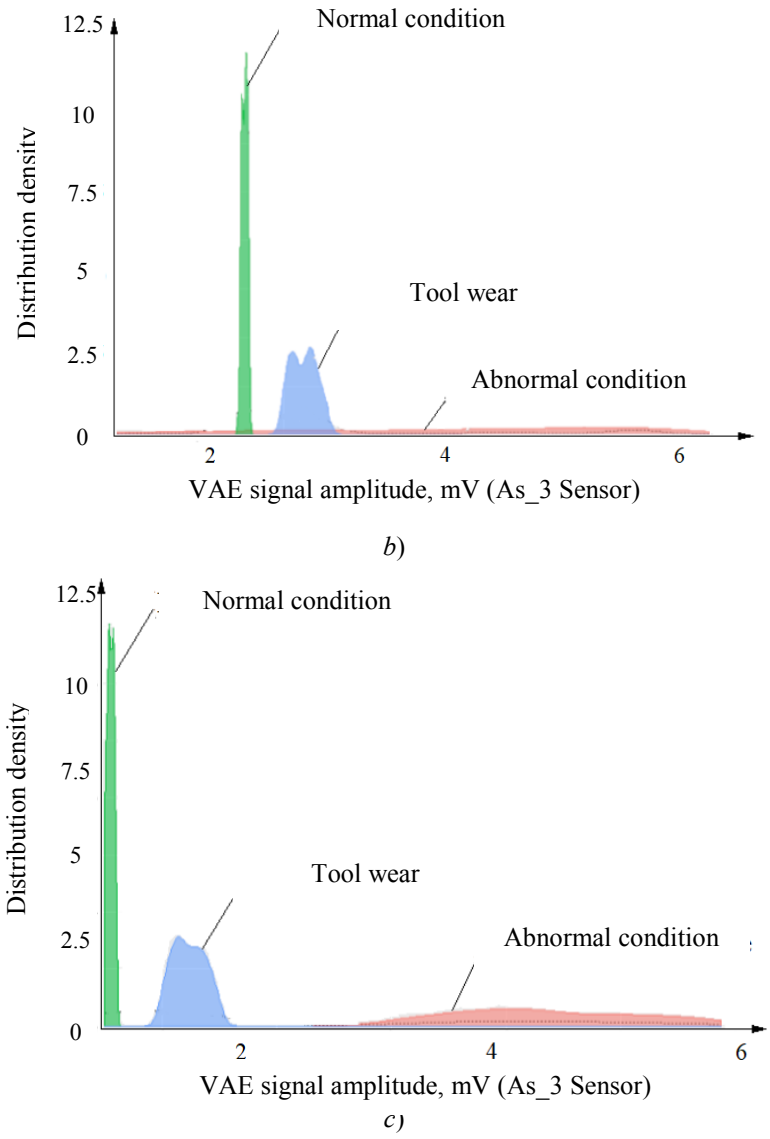


Fig. 5. Distribution of amplitude values of sensor signal: As_1 (*a*), As_2 (*b*), As_3 (*c*)

Neural network models as part of the digital twin are applicable not only under real production conditions, but also when choosing and justifying the purchase of equipment. At this stage, it is possible to select the optimal characteristics of the equipment and debug the planned process on digital twins.

To assess the current condition of the equipment, a classification model was developed using decision trees [5]. Values of the VAE signal amplitude at different points of time from three sensors located on the tool (As_1), at the front spindle support (As_3) and at the rear center (As_2) were selected as attributes for decision-making. Three conditions were used as classes: Normal, Tool wear, and Abnormal. Any condition, when the quality of the treated surface and the operation of the equipment were unsatisfactory, was implied as abnormal. The criterion for tool wear was considered bevel on the rear surface (h = 0.15 mm). Upon receiving the learning sample, an exploratory data analysis was performed.

Fig. 5 shows the correspondence of the distribution of the VAE signal amplitude values of the As_1 sensor to the three groups of conditions. The data obtained from this sensor differentiates well the normal condition and tool wear. However, the values corresponding to the abnormal condition have intersections with data from the other groups. Therefore, the information from the As_1 sensor cannot be used with high accuracy to classify conditions.

Fig. 5, *b,* characterizes the distribution of signal amplitude values from the As_2 sensor. The data obtained have low information value due to two shortcomings:

- intersections of the values of the abnormal condition group with other classes;

- weak differentiation of the normal condition and tool wear (this is due to their close location).

Fig. 5, *c,* shows the distribution of signal amplitude values from the As_3 sensor. The information from this sensor identifies well the values of the abnormal condition, while there are no intersections of the values of all classes of conditions. Thus, the exploratory data analysis enables to draw a conclusion on the possibility of obtaining a statistical model for the classification of the CNC machine condition.

The CART algorithm was used to train the model [3]. It divides the data into specific subsets at different levels. The objective minimization function of informational entropy is used as a separation criterion. Hence, each new level has a more specific content of attributes. A set of all levels obtained is represented as a hierarchical tree structure consisting of nodes and branches, and the branching conforms to the standard logical rules.

From the entire set of attributes, a value of the division condition for the root node is chosen, which better minimizes the informational entropy values. Then, branches are built from each node received, and a similar division into new nodes takes place. Such branching may continue until all values of the training sample are classified. In general, such a tree will represent a complex branched structure that ideally classifies the current sample. However, it may show very low accuracy when working with new data. Therefore, it is necessary to optimize the tree size. It can be carried out using a cross-validation, on the basis of which several candidate trees are built, and, in total, the tree that showed the best result is selected. Alongside with this, pruning is performed until this procedure increases dramatically the error.

Fig. 6 shows the model obtained for classifying the equipment condition based on the readings of the VAE sensors using a decision tree.

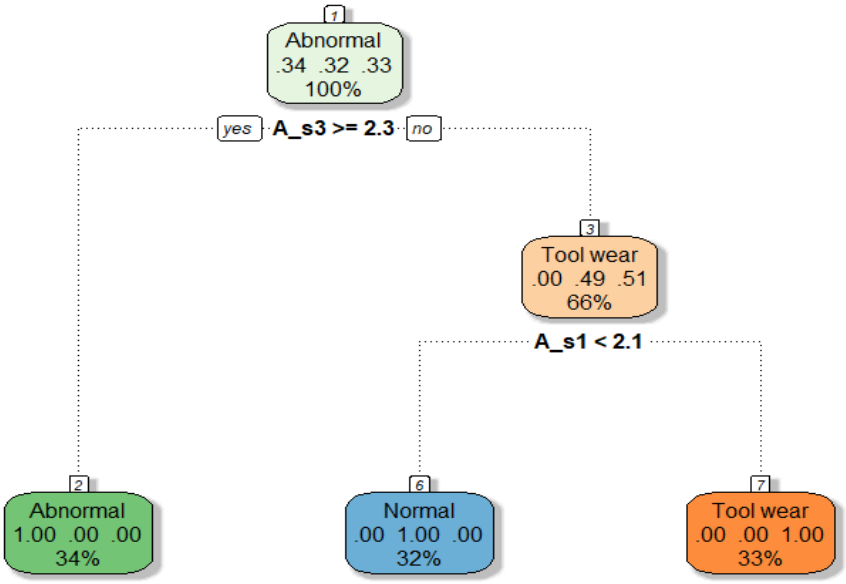


Fig. 6. Decision tree for classifying equipment condition

The first condition chosen is the VAE signal amplitude readings from the As_3 sensor. If the VAE amplitude is greater than or equal to 2.3, then the condition is uniquely abnormal. This judgment is supported by the exploratory data analysis. In the contrary case, the subsequent splitting occurs, and the data from the As_1 sensor determines the classification. If the VAE amplitude of the As_1 sensor is less than 2.1, the condition is considered normal; if it is more, then we are talking about the tool wear.

As can be seen from Fig. 6, the decision-making does not take into account readings of the As_2 sensor. Low information content of the signal from this sensor was also detected under the exploratory analysis.

High accuracy of the determination of the abnormal operation mode based on the analysis of the signal from the As_3 sensor may be related to the place of its installation (front spindle support). Most likely, the dynamic processes in the spindle assembly caused this condition, which also affected the machining quality. On considering the nonlinear processes of oscillations dispersion in the elastic machine system, this information reached other sensors in already heavily distorted form.

At this, the As_1 sensor mounted on the tool specifies its wear with high accuracy. This can be explained by the dominant role of the cutter in the general oscillatory system.

The As_2 sensor mounted on the rear center records information for classifying the condition with the least efficiency. This can be explained by the distance of As_2 from the main dynamic processes.

Hence, the described method can be applied not only to the task of classifying the equipment condition, but to the determination of the optimal number of sensors and their locations. However, it is necessary to consider the possibility of restructuring dynamic processes in the cutting system when changing processing conditions and natural tear wear of machine parts.

Discussion and Conclusions. The use of digital equipment twins under manufacturing planning reveals bottlenecks in the technological operations, improves product quality, and reduces tool breakage and equipment failure risks. Digital twins can optimize the processing modes taking into account the technical and dynamic condition of each production unit. This approach provides a highly accurate assessment of the production facilities under the development of the production program. Besides that, real-time equipment faults are detected through the intelligent data analysis of the distributed sensor system.

References

1. Kabaldin, Yu.G., et al. Iskusstvennyy intellekt i kiberfizicheskie mekhanoobrabatyvayushchie sistemy v tsifrovom proizvodstve. [Artificial intelligence and cyber-physical processing systems in digital manufacturing.] Nizhny Novgorod: NNSTU Publ. House, 2018, 271 p. (in Russian).

2. Kabaldin, Yu.G., et al. Organizatsiya i upravlenie mekhanoobrabatyvayushchim tsifrovym proizvodstvom. [Organization and management of digital machining production.] Vestnik Mashinostroeniya, 2018, no. 11, pp. 19–27 (in Russian).

3. Kabaldin, Yu.G., Bilenko, S.V., Sery, S.V. Upravlenie dinamicheskimi protsessami v tekhnologicheskikh sistemakh mekhanoobrabotki na osnove iskusstvennogo intellekta. [Dynamic processes control in technological machining systems based on artificial intelligence.] Komsomolsk-on-Amur: KoASTU Publ. House, 2003, 201 p. (in Russian).

4.Frankel, A., Larsson, J. Est' sposob luchshe: tsifrovoy dvoynik povysit effektivnost' protsessov konstruktorsko-tekhnologicheskogo proektirovaniya i proizvodstva. [There is a better way: a digital twin will increase the efficiency of design and engineering.] CAD/CAM/CAE Observer, 2016, no. 3, pp. 36–40 (in Russian).

5.Shitikov, V.K., Mastitskiy, S.E. Klassifikatsiya, regressiya i drugie algoritmy Data Mining s ispol'zovaniem R. [Classification, regression and other Data Mining algorithms using R.] Available at: http://www.ievbras.ru/ecostat/Kiril/R/DM/DM_R.pdf (accessed: 14.02.19) (in Russian).

6.White, T. Hadoop. Podrobnoe rukovodstvo. [Hadoop. The Definite Guide.] St.Petersburg: Piter, 2013, 672 p. (in Russian).

Received 21.01.2019
Submitted 22.01.2019
Scheduled in the issue 15.02.2019

Authors:

Kabaldin, Yury G.,
professor of the Technology and Equipment of Mechanical Engineering Department, Nizhny Novgorod State Technical University n.a. R.E. Alekseev, (24, ul. Minina, Nizhny Novgorod, 603950, RF), Dr.Sci. (Eng.), professor,
ORCID: https://orcid.org/0000-0003-4300-6659
Uru.40@mail.ru

Shatagin, Dmitry A.,
senior lecturer of the Technology and Equipment of Mechanical Engineering Department, Nizhny Novgorod State Technical University n.a. R.E. Alekseev, (24, ul. Minina, Nizhny Novgorod, 603950, RF),
ORCID:http://orcid.org/0000-0003-1293-4487,
dmitsanych@gmail.com

Anosov, Maxim S.,
senior lecturer of the Technology and Equipment of Mechanical Engineering Department, Nizhny Novgorod State Technical University n.a. R.E. Alekseev, (24, ul. Minina, Nizhny Novgorod, 603950, RF),
ORCID:http://orcid.org/0000-0001-8150-9332
ansv-maksim@rambler.ru

Kuzmishina, Anastasia M.,
senior lecturer of the Technology and Equipment of Mechanical Engineering Department, Nizhny Novgorod State Technical University n.a. R.E. Alekseev, (24, ul. Minina, Nizhny Novgorod, 603950, RF),
ORCID:https://orcid.org/0000-0001-6966-3417,
Foxi-16@mail.ru

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 56–62. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 56–62. ISSN 1992-5980 eISSN 1992-6006

МАШИНОСТРОЕНИЕ И МАШИНОВЕДЕНИЕ
MACHINE BUILDING AND MACHINE SCIENCE

UDC 62-787 https://doi.org/10.23947/1992-5980-2019-19-1-56-62

Risk-based approach in "personnel-machinery-production environment" system at the facilities running tower cranes*

E. V. Yegelskaya[1], A. A. Korotkiy[2], E. A. Panfilova[3], A. A. Kinzhibalov[4]**
[1,2,3,4] Don State Technical University, Rostov-on-Don, Russian Federation

Риск-ориентированный подход в системе «персонал-механизмы-производственная среда» на объектах, эксплуатирующих башенные краны***

Е. В. Егельская[1], А. А. Короткий[2], Э. А. Панфилова[3], А. А. Кинжибалов[4]**
[1,2,3,4]Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация

Introduction. The paper discusses the applicability of a risk-based approach in the "personnel-machinery-production environment" system at the facilities running tower cranes through the introduction of IT-technologies that provide open communication of the staff, management team of the operating organizations, and the National Supervisory Authority, to reduce the accident rate.

Materials and Methods. An example of a hazardous production facility running tower cranes is given. Materials on the analysis of operational status of tower cranes within the framework of the current legislation in the field of industrial safety are used. The necessity for innovations and transformations, one of whose methods is the risk-based approach allowing for the implementation of all required levels of control, is identified.

Research Results. The stages of introduction of the risk-oriented approach for tower crane operators implemented through IT-technologies using Web-applications on safety management under the tower crane operation based on the hazard analysis and risk assessment in gamut, algorithmically associated with the electronic block key of its local security system, are determined.

Discussion and Conclusions. The application of a risk-based approach in the "personnel-machinery-production environment" system at the facilities running tower cranes through the introduction of IT-technologies will ensure the proper operation of all departments and each participant of the operation, as well as provide Supervisory bodies with an opportunity to access information on the operation for the implementation of control and supervision functions remotely.

Введение. В статье рассматривается возможность применения риск-ориентированного подхода в системе «персонал-механизмы-производственная среда» на объектах, использующих башенные краны посредством внедрения IT-технологий, обеспечивающих открытое взаимодействие персонала, руководящего состава эксплуатирующих организаций и государственных надзорных органов в целях снижения уровня аварийности.

Материалы и методы. Приведен пример опасного производственного объекта, эксплуатирующего башенные краны. Использованы материалы анализа состояния эксплуатации башенных кранов в условиях действующего законодательства в области промышленной безопасности, выявлена необходимость инноваций и преобразований, одним из составляющих методов которых является риск-ориентированный подход, позволяющий реализовать все необходимые уровни контроля.

Результаты исследования. Определены этапы внедрения риск-ориентированного подхода для машинистов башенного крана, реализованного посредством IT-технологий с использованием Web-приложений по управлению безопасностью при эксплуатации башенного крана на основе анализа опасностей и оценки риска в цветовой гамме, алгоритмически связанной с электронным ключом блокировки его локальной системы безопасности.

Обсуждение и заключения. Применение риск-ориентированного подхода в системе «персонал-механизмы-производственная среда» на объектах, эксплуатирующих башенные краны посредством внедрения IT-технологий, позволит обеспечить надежное функционирование всех подразделений и каждого участника работ по эксплуатации, а также предоставит возможность надзорным органам получать доступ к информации по эксплуатации для осуществления контрольно-надзорных функций дистанционно.

* The research is done within the frame of the independent R&D.
** E-mail: korot@novoch.ru, egelskaya72@mail.ru, korotkaya_elvira@mail.ru, kinzha@gmail.com
*** Работа выполнена в рамках инициативной НИР.

Keywords: risk-based approach, National Supervisory Authority, risk assessment, tower cranes, subdivisions of operating enterprise, crane operator.

Ключевые слова: риск-ориентированный подход, государственный надзорный орган, оценка риска, башенные краны, подразделения эксплуатирующего предприятия, машинист крана.

For citation: E.V. Yegelskaya, et al. Risk-based approach in "personnel-machinery-production environment" system at the facilities running tower cranes. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 56–62. https://doi.org/10.23947/1992-5980-2019-19-1-56-62

Образец для цитирования: Егельская, Е. В. Риск-ориентированный подход в системе «персонал-механизмы-производственная среда» на объектах, эксплуатирующих башенные краны / Е. В. Егельская, А. А. Короткий, Э. А. Панфилова, А. А. Кинжибалов // Вестник Донского гос. техн. ун-та. — 2019. — Т. 19, №1. — С. 56–62. https://doi.org/10.23947/1992-5980-2019-19-1-56-62

Introduction. The first decades of the XXI century were marked by serious achievements in science and technology, modernization and improvement of digital technologies, production processes and engineering capacities of the production potential of Russia, which serves to strengthen the economy of the state and to improve the living standards of the citizens.

Against the background of the above achievements and prospects, issues of the operational safety and the professional competence of personnel at the production facilities, as well as supervision of the safe operation of hazardous production facilities (HPF), are challenge problems.

The HPF condition as a whole is estimated by such indicators as accident rate and injury. The level of safety of the HPF is under the state control and is supported by Federal Law No. 116 [1]. Since its adoption in 1997, there has been a transition from the deterministic to probabilistic risk assessment. New techniques for analyzing technically challenging and hazardous objects are introduced [2]; it is supposed to evaluate them by the parameters of risk, survivability and safety (Fig. 1).

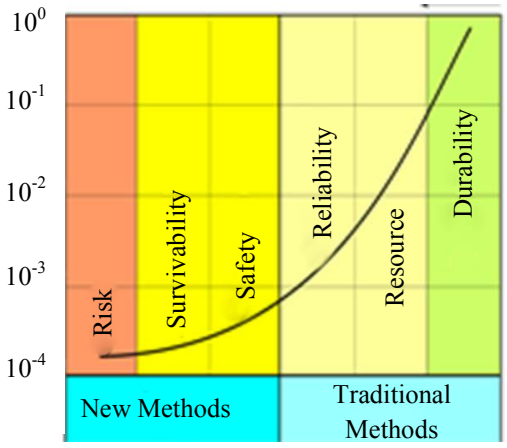


Fig. 1. History of development of methods for calculating HPF

Research Results. Consider HPF running lifting devices, in particular, tower cranes. At present, such an object is ranged in the fourth hazard class (low hazard). In this regard, it is exempted from periodic inspections by the federal executive authority in the field of industrial safety, and the responsibility for complying with the industrial safety requirements rests in full with the owner or operator. The discussion of adequacy of the internal control over compliance with industrial safety requirements without the participation of supervisory authorities was held at different levels.

The recent rise in accidents and injuries at the facilities running tower cranes, has led to the tightening of administrative measures, namely:

- Order of Rostekhnadzor No. 146 of April 12, 2016, changes the Federal norms and rules for lifting devices under the operation of tower cranes. Article No. 141 assigned for the participation of an inspector of Rostechnadzor in the commission for decision-making on the possibility of bringing tower cranes in operation[3];
- Instruction of the Government of the Russian Federation No. АХ-П9-682 of February 8, 2017, "On the organization and conduct of unscheduled inspections of the enterprises running tower cranes from 2017 to 2019" regulates the procedure of random checks.

Additionally, the draft law proposes to classify enterprises using tower cranes to the third hazard class. This will allow for scheduled oversight activities against these facilities every three years. The operating organizations themselves also consider this measure appropriate.

The need for transformation and innovation in the field of industrial safety has come to a head. Significant measures are as follows: the introduction of a risk-based approach during the organization and implementation of all types of state control (supervision) from 2018 (introduced by Federal Law No. 246-FZ of July 13, 2015) [4]; using the information and communication technologies to perform the duties in an electronic form during the implementation of state control (supervision). The regulating document is the Decree of the Government of the Russian Federation No. 482 of April 21, 2018, "On the State Information System of "Typical Cloud Solution to Automation of Control (supervisory) Activities" [5].

The following tasks of the state policy in the field of industrial safety are formulated, among other documents, by the Decree of President of the Russian Federation "On the Fundamentals of the State Policy of the Russian Federation in the Field of Industrial Safety for the Period to 2025 and for the Future" of May 06, 2018: introduction of a risk-oriented approach under organizing the federal state control (supervision) in the field of industrial safety; annulment of obsolete, redundant and duplicate industrial safety requirements [6].

The operation of tower cranes is carried out, as a rule, at construction sites geographically distant from the workplaces of managers and experts, which allows the crane operator to make autonomous decision on the work commencement and performance. Periodic on-site visits of specialists responsible for the operable state do not allow for an objective situation assessment, which leads to numerous emergencies and injuries.

The functions of a tower crane driver, indicated in the production instruction, presuppose a certain algorithm of his actions before work starting, related to the determination of the technical readiness of the crane for safe operation.

The statistics of most accidents with tower cranes and the results of their investigations reflect the failed state of cranes (nonoperable condition of separate units), which is not recorded in the duty log during the inspection of the crane operator. Consequently, a specialist responsible for the operational condition of the lifting device has no information of it because of non-compliance with the instructions by a crane operator. The enterprise management learns about the incidents that have already happened as fait accompli — during the investigation of the causes of accidents (injuries).

Multi-level control and interaction of all participants engaged in the industrial safety under the operation of tower cranes (Fig. 2) can be considered the only solution to safety problems.

