

# **СОДЕРЖАНИЕ**

---

---

**Том 35, номер 3, 2021**

---

---

## **Зрительная система**

An overview of the visual acuity assessment. 1. Primary measures  
and various notations

*G. I. Rozhkova, M. A. Gracheva, G. V. Paramei*

179

Комплексный анализ функциональных показателей  
бинокулярной зрительной системы у школьников

*H. H. Васильева*

199

Окуломоторная активность при решении зрительных  
когнитивных задач в различных временных условиях

*A. И. Талеева, Н. В. Звягина*

217

## **Слуховая система**

Различение звуковых сигналов с гребенчатыми спектрами  
в присутствии дополнительных сигналов

*M. С. Томозова, Д. И. Нечаев, О. Н. Милехина, А. Я. Супин*

228

## **Техническое зрение**

The role of projective transformations in image normalization

*I. A. Konovalenko, P. P. Nikolaev*

236

Правила для авторов журнала “Сенсорные системы”

260

# **Contents**

---

---

**Vol. 35, No. 3, 2021**

---

---

## **Visual system**

An overview of the visual acuity assessment.1. Primary measures  
and various notations

*G. I. Rozhkova, M. A. Gracheva, G. V. Paramei* 179

Analysis of the functions of binocular visual systems in schoolchildren.

*N. N. Vasilyeva* 199

Oculomotor activity in solving visual cognitive tasks under different  
time conditions

*A. I. Taleeva, N.V. Zvyagina* 217

---

## **Auditory systems**

Discrimination of rippled sound signals with overlapping  
additional signals

*M. S. Tomozova, D. I. Nechaev, O. N. Milekhina, A. Ya. Supin* 228

---

## **Technical vision**

The role of projective transformations in image normalization

*I. A. Konovalenko, P. P. Nikolaev* 236

---

Guidelines for authors

260

---

## AN OVERVIEW OF THE VISUAL ACUITY ASSESSMENT. 1. PRIMARY MEASURES AND VARIOUS NOTATIONS

© 2021 г. G. I. Rozhkova<sup>1,\*</sup>, M. A. Gracheva<sup>1</sup>, and G. V. Paramei<sup>2</sup>

<sup>1</sup> Institute for Information Transmission Problems RAS, 127051 Moscow, B. Karetny, 19/1, Russia

<sup>2</sup> Department of Psychology, Liverpool Hope University Hope Park, L16 9JD Liverpool, UK

\*E-mail: gir@iitp.ru

Received March 17, 2021; Revised April 8, 2021; Accepted April 25, 2021

The paper presents a brief overview and analysis of various approaches to assessment of visual acuity (VA). It aims to explicate the reasons of disparity of opinions among experts on the methods of assessing VA by reflecting on advantages and shortcomings of individual methods and the corresponding VA notations. In the course of time, the number of methods and procedures of VA assessing was increasing; they became more diverse and complex. In parallel, the number of professionals who assessed VA for different purposes was also increasing. Such state of affairs has resulted in certain difference between the views on the interpretation of outcomes of VA measurement and to introducing different VA notations. Currently, however, reaching a consensus among experts is becoming crucial since numerous research projects require cooperation of professionals from different areas and, also, involve international collaboration. In search of the common ground for the consensus and the unified terminology and notation, it is reasonable to appraise the crux of the VA assessment problem and, as well, rationales for a great variety of current viewpoints on solving the problem in practice. An overview of contemporary approaches to assessment of VA points out that there exists the sole primary measure of VA that can be obtained by a direct measurement and expressed by the base unit of spatial metrics – the minimum angle of resolution (MAR,  $\alpha_m$ ) – defined as the minimum angle at which two points are just perceived as separate. There exists one more measure, in place of MAR, – the critical spatial frequency ( $F_c$ ) – which can be gauged directly when one employs gratings of varying spatial frequency to estimate the maximum (critical) spatial frequency above which periods of a grating can no longer be distinguished. It is reasonable to consider all other VA measures as secondary, or derived, since these are calculated as functions of  $\alpha_m$ . Introduction of various secondary measures, such as Snellen fraction, decimal units, logMAR, visual efficiency (VE), visual acuity rating (VAR), and others, was stipulated by the demand of developing alternative notations of VA, which are more convenient and comprehensible, than  $\alpha_m$ , for practitioners who assess VA in applied areas for various purposes, such as screening, diagnostics, monitoring, rehabilitation, disability determination, population statistical evaluations, designing of new VA tests, etc. We conclude that, in view of substantial differences of the purposes, requirements and criteria among experts in various areas, the quest for a unique measure of VA, which would be optimal in all contexts, is probably unresolvable, as is establishing of the ultimate, “gold standard”, practical method of VA assessment.