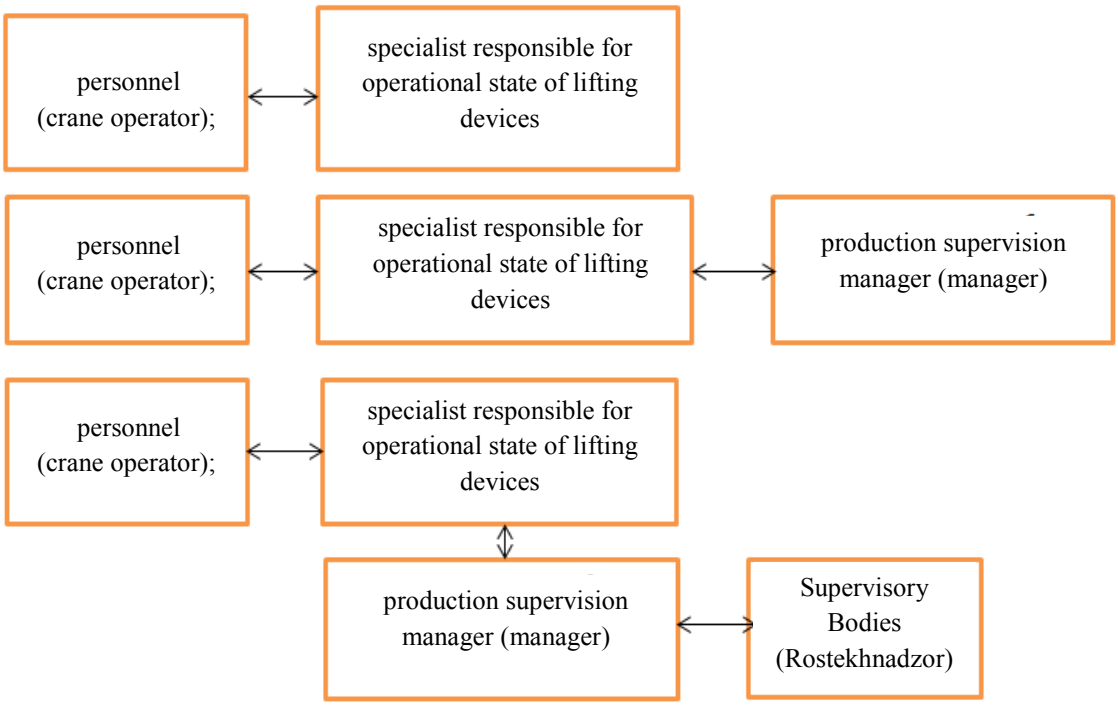


Fig. 2. Personnel and managers interaction algorithm

The application of a risk-based approach provides the methods of organization and implementation of all levels of control, including measures of preventive control of the compliance with safety requirements concerning facilities operating tower cranes. This approach can be implemented through IT technology using Web applications. These applications help to manage safety when operating a technical facility based on hazard analysis and risk assessment in color gamut. An algorithmic connection with the electronic block key of the local safety system is made. Thus, the compliance with the standards established by law in the field of HPF industrial safety is ensured through the remote technologies.

One of the steps of introducing a risk-based approach will be the implementation of electronic checklists for meeting industrial safety requirements through individual mobile devices. At this, information on the fulfillment of the security requirements is sent in real-time via the Internet to individual mobile devices to the manager, specialists and service personnel (Fig. 3).

Fig. 3. Electronic checklists in the form of mobile applications

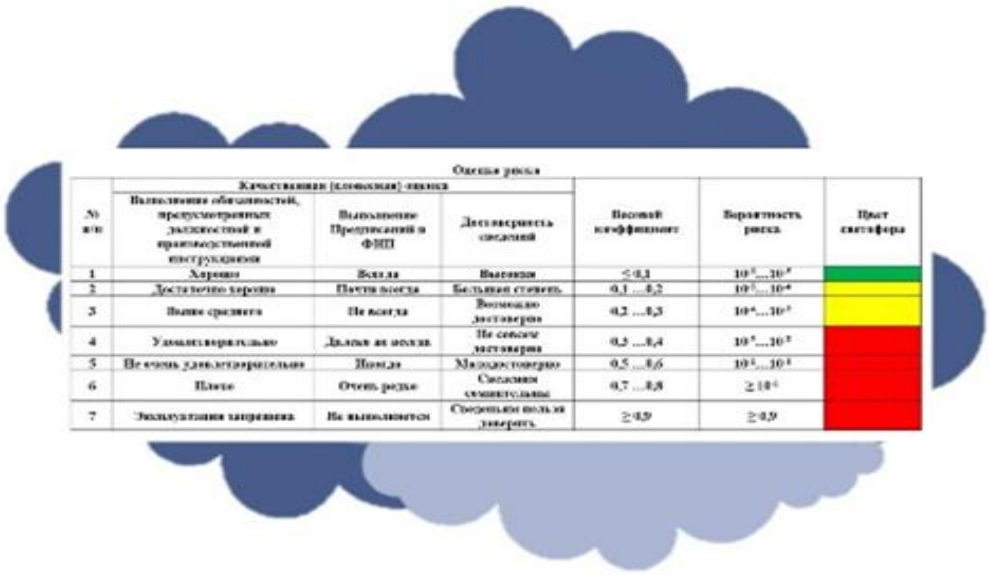
The next step involves monitoring and confirming the fulfillment of security requirements, determined by the duty and working instructions through individual mobile devices, where prompts in the form of tables on the controlled action and its scheduled time should popup (Fig. 4.)

Fig. 4. Mobile screen with popup checklists

The next step includes processing of the language data from managers, specialists and service personnel, as well as digital information obtained from the coordinate protection of the crane in the cloud space using a risk assessment algorithm based on the theory of fuzzy sets of L. A. Zade.

The software to assess risk under the crane operation is displayed as a color gamut, for example, of three colors according to the scheme:

- red - "work prohibited";
- yellow - "work permissive, actions required";
- green - "work permissive ".



№ п/п	Качественная (словесная) оценка			Весовой коэффициент	Вероятность риска	Цвет светофора
	Выполнение обязанностей, предусмотренных должностной и производственной инструкциями	Выполнение Предписаний и ФИП	Достоверность сведений			
1	Хорошо	Всегда	Высокая	<0,1	$10^{-5}...10^{-4}$	
2	Достаточно хорошо	Почти всегда	Большая степень	0,1 ...0,2	$10^{-4}...10^{-3}$	
3	Выше среднего	Не всегда	Возможно достоверно	0,2 ...0,3	$10^{-3}...10^{-2}$	
4	Удовлетворительно	Далеко не всегда	Не совсем достоверно	0,3 ...0,4	$10^{-2}...10^{-1}$	
5	Не очень удовлетворительно	Никогда	Малодостоверно	0,5 ...0,6	$10^{-1}...10^{0}$	
6	Плохо	Очень редко	Сведения сомнительны	0,7 ...0,8	$\geq 10^{0}$	
7	Эксплуатация запрещена	Не выполняется	Сведения нельзя проверить	≥0,9	≥0,9	

Fig. 5. Risk assessment under crane operation in color gamut

Crucially, the access to information on existing violations should be implemented using an appropriate login and password, which is confidential information. This will enable to establish causes of the unfavorable cases (accident, incident), as well as a concrete employee who has failed to fulfill his duty or working instructions (Fig. 6).

Fig. 6. Steps of displaying risk-based approach algorithm on mobile devices

One of the stages of the implementation of the risk-based approach is to monitor the crane operation by the video surveillance signal through a remote CCTV camera – a “black-box” recorder. Its signal is transmitted via the Internet to the cloud space where the video information is stored (Fig. 7).

Fig. 7. Displaying panorama of object on mobile device using video surveillance

Conclusion. The application of the risk-based approach in the "personnel-machinery-production environment" system at the facilities running tower cranes through the introduction of IT-technologies will ensure reliable operation of all departments and each participant of the operation. This will provide supervisors with the ability to remotely access to operating information for the implementation of the control and supervisory functions. Reducing the number of accidents and injuries during the tower cranes operation can be achieved only under the condition of open interaction of the personnel, management and supervisory authorities.

References

1. O promyshlennoy bezopasnosti opasnykh proizvodstvennykh ob"ektov: federal'nyy zakon ot 21.07.1997 № 116-FZ (red. ot 07.03.2017) (s izm. i dop., vstup. v silu s 25.03.2017). [On industrial safety of hazardous production facilities: Federal Law no. 116-FZ, of July 21, 1997 (as amended on March 7, 2017) (valid since March 25, 2017).] Electronic fund of legal and regulatory technical documentation. Available at: http://docs.cntd.ru/document/9046058 (accessed: 15.02.2019) (in Russian).

2. Makhutov, N.A., Gadenin, M.M. Struktura osnovnykh raschetov dlya opredeleniya iskhodnogo i ostatochnogo resursa bezopasnoy ekspluatatsii. [Structure of the main calculations for determination of the initial and remaining safe service.] Safety and Emergencies Problems, 2018, no. 2, pp. 21–33 (in Russian).

3. O vnesenii izmeneniy v Federal'nye normy i pravila v oblasti promyshlennoy bezopasnosti «Pravila bezopasnosti opasnykh proizvodstvennykh ob"ektov, na kotorykh ispol'zuyutsya pod"emnye sooruzheniya» utverzhdennye prikazom Federal'noy sluzhby po ekologicheskomu, tekhnologicheskomu i atomnomu nadzoru ot 12 noyabrya 2013 goda № 533. [On amendments to the Federal Norms and Rules in Industrial Safety: "Safety Rules for Hazardous Production Facilities Using Lifting Devices" approved by Order of the Federal Service for Environmental, Technological and Nuclear Supervision, of November 12, 2013, no. 533.] Electronic fund of legal and regulatory technical documentation. Available at: http://docs.cntd.ru/document/420351983 (accessed: 15.02.2019) (in Russian).

4. O vnesenii izmeneniy v Federal'nyy zakon «O zashchite prav yuridicheskikh lits i individual'nykh predprinimateley pri osushchestvlenii gosudarstvennogo kontrolya (nadzora) i munitsipal'nogo kontrolya». [On amendments to the Federal Law "On protection of the rights of legal entities and individual entrepreneurs during state control (supervision) and municipal control".] Electronic fund of legal and regulatory technical documentation. Available at: http://docs.cntd.ru/document/552008582 (accessed: 17.02.2019) (in Russian).

5. O gosudarstvennoy informatsionnoy sisteme "Tipovoe oblachnoe reshenie po avtomatizatsii kontrol'noy (nadzornoy) deyatel'nosti" (s izmeneniyami na 20 noyabrya 2018 goda). [On state information system "Typical cloud solution to the automation of control (supervisory) activities" (as amended on November 20, 2018).] Electronic fund of legal and regulatory technical documentation. Available at: http://docs.cntd.ru/document/557244991 (accessed: 18.02.2019) (in Russian).

6. Ob Osnovakh gosudarstvennoy politiki Rossiyskoy Federatsii v oblasti promyshlennoy bezopasnosti na period do 2025 goda i dal'neyshuyu perspektivu: ukaz Prezidenta Rossiyskoy Federatsii ot 6 maya 2018 g. №198. [On Fundamentals of the state policy of the Russian Federation in the field of industrial safety for the period up to 2025 and for the long run: Decree of President of the Russian Federation of May 6, 2018, no. 198.] Electronic fund of legal and

regulatory technical documentation. Available at: http://docs.cntd.ru/document/557306107 (accessed: 18.02.2019) (in Russian).

7. Korotkiy, A.A., Kinzhibalov, A.V., Kinzhibalov, A.A. Monitoring proizvodstvennogo kontrolya, avariynosti i opasnosti OPO IV klassa pri ekspluatatsii bashennykh kranov. [Monitoring of industrial control, danger and emergency at the operation of tower cranes.] Monitoring. Science and Technologies, 2017, no. 4(33), pp. 80–85 (in Russian).

8. Korotkiy, A.A., Panfilov, A.V., Kinzhibalov, A.V., Kinzhibalov, A.A. Sovershenstvovanie sovremennykh sistem bezopasnosti bashennykh kranov na osnove tsifrovykh tekhnologiy v usloviyakh risk-orientirovannogo nadzora. [Improvement of modern security systems of tower cranes based on digital technologies in conditions of risk-based supervision.] Science and Business: Ways of Development, 2018, no. 7, pp. 46–54 (in Russian).

9. Yegelskaya, E.V. Otsenka riska chelovecheskogo faktora v sisteme «personal- pod"emnye mekhanizmy-proizvodstvennaya sreda» na predpriyatiyakh mashinostroeniya: avtoref. dis. … kand. tekhn. nauk. / [Human factor risk assessment in the "personnel-lifting gear-production environment" system at mechanical engineering enterprises: Cand.Sci. (Eng.), diss., author's abstract.] Rostov-on-Don, 2015, 20 p. (in Russian).

Received 10.10.2018
Submitted 12.10.2018
Scheduled in the issue 15.01.2019

Authors:

Yegelskaya, Elena V.,
associate professor of the Transport Systems Operation and Logistics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Eng.),
ORCID: http://orcid.org/0000-0003-3864-9254
egelskaya72@mail.ru

Korotkiy, Anatoly A.,
Head of the Transport Systems Operation and Logistics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Eng.), professor,
ORCID: http://orcid.org/0000-0001-9446-4911
korot@novoch.ru

Panfilova, Elvira A.,
associate professor of the Transport Systems Operation and Logistics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Phylosophy),
ORCID: http://orcid.org/0000-0002-8485-5983
korotkaya_elvira@mail.ru

Kinzhibalov, Alexander A.,
postgraduate of the Transport Systems Operation and Logistics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: http://orcid.org/0000-0002-1742-9407
kinzha@gmail.com

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 63–73. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 63–73. ISSN 1992-5980 eISSN 1992-6006

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

UDC 004.932.72'1 https://doi.org/10.23947/1992-5980-2019-19-1-63-73

Deep convolution neural network model in problem of crack segmentation on asphalt images*

B. V. Sobol[1], A. N. Soloviev[2], P. V. Vasiliev[3], L. A. Podkolzina[4]**
[1, 2, 3, 4] Don State Technical University, Rostov-on-Don, Russian Federation

Модель глубокой сверточной нейронной сети в задаче сегментации трещин на изображениях асфальта †***

Б. В. Соболь[1], А. Н. Соловьев[2], П. В. Васильев[3], Л. А. Подколзина[4]**
[1, 2, 3, 4] Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация

Introduction. Early defect illumination (cracks, chips, etc.) in the high traffic load sections enables to reduce the risk under emergency conditions. Various photographic and video monitoring techniques are used in the pavement managing system. Manual evaluation and analysis of the data obtained may take unacceptably long time. Thus, it is necessary to improve the conditional assessment schemes of the monitor objects through the autovision.

Materials and Methods. The authors have proposed a model of a deep convolution neural network for identifying defects on the road pavement images. The model is implemented as an optimized version of the most popular, at this time, fully convolution neural networks (FCNN). The teaching selection design and a two-stage network learning process considering the specifics of the problem being solved are shown. Keras and TensorFlow frameworks were used for the software implementation of the proposed architecture.

Research Results. The application of the proposed architecture is effective even under the conditions of a limited amount of the source data. Fine precision is observed. The model can be used in various segmentation tasks. According to the metrics, FCNN shows the following defect identification results: IoU - 0.3488, Dice - 0.7381.

Discussion and Conclusions. The results can be used in the monitoring, modeling and forecasting process of the road pavement wear.

Keywords: artificial neural networks, defect identification, segmentation, road pavement, cracks, IoU, Dice.

Введение. Своевременное устранение дефектов (трещин, сколов и пр.) на участках повышенной нагрузки дорожного полотна позволяет снизить риск возникновения аварийных ситуаций. В настоящее время для контроля состояния дорожного покрытия применяются различные методы фото- и видеонаблюдения. Оценка и анализ полученных данных в ручном режиме могут занять недопустимо много времени. Таким образом, необходимо совершенствовать процедуры осмотра и оценки состояния объектов контроля с помощью технического зрения.

Материалы и методы. Авторами предложена модель глубокой сверточной нейронной сети для идентификации дефектов на изображениях дорожного покрытия. Модель реализована как оптимизированный вариант наиболее популярных на данный момент полностью сверточных нейронных сетей (FCNN). Показано построение обучающей выборки и двухэтапный процесс обучения сети с учетом специфики решаемой задачи. Для программной реализации предложенной архитектуры использовались фреймворки Keras и TensorFlow.

Результаты исследования. Применение предложенной архитектуры эффективно даже в условиях ограниченного объема исходных данных. Отмечена высокая степень повторяемости результатов.

Модель может быть использована в различных задачах сегментации. Согласно метрикам, FCNN показывает следующие результаты идентификации дефектов: IoU — 0,3488, Dice — 0,7381.

Обсуждение и заключения. Полученные результаты могут быть использованы в процессе мониторинга, моделирования и прогнозирования процессов износа дорожных покрытий.

Ключевые слова: искусственные нейронные сети, идентификация дефектов, сегментация, дорожное покрытие, трещины, IoU, Dice.

*The research is done on RFFI grants nos. 19-08-00074, 18-31-00024, 16-01-00390.
**E-mail: b.sobol@mail.ru, solovievarc@gmail.com, lyftzigen@mail.ru, podkolzinalu@gmail.com
*** Работа выполнена при поддержке грантов РФФИ 19-08-00074, 18-31-00024, 16-01-00390.

For citation: B.V. Sobol, et al. Deep convolution neural network model in problem of crack segmentation on asphalt images. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 63–73 https://doi.org/10.23947/1992-5980-2019-19-1-63-73

***Образец для цитирования*:** Модель глубокой сверточной нейронной сети в задаче сегментации трещин на изображениях асфальта / Б. В. Соболь [и др.] // Вестник Донского гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 63–73. https://doi.org/10.23947/1992-5980-2019-19-1-63-73

Introduction. The road pavement wear requires regular monitoring. Efficient monitoring strategies allow for the timely detection of problem areas. This approach improves significantly the efficiency of road maintenance, reduces maintenance costs and provides continuous operation. Technologies for identifying critical features of the condition of the road surface have developed from the manual photofixation methods to the high-speed digital technology [1].

Russia is one of the five countries with the longest automobile roads. To provide photo- and video monitoring of such a large-scale infrastructure special systems are required. They should be reliable and easy-to-use on-line systems. It seems clear that in this case it is not about solutions that involve the data analysis in manual mode. Such an approach is unacceptable due to the considerable time spent on information processing and low quality of the analysis.

The authors of the paper propose a new technological solution in the field of machine learning. Its implementation enables to automate the process of road surface quality assessment. To this end, the convolution neural network is trained on the data marked up manually. The system learns to recognize and evaluate the basic failure modes of the monitor objects by this means.

Research Digest. Many papers describe the improvement of algorithms for detecting defects in the structural components and infrastructure facilities. To solve this problem, computer vision capabilities are widely used. Their continuous improvement is supported by the development of sensing technologies, hardware and software. However, it should be recognized that at present, computer vision is used sparingly. This is due to many factors, including:

- heterogeneity of defects,
- variety of types of surfaces,
- complexity of the background,
- junctions.

The authors of a number of publications investigate automated methods for detecting cracks in images and propose their own solutions [2–9]. Some works consider the specifics of monitoring objects of road transportation infrastructure [10, 11], as well as bridges and structural systems [12, 13].

Until recently, mostly manual monitoring techniques were used to solve these problems, such as:

- morphological operations [13],
- analysis of geometric features [6],
- application of Gabor filters [14],
- wavelet transforms [15],
- building of histograms of oriented gradients (HOG) [16],
- texture analysis,
- machine learning [4].

However, currently, the listed tool is used more and more seldom. It is replaced by a global spread of neural network technologies and machine learning supported by the computational capabilities of graphic processors.

A convolution neural network (CNN) is a multi-layered artificial neural network architecture designed specifically for imaging [17]. In this case, a subsampling layer (pooling layer) allows local capable fields to be realized through convolution layers and invariance with respect to small geometric strains.

This architecture demonstrates outstanding results in solving the following recognition problems:

- handwritten numbers [18],
- house numbers based on the Google Street View house number (SVHN) dataset [19],
- road signs [20].

The computation power rise of graphics processors allows for the use of deeper architectures of machine learning models [21]. Now it is possible to avoid retraining [22]. This is facilitated by the development of such modern techniques as data augmentation, regularization, etc.

Improving convolution neural networks opens up the possibility of more efficient study and generalization of the features of images (for example, image classification [23], object search [24], vehicle detection [25]).

The flexibility and prospects of deep learning for the tasks of automatic detection of pavement cracks is shown in [26, 27].

In [28], the application of neural networks for the automatic detection and classification of cracks in asphalt is considered. The authors propose to use the universe mean and dispersion of the grayscale values. With consideration to these indicators, the image is divided into fragments, and then each cell is classified as a crack. Motivation for using the fall weight deflectometers (FWD) to assess asphalt cracks was shown. In 98% of cases, the system detects effectively a crack in the image.

In [29], the use of a neural network for the detection of defects was investigated. The advantages of the method of clustering pixels as objects were clarified. It enables to increase identification accuracy and reduce noise.

In [30], the authors used a deep learning architecture that includes the VGG-16 model. It was previously trained to identify features that allow distinguishing between classes of images. The model demonstrated excellent recognition quality even under working with images from the areas unknown to it.CNN VGG-16 was used as a deep feature generator of pavement images. The authors taught only the last layer of the classifier. They experimented with different models of machine learning, and showed their strengths and weaknesses.

In [31], the use of CNN in an applied robotics problem is shown. It is referred to the autonomous detection and assessment of cracks and damages in the sewer pipe. CNN filters the data and localizes the cracks, which provides obtaining a characteristic of their geometrics.

The objective of [32] is to automate the sequential detection of chips and to give numeric representation of damage in metro networks. For that, an integrated model that implements a hybrid algorithm and an interactive 3D representation is created. Chipping depth prediction is supported by regression analysis.

The paper [33] provides an overview and assessment of promising approaches that automatically detect cracks and corrosion in the civil infrastructure systems.

In [34], an effective CNN-based architecture for detecting pavement cracks on a three-dimensional asphalt surface is described. The CrackNet architecture provides high precision of the data processing through an ingenious method of representing the road surface geometry. CrackNet consists of five layers and includes more than a million of trainable parameters. The experiments using 200 test 3D-images have shown that the CrackNet accuracy can reach 90.13%.

Offered Method. To identify defects in pavement images, it is necessary to determine what a defect is and what is not. In other words, the image segmentation should be carried out, and the appropriate classes should be identified. Recently, this type of problem has been effectively solved through purposely developed architectures of the convolution neural networks, such as SegNet [35] and U-Net [36].

The specifics of pavement images involve a small range of gray shades and a minor difference between the background and the target object. Also, the task is complicated by noise, defects and the occurrence of foreign objects in the image.

Various data sets are used to learn neural networks [7, 37]. These sets include original pavement images and their corresponding mask images with or without defects. Images with defects on the road surface are specific, so the authors offer their own simplified model of a deep convolution neural network. For image segmentation, a fully convolution neural network (FCNN) [38] with an encoder-decoder structure is proposed. The image of the road surface is fed to the system input, and the output is a binary image. A segmented image showing the presence or absence of defects is resulted.

Architecture of Deep Convolution Neural Network. Fig. 1 shows the architecture of the proposed deep convolution neural network.

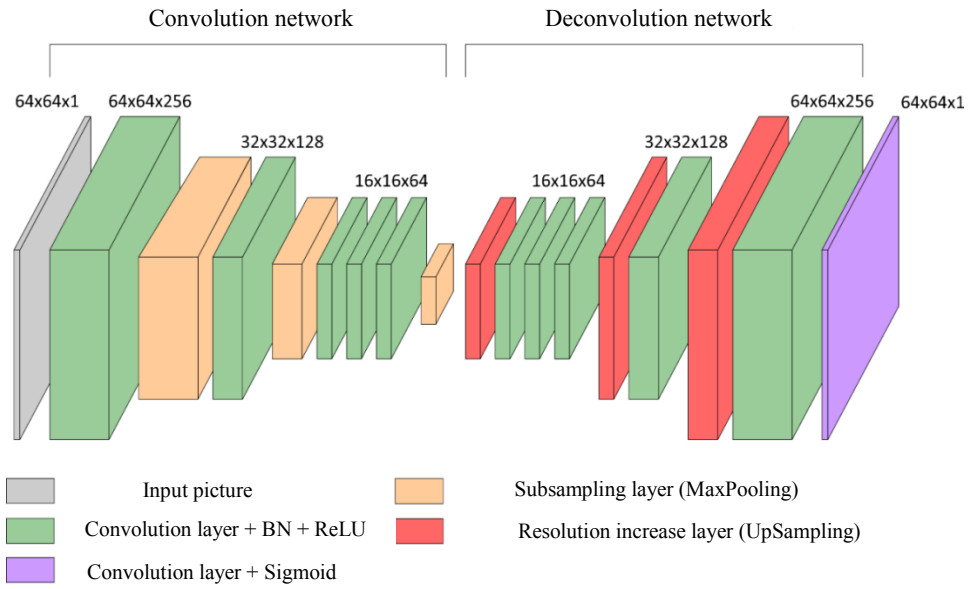


Fig. 1. Architecture of proposed network

The neural network consists of two parts, convolution and deconvolution. The convolution part converts the input image into a multi-dimensional representation of features. In other words, it performs the function of extracting features. The deconvolution network serves as a generator, which creates a segmented image based on the characteristics obtained from the convolution network. The last convolution network layer with a sigmoid activation function generates a segmented image — a map of probabilities of the defect of the same size as that in the input picture.

The first part of the network consists of five convolution layers with filter sets (256, 128, 64, 64, 64). The "batch normalization" (BN) tool [39] is used. ReLU, rectified linear unit, is used as the activation functions. Then, of subsampling (pooling) layers with a 2×2 window follow. In passing through this layer, the image is decreased by half. The second part of the network is a mirror image of the first. The image size needs to be restored to the original one and to form a probability map based on the input image features. To this end, upsampling layers are used in combination with the convolution layers. The proposed neural network has 10 convolution layers and 929665 trainable parameters.

Dataset Generation. The CrackForest data set [7] is used for training the constructed model. Its augmentation (artificial increase of the dataset) is carried out, since the training and operation of the neural network is based on the path-based approach, which involves the use of randomly clipping elements of the original images.

So, the data set consists of 117 images. It is divided into training, test and validation samples. Fragments of 64×64 are randomly selected for each image of the training and test samples. The studies have shown that gamma correction of images improves the neural network quality within the framework of the task. Each image fragment undergoes rotation, reflection, and deformation. The optimal ratio of fragments with and without defect was established as 95% to 5%. At this, defects that occupy at least 5% of the image area are considered. The sample size affects the learning process and the network quality. The optimal ratio was established as follows: 15 200 fragments of the training sample and 3 968 – of the test one. Fig. 2 shows the images and the corresponding masks used for training the neural network.

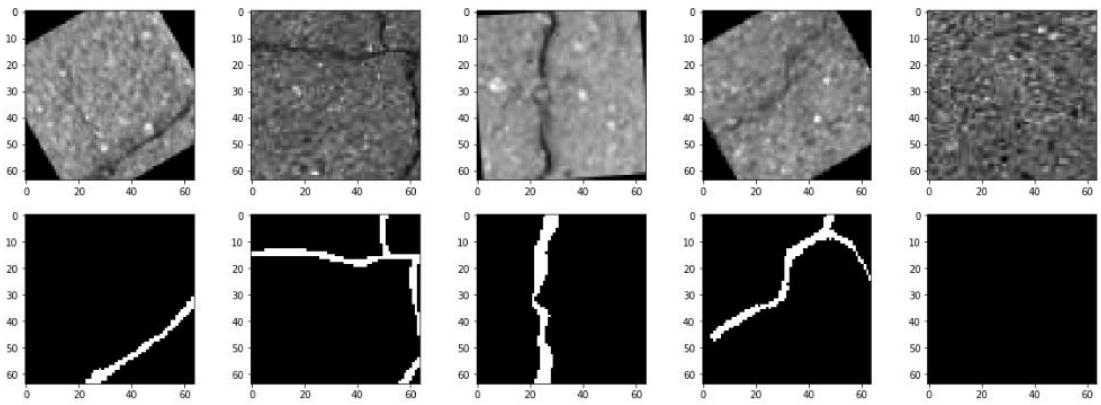


Fig. 2. Images and binary masks obtained from data augmentation

Neural Network Training. Intersection metrics between two detections (intersection over union, IoU, Jaccard index) and an equivalent binary similarity measure (Sörensen-Dice coefficient) are used to train and evaluate the neural network. The function $1 - J$ is used as cost:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}, S(A,B) = \frac{2|A \cap B|}{|A| + |B|}.$$

The initialization of weights in the layers of the neural network is carried out through the Glorot method [40]. To reduce the internal covariance shift, the batch is normalized through normalizing the input distributions of each layer. Adam-algorithm (stochastic optimization method) is used for learning [41].

At the first stage, the neural network is trained on a small amount of data (30% of the core set) during 5 epochs. At the second stage, the network is trained on the full amount of data during the required number of epochs. The learning rate varies with each epoch according to the established dependence (Fig. 3).

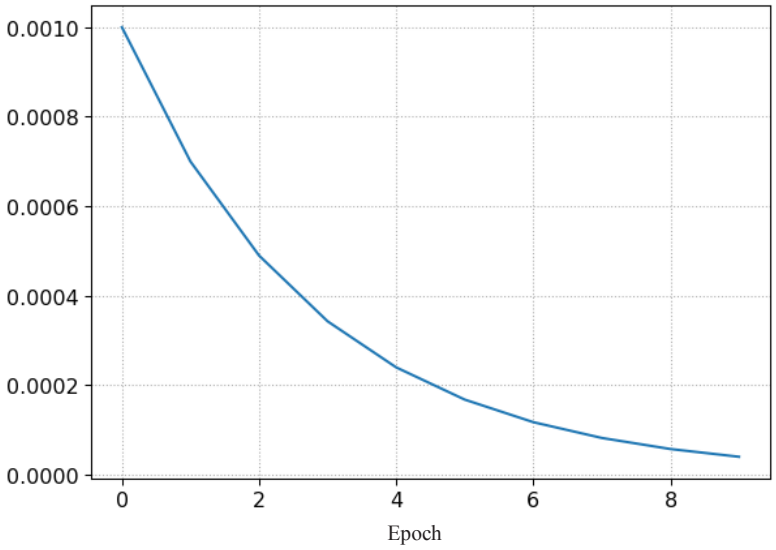


Fig. 3. Variation of network learning rate depending on epoch

As part of the task, it was established that the optimal number of training epochs is 25 (5 epochs at the first stage of training, and 20 – at the second one). With more epochs, the neural network accuracy did not change crucially.

Keras and TensorFlow frameworks were used for the implementation of the developed architecture of deep CNN.

Research Results. After training the neural network, it is validated on test data. Each image fragment is fed to the network input, and the output is a generated map of the defect probabilities. Fig. 4 shows the results of the trained network performance and their comparison with the true values from the test sample.

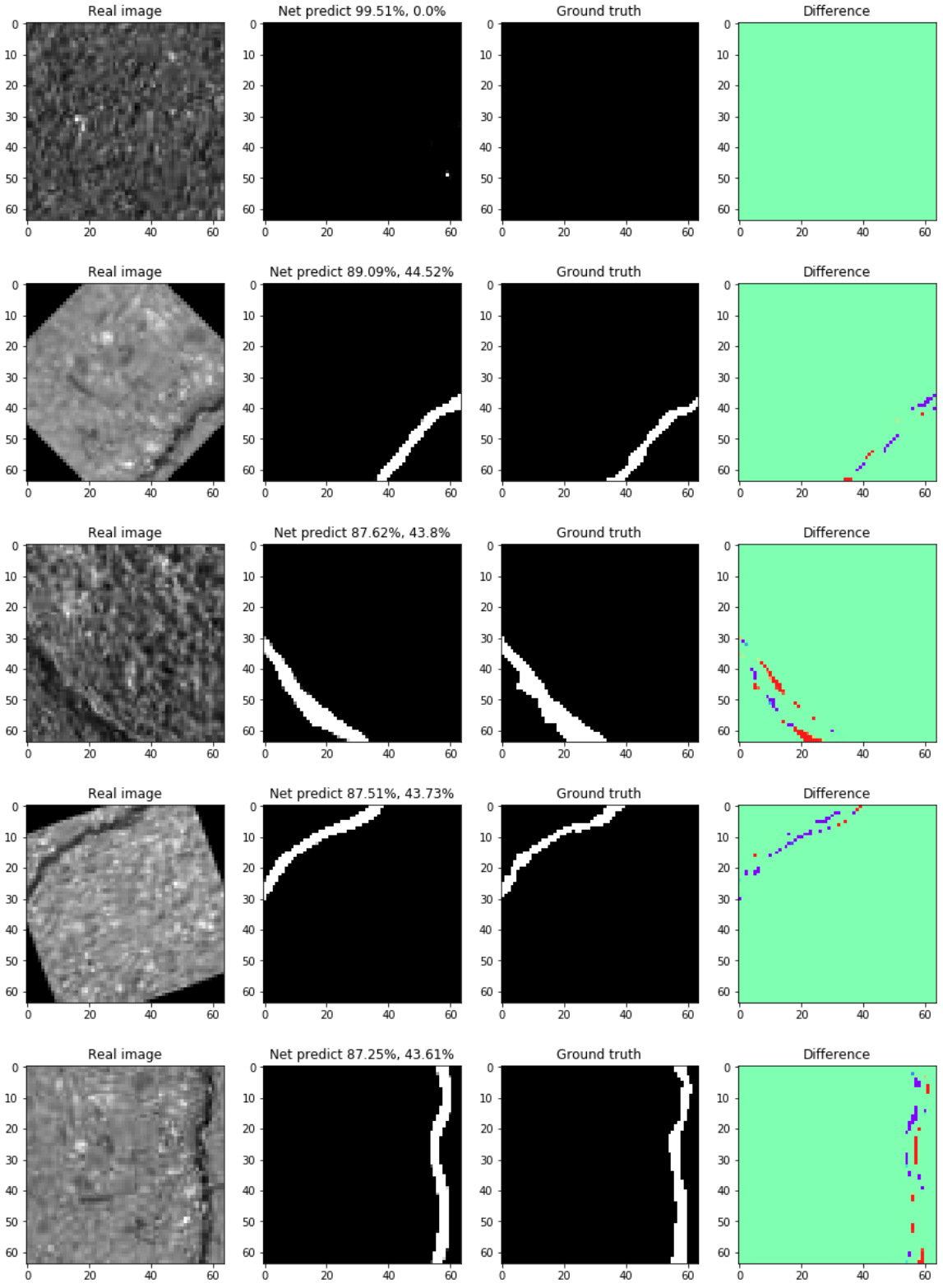


Fig. 4. Results of trained neural network performance

In Fig. 4, the data is produced in four columns:

1) image under study,

2) neural network output,

3) manually identified defect,

4) difference between 2) and 3).

Assume that the networks are compared with the true values. The values of IoU and Dice metrics are conditioned by the specific ratios of the following factors:

- area of the defect and area of the whole image,

- binary (one-bit) mask and real (4-byte) generated image.

It should be noted that when using IoU for defectless fragments, the metric values are 0 (Fig. 5).

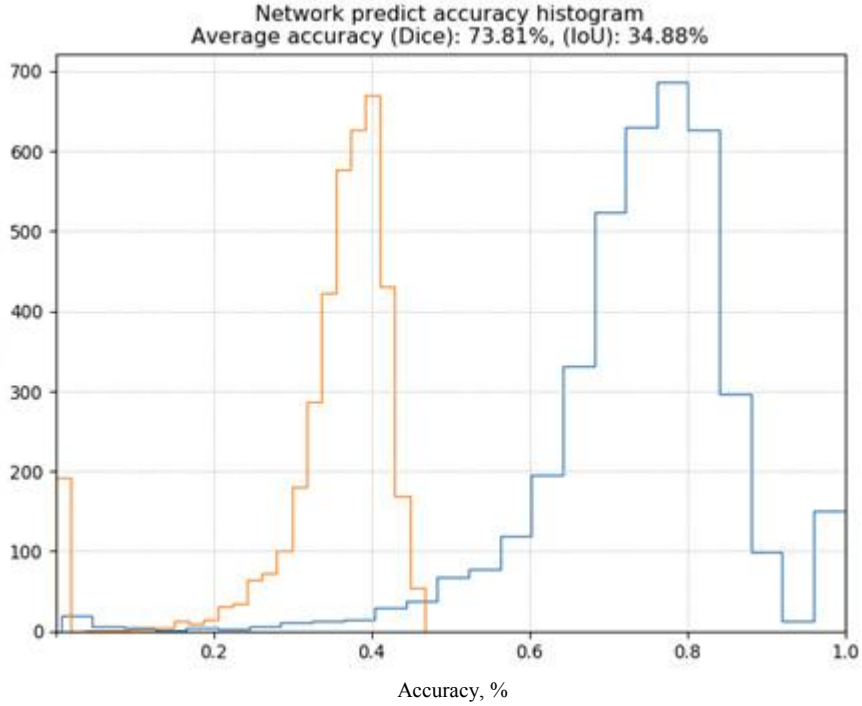


Fig. 5. Number of images identified with specified degree of accuracy

The prepared data set quality affects drastically the learning and the neural network output. In some cases, the neural network indicates a defect, although the true image does not have it, or vice versa. This affects the overall assessment of the model quality. In general, the accuracy evaluation of the neural network according to the proposed metrics can be subjective, so you should not take the data from Fig. 4 as absolute.

Some models of FCN networks are evaluated in the framework of the presented paper. The results are shown in Table 1.

Table 1

Accuracy of some neural networks

Network Architecture	Accuracy
10 layers (256, 128, 64, 64, 64, …), 929 665 parameters	Dice: 73.81 %, IoU: 34.88 %
16 layers (32, 32, 16, 16, 16, 8, 8, 8, …), 43 441 parameters	Dice: 70.40 %, IoU: 33.24 %
12 layers (32, 32, 16, 16, 8, 8, …), 37 537 parameters	Dice: 67.57 %, IoU: 32.12 %

Here, the number of filters on the first part of the network is indicated in brackets. The number of filters on the second part of the network is flipped (see Fig. 1).

A sliding window method with a specified step that regulates the processing speed and detailing is used to process high-resolution images. This is how the resulting defect probability map for the whole image is developed. Several images from the validation set and the result of their processing by the neural network are shown in Fig. 6.

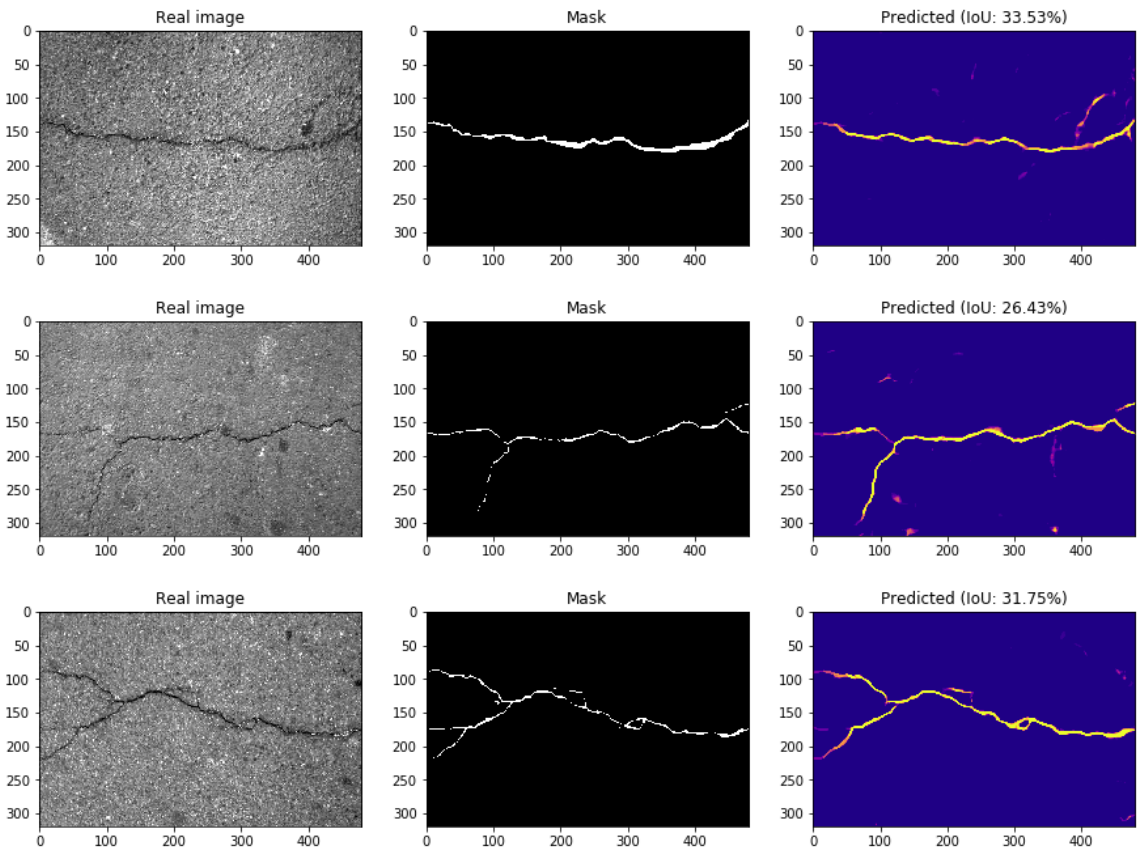


Fig. 6. Validation images processed by trained FCN

Discussion and Conclusions. A model of deep convolution neural network is proposed for identifying defects in pavement images. The model is implemented as a simplified and optimized version of the most popular on-the-day FCN networks. The techniques for constructing a training set and a two-stage network learning process considering the specifics of the current problem are described. The work done shows that the use of such architectures is successful with a small amount of the input data. Fine precision is registered. The described model can be used in various segmentation tasks. According to the metrics, FCNN shows the following results: IoU - 0.3488, Dice - 0.7381.

References

1. Quality Management of Pavement Condition Data Collection. National Academies of Sciences, Engineering, and Medicine. Washington: The National Academies Press, 2009, 144 p. DOI: https://doi.org/10.17226/14325 .

2. Mahler, D.S., et al. Pavement Distress Analysis Using Image Processing Techniques. *Computer-Aided Civil and Infrastructure Engineering*, 1991, vol. 6, iss. 1, pp. 1–14. DOI: https://doi.org/10.1111/j.1467-8667.1991.tb00393.x .

3. Tizhoosh, H. R., Krell, G., Michaelis, B. Locally adaptive fuzzy image enhancement. Computational Intelligence Theory and Applications. Fuzzy Days 1997. Lecture Notes in Computer Science. Berlin; Heidelberg: Springer, 1997, vol. 1226, pp. 272–276. DOI: https://doi.org/10.1007/3-540-62868-1_118 .

4. Zou, Q., et al. Crack Tree: Automatic crack detection from pavement images. *Pattern Recognition Letters*, 2012, vol. 33, pp. 227–238. DOI: 10.1016/j.patrec.2011.11.004 .

5. Xu, W., et al. Pavement crack detection based on saliency and statistical features. IEEE International Conference on Image Processing. 2013. Melbourne: IEEE, 2013, pp. 4093–4097. DOI: https://doi.org/10.1109/ICIP.2013.6738843 .

6. Oliveira, H., Correia, P.L. CrackIT — An Image Processing Toolbox for Crack Detection and Characterization. IEEE International Conference on Image Processing — ICIP 2014. Paris: IEEE, 2014, pp. 798–802. DOI: http://doi.org/10.1109/ICIP.2014.7025160 .

7. Shi, Y., et al. Automatic Road Crack Detection Using Random Structured Forests. *IEEE Transactions on Intelligent Transportation Systems*, 2016, vol. 17, iss. 12, pp. 3434–3445. DOI: https://doi.org/10.1109/TITS.2016.2552248 .

8. Lee, B.J., Lee, H. Position-Invariant Neural Network for Digital Pavement Crack Analysis. *Computer-Aided Civil and Infrastructure Engineering,* 2004, vol. 19, iss. 2, pp. 105–118. DOI: http://dx.doi.org/10.1111/j.1467-8667.2004.00341.x .

9. Sun, B.-C., Qiu, Y.-J. Automatic Identification of Pavement Cracks Using Mathematic Morphology. First International Conference on Transportation Engineering. Chengdu: ASCE, 2007, pp. 1783–1788. DOI: https://doi.org/10.1061/40932(246)292 .

10. Chambon, S. Detection of road cracks with multiple images. International Joint Conference on Computer Vision Theory and Applications, VISAPP. Angers: Springer, 2010, 7 p.

11. Zhang, L., et al. Road crack detection using deep convolution neural network. IEEE International Conference on Image Processing (ICIP-2016). Phoenix: IEEE, 2016, pp. 3708–3712. DOI: https://doi.org/10.1109/ICIP.2016.7533052 .

12. Prasanna, P., et al. Automated Crack Detection on Concrete Bridges. *IEEE Transactions on Automation Science and Engineering*, 2014, vol. 13, iss. 2, pp. 591–599. DOI: https://doi.org/10.1109/TASE.2014.2354314 .

13. Jahanshahi, M.-R., et al. An innovative methodology for detection and quantification of cracks through incorporation of depth perception. *Machine Vision and Applications*, 2013, vol. 24, pp. 227–241. DOI: https://doi.org/10.1007/s00138-011-0394-0 .

14. Medina, R.J., et al. Enhanced automatic detection of road surface cracks by combining 2D/3D image processing techniques. IEEE International Conference on Image Processing (ICIP-2014). Paris: IEEE, 2014, pp. 778–782. https://doi.org/10.1109/ICIP.2014.7025156 .

15. Chanda, S., et al. Automatic Bridge Crack Detection — A Texture Analysis-Based Approach. Artificial Neural Networks in Pattern Recognition (ANNPR): Lecture Notes in Computer Science. Cham: Springer, 2014, vol. 8774, pp. 193–203. DOI: https://doi.org/10.1007/978-3-319-11656-3_18 .

16. Kapela, R., et al. Asphalt Surfaced Pavement Cracks Detection Based on Histograms of Oriented Gradients. 22-nd International Conference Mixed Design of Integrated Circuits & Systems. Moscow: ALT Linux, 2015, pp. 579–584. DOI: https://doi.org/10.1109/MIXDES.2015.7208590 .

17. LeCun, Y., et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, vol. 86, iss. 11, pp. 2278–2324. DOI: https://doi.org/10.1109/5.726791 .

18. Ciresan, D., Meier, U., Schmidhuber, J. Multi-column deep neural networks for image classification. Proc. of Conference on Computer Vision and Pattern Recognition (CVPR). Providence: IEEE, 2012, pp. 3642–3649. DOI: https://arxiv.org/abs/1202.2745 .

19. Goodfellow, I.J., et al. Multi-digit number recognition from street view imagery using deep convolution neural networks. Proc. of International Conference on Learning Representations (ICLR). Banff: Deep Learning, 2014, pp. 1–12. DOI: https://arxiv.org/abs/1312.6082 .

20. Ciresan, D., et al. A committee of neural networks for traffic sign classification. Proc. of International Joint Conference on Neural Networks (IJCNN). San Jose: IEEE, 2011, pp. 1918–1921. DOI: http://doi.org/10.1109/IJCNN.2011.6033458 .

21. Arel, I.,Rose, D.C., Karnowski, T.P. Deep machine learning — a new frontier in artificial intelligence research. *IEEE Computational Intelligence Magazine*, 2010, vol. 5, iss. 4, pp. 13–18. DOI: http://doi.org/10.1128/IAI.02190-14 .

22. Simard, P.Y., Steinkraus, D., Platt, J.C. Best practices for convolution neural networks applied to visual document analysis. Proc. of International Conference on Document Analysis and Recognition (ICDAR). Edinburgh: IEEE, 2003, pp. 958–963. DOI: http://doi.org/10.1109/ICDAR.2003.1227801 .

23. Szegedy, C., et al. Going deeper with convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015, pp. 1–9. DOI: http://doi.ieeecomputersociety.org/10.1109/CVPR.2015.7298594 .

24. Girshick, R., et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. IEEE Conference on Computer Vision and Pattern Recognition. Columbus IEEE, 2014, pp. 580–587. DOI: https://doi.org/10.1109/CVPR.2014.81 .

25. Chen, X., et al. Vehicle Detection in Satellite Images by Hybrid Deep Convolution Neural Networks. IEEE Geoscience and Remote Sensing Letters, 2014, vol. 11, iss. 10, pp. 1797–1801. DOI: 10.1109/ACPR.2013.33.

26. Some, L. Automatic image-based road crack detection methods. Stockholm: Royal Institute of Technology, 2016, 61 p.

27. Xie, D., Zhang, L., Bai, L. Deep learning in visual computing and signal processing. Applied Computational Intelligence and Soft Computing, 2017, vol. 2017 (10), pp. 1–13. DOI: https://doi.org/10.1155/2017/1320780 .

28. Saar, T., Talvik, O. Automatic Asphalt pavement crack detection and classification using Neural Networks. 12th Biennial Baltic Electronics Conference. Tallinn: IEEE, 2010, pp. 345–348. DOI: http://dx.doi.org/10.1109/BEC.2010.5630750 .

29. Meignen, D., Bernadet, M., Briand, H. One application of neural networks for detection of defects using video data bases: identification of road distresses. Database and Expert Systems Applications: Proc. 8th International Workshop in Toulouse, France. Berlin; Heiderberg: Springer-Verlag, 1997, pp. 459–464. DOI: https://doi.org/10.1109/DEXA.1997.617332 .

30. Gopalakrishnan, K., et al. Deep Convolution Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, 2017, vol. 157, pp. 322–330. DOI: https://doi.org/10.1016/j.conbuildmat.2017.09.110 .

31. Browne, M., Ghidary, S.S. Convolution Neural Networks for Image Processing: An Application in Robot Vision. AI 2003: Advances in Artificial Intelligence. Berlin; Heidelberg: Springer, 2003, vol. 2903, pp. 641–652. DOI: https://doi.org/10.1007/978-3-540-24581-0_55 .

32. Dawood, T., Zhu, Z., Zayed, T. Machine vision-based model for spalling detection and quantification in subway networks. *Automation in Construction*, 2017, vol. 81, pp. 149–160. DOI: http://dx.doi.org/10.1016/j.autcon.2017.06.008 .

33. Jahanshahi, M.R. et al. A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance. 2009, vol. 5, iss. 6, pp. 455–486. DOI: https://doi.org/10.1080/15732470801945930 .

34. Zhang, A., et al. Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces Using a Deep-Learning Network. Computer-Aided Civil and Infrastructure Engineering. 2017, vol. 32, iss. 10, pp. 805–819. DOI: https://doi.org/10.1111/mice.12297 .

35. Badrinarayanan, V., Kendall, A., Cipolla, R. SegNet: A Deep Convolution Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, iss. 12, pp. 2481–2495. DOI: 10.1109/TPAMI.2016.2644615.

36. Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolution Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. Cham: Springer, 2015, vol. 9351, pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28.

37. Eisenbach, M., et al. How to Get Pavement Distress Detection Ready for Deep Learning? A Systematic Approach. IEEE International Joint Conference on Neural Networks (IJCNN). Anchorage: IEEE, 2017, pp. 2039–2047. DOI: http://doi.org/10.1109/IJCNN.2017.7966101 .

38. Shelhamer, E., Long, J., Darrell, T. Fully Convolution Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2017, vol. 39, iss. 4, pp. 640–651. DOI: 10.1109/TPAMI.2016.2572683.

39. Ioffe, S., Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Computing Research Repository*, 2015, 9 p. DOI: abs/1502.03167.

40. Glorot, X., Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. *AISTATS*, 2010, vol. 9, pp. 249–256.

41. Diederik, P. K., Ba, J. Adam: A Method for Stochastic Optimization. International Conference on Learning Representations. Banff: IEEE, 2014, 15 p.

Received 02.11 .2018
Submitted 02.11.2018
Scheduled in the issue 15.01.2019

Authors:

Sobol, Boris V.,
Head of the Information Technologies Department,
Don State Technical University (1, Gagarin sq.,
Rostov-on-Don, 344000, RF), Dr.Sci. (Eng.),
professor,
ORCID: http://orcid.org/0000-0003-2920-6478
b.sobol@mail.ru

Soloviev, Arkady N.,
Head of the Theoretical and Applied Mechanics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Phys.-Math.), professor,
ORCID: http://orcid.org/0000-0001-8465-5554
solovievarc@gmail.com

Vasiliev, Pavel V.,
Senior lecturer of the Information Technologies Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: http://orcid.org/0000-0003-4112-7449
lyftzeigen@mail.ru

Podkolzina, Lubov A.,
Post-graduate student of the Information Technologies Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: http://orcid.org/0000-0001-9476-5802
podkolzinalu@@gmail.com

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 74–80. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 74–80. ISSN 1992-5980 eISSN 1992-6006

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

UDC 62-50 https://doi.org/10.23947/1992-5980-2019-19-1-74-80

Production machines maintenance based on digitalization*

A. K. Tugengold[1], R. N. Voloshin[2], A. R. Yusupov[3], T. N. Kruglova[4]**
[1, 2, 3] Don State Technical University, Rostov-on-Don, Russian Federation
[4] Platov South-Russian State Polytechnic University (NPI), Novocherkassk, Russian Federation

Техническое обслуживание технологических машин на базе цифровизации***

А. К. Тугенгольд[1], Р. Н. Волошин[2], А. Р. Юсупов[3], Т. Н. Круглова[4]**
[1, 2, 3] Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация
[4] Южно-Российский государственный политехнический университет (НПИ) имени М. И. Платова, г. Новочеркасск, Российская Федерация

Introduction. Digital data and analytics transform the role of the production equipment maintenance. Analytical information of sensors placed on the product allows continuous monitoring of the production machines operation and their timely servicing. Thus, defects in technical equipment are identified, the analysis of which enables to develop algorithms for monitoring and forecasting, and to prevent equipment from overshooting the limits of the safe operation.
Materials and Methods. Basic digitalization principles and the digital images structure are presented. A mathematical method is used to describe the digital image vector and the control system algorithm.
Research Results. The achievements of the known systems of maintenance and digitalization of various machines are summarized. The application of a dynamic digital image made it possible to determine the desired levels of the production facilities maintenance. An optional version of monitoring the equipment state within the framework of the production digitalization concept is shown. It is based on the proposed algorithm for an autonomous control of the process state.
Discussion and Conclusions. The construction of machine digital images in accordance with the main stages of its life cycle is described. The task of automated maintenance of machine tools based on digitalization is considered.

Keywords: monitoring, digitalization, autonomous control, maintenance, digital image.

Введение. Цифровые данные и аналитика преобразуют роль технического обслуживания производственного оборудования. Аналитическая информация датчиков, размещенных на изделии, позволяет непрерывно наблюдать функционирование технологических машин и своевременно обслуживать их. Так, выявляются дефекты технического оснащения, анализ которых позволяет разрабатывать алгоритмы мониторинга, прогнозирования и предупреждать выход оборудования за пределы надежной работы.
Материалы и методы. Представлены основные принципы цифровизации и структура построения цифровых образов. Используется математический метод описания вектора цифровых образов и алгоритмизации системы управления.
Результаты исследования. Обобщены достижения известных систем технического обслуживания и цифровизации различных машин. Использование динамического цифрового образа позволило определить необходимые уровни для поддержания работоспособности технологических объектов. Показан возможный вариант мониторинга состояния оборудования в рамках концепции цифровизации производства. Он основан на предложенном алгоритме автономного управления технологическим состоянием.
Обсуждение и заключение. Описано построение цифровых образов станка в соответствии с основными стадиями его жизненного цикла. Рассмотрена задача автоматизированного поддержания работоспособности станков на базе цифровизации.

Ключевые слова: мониторинг, цифровизация, автономное управление, обслуживание, цифровой образ.

* The research is done within the frame of the independent R&D.
** E-mail: akt0@yandex.ru, r.voloshin2909@gmail.com, sthedgehog@icloud.com, kruglovatanya@rambler.ru
*** Работа выполнена в рамках инициативной НИР.

For citation: A.K. Tugengold, et al. Production machines maintenance based on digitalization. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 74–80. https://doi.org/10.23947/1992-5980-2019-19-1-74-80

Образец для цитирования: Техническое обслуживание технологических машин на базе цифровизации / А. К. Тугенгольд [и др.] // Вестник Дон. гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 74–80. https://doi.org/10.23947/1992-5980-2019-19-1-74-80

Introduction. At the present stage of development of engineering and technology, issues of the production digitalization, in particular, the life-cycle management of production machines, the independent management of their operation and maintenance, are particularly sensitive. According to K. Schwab, Head of the World Economic Forum, digital data and analytics transform the role of servicing [1]. We are talking about the analytical information of sensors placed on the product. It helps to conduct continuous monitoring, and to study equipment operation and its maintenance. Thus, defects in the equipment are identified, the analysis of which enables to develop monitoring algorithms, to make forecasts, and to prevent equipment from going beyond the limits of reliable operation. This approach generally improves the efficiency of production processes. Description of the product operation based on digitalization allows for the creation of digital twins of machines for various purposes. The concept of "digital twin" includes a temporal information copy of an object, artificial intelligence, information technologies, and software. All these components are involved in the creation and support of intelligent digital models (DM) of complex technical products.

Publications devoted to this topic represent research in the field of intellectual monitoring and control of the technical condition of machines and machine tool systems [2–14].

Metal-cutting machines are provided with new properties that allow them to meet the growing demands for speed and accuracy of cutting, reliability and safety of operation under the conditions of fast-speed processes. As in mechatronics in general, in machine tools, a number of organizational and technology principles related to digitalization is implemented. Some of them are listed below.

Openness and controllability. A machine is an open, self-controlled system connected with the outdoor environment. Information on the state of the equipment, the outdoor environment is used for control, and the object behavior is simulated.

Originality of properties of the entire machine system. The characteristic of a system object extends further than a sum of features of its components. The system obtains properties that none of its components possesses. In other words, the functions of the components can synthesize new properties of the system.

Self-sufficiency of the system. The system and its elements can function independently of external systems.

Intelligent management. Procedures for controlling an object (machine) are based on an analysis of its states and manufacturing situations using information technologies and knowledge processing mechanisms.

Information and intellectual support of the machine operation includes procedures for maintenance and repair. The development of digital automated diagnostic and control systems involves the creation of specific software for information processing integrated with an external network communication environment.

Main Part. In the presented paper, the digital embodiment of an object will be called its digital image (DI).

The machine DI corresponds to various methods of constructing digital models (Digital model). The model selection is determined by the object life cycle stage – from design to disposal (Fig. 1).

.

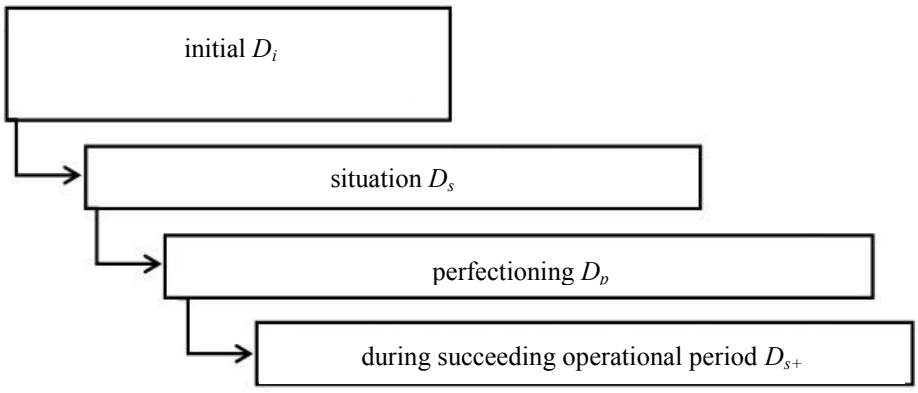


Fig. 1. Stages of representation of machine digital images

According to the notation of Fig. 1 and considering the sequence of transformations, an ordered set of digital models of the machine state during the life cycle is as follows:

$$D_m = (D_i, D_s, D_p, D_{s+}).$$

According to the standard rules and test methods, the work on creating D_i includes the formation of an initial database for assessing the machine quality. These include the following: performance accuracy factors, dynamic characteristics, thermal deformations, reliability rating estimates. The database (DB) of the initial state of the machine forms the basis of the temporal machine DI built on the principles of the e-Mind Machine (e-MM) [10, 11].

If we are talking about the machine in use, its state is considered under processing typical test parts or on test modes subject to the technical requirements and a test program. The structure of the machine DI in the operation processes (D_s), considering the e-MM, is shown in Fig. 2.

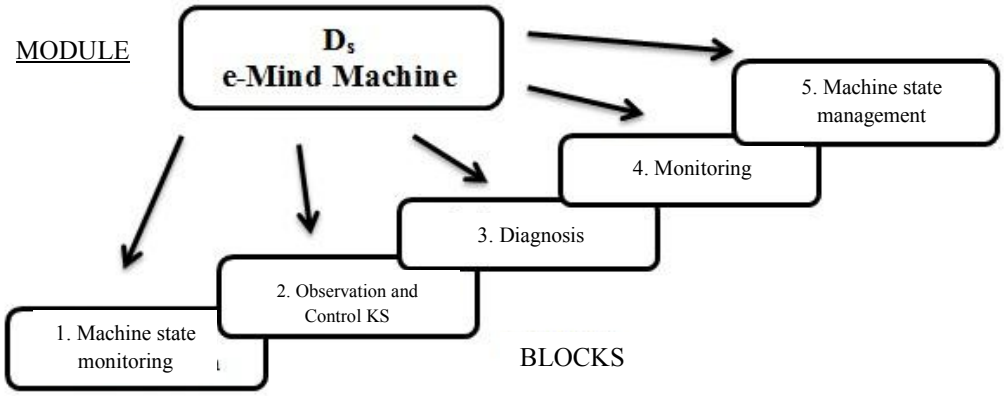


Fig. 2. Structure of machine dynamic DI at service stage (D_s)

The DI state specificity is determined by the alternation of the operation, maintenance and repair processes. The DI intelligence depends on the knowledge system (KS) located in the blocks of the machine temporal DI.

Solutions based on the embedded expert systems (ES) and information obtained from the diagnostic and monitoring units are synthesized in the modules for controlling the machine technical state. They identify the measured parameters of the machine, the process and its result, which enables to diagnose and evaluate the state of processes and devices. In this case, a situational assessment of the system state, which is necessary for the management, adequate to the current situation, is performed, and the results are predicted. Based on the decisions made, control actions are generated corresponding to the block assignments. Considering the huge amounts of stored data, optimizing analysis techniques can be used under the system operation — for example, online diagnostics with the aid of Mel-frequency cepstral coefficients [12].

At this, it is important to apply not quantitative, but qualitative assessments and concepts based on fuzzy procedures of processing, accumulating and applying knowledge and meta-knowledge. It is advisable to use temporal ES. First, they work in real time. Secondly, training and adaptation algorithms are built into such ES, and they are able to improve the performance of the machine subsystems.

Under the intelligent control with the KS units during the machine operation period, complete information on the DI state can be represented as an ordered set (vector) of possible states of the machine system and control:

$$S = \left(S_w, S_t, S_v, S_z, S_u\right).$$

Here, S_w, S_t, S_v, S_z, S_u are, respectively, sets of states of the machining process, the tool, the technical state of the machine components, the state of the product (workpiece - part) and control. The set of admissible vectors (S) is called the admissible set of DI states, and we denote it as $\tilde{S}$. KS synthesizes sets of time-ordered regulated actions: $U_p \in \tilde{U}$. They provide transformation of the vector of the DI initial condition $S^* \in \tilde{S}$ from a hypothetically identified admissible set of states ($\tilde{S}$) to the specified goal state $S_g \in \tilde{S}$. This search is based on the analysis of knowledge about the functionality of a specific machine tool system. Actually, KS should realize λ_S indication:

$$\lambda_S : \tilde{S} \times \tilde{S} \to \tilde{U} \,|\, U_p = \lambda_S\left(S^*, S_g\right),$$

which, out of the set of admissible controls ($\tilde{U}$), finds such control ($U_p \in \tilde{U}$) that corresponds to the of the initial state vector ($S^* \in \tilde{S}$) and to the specified goal state ($S_g \in \tilde{S}$). In this case, $U_p \in \tilde{U}$ control enables the transition of the machine from $S^* \in \tilde{S}$ state to $S_g \in \tilde{S}$ state. The management apparatus ($U_p \in \tilde{U}$) is based on the formalisms of fuzzy alge-

bra and fuzzy sets. This approach allows for the formulation of a plausible hypothesis about the organization of the appropriate behavior of the machine tool system. In this case, a coherent transition from one state to another is observed, and at each transition, information about changes in the machine DI is generated [15]:

$$S_{j-1} \xrightarrow{U_P} \left(S_j, I_j\right).$$

Here, S_{j-1} is DI state at the beginning of *j*-th transition; S_j is DI state obtained as a result of *j*-th transition; I_j are changes to be made in the descriptions of DI state obtained as a result of *j*-th transition; U_p is control ensuring the transition from S_{j-1} state to S_j state.

The current state (S_j) obtained as a result of the *j*-th transition is determined by the state of:

- product (workpiece - part) (S_{zj}),
- tool subsystems (S_{ij}),
- object (S_{vj}).

That is

$$S_{zj} = \left(S_{zj}, S_{ij}, S_{vj}\right).$$

Models of technical condition management, built on this KS base, are the framework for building a digital system for autonomous control of machine tools.

Developers and researchers pay serious attention to the automated maintenance of machine operability, especially, under the conditions of computerization of production and improvement of information support [1].

The digitalization systems developed for the technological machines maintenance require the application of integrated and/or remote software and hardware support components. These systems are created to fulfill the following tasks:

- to obtain information on the equipment state in real time,
- to forecast the development of the state of various devices/units (D/U),
- to give the personnel warning of emergency and other dangerous equipment conditions,
- to carry out self-maintenance and troubleshooting,
- to conduct maintenance or updating of the control programs in the process of equipment operation.

Within the framework of the e-MM concept, an automated autonomous control of the technical state of machine tools (AUTS) was created [16]. AUTS means independent automated control. Special tools and informational links allow it to assess the state and role of the machine. Based on these estimates, the AUTS provides certain signals or acts on a specific machine device to maintain or restore its operability.

The AUTS features are listed below.

First, the system should be equipped with a sufficient number of sensors that transmit reliable information on the status of units and devices.

Secondly, when recognizing the obtained data, it is necessary to exclude noise and determine the parameters informatively describing the state of the equipment.

Thirdly, the forecasting systems should flexibly adjust the indicators in real time, focusing on the information obtained from the diagnostic unit.

Fourthly, the decision should include the possibility of reducing the risk from the machinery breakdown or wear reduction of the equipment components, and it should comply with the selected treatment program with minimal departure from the optimal process.

In view of the above, the structure of the system and the autonomous control of the machines technical state is developed [15]. The algorithm of the state monitoring system is presented in Fig. 3.

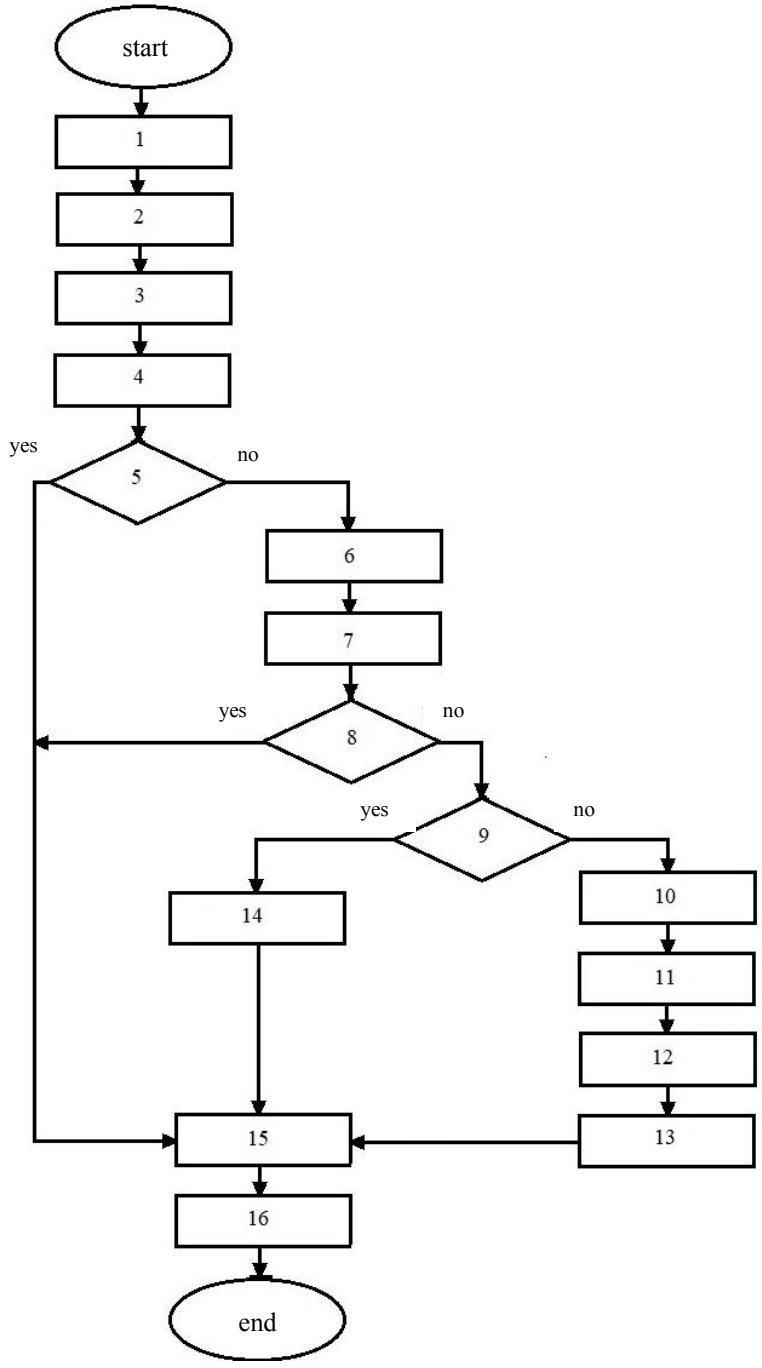


Fig. 3. Flowchart of state monitoring system (AUTS)

The algorithm provides the following actions.

1. 1. Search for information on the D/U inheriting states (i.e., at the preceding stage – n - 1) in the knowledge system, and basing the *n* stage states into the monitoring unit DB.

2. Obtaining current operational data of diagnostics of the D/U state at *n* stage.

3. Creation (change) of the current image of the D/U state in the working memory of the knowledge system of $n - KS_n$ stage of the machine condition monitoring unit.

4. Determination of the ES belonging to the KS of the monitoring unit, and tendencies of changes in the D/U state.

5. Comparison of the obtained state estimates with the parameters of the fuzzy boundaries of the D/U operability state extracted from the database.

6. Predictable assessment of the preservation of permissible state parameters within the limits of the D/I operability under the condition of work on the input control program of the CNC device.

7. Estimation of the accessibility of fail-safety under the condition of correction of the processing program components.

8. Decision-making on the accessibility of fail-safety.

9. Decision-making on technical condition management.

10. Correction of the D/I state without processing interrupt.

11. Part process shutdown for the correction (restoration) of the state.

12. Selection of the offline means to prevent (eliminate) malfunction of the D/U or external services.

13. Conducting off-line servicing.

14. Support call to the external services for troubleshooting of the technical condition.

15. Diagnosis and assessment of the D/I status after maintenance.

16. Logging and input of the list of the operations performed and the resulting assessments of the D/I state to the KS. Accumulation of experience in assessing the technical condition and autonomous state control in the KS.

In this case, the decision-making and autonomous control of actions to maintain working capacity is made on the basis of the presented approach to the digitalization of machine tools, and this is the key feature of the unit for monitoring the object state and of the algorithm for monitoring the AUTS state.

Conclusion. The construction of machine digital images in accordance with the main stages of its life cycle is proposed. The automated maintenance of machine tools based on digitalization through the creation of autonomous technical condition control systems is considered. These systems transmit information on the equipment state in real time, signalize purposefully about the equipment state, and maintain (restore) its performance.

References

1. Schwab, K. Chetvertaya promyshlennaya revolyutsiya. [The fourth industrial revolution.] Moscow: "Eksmo", 2016, 138 p. (Top Business Awards) (in Russian).

2. Tugengold, A.K., Lysenko, A.F., Statovoy, D.A. Sistema znaniy v vide intellektual'noy elektronnoy tekhnicheskoy dokumentatsii dlya mnogooperatsionnykh stankov. [Knowledge-based system by way of intelligent electronic technical documentation for multi-operation machines.] Vestnik Mashinostroeniya, 2015, no. 11, pp. 14–17 (in Russian).

3. Kemerait, R.C. New cepstral approach for prognostic maintenance of cycle machinery. Proceedings of the IEEE Southeast Conference in Tampa, 1987, pp. 256–262.

4. Fisher, C., Baines, N.C. Multi-sensor condition monitoring systems for gas turbines. Journal of Condition Monitoring, 1988, no. 1, pp. 57–68.

5. Tugengold, A.K., Voloshin, R.N., Yushchenko, S.V. Monitoring sostoyaniya mnogooperatsionnykh stankov na baze kontseptsii e-MindMachine. [Monitoring of multioperational machines based on the concept of e-MIND MACHINE.] Vestnik of DSTU, 2016, vol. 16, no. 1 (84), pp. 77–86 (in Russian).

6. Tugengold, A.K., Voloshin, R.N. Gibkiy monitoring mekhatronnykh tekhnologicheskikh mashin. [Flexible monitoring of mechatronic technological machines.] Vestnik of DSTU, 2016, no. 4, pp. 51–58 (in Russian).

7. Byington, C.S., et al. Shaft coupling model-based prognostics enhanced by vibration diagnostics. Insight, 2009, vol. 51, pp. 420–425 (Non-Destructive Testing and Condition Monitoring).

8. Tan, C. K., Irving, P., Mba, D. A comparative experimental study on the diagnostic and prognostic capabilities of acoustics emission, vibration and spectrometric oil analysis for spur gears. Mechanical Systems and Signal Processing, 2007, no. 21, pp. 208–233.

9. Tugengold, A.K., et al. Monitoring sostoyaniya stankov i stanochnykh sistem. [Monitoring the state of machines and machine tools.] STIN, 2017, no. 3, pp. 11–17 (in Russian).

10. Tugengold, A.K., et al. Monitoring and Control of Tools in Multifunctional Machine Tools. Russian Engineering Research, 2017, vol. 37, no. 5, pp. 440–446.

11. Tugengold, A.K., et al. Monitoring of Machine Tools. Russian Engineering Research, 2017, vol. 37, no.8, pp. 440–446.

12. Tugengold, A.K., Izyumov, A.I. Printsipy kontseptual'nogo podkhoda k sozdaniyu podsistemy «INSTRUMENT» v smart-pasporte mnogooperatsionnogo stanka. [Principles of conceptual approach to creating TOOL subsystem for multioperation machine smart-passport.] Vestnik of DSTU, 2014, vol. 14, no. 2, pp. 74–83 (in Russian).

13. Tugengold, A.K. Smart-Passport otkrytoy mekhatronnoy tekhnologicheskoy sistemy. Kontent. [Smart-Passport of open mechatronic technology system. Content.] Saarbrucken: Lambert Academic Publishing, 2013, 83 p. (in Russian).

14. Tsifrovoy dvoynik (Digital Twin) [Digital Twin.] CADFEM CIS. Available at: https://www.cadfem-

cis.ru/products/ansys/systems/digital-twin/ (accessed: 01.02.18.) (in Russian).

15. Tugengold, A.K., et al. Upravlenie tekhnicheskim sostoyaniem stankov. [Machine Condition Management.] STIN, 2018, no. 7, pp. 8–15 (in Russian).

16. Tugengold, A.K. Smart-pasport mekhatronnogo tekhnologicheskogo ob"ekta. Kontsept. [Smart-Passport of mechatronic production facility. Concept.] Vestnik of DSTU, 2012, no. 7, pp. 33–41 (in Russian).

Received 16.11.2018
Submitted 18.11.2018
Scheduled in the issue 15.01.2019

Authors:

Tugengold, Andrey K.,
professor of the Robotics and Mechatronics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Eng.), professor,
ORCID: http://orcid.org/0000-0003-0551-1486
akt0@yandex.ru

Voloshin, Roman N.,
postgraduate student of the Robotics and Mechatronics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: http://orcid.org/0000-0001-6147-2907
r.voloshin2909@gmail.com

Yususpov, Alexander R.,
graduate student of the Robotics and Mechatronics Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF),
ORCID: http://orcid.org/0000-0003-2179-616X
sthedgehog@icloud.com

Kruglova, Tatyana N.,
postdoctoral student of the Automation and Robotization of Agroindustrial Complex and Biosystems Engineering Department, Platov South-Russian State Polytechnic University (NPI) (132, ul. Prosveshcheniya, Novocherkassk, Rostov Region, 346428, RF),
ORCID: https://orcid.org/0000-0002-2730-0498
kruglovatanya@rambler.ru

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 81–85. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 81–85. ISSN 1992-5980 eISSN 1992-6006

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

UDC 519.87 https://doi.org/10.23947/1992-5980-2019-19-1-81-85

On efficiency of methods and algorithms for solving optimization problems considering objective function specifics*

E. N. Ostroukh[1], Yu. O. Chernyshev[2], L. N. Evich[3], P. A. Panasenko[4]**
[1,2,3]Don State Technical University, Rostov-on-Don, Russian Federation
[4]Krasnodar Higher Military School named after General of the Army S. M. Shtemenko, Krasnodar, Russian Federation

К вопросу эффективности методов и алгоритмов решения оптимизационных задач с учетом специфики целевой функции***

Е. Н. Остроух[1], Ю. О. Чернышев[2], Л. Н. Евич[3], П. А. Панасенко[4]**
[1,2,3]Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация
[4] Краснодарское высшее военное училище имени генерала армии С. М. Штеменко, г. Краснодар, Российская Федерация

Introduction. The estimation of efficiency of methods and algorithms for solving optimization problems with a vector criterion and a set of nonlinear constraints is considered. The approach that allows proceeding to an optimization problem with a single objective function (i.e., an unconditional optimization problem) after equivalent transformations is described. However, the objective function obtained in this way has properties (nonlinearity, multimodality, ravine, high dimension) that do not allow classical methods to be used to solve it. The presented work objective is to develop hybrid methods, based on combinations of the algorithms inspired by wildlife with other approaches (gravitational and gradient) for the solution to this problem.

Materials and Methods. New methods to solve the specified problem are developed. A computer experiment was conducted on a number of test functions; its analysis was performed, showing the efficiency of various combinations on various functions.

Research Results. The efficiency of hybrid algorithms that combine the following approaches is evaluated: genetic and immune; methods of swarm intelligence and genetic and immune; immune and swarm with gravity and gradient.

Discussion and Conclusions. The hybrid algorithms in optimization problems are studied. In particular, decisions can be made on their basis under the management of compound objects in the military and industrial sectors, in the creation of innovative projects related to the digital economy. It is established that the type of the objective function affects the result much more than the combination of algorithms.

Введение. Статья посвящена оценке эффективности методов и алгоритмов решения оптимизационных задач с векторным критерием и системой нелинейных ограничений. Описан подход, позволяющий после проведения эквивалентных преобразований перейти к оптимизационной задаче с одной целевой функцией (т. е. к задаче безусловной оптимизации). Однако полученная таким способом целевая функция обладает свойствами (нелинейность, мультимодальность, овражность, большая размерность), не позволяющими использовать для ее решения классические методы.

Цель представленного исследования — разработать для решения данной задачи гибридные методы, основанные на комбинациях алгоритмов, инспирированных живой природой, с другими подходами (гравитационным и градиентным).

Материалы и методы. Созданы новые методы для решения указанной задачи. Проведен компьютерный эксперимент на ряде тестовых функций, выполнен его анализ, показывающий эффективность различных комбинаций на различных функциях.

Результаты исследования. Оценена эффективность гибридных алгоритмов, которые комбинируют следующие подходы: генетический с иммунным; методы роевого интеллекта с генетическими и иммунными; иммунные и роевые с гравитационным и градиентным.

Обсуждение и заключение. Изучены возможности гибридных алгоритмов в оптимизационных задачах. В частности, на их основе могут приниматься решения при управлении сложными объектами в военной и промышленной сферах, при создании инновационных проектов, связанных с цифровой экономикой. Установлено, что вид целевой функции влияет на результат гораздо более существенно, чем комбинация алгоритмов.

*Работа выполнена при поддержке грантов РФФИ: 16-01-00391, 16-01-00390, 18-01-00314.
*** The research was supported by the RFFI grants nos. 16-01-00391, 16-01-00390, 18-01-00314.
**E-mail: Eostr@Donpac.Ru, Myvnn@list.ru, Bkln@mail.ru, we_panasenko_777@mail.ru

Keywords: combination, hybrid, bio-inspired algorithm, swarm intelligence, gradient-based algorithm, gravity search algorithm, efficiency, convergence.

Ключевые слова: комбинация, гибрид, биоинспирированный алгоритм, роевой интеллект, градиентный алгоритм, гравитационный алгоритм, эффективность, сходимость.

For citation: E.N. Ostroukh, et al. On efficiency of methods and algorithms for solving optimization problems considering objective function specifics. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. C. 81–85. https://doi.org/10.23947/1992-5980-2019-19-1-81-85

Образец для цитирования: К вопросу эффективности методов и алгоритмов решения оптимизационных задач с учетом специфики целевой функции / Е. Н. Остроух [и др.] // Вестник Донского гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 81–85. https://doi.org/10.23947/1992-5980-2019-19-1-81-85

Introduction. A general optimization problem is multicriteria with a number of restrictions in the form of equations and inequalities. To solve a multicriteria task, the following types of algorithms are used:

- a posteriori (sequential assignment algorithms);
- aprior;
- sensing;
- adaptive;
- approximation algorithms based on estimated accuracy of Pareto frontier.

The listed approaches enable to narrow down the initial set of feasible solutions to one or several points [1, 2].

Let us identify the features of the methods mentioned above. The use of a posteriori, aprior, and adaptive algorithms reduces the initial problem with a vector criterion to the problem with a single objective function, which can be solved through bioinspired algorithms or their combinations. The sensing method requires a lot of computational resources. The approximation approach is characterized by high efficiency, which is due to the possibility of parallelizing the computational process [2].

The solution to the unconditional optimization problem is based on one of two algorithms: penalty functions and (or) sliding tolerance. Both of these approaches are well represented in [2] and [3]. The task is to find the global optimum of the unconstrained function.

Note features of the objective function which helps to describe a large class of technical and economic problems. The objective function is often non-linear, in most cases it is not differentiable, not unimodal, and it has a complex topology of the tolerance region. Therefore, to find the global optimum of the presented problem, the authors of this paper have created hybrid algorithms. At this, pairs can combine:

- two bioinspired algorithms;
- a bioinspired and a classical algorithms;
- a bioinspired algorithm and one based on physical laws.

Thus, a hybrid based on Fourier series and a firefly algorithm was proposed in [4]. In [5], gradient and immune algorithms are combined, and in [6], swarm and gravity algorithms are paired. The combination of genetic and swarm approaches [7, 8, 9] was also considered for solving various optimization problems with an objective function (including the problems of learning neural networks of various topology).

The efficiency criterion of the bioinspired algorithm should be considered the number of iterations (steps) at which:

- the algorithm finds a result close enough to the optimum,
- the number of steps (time) is acceptable,
- the required accuracy of the algorithm is provided.

Materials and Methods. The combinations of bioinspired algorithms developed by the authors [4–10] were tested on Rosenbrock, Rastrigin, Griewank, and Schwefel [2] functions. It should be noted that the combinations reinforced the advantages of each algorithm of the pair and leveled their shortcomings. For each function, the hybrid showed the best performance and accuracy compared to single algorithms. However, the hybrids that produced the most effective result on one test function work much more poorly with other test functions. This observation leads to the following conclusion: the efficiency of bioinspired algorithms and their combinations with other algorithms depends much more on the objective function than on the features of the combination of algorithms. The same conclusion can be drawn from the NFL theorem (short for “no free lunch”) [11].

Consider that it is necessary to find the global optimum of some function that has the properties described above. There is a bank of bioinspired algorithms and (or) their combinations. And there is a bank with an objective function that does not have the unimodality property, with a large number of variables and parameters. What algorithm can solve the optimization problem for this function as smoothly as possible? For what types of functions will this algo-

rithm show the best result, and for which of them will it be inappropriate? It is obvious that it is impossible to develop such an algorithm. Algorithms may perform well on one function and be completely unacceptable for other functions.

Concerning the structure of bioinspired algorithms, it should be noted that a large number of empirical parameters used by them does not enable to evaluate the efficiency of such algorithms and their combinations in advance. (The opposite example is deterministic algorithms that successfully solve problems with linear, quadratic, strictly convex, unimodal functions).

For alternative estimates of bioinspired algorithms, public libraries with test problems are used, which allow comparing the known and new algorithms and their combinations [11, 12]. Sobolev Institute of Mathematics is one of the largest collections of test problems [12]. It presents various approaches to solving complex NP-complete optimization problems considering estimates of computational complexity. Among other libraries, narrow-focus ones can be noted: an open-source library for GAlib genetic algorithms [13]; libraries for building EAlib [14], Perl [15] evolutionary algorithms, and Java frameworks [16].

Research Results. The study on the features of genetic and population algorithms, performed by the authors, makes it possible to define the following advantages and disadvantages.

Genetic algorithms (GA) are based on the identity of the behavior principles of biological and technical systems. The GA use the principle of choosing the best decisions from the population available, which allows us to find the optimal solution to the problem. These algorithms give good diversity, as far as the information on solution point sets is processed in parallel. In these points, the optimum is based on the application of the objective function, and not on its various increments.

Ant colony optimizations simulate the principles of vital activity of the ant colony. There, the principle of autonomous functioning of each agent is combined with the activities of the colony as a whole, which allows solving rather complex optimization problems. The combination of ant colony algorithms with local search algorithms enables to quickly find the starting points of the optimum search. Such algorithms give the best results for large-scale problems. In terms of efficiency, they are very close to the problem-oriented and metaheuristic algorithms. This approach has shown good results in solving various practical tasks, for example, problems of traveling salesman, of optimal design of electronics, etc. Due to the possibility of using adjustable parameters, ant colony algorithms are applied in solving distribution and transport problems.

The following positive features of ant colony algorithms should be noted:

- for some tasks, they provide a more efficient solution than genetic ones or algorithms based on neural networks;
- the genetic algorithm stores information only about the previous generation, whereas ant algorithms store information about the entire colony, which is more efficient;
- random routing in the ant colony algorithm enables to exclude non-optimal initial solutions;
- the selection of parameters responsible for changing the optimization step enables to successfully use this algorithm in dynamic applications.

The disadvantages of ant colony algorithms include the following:

- theoretical analysis is difficult due to a sequence of random decisions, which is caused by changes in probability distributions during iterations;
- the algorithm convergence time cannot be predetermined, and to solve this problem, the ant algorithm is supplemented by local search methods;
- free parameters for work adjustment under solving a specific task are determined only experimentally.

When identifying extremes for composite multidimensional nonmonotonic functions, the swarm algorithms with equal probability determine the optimal element (an element with given properties) at any iteration. These algorithms are effectively used to optimize nonmonotonic functions in NP-complete problems, including distribution and transport ones. They are searching the only optimal element that determines the extremum (or sets of such elements). The swarm algorithm implements a search in the neighborhood of the best and selected sites in parallel. At this, it does not have any of the drawbacks of evolutionary methods, for example, it does not require a significant amount of memory to store a population of solutions.

Discussion and Conclusion. Analysis of the known hybrid algorithms has shown the following. One combination can give a good result on some test function (for example, Rosenbrock), but this hybrid will be significantly less efficient than other combinations on the other function. Hence, it is fair to say that the type of the objective function has much more significant impact on the result than a combination of algorithms. Therefore, the study of the form and type of the objective function enables to select the best combination.

References

1. Podinovskiy, V.V., Nogin, V.D. Pareto-optimal'nye resheniya zadach. [Pareto optimal problem solving.] Moscow: Fizmatlit, 2007, 255 p. (in Russian).

2. Karpenko, A.P. Sovremennye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennye prirodoy. [Modern search engine optimization algorithms. Algorithms inspired by nature.] Moscow: Bauman University Publ. House, 2014, 446 p. (in Russian).

3. Agibalov, O.I., Zolotarev, A.A., Ostroukh, E.N. Uslovnaya i bezuslovnaya optimizatsiya pri reshenii bioinspirirovannymi algoritmami. [Conditional and unconditional optimization under solving bioinspired algorithms.] Intellektual'nye tekhnologii i problemy matematicheskogo modelirovaniya: mat-ly vseros. konf. [Intellectual technologies and problems of mathematical modeling: Proc. All-Russian Conf.] Rostov-on-Don: DSTU Publ. Centre, 2018, pp. 21–22 (in Russian).

4. Ostroukh, E.N., Klimova, D.N., Markin, S. Reshenie zadach biznes-prognozirovaniya na osnove ryadov Fur'ye i algoritma svetlyachkov. [Solving business forecasting problems based on Fourier series and firefly algorithm.] Sistemnyy analiz, upravlenie i obrabotka informatsii: tr. VIII mezhdunar. seminara. [System analysis, management and information processing: Proc. VIII Int. Workshop.] Rostov-on-Don: DSTU Publ. Centre, 2017, pp. 153–158 (in Russian).

5. Agibalov, O.I. Optimizatsiya mnogomernykh zadach na osnove kombinirovaniya determinirovannykh i stokhasticheskikh algoritmov. [Combination of deterministic and stochastic algorithms for multidimensional tasks optimization.] Modern High Technologies, 2017, no. 9, pp. 7–11 (in Russian).

6. Ostroukh, E.N., Evich, L.N., Panasenko, P.A. Razrabotka gibridnogo algoritma resheniya optimizatsionnykh zadach prinyatiya resheniy i upravleniya. [Development of a hybrid algorithm for solving optimization decision-making problems and control.] Iskusstvennyy intellekt: problemy i puti ikh resheniya: mat-ly konf. [Artificial intelligence: problems and solutions: Proc. Conf.] Moscow, 2018, pp. 165–168. Available at: https://docplayer.ru/74455293-Programma-konferencii-iskusstvennyy-intellekt-problemy-i-puti-resheniya-2018.html (accessed 08.02.19) (in Russian).

7. Evich, L.N., Ostroukh, E.N., Panasenko, P.A. Razrabotka gibridnogo algoritma resheniya optimizatsionnoy zadachi s nelineynoy tselevoy funktsiey. [Development of hybrid algorithm for solution of optimization problem with nonlinear target function.] International Research Journal, 2018, no. 1 (1), pp. 61–65 (in Russian).

8. Evich, L.N., Ostroukh, E.N., Panasenko, P.A. Metody resheniya zadach optimizatsii s mul'timodal'noy tselevoy funktsiey na osnove gibridnykh algoritmov.[Development of hybrid algorithm for solution of optimization problem with nonlinear target function.] Prom-Inzhiniring (ICIE-2018): tr. IV mezhdunar. nauch.-tekhn. ko [Prom-Engineering (ICIE-2018): Proc. IV Int. Sci.-Tech. Conf.] Moscow, 2018, no. 1-1 (67), pp. 61-65 (in Russian).

9. Ostroukh, E.N., et al. Issledovanie kombinirovannogo algoritma pri obuchenii trekhsloynykh neyronnykh setey razlichnoy topologii. [Investigation of a combined algorithm for learning three-layer neural networks of different topologies.] Software & Systems, 2018, vol. 31, no. 4, pp. 673–676 (in Russian).

10. Wolpert, D.H., Macready, W.G. The no free lunch Theorems for optimization. *IEEE Transactions* on Evolutionary Computation, 1997, vol. 1, no. 1, pp. 67–82.

11. Rodzin, S.I. Algoritmy biostokhasticheskoy optimizatsii: dostizheniya, problemy teorii, trudoemkost'. [Algorithms of bio-stochastic optimization: achievements, problems of theory, complexity.] IS&IT'18: tr. kongressa po intellekt. sistemam i inform. tekhnologiyam. [IS&IT'18: Proc. Congress on intelligent systems and information technologies.] Taganrog: S.A. Stupin's Publ. House, 2018, vol. 2, pp. 141–158 (in Russian).

12. Discrete Location Problems. Benchmark library. Sobolev Institute of Mathematics; Russian Foundation for Basic Research. Available at: http://www.math.nsc.ru/AP/benchmarks/index.html (accessed: 11.05.18).

13. Wall, M. GAlib: A C++ Library of Genetic Algorithm Components. Massachusetts Institute of Technology. Available at: http://lancet.mit.edu/ga/ (accessed: 08.11.18).

14. EALib: An Evolutionary Algorithms Library. GitHub Inc. Available at: http://github.com/dknoester/ealib/ (accessed: 08.11.18).

15. Merelo, J. J. Library for doing evolutionary computation in Perl. Available at: http://cpeal.sourceforge.net/ (accessed: 08.11.18).

16. Dyer, D. W. Watchmaker Framework for Evolutionary Computation. Available at: http://watchmaker.uncommons.org (accessed: 08.11.18).

Received 02.11 .2018
Submitted 02.11.2018
Scheduled in the issue 15.01.2019

Authors:

Ostroukh, Evgeny N.,
associate professor of the Information Technologies Department, Don State Technical University, (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Eng.), associate professor,
ORCID:http://orcid.org/0000-0002-1384-0469,
eostr@donpac.ru

Evich, Lyudmila N.,
associate professor of the Mass Communications and Multimedia Technologies Department, Don State Technical University, (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Phys -Math.), associate professor,
ORCID:http://orcid.org/0000-0002-7886-0954,
evichlng@gmail.com

Chernyshev, Yury O.,
professor of the Production Automation Department, Don State Technical University, (1, Gagarin sq., Rostov-on-Don, 344000, RF), Dr.Sci. (Eng.), professor,
ORCID:http://orcid.org/0000-0002-4901-1101
myvnn@list.ru

Panasenko, Pavel A.,
senior assistant of Head of the Department of Organizing Scientific Research and Training of Academic and Teaching Staff, Krasnodar Higher Military School named after General of the Army S. M. Shtemenko (4, ul. Krasina, Krasnodar, 350063, RF), Cand.Sci. (Eng.),
ORCID:https://orcid.org/0000-0003-1264-1481
we_panacenko_777@mail.ru

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 86–92. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 86–92. ISSN 1992-5980 eISSN 1992-6006

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

UDC 519.683.4 https://doi.org/10.23947/1992-5980-2019-19-1-86-92

Arithmetic coder optimization for compressing images obtained through remote probing of water bodies [***]

R. V. Arzumanyan[1**]
[1] Institute of Computer Technology and Information Security, Southern Federal University, Taganrog, Russian Federation

Оптимизация арифметического кодера для сжатия изображений, полученных при дистанционном зондировании водных объектов [*]

Р. В. Арзуманян[1**]
[1] Институт компьютерных технологий и информационной безопасности Южного федерального университета, г. Таганрог, Российская Федерация

Introduction. The fast program algorithm of arithmetic coding proposed in the paper is for the compression of digital images. It is shown how the complexity of the arithmetic coder algorithm depends on the complexity measures (the input size is not considered). In the course of work, the most computationally complex parts of the arithmetic coder algorithm are determined. Performance optimization of their software implementation is carried out. Codecs with the new algorithm compress photo and video records obtained through the remote probing of water bodies without frame-to-frame difference.

Materials and Methods. In the presented paper, a selection of satellite images of the Azov Sea area was used. At this, the software algorithm of the arithmetic coder was optimized; a theoretical study was conducted; and a computational experiment was performed.

Research Results. The performance of the software implementation of the arithmetic coder is increased by the example of the VP9 video codec. Numerous launches of reference and modified codecs were made to measure the runtime. Comparison of the average time of their execution showed that the modified codec performance is 5.21% higher. The overall performance improvement for arithmetic decoding was 7.33%.

Discussion and Conclusions. Increase in the speed of the latest digital photo and video image compression algorithms allows them to be used on mobile computing platforms, also as part of the onboard electronics of unmanned aerial vehicles. The theoretical results of this work extend tools of the average-case complexity analysis of the algorithm. They can be used in case where the number of algorithm steps depends not only on the input size, but also on non-measurable criteria (for example, on the common RAM access scheme from parallel processors).

Введение. Предложенный в статье быстрый программный алгоритм арифметического кодирования предназначен для сжатия цифровых изображений. Показано, каким образом сложность алгоритма арифметического кодера зависит от критериев сложности (при этом размер входа не учитывается). В процессе работы определены наиболее вычислительно сложные части алгоритма арифметического кодера. Выполнена оптимизация производительности их программной реализации.

Кодеки с новым алгоритмом сжимают без учета межкадровой разницы фото- и видеоматериалы, полученные при дистанционном зондировании водных объектов.

Материалы и методы. В представленной научной работе использована подборка спутниковых снимков акватории Азовского моря. При этом оптимизирован программный алгоритм арифметического кодера, проведено теоретическое исследование, выполнен вычислительный эксперимент.

Результаты исследования. Увеличена производительность программной реализации арифметического кодера на примере видеокодека VP9. Для измерения времени выполнения произведены многочисленные запуски эталонного и модифицированного кодеков. Сравнение среднего времени их исполнения показало, что производительность модифицированного кодека на 5,21 % выше. Прирост общей производительности для арифметического декодирования составил 7,33 %.

Обсуждение и заключения. Увеличение скорости работы новейших алгоритмов сжатия цифровых фото- и видеоизображений позволяет применять их на мобильных вычислительных платформах, в том числе в составе бортовой электроники беспилотных летательных аппаратов. Теоретические результаты данной работы расширяют методы анализа сложности алгоритма в среднем случае. Они могут использоваться в ситуации, когда количество шагов алгоритма зависит не только от размеров входа, но и от неизмеримых критериев (например, от схемы обращения к общей оперативной памяти со стороны параллельных процессоров).

[*] The research is done within the frame of RSF project no. 17-11-01286.
[**] E-mail: roman.arzum@gmail.com
[***] Работа выполнена в рамках проекта РНФ 17-11-01286.

Keywords: arithmetical coding, performance optimiza tion, image compression, average-case algorithm complexity, video codec.

For citation: R.V. Arzumanyan. Arithmetic coder optimization for compressing images obtained through remote probing of water bodies. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 86–92. https://doi.org/10.23947/1992-5980-2019-19-1-86-92

Ключевые слова: арифметическое кодирование,оптимизация производительности, сжатие изображений, сложность алгоритма в среднем, видеокодек.

***Образец для цитирования*:** Арзуманян, Р. В. Оптимизация арифметического кодера для сжатия изображений, полученных при дистанционном зондировании водных объектов / Р. В. Арзуманян // Вестник Донского гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 86–92. https://doi.org/10.23947/1992-5980-2019-19-1-86-92

Introduction. Monitoring of the condition of the water body is often carried out through unmanned aerial vehicles (UAV), conducting aerial photography in the visible and infrared bands. The performance capabilities of the mobile camera, which the UAV has, impose a number of restrictions on the equipment that processes and stores the footage until the UAV returns. Specifically, the following factors should be considered.

1. Energy efficiency of the equipment that encodes the footage since it depends on the off-line operation time of the UAV.

2. Coding gain of photo and video data during the flight. High resolution images occupy a significant amount of the software memory, and this limits the amount of information that a UAV can accumulate.

Virtually all equipment for photo and video hardware supports the most common codec for compressing JPEG images. However, it is inferior to the most modern HEVC and VP9 codecs, which not only support video sequences, but also allow for better compression of single images. Thus, GoogleVP9 shows similar JPEG visual quality by the SSIM metric (structural similarity) and at the same time compresses images by 25–34% stronger [1].

Compared to the same JPEG, HEVC codec enables to improve the compression rate by 10–44% in PSNR metric (peak signal-to-noise ratio) [2]. However, a high compression ratio with a smaller bitstream size results in greater computational complexity of the HEVC and VP9 codecs [3, 4]. The architecture of the JPEG codec is presented in general in Fig. 1.

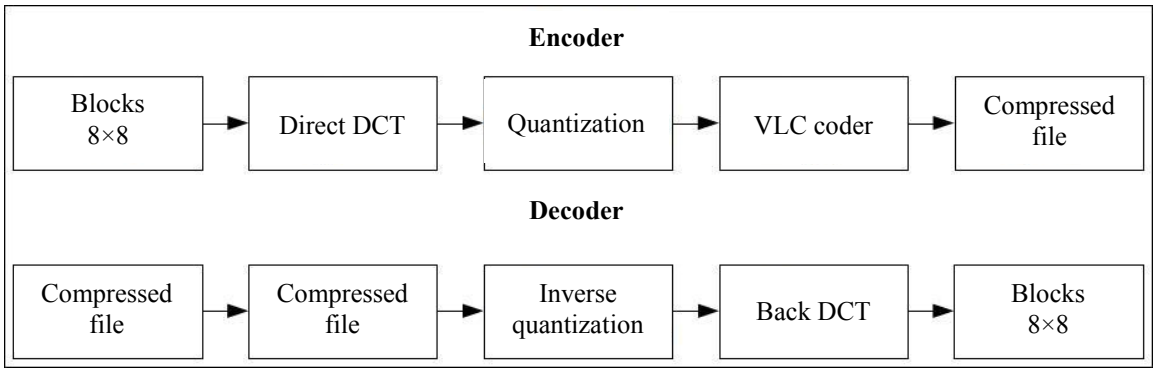


Fig. 1. JPEG codec flow chart

Here, VLC means compression through variable-length codes (variable length coding). The input frame is divided into blocks of fixed size (8×8). Each of them is subjected to direct discrete cosine transform (DCT), quantization of transform coefficients and subsequent entropic compression using the Huffman algorithm [4]. Discrete cosine transform is performed in the integer form [5]. Since the adoption of the JPEG standard in 1992, a number of fast algorithms have been developed. They allow the conversion to be carried out entirely in the CPU registers. Huffman's entropy compression is also not a complicated problem, so, even mobile processors in program mode can compress and decode JPEG images [6].

HEVC [6] and VP9 are fundamentally similar, and they are hybrid block codecs with splitting a frame into blocks of indeterminate length, intraframe prediction, discrete transform, and subsequent filtering to eliminate blocking artifacts. The blockdiagram of the HEVC codec is shown in Fig. 2.

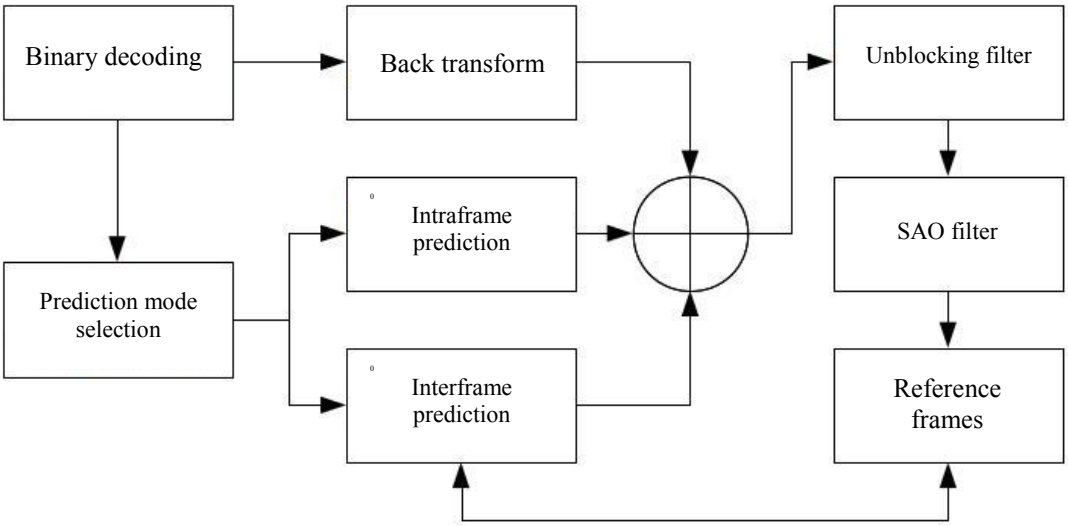


Fig. 2. HEVC codec flowchart

Here, the abbreviation SAO denotes the sample adaptive offset. In addition to the above image reconstruction algorithms, both codecs use context-adaptive binary arithmetic entropy coding, which is much more complicated than Huffman compression. With a high level of visual quality, it is arithmetic coding that occupies a significant part of the total decoder operation time. In general, the flowchart of the arithmetic coder is shown in Fig. 3.

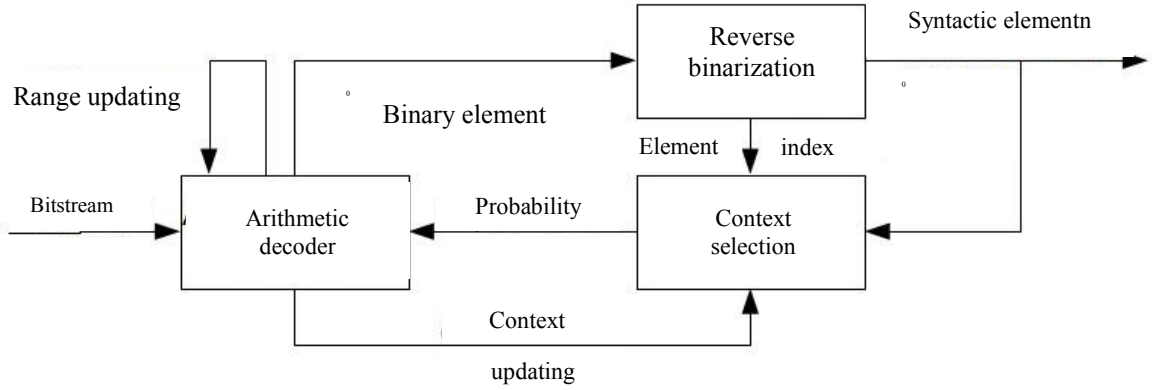


Fig. 3. Arithmetic coder flowchart

Codecs of a new generation are supposed to be used, among other things, for compressing images obtained through remote sensing of natural objects. In this case, it is necessary to optimize the operation of hybrid block codecs to ensure the processing of photo and video UAV data in real time. It is necessary to put more focus on the optimization of binary context-adaptive arithmetic coding, due to its essential complexity.

The objective of this study is to speed up the operation of the arithmetic decoder on mobile processors of the ARM architecture. This will improve the performance of Google VP9 video codecs and will allow for their application in the UAV on-board electronics for remote sensing of water bodies. The use of enhanced tools for compressing photos and videos obtained from aerial photography enables to increase the amount of stored data, improve their visual quality and resolution.

Main Part. Core components of a modern video codec include the binarization of syntactic elements of the bitstream and the adaptive binary encoding of these elements into the bitstream. This stage is not safe for the vectorization and parallelization, but it can be optimized through the statistical analysis of the input data. The main way to predict the running time of a program is to analyze the complexity of the corresponding algorithm. Distinction is provided between complexity at the best, worst and average case. $\supset x$ is input data of A algorithm which is used to calculate the output of y algorithm. The time function of the algorithm is denoted by $C_A^T(x)$, and the memory cost function – by $C_A^S(x)$. In the worst case, we will call $T_A(n) = \max_{\|x\|=n} C_A^T(x)$ functions of numeric argument the time and space complexity A

$$S_A(n) = \max_{\|x\|=n} C_A^S(x).$$

Consider a finite set of n size inputs:

$$X_n = \{x: \|x\| = n\}.$$

$\forall x \in X_n$ corresponds to the probability:

$$P_n(x) \in [0,1]: \sum_{x \in X_n} P_n(x) = 1.$$

The following expectation is called the average-case complexity:

$$T_A = \sum_{x \in X_n} P_n(x) C_A^T(x),$$

$$S_A = \sum_{x \in X_n} P_n(x) C_A^S(x).$$

The described approach is classical for analyzing average-case complexity of the algorithm, and it is described in detail in [7, 8]. Note the reasons why the application of this method in practice may be difficult or impractical.

1. Difference between the number of steps of the algorithm in theory and the number of processor cycles required to perform the step in practice. Thus, the majority of modern central processors produce addition and multiplication per cycle, while the remainder of the division is calculated for dozens of cycles.

2. Hardware features of the memory system. Modern computers use a multi-level memory hierarchy. Its components operate at different speeds. Memory calls take considerably longer than register operations.

3. Optimizing compilers and hardware planners. When building executable files, optimizing compilers transform dramatically the code without changing the program state engine. CPU hardware schedulers change the procedure of executing instructions for greater performance and predict conditional transitions, while cache controllers read from memory in blocks.

4. In case of the software algorithms implementation on the general-purpose processors, software components interfere. For example, the task scheduler shares processor time, and parallel processes that have multiple threads can run on a variable number of processor cores.

The proposed modification of the algorithm complexity serves as a theoretical addition to practical tools for measuring performance, such as, for example, profiling and instrumentation of the program code. For the first time, the method of splitting the algorithm inputs into complexity classes was presented in [9]. Consider A algorithm and a set of all possible inputs:

$$G: \{g_1, g_2 \ldots\},$$

and also all possible samples from G, different in size and composition:

$$g_i: \{g_i^1, g_i^2 \ldots\}.$$

A set of criteria for the complexity measure of the algorithm implementation (for example, the number of processor cycles, run-time, etc.):

$$\propto_i: g_i \to \mathbb{R}.$$

A set of complexity measures:

$$A: \{\propto_1, \propto_2 \ldots\}.$$

This set has the following properties.

1. $\forall \propto_1, \propto_2 \in A: \propto_1 \neq \propto_2$ — all A elements are different.
2. $\forall \propto_i \in A$ splits G into the set of equivalence complexity classes

$$G(\alpha_i) = \{g_i^1 \cap g_i^2 \ldots\}.$$

3. All samples from $G(\alpha_i)$ are equally complex:

$$\sqsupset \alpha_i \in A, g_i^k \in G(\alpha_i), \alpha_i: g_i^k \to r_i, k \in [0, \|G(\alpha_i)\|], r_i \in \mathbb{R}.$$

Thus, all elements of A are different, and they can be reordered so that the complexity function of the criterion is nondecreasing throughout the whole set of criteria. The expected complexity is similar to the estimates of the average-case complexity of the algorithm for the discrete and continuous probability of complexity. For the discrete case:

$$R(A) = \sum_{\alpha_i \in A} r_i p_i,$$

for the continuous case:

$$R(A) = \int_A r \, dF(r).$$

The method in question is applicable for analyzing the complexity of an arithmetic codec. The process of entropy compression [10, 11] can be divided into component parts.

1. Binarization or conversion of the coded character (syntactic element of the compressed bitstream) into a string made up of zeros and ones (bit string).

2. Context modeling for compressing syntactic elements in the normal mode. This step is not performed for syntactic elements whose statistical distribution is close to normal, and they are coded in the bypass mode.

3. Arithmetic coding of a bit string.

Consider in more detail the arithmetic decoding scheme of the Google VP9 codec, namely the part that is associated with the subexponential coding of syntactic elements.

Imagine an algorithm of the subexponential coding in general [12]. The first step includes the calculation of the variables:

$$b = \begin{cases} k: n < 2^k \\ \lfloor \log_2 n \rfloor: n \geq 2^k, \end{cases}$$

$$u = \begin{cases} 0: n < 2^k \\ b - k + 1: n \geq 2^k, \end{cases}$$

where *k* is parametrical value, it is 4 for the Google VP9 codec.

In the second step, the unary code $u(u+1)$ bit is complemented by n low-order bits. The code length is equal to:

$$u + 1 + n = \begin{cases} k + 1: n < 2^k \\ 2\lfloor \log_2 n \rfloor - k + 2: n \geq 2^k. \end{cases}$$

Hence, literal decoding is reduced to decoding the bits that compose it in the loop. To optimize the performance of this algorithm, it is important to know the probability distribution of literal lengths. Literals occupying the largest number of bits in a compressed bitstream (such as inverse transform coefficients and motion vectors) are coded in series, so there is a high probability that the distribution of literal lengths in a compressed bitstream will be constant with multiple repeats of elements of the same value. To test this hypothesis, experimental data on the distribution of literal lengths in a set of satellite images of the Sea of Azov is acquired (Table 1).

Table 1

Literal length, bit	1	2	3	4	5	6
Probability, %	0.94	0	67.35	18.25	0	13.46

The literals of 3, 4, and 6 bits in length are most probable. The maximum possible literal length for this sequence is just 6 bits. This fact is essential for software optimization of the function of subexponential decoding of a literal. As part of optimizing a real codec, the run-time criterion is of prime interest. To obtain the set of difficulties: $\{r_0, \ldots, r_4\}$, we will profile the program performance.

A set of unique elements (*R*) will make up a set of criteria of the run-time complexity (A).

The following approaches that are based on the data obtained are applied for the optimization.

1. Memoization of calculating the literal length to decode a series of literals of equal length.
2. Unwinding of a cycle of subexponential decoding of a literal.
3. More efficient algorithm for calculating the number of bits in a literal.
4. More efficient use of the processor registers immediately within the arithmetic decoding function.

The implementation of points 1 and 4 is fairly obvious; therefore, we consider in more detail points 2 and 3. Compiling of a decoded literal by bits decoded from a compressed bitstream occurs within the subexponential decoding function. In this case, the bottleneck is a loop with varying number of iterations. It can be replaced with a switch-case set without break at the end. This technique is known as the Duff's device method. It allows replacing several loop iterations through sequentially execution of the instructions without the need for conditional transitions. The bit shift amount is a constant that does not need to be read from the register - loop counter.

Code Listing 1: The original literal decoding function

```
static int vp9_read_literal(vp9_reader *br, int bits)
{
int z = 0, bit;
 for (bit = bits – 1; bit >= 0; bit –)
   z |= vp9_read_bit(br) << bit;
return z;
}
```

Code Listing 2: Modified literal decoding function

```
static int vp9_read_literal(vp9_reader *br, int bits) {
 register int z = 0;
 switch(bits -1){
   case 6: z |= vp9_read(br, 128) << 6;
   case 5: z |= vp9_read(br, 128) << 5;
   case 4: z |= vp9_read(br, 128) << 4;
```

```
  case 3: z |= vp9_read(br, 128) << 3;
  case 2: z |= vp9_read(br, 128) << 2;
  case 1: z |= vp9_read(br, 128) << 1;
  case 0: z |= vp9_read(br, 128);
break;
}
return z;
}
```

Another bottleneck is the calculation of the number of literal bits in the while loop [13]. This is a worse solution because the number of loop iterations is unpredictable. Instead, a fast bit-counting algorithm was used [14, 15], which performs the calculation for a fixed number of steps without conditional transitions.

Code Listing 3: Fast counting of the number of bits in a literal

```
Unsig ned intv; // 32-bit argument
Register unsig ned intr; // variable for the number of bits
register unsigned int shift;
r = (v > 0xFFFF) << 4;
v >>= r;
shift = (v > 0xFF) << 3;
v >>= shift;
r |= shift;
shift = (v > 0xF) << 2;
v >>= shift;
r |= shift;
shift = (v > 0x3) << 1;
v >>= shift;
r |= shift;
r |= (v >> 1);
```

Numerous starts of the reference and modified codecs were made to measure the runtime. In this case, their average run-time was compared. It has been found that the performance of the modified codec is 5.21% higher. The increase in overall performance for arithmetic decoding was 7.33%.

Conclusions. The operation of the arithmetic coder as a component of the video codec has been optimized using the example of the Google VP9 standard. To solve this problem, a modification of the method for analyzing the average-case complexity of algorithm has been proposed. The approach is based on the partitioning of the set of inputs into equivalence complexity classes. The considered method enables to predict the average-case complexity of the algorithm when the number of steps of the algorithm and the time of its execution depend on difficult-to-measure parameters, which is typical of the context-adaptive arithmetic coding. The proposed method has been applied to optimize the speed of the arithmetic binary coder (using the example of the Google VP9 codec) for the image compression problems obtained under remote sensing of water bodies. The research results make it possible to apply advanced methods of compressing photo and video data obtained through the aerial photography of water bodies. Thus, it is possible to increase an amount of the accumulated data, to improve the visual quality and resolution of the footage by 25–34% (according to the SSIM visual quality metric) and to increase the speed of the arithmetic coder by 7%.

References

1. WebP Compression Study. Google Developers. Available at: https://developers. google.com/ speed/webp/docs/webp_study (accessed 01.02.19).

2. Nguyenand, T., Marpe, D. Objective Performance Evaluation of the HEVC Main Still Picture Profile. *IEEE Transactions on Circuits and Systems for Video Technology*, 2015, vol. 25, no. 5, pp. 790–797.

3. Blahut, R. Bystrye algoritmy tsifrovoy obrabotki signalov. [Fast algorithms for signal processing.] Moscow: Mir, 1989, 448 p. (in Russian).

4. Wallace, G.K. The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, 1992, vol. 38, no. 1, pp. XVIII–XXXIV.

5. Dvorkovich, A.V., Dvorkovich, V.P. Tsifrovye videoinformatsionnye sistemy (teoriya i praktika). [Digital video information systems (theory and practice).] Moscow: Tekhnosfera, 2012, 1009 p. (in Russian).

6. Asaduzzaman, A., Suryanarayana, V. R., Rahman, M. Performance-power analysis of H.265/HEVC and H.264/AVC running on multicore cache systems. Intelligent Signal Processing and Communications Systems. Available at: https://ieeexplore.ieee.org/document/6704542 (accessed 01.02.19).

7. Sedgewick, R., Wayne, K. Algorithms. Fourth edition. Upper Saddle River: Addison-Wesley, 2016, 960 p.

8. Cormen, T.H., et al. Introduction to Algorithms. 3rd edition. Cambridge; London: The MIT Press, 2009, 1296 p.

9. Welch, W.J. Algorithmic complexity: three NP — hard problems in computation all statistics. Journal of Statistical Computation and Simulation, 1982, vol. 15, no. 1, pp. 17–25.

10. High efficiency video coding. Fraunhofer Heinrich Hertz Institute. Available at: http://hevc.info/ (accessed: 01.02.19).

11. Sze, V., Budagavi, M. Parallelization of CABAC transform coefficient coding for HEVC. Semantic Scholar. Allen Institute for Artificial Intelligence Logo. Available at: https://www.semanticscholar.org/ paper/Parallelization-of-CABAC-transform-coefficient-for-Sze-Budagavi/0653a22ff7b82bdd0130cea8b597a7024ab46882 (accessed: 01.02.19).

12. Salomon, D., Motta, G. Handbook of data compression. London; Dordrecht; Heidelberg; New York: Springer-Verlag, 2010, 1360 p.

13. Anderson, S. E. Bit Twiddling Hacks. Available at: https://graphics.stanford.edu/~seander/bithacks.html (accessed 01.02.19).

14. Gervich, L.R., Shteinberg, B.Y., Yurushkin, M.V. Programmirovanie ekzaflopsnykh sistem. [Exaflops systems programming.] Open Systems. DBMS. 2013, vol. 8, pp. 26–29 (in Russian).

15. Warren, Jr., G.S. Algoritmicheskie tryuki dlya programmistov. [Algorithm programming tricks.] 2nd ed. Moscow: Williams, 2013, 512 p. (in Russian).

Received 02.11.2018
Submitted 03.11.2018
Scheduled in the issue 15.01.2019

Author:

Arzumanyan, Roman V.,
postgraduate of the Intelligent Multiprocessor Systems Department, Institute of Computer Technology and Information Security, Southern Federal University (22, ul. Chekhova, Taganrog, Rostov Region, 347922, RF),
ORCID: https://orcid.org/0000-0003-3370-5093
roman.arzum@gmail.com

Vestnik of Don State Technical University. 2019. Vol. 19, no. 1, pp. 93-100. ISSN 1992-5980 eISSN 1992-6006
Вестник Донского государственного технического университета. 2019. Т. 19, № 1. С. 93-100. ISSN 1992-5980 eISSN 1992-6006

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

UDC 62-50 https://doi.org/10.23947/1992-5980-2019-19-1-93-100

Method of terminal control in ascent segment of unmanned aerial vehicle with ballistic phase *

N. Y. Polovinchuk[1], S. V. Ivanov[2], M. Y. Zhukova[3], D. G. Belonozhko[4] **
[1] Moscow State Technical University of Civil Aviation, Rostov Branch, Rostov-on-Don, Russian Federation
[2,4] Krasnodar Higher Military School named after army general S. M. Shtemenko, Krasnodar, Russian Federation
[3] Don State Technical University, Rostov-on-Don, Russian Federation

Способ терминального управления на участке выведения беспилотного летательного аппарата с баллистической фазой полета ***

Н. Я. Половинчук[1], С. В. Иванов[2], М. Ю. Жукова[3], Д. Г. Белоножко[4]**
[1] Ростовский филиал московского государственного технического университета гражданской авиации, г. Ростов-на-Дону, Российская Федерация
[2,4] Краснодарское высшее военное училище им. С. М. Штеменко, г. Краснодар, Российская Федерация
[3] Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация

Introduction. The solution to the problem on the centroidal motion control synthesis (guidance problem) of an unmanned aerial vehicle (UAV) with long-range capabilities in the boost phase is considered. Control condition requires optimum fuel consumption. The principle of dynamic programming considering the restrictions to the vector modulus of the thrust output is used to solve the problem. The implementation of terminal guidance requires the formation of control as a function of the object state at the end of the ascent phase. The attainment of these boundary conditions determines the further transition to the ballistic flight phase.

Materials and Methods. Bellman's principle of dynamic programming is the most reasonable from the point of view of the implementability of the computationally efficient on-board algorithms and the solution to the problems in the form of synthesis. With natural scarcity of thrust and energy resources on board, this principle enables to obtain solutions free from the switching functions. In this case, the optimal control is a smooth function (without derivative discontinuity) of the current and final parameters of the UAV.

Research Results. A new algorithmic method for the synthesis of terminal motion control is developed. Its difference is that the UAV movement control in the ascent phase is formed by the function of the motion actual and terminal parameters. This ensures movement along an energetically optimal trajectory into the given region of space. The problem solution results enable to build closed terminal guidance algorithms for

Введение. Статья посвящена решению задачи синтеза управления движением центра масс (задача наведения) беспилотного летательного аппарата (БЛА) с большой дальностью полета на разгонном участке. Условие управления: оптимальный расход топлива. Для решения задачи используется принцип динамического программирования с учетом ограничений на модуль вектора тяги двигателя. Реализация терминального наведения требует формирования управления как функции состояния объекта в конце участка выведения. Достижение этих граничных условий определяет дальнейший переход к баллистической фазе полета.

Материалы и методы. Принцип динамического программирования Беллмана является наиболее рациональным с точки зрения реализуемости эффективных в вычислительном отношении бортовых алгоритмов и решения задачи в форме синтеза. При естественной ограниченности величины тяги и энергетических ресурсов на борту данный принцип позволяет получить решения, не содержащие функции переключения. Оптимальное управление в этом случае является гладкой функцией (без разрыва производной) текущих и конечных параметров БЛА.

Результаты исследования. Разработан новый алгоритмический способ синтеза терминального управления движения. Его отличие в том, что управление движением БЛА на разгонном участке траектории формируется функцией текущих и конечных параметров движения. Таким образом обеспечивается движение по энергетически оптимальной траектории в заданную область пространства. Результаты решения задачи позволяют строить замкнутые алгоритмы терминального наведения для разгонного участка траектории БЛА с большой дальностью полета. Такие алгоритмы обладают

* The research is done within the frame of the independent R&D.
** E-mail: npolovinchuk@mail.ru, sta399@yandex.ru, marg88@list.ru, staeer@rambler.ru
*** Работа выполнена в рамках инициативной НИР.

the boost phase of the UAV trajectory with long-range capabilities. Such algorithms have good convergence and injection accuracy due to the prediction of parameters during the flight at a shorter time interval.

Discussion and Conclusions. The most preferred is the principle of dynamic programming. It should be used when solving the problem on the centroidal motion control synthesis (guidance problem) of the UAV with long-range capabilities in the boost phase.

хорошей сходимостью и точностью выведения за счет прогнозирования параметров в процессе полета на сокращающемся интервале времени.

Обсуждение и заключения. Наиболее предпочтительным представляется принцип динамического программирования. Именно его следует использовать при решении задачи синтеза оптимального по расходу топлива управления движением центра масс (задача наведения) БЛА с большой дальностью полета на разгонном участке.

Keywords: unmanned aerial vehicle (UAV), terminal guidance, direction cosines, pitching angle, angle of attack, boundary conditions, boost phase, ballistic flight phase.

Ключевые слова: беспилотный летательный аппарат (БЛА), терминальное наведение, направляющие косинусы, угол тангажа, угол атаки, граничные условия, разгонный участок, баллистическая фаза полета.

For citation: N.Y. Polovinchuk, et al. Method of terminal control in ascent segment of unmanned aerial vehicle with ballistic phase. Vestnik of DSTU, 2019, vol. 19, no. 1, pp. 93-100. https://doi.org/10.23947/1992-5980-2019-19-1-93-100

***Образец для цитирования*:** Способ терминального управления на участке выведения беспилотного летательного аппарата с баллистической фазой полета / Н. Я. Половинчук [и др.] // Вестник Донского гос. техн. ун-та. — 2019. — Т. 19, № 1. — С. 93-100. https://doi.org/10.23947/1992-5980-2019-19-1-93-100

Introduction. Currently, the capabilities and scope of applicability of the unmanned aerial vehicles (UAV) have increased significantly. This is primarily due to the UAV flying range extension. Vehicles with a ballistic phase of flight should be provided with the control in various areas, including the UAV positioning. For the control design, it is reasonable to apply the principle of terminal guidance. The control should be developed as a function of the terminal motion variables, and not as a temporal function.

A great many publications are devoted to this problem solution, but the task described above holds relevance. In particular, the development of highly computationally efficient algorithmic methods of terminal guidance is of interest. At the same time, the features universal for various types of launch vehicles should be considered. They have adaptive features and in a certain sense meet the optimality requirements.

Considering the guidance task, such a control, which uses the minimum amount of fuel, is valuable. The synthesis of optimal control is based on the use of the Bellman dynamic programming method [1].

Materials and Methods. The solution to the problem of the synthesis of optimal control of the UAV motion in the boost phase of power-on flight has been studied in many papers. However, the optimal control solution obtained in most cases is reduced to the implementation of a time or parametric program. We will solve the problem of finding the optimal control in the UAV ascent phase in the following formulation. The UAV motion parameters are known as X_0, Y_0, Z_0 coordinates of the current path point, obtained through solving the problem of navigation. $X_к$, $Y_к$, $Z_к$ parameters of the final point satisfy the boundary state that fixes the transition to the ballistic phase of flight determined by the hypersurface in the phase space. The condition of $S_k(X_к, Y_к, Z_к) = 0$ is satisfied by a whole set of finite parameters. It is required to synthesize optimal control in the task of UAV guidance, which ensures its transfer from the initial state to the hypersurface of the final conditions.

The optimality criterion is the amount of fuel consumed in the boost phase:

$$m(t) = \int_0^t \dot{m}(t)dt\,, \tag{1}$$

where $\dot{m}(t)$ is fuel mass flow rate.

The following system of differential equations is adopted as a mathematical model of the centroidal motion:

$$\dot{\bar{\mathrm{R}}}(\mathrm{t}) = \bar{\mathrm{V}}(\mathrm{t})\,, \tag{2}$$
$$V(t) = \dot{W}(t)\bar{E}_w(t) + \bar{g}(\bar{r})\,,$$

where $\bar{R}(t)$ is radius-vector, $\bar{g}(\bar{r})$ is terrestrial attraction vector, $\dot{W}(t)$ is module of control acceleration vector.

The module of the control acceleration vector ($\dot{W}(t)$) is a specified temporal function, and it is determined by the UAV engine performance. Unknown is the unit vector of control acceleration – $\bar{E}_w(t)$. When solving the optimization problem, it will determine the required properties of the UAV guided motion in the boost phase. There is no need in the

thrust vector control (in a sense, this is apparent acceleration amount – $\dot{W}(t)$). A more rational approach is to maximize this value for the engine booster. This ensures the application of the solutions obtained for the case of using solid fuel engines. The UAV engine performance is stable enough, and at $\dot{m}(t) = const$ constant fuel-flow rate, the optimal control is determined by the combustion duration. In this case, the functional (1) will be a function of the upper limit of integration. Hence, the problem of minimizing the fuel amount turns into an equivalent task of minimizing the flight time, and speed-of-response will be the optimality criterion. The task of the synthesis is to find the orientation of the thrust vector of the UAV engine, which is determined by the direction cosines of the thrust vector ($\bar{P}(t)$) as a function of the current parameters and the final state.

We admit two assumptions. The first is as follows. Since a large part of the boost phase lies outside the dense atmosphere, we will not consider the angle rate of evolution of the thrust vector in space (longitudinal axis of the UAV) as constraint. The second assumption is the following. Assume that the UAV motion is passing in a predetermined plane.

The UAV final state for the transition to the ballistic phase of flight is fixed by satisfying the following boundary condition [2]:

$$\begin{aligned} S_{1k} &= \left(V_{xk}y_k - x_kV_{yk}\right)\left[V_{xk}\left(y_k - y_ц\right) - V_{yk}\left(x_k - x_ц\right)\right] - \\ &- \pi_0\left(x_ц^2 + y_ц^2\right)^{1/2}\left[1 - \frac{x_kx_ц + y_ky_ц}{\left(x_ц^2 + y_ц^2\right)^{1/2}\left(x_k^2 + y_k^2\right)^{1/2}}\right] = 0. \end{aligned} \tag{3}$$

Here, $x_k, y_k, x_ц, y_ц$ are, respectively, the coordinates of the initial point and the point of the starting of the UAV operation on the final path segment. The current value of the boundary condition ($S_{1k}(t_0)$) is a measure of non-compliance with the condition (3). A mathematical notation for this condition corresponds to the hypersurface, which is a smooth function of phase coordinates and describes the entire family of possible UAV ascent trajectories [2].

The boundary conditions are specified for the central field of attraction. For this case, the equations of the UAV motion will have the following form:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\frac{\pi_0}{r^3}x_1 + \frac{1}{m}P\cos\alpha_1, \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= -\frac{\pi_0}{r^3}x_3 + \frac{1}{m}P\cos\alpha_2. \end{aligned} \tag{4}$$

Here, $\pi_0 = f \cdot M_з$ is constant of the central field of the earth's attraction equal to the product of the gravitational constant (f) and the mass of the Earth ($M_з$); $r = (x_1^2 + x_3^2)^{1/2}$, $x_1 = x$, $x_2 = V_x$, $x_3 = y$, $x_4 = V_y$; P is thrust vector value; X_A is aerodynamic drag force; Y_A is aerodynamic lift.

Since the UAV is fitted up with a steerable thruster, its direction cosines will determine its pointing. In this case, the handling constraint will be determined by the ratio:

$$\left\|\bar{P}(t)\right\| = \left\{\left[P(t)\cos\alpha_1(t)\right]^2 + \left[P(t)\cos\alpha_2(t)\right]^2\right\}^{1/2} \le \left|\bar{P}^0(t)\right|. \tag{5}$$

The handling constraint (5) is "hypersphere restriction". This implies the solution in which optimal control is not piecewise constant, without switching. In this case, the time optimal ascent trajectory in the phase space has no "angles" and no discontinuities of the derivative. The time optimal control will be a nonlinear, continuous function of the boundary conditions (3) [3].

The synthesis task is formulated as follows. The control object from an arbitrary current state, taken as the initial one and determined by the current value (S_{1k}) at the time (t_0), is transferred to the hypersurface of the boundary condition $S_{1k} = 0$ at the time (t_k). This takes into account the control acceleration amount (engine thrust module). It is required to find the optimal control in the form of synthesis, which provides such a transfer in the shortest possible time.

The plant state at the final instance satisfies the boundary condition $S_{ik}[X(t_k)] = 0$ and determines the moment of transition to the ballistic phase of flight.

In accordance with R. Bellman's dynamic programming principle [4], the necessary and sufficient condition for optimality of the formulated problem will be the ratio:

$$\min_{u \in U}\left[\sum_{i=1}^{n}\frac{\partial T^0}{\partial x_i}f_i(\bar{x},\bar{u})\right] = -1. \tag{6}$$

Considering the mathematical model of the object specified by the system (4), the dynamic programming equa-

tion [4] with the optimality criterion, is determined by the relation:

$$\min_{u\in U}\left[x_2\frac{\partial T^0}{\partial x_1}-\left(\frac{\pi_0}{r^3}x_1-\frac{1}{m}P_x\right)\frac{\partial T^0}{\partial x_2}+x_4\frac{\partial T^0}{\partial x_3}-\left(\frac{\pi_0}{r^3}x_3-\frac{1}{m}P_y\right)\frac{\partial T^0}{\partial x_4}\right]=-1, \tag{7}$$

where $P_x = P\cos\alpha_1, P_y = P\cos\alpha_2$.

Due to the methodology of Bellman's dynamic programming, the minimization can be carried out through the application of the Schwartz inequality [3] to the relation (5). This will significantly simplify the solution to the optimization problem. Then, the expression (7) with regard to inequality (5) will be determined by the relation:

$$\min_{u\in U}\left(\frac{1}{m}P_x\frac{\partial T^0}{\partial x_2}+\frac{1}{m}P_y\frac{\partial T^0}{\partial x_4}\right)=-P^0(t)\left[\left(\frac{1}{m}\frac{\partial T^0}{\partial x_2}\right)^2+\left(\frac{1}{m}\frac{\partial T^0}{\partial x_4}\right)^2\right]^{1/2}. \tag{8}$$

In this case, the Hamilton-Jacobi equation will be as follows:

$$\begin{gathered}x_2\frac{\partial T^0}{\partial x_1}-\frac{\pi_0}{r^3}x_1+x_4\frac{\partial T^0}{\partial x_3}-\frac{\pi_0}{r^3}x_3-P^0(t)\times\\ \times\left[\left(\frac{1}{m}\frac{\partial T^0}{\partial x_2}\right)^2+\left(\frac{1}{m}\frac{\partial T^0}{\partial x_4}\right)^2\right]^{1/2}=-1.\end{gathered} \tag{9}$$

The equation (9) is solved considering the given boundary condition:

$$(x_1, x_2, x_3, x_4)\in S^*_{1k}. \tag{10}$$

The calculation of the partial derivatives of S for phase variables (x_1, x_2, x_3, x_4) gives the following dependences:

$$\begin{aligned}\frac{\partial S^0}{\partial x_1}&=\left(-2x_2x_3x_4+x_2x_4x_{3ц}+2x_4^2x_1-x_4^2x_{1ц}\right),\\ \frac{\partial S^0}{\partial x_2}&=\left(2x_3^2x_2-2x_1x_4x_3-2x_2x_3x_{3ц}+x_1x_4x_{3ц}+x_4x_3x_{1ц}\right),\\ \frac{\partial S^0}{\partial x_3}&=\left(2x_2^2x_3-2x_2x_4x_1-x_2^2x_{3ц}-x_2x_4x_{1ц}\right),\\ \frac{\partial S^0}{\partial x_4}&=\left(2x_1^2x_4-2x_1x_2x_3+2x_4x_1x_{1ц}+x_2x_3x_{3ц}+x_2x_3x_{1ц}\right).\end{aligned} \tag{11}$$

Through substituting the relations (11) into (9) and performing simple transformations, we obtain the dependence:

$$\begin{gathered}\{\frac{\pi_0}{r^3}(x_1-x_3)-2x_2x_4^2x_{1ц}-\frac{P^0}{m}\times\\ \times[\left(2x_3^2x_2-2x_1x_4x_3-2x_2x_3x_{3ц}+x_1x_4x_{3ц}+x_4x_3x_{1ц}\right)^2+\\ +\left(2x_1^2x_4-2x_1x_2x_3+2x_4x_1x_{1ц}+x_2x_3x_{3ц}+x_2x_3x_{1ц}\right)^2]^{1/2}\}\frac{\partial T^0}{\partial S_1}=-1\end{gathered} \tag{12}$$

with the boundary conditions

$$T^0(S_1)=0, \text{ at } (x_1, x_2, x_3, x_4)\in S^*_{1k}. \tag{13}$$

The expression in curly brackets in (12), denoted by $S'(x_i)$, can be written in a compact form:

$$S'(x_i)\cdot\frac{\partial T^0}{\partial S_1}=-1. \tag{14}$$

The equation (14) with boundary conditions (13) can be solved in various ways, for example, by the method of characteristics [5]. However, it is more rational to use the following method.

The expression (11) defines the optimal control structure [6]:

$$\begin{aligned}\bar{P}_{opt}(\bar{x})&=-PE^{onm}(\bar{x}),\\ \dot{\bar{W}}^{onm}(\bar{x})&=-\dot{W}E^{onm}(\bar{x}),\end{aligned} \tag{15}$$

where $E(\bar{x})$ is unit thrust vector.

The ratio (15) in the scalar form:

$$\begin{aligned}P_x^{opt}&=P\cos\alpha_1,\\ P_y^{opt}&=P\cos\alpha_2,\end{aligned} \tag{16}$$

$$\cos\alpha_1 = \frac{\dfrac{\partial T^0}{\partial x_2}}{\left[\left(\dfrac{\partial T^0}{\partial x_2}\right)^2 + \left(\dfrac{\partial T^0}{\partial x_4}\right)^2\right]^{1/2}},$$
$$\cos\alpha_2 = \frac{\dfrac{\partial T^0}{\partial x_4}}{\left[\left(\dfrac{\partial T^0}{\partial x_2}\right)^2 + \left(\dfrac{\partial T^0}{\partial x_4}\right)^2\right]^{1/2}}. \quad (17)$$

The equation (9) and relations (17) can be put in a more convenient form for the following reasons. According to the formation of a set of perturbed trajectories, through varying the control, it is possible to construct a hypersurface with equal ascent time, that is, an isochronous surface, in the phase space. Indeed, for each phase trajectory point, we will calculate the optimality criterion value: $J(t_1), J(t_2), \ldots, J(t_k)$.

Hence, we obtain a set of trajectories for each $t \in [t_н, t_k]$. Owing to the continuity of the dependence of $x(t)$ and J on the variable control, a set of trajectories forms a surface in the phase space (X). This boundary surface formed by the set of vectors $x[t_i, J(t_i)]$, is convex and smooth. For transition conditions to the ballistic flight phase, the isochronous surface has a tangency point with the hypersurface of the boundary conditions.

Under the qualitative implementation of the optimal control, the distance in the phase space between the hypersurface of the boundary conditions $S_k = 0$ and the isochronous surface $T(x, R_ц, t) = 0$ will decrease. During some time, the two surfaces will have a common point (Fig. 1).

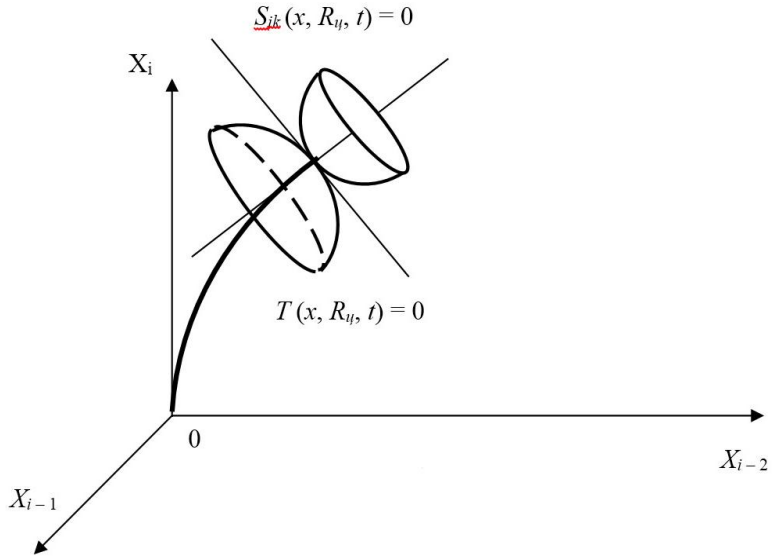


Fig. 1. Geometric interpretation of hypersurfaces in phase space under guidance in boost phase

At the point S_k and T^0, they have a common tangency, that is, their gradients coincide in the phase space. In this way:

- existence of the optimal trajectory $x_{opt}(t)$ and optimal control $u_{opt}(t)$ is noted;
- termination of the boost phase of the trajectory ($t = t_k$) and the transition to the ballistic flight phase are determined.

At the tangency point, the isochronous surface and the hypersurface of the boundary conditions have a common tangency and a normal. Mathematically, the condition for the existence of a common normal is determined by the expression [2]:

$$\left\{\frac{\partial T^0}{\partial x_i}\right\} = A\left\{\frac{\partial S_{ik}}{\partial x_i}\right\}, \quad (18)$$

where A is constant determined from the analysis of the convexity of both hypersurfaces.

From the relations (11) and (18), we transform (17). We obtain the expression for the direction cosines of the thrust vector as a function of the current and final motion parameters:

$$A = \frac{\left(-2x_2x_3x_4 + x_2x_4x_{3ц} + 2x_4^2x_1 - x_4^2x_{1ц}\right)}{\left[\left(-2x_2x_3x_4 + x_2x_4x_{3ц} + 2x_4^2x_1 - x_4^2x_{1ц}\right)^2 + \left(2x_4x_1^2 - 2x_1x_2x_3 + 2x_1x_4x_{1ц} + x_2x_3x_{3ц} + x_2x_3x_{1ц}\right)^2\right]^{1/2}}, \quad (19)$$

$$B = \frac{\left(2x_1^2x_4 - 2x_2x_1x_3 + 2x_4x_1x_{1ц} + x_2x_3x_{3ц} + x_2x_3x_{1ц}\right)}{\left[\left(-2x_2x_3x_4 + x_2x_4x_{3ц} + 2x_4^2x_1 - x_4^2x_{1ц}\right)^2 + \left(2x_4x_1^2 - 2x_1x_2x_3 + 2x_1x_4x_{1ц} + x_2x_3x_{3ц} + x_2x_3x_{1ц}\right)^2\right]^{1/2}}, \quad (20)$$

where $A = \cos\alpha_1$, $B = \cos\alpha_2$.

From the relations (19), (20), it is easy to obtain a parameter natural for this UAV type to determine the orientation vector of the control acceleration (in a certain sense, of the thrust vector), the pitching angle:

$$\vartheta(S) = arctg\frac{\cos\alpha_2}{\cos\alpha_1} = \frac{\left(2x_1^2x_4 - 2x_2x_1x_3 + 2x_4x_1x_{1ц} + x_2x_3x_{3ц} + x_2x_3x_{1ц}\right)}{\left(-2x_2x_3x_4 + x_2x_4x_{3ц} + 2x_4^2x_1 - x_4^2x_{1ц}\right)}. \quad (21)$$

In our case, the UAV motion takes place in the given plane, and the yaw angle is $\varphi(S) = 0$. Using the proposed methodology for the synthesis of terminal optimal control, an algorithm for calculating the pitching angle was developed [2], and the computational simulation of the UAV flight was carried out [7].

Research Results. The numerical studies have been conducted using the software that implements the proposed method. We are talking about the algorithmic software for terminal guidance of ballistic aircraft based on the solution to boundary problems of ballistics. The corresponding computer program was registered in 2013.

When modeling, a hypothetical accelerating tool was used with the characteristics and initial conditions for the UAV launch ascent, given in [8].

The initial conditions of the simulation and the calculation results of the early ascent are given in Table 1.

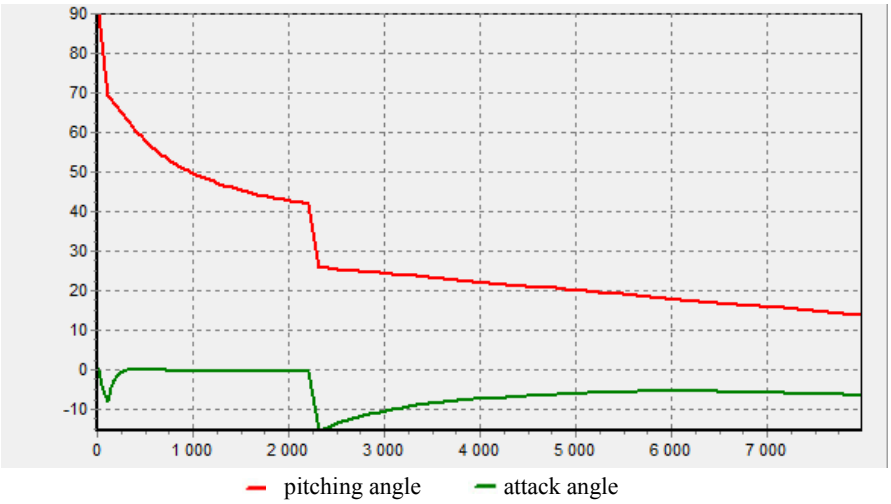
Table 1

Initial conditions for UAV launch ascent

t	x/vx	y/vy	z/vz	wx/tang	wy/tet	w1/alf
0.0000	0.0 5.6323	0.0 –0.0000	0.0 322.6757	0.0000 90.00	0.0000 0.00	0.0000 0.000
0.5537	3.1 5.6323	3.1 11.1626	178.7 322.6756	–0.0000 90.00	16.6000 1.98	16.6000 –0.048
3.4580	34.7 21.4565	120.9 69.4705	1115.8 322.6727	15.8242 69.34	103.4286 12.13	105.3386 –7.881
13.0193	832.4 155.6274	1683.4 259.0827	4200.8 322.6335	150.0000 60.00	386.9131 35.91	419.3209 0.011
23.0193	3384.5 368.3725	5300.4 464.9493	7426.8 322.5441	362.7748 50.02	690.8797 43.59	790.8281 –2.076
33.0193	8469.9 662.2567	11009.1 679.2317	10651.6 322.4053	656.7462 43.44	1003.1187 42.79	1219.9140 –2.582
43.0193	16923.8 1044.7345	18956.0 915.1428	13874.7 322.2176	1039.4134 38.74	1336.7768 40.11	1727.7590 –2.677
50.9898	26729.1 1429.1498	27089.4 1131.2610	16442.2 322.0334	1424.0909 36.01	1630.6058 37.94	2211.8627 –2.501

Fig. 2 shows the pitching-angle profile in the boost flight phase.

Dependence of pitching and attack angles on apparent speed rate



Dependence of UAV entrance angle on apparent speed rate

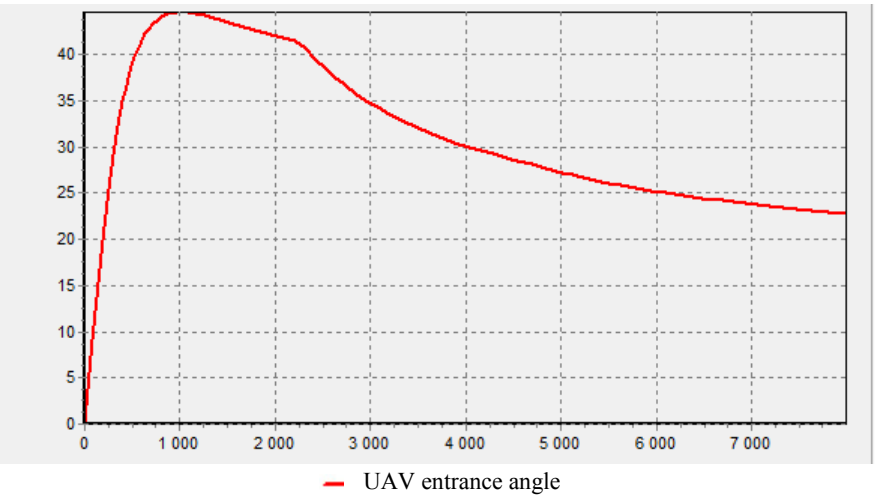


Fig. 2. Variation of UAV pitching and entrance angles

Fuller information on the simulation results given in [8] is as follows.

1. The resulting control is suboptimal due to the use of boundary conditions in an analytical form. To improve the accuracy of the boundary conditions in the navigation algorithm, it is necessary to introduce an up-date algorithm, which is based on attracting more accurate models of the Earth's gravitational field. Hence, the launch ascent accuracy will be improved.

2. The effect of random disturbances is compensated by the adaptive properties of terminal guidance, as well as by predicting the motion parameters and the formation of control at a decreasing time interval each time until the end of the launch ascent process.

3. The obtained estimates under modeling allowed us to rationally select the discreteness of the navigation and guidance algorithms and thereby limit the range of requirements for the FMC during its implementation.

4. The implementability of guidance algorithms based on the developed methodology on modern onboard computers creates no difficulties. The required response rate is $(1–1.5)\times10^6$ k.o./s.

Discussion and Conclusions. Thus, the principle of dynamic programming seems to be most preferable. It should be used when solving the problem on the centroidal motion control synthesis (guidance problem) of the UAV with long-range capabilities in the boost phase. The well-known remark on the applicability of dynamic programming, the so-called "curse of dimensionality", is inappropriate in the task of developing control as a function of the final state [5]. In addition, the use of boundary conditions in an analytical form simplifies essentially the formation of suboptimal control and enables to change strategically the flight missions. This scales up the applicability of this algorithm for UAV of various purposes.

References

1. Polovinchuk, N.Y., Ardashov, A.A. Proektirovanie sistem upravleniya raket-nositeley i mezhkontinental'nykh ballisticheskikh raket. [Design of control systems for launch vehicles and intercontinental ballistic missiles.] Rostov-on-Don: RVIRV, 2010, 242 p. (in Russian).
2. Mogilevskiy, V.D. Navedenie ballisticheskikh letatel'nykh apparatov. [Ballistic targeting.] Moscow: Mashinostroenie, 197, 208 p. (in Russian).
3. Athans, M., Falb, P. Optimal'noe upravlenie. [Optimal control.] Moscow: Mashinostroenie, 1968, 764 p. (in Russian).
4. Bellman, R. Dinamicheskoe programmirovanie. [Dynamic programming.] Moscow: Mir, 1965, 286 p. (in Russian).
5. Bryson, A., Yu-Chi Ho. Prikladnaya teoriya optimal'nogo upravleniya. [Applied Optimal Control.] Moscow: Mir, 1972, 544 p. (in Russian).
6. Petrov, B.N., et al. Bortovye terminal'nye sistemy upravleniya. [Onboard terminal control systems.]. Moscow: Mashinostroenie, 1983, 200 p. (in Russian).
7. Polovinchuk, N.Y., Shcherban, I.V. Metody i algoritmy terminal'nogo upravleniya dvizheniem letatel'nykh apparatov. [Methods and algorithms for terminal motion control of aircraft.] Moscow: RF MD Print.-Publ., 2004, 290 p. (in Russian).
8. Polovinchuk, N.Y., Ivanov, S.V., Kotelnitskaya, L.I. Sintez upravleniya manevrom ukloneniya bespilotnym letatel'nym apparatom s uchetom terminal'nykh ogranicheniy. [Synthesis of evasive maneuver control of unmanned aerial vehicle for terminal restrictions.] Vestnik of DSTU, 2018, vol. 18, no. 2, pp. 190–200 (in Russian).

Received 22.09.2018
Submitted 24.09.2018
Scheduled in the issue 15.01.2019

Authors:

Polovinchuk, Nikolay Y.,
professor of the Airborne Electrical Systems and Navigation Instrumentation Department, Moscow State Technical University of Civil Aviation, Rostov Branch (262c, Sholokhov pr., Rostov-on-Don, RF), Cand.Sci. (Eng.), professor ,
ORCID: http://orcid.org/0000-0003-0002-5120
npolovinchuk@mail.ru

Ivanov, Stanislav V.,
associate professor of the Production Automation Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Eng.),
ORCID: https://orcid.org/0000-0002-3237-0415
sta399@yandex.ru

Zhukova, Maria Y.,
postgraduate of the Computer and Automated Systems Software Department, Don State Technical University (1, Gagarin sq., Rostov-on-Don, 344000, RF), Cand.Sci. (Eng.),
ORCID: http://orcid.org/0000-0001-9555-3756
marg88@list.ru

Belonozhko, Dmitry G.,
adjunct, Krasnodar Higher Military School named after army general S. M. Shtemenko (4, ul. Krasina, Krasnodar, 350035, RF),
ORCID: http://orcid.org/0000-0002-7185-9475
staeer@rambler.ru