**Key words:** human vision, visual acuity, scales of measurement, MAR, Snellen fraction, decimal units, log-MAR, VAR

**DOI:** 10.31857/S0235009221030033

### INTRODUCTION: ISSUES IN ASSESSMENT OF VISUAL ACUITY

In a broader sense, visual acuity (VA) is considered as the ability to detect and recognize small objects, and to discern their elements. A colloquial term implying a good VA is *clear vision*. It is hardly possible to ascertain when the necessity of VA assessment was realized and became practicable for the first time. It is known though that in ancient Persia more than a thousand years ago, there existed ‘the Test’, or ‘the Riddle’, that was used to gauge warriors’ eyesight based on viewing the constellation of Ursa Major (the Great Bear) on a clear night (Bohigian, 2008). ‘The Test’ was to dis-

cern, in the constellation handle, a double star: if one was able to see with the naked eye Alcor, faint companion of Mizar, one has passed the Test. Adopted by nomadic Arabs, ‘the Arabic Test’ was used in antiquity: in the Roman army it was necessary to pass this test to become an archer.

For centuries past, methods and procedures developed for gauging vision greatly increased in number and complexity in response to considerable variegation of human activities that required satisfactory VA, whereby its assessment became obligatory. It was a long way from testing the visual ability of an ancient warrior to discern the double stars in the night sky to

the modern routine VA assessment by means of special test stimuli either presented as charts or generated on displays of computerized setups. In the last six decades, we observe a significant increase in the number of researchers, experts and practitioners, who assess VA for different purposes and are accustomed to different methods as well as to differing VA notations. The diversity of aims and purposes – and, accordingly, the methods used – resulted in considerable loss of consensus among the researchers performing VA assessment and analysis. Currently, however, such a consensus became crucial since many projects involve collaboration of professionals from various national institutions and, as well, international cooperation.

The claims of precise VA measurement and standardization of VA assessment appear in literature regularly since 1950s (e.g., Ogle, 1953; Sloan, 1980; Lovie-Kitchin, 1988; Elliott, Sheridan, 1988; Siderov, Tiu, 1999; Lovie-Kitchin, Brown, 2000; Beck et al., 2003; Rosser et al., 2003; Koskin, 2009; Rozhkova, Malykh, 2017; Rozhkova, 2018). Unexpectedly, the task of coordinating different approaches to VA assessment appeared to be not a simple one and, perhaps, not fully resolvable. To find the common ground for the consensus and unification of the terminology, it seems reasonable to begin with an analysis of the crux of the VA assessment task and the rationale behind a great variety of the existing views on its implementation in practice.

Regular procedures of VA assessment in an optometrist's office, by viewing charts with letters of different size, are familiar to everyone. However, a universal and generally accepted method for quantitative assessment of VA has not been found so far. Moreover, progress in academic and clinical research makes it more and more unlikely that such generic method can be arrived at.

Despite considerable efforts, abundant theoretical and practical problems remain unresolved with regard to the best way of performing VA measurements. Along with an obvious complexity of some optical, physiological and psychological problems related to the vision function, one of the most likely reasons of this unfortunate situation is a significant variety of the objectives of professionals working in different areas. The issues of the VA assessment method and measures, which we consider below, can be categorized based on the answers to the following key questions:

- Who requires “a good vision”?
- What for does VA need to be assessed?
- What is a proper representation, or notation, of the VA assessment outcome?
- How will the VA data be used in practice?

**WHO?** In everyday life, a “clear vision” is crucial for any human regulatory function, such as self-care, orientation in space, successful communication, efficient education, professional performance, competitive sports, etc. It is apparent that proper information

about one's visual ability is required both for the persons themselves and for people interacting with them – parents, teachers, doctors, professionals, or designers of visually perceived products (books, movies, TV-broadcasts, social media), etc. Notably, from a functional viewpoint, different types of visual behavior imply different visual tasks, which vary in the required accuracy and speed. The visual tasks include detection and recognition of individuals and objects, determining their location and properties (size, shape, direction of movement), forecasting incidents and accidents (falls, collisions, fatal errors in recognizing dangerous objects), reading and writing, etc. It is apparent that a “good vision” could have different meaning in different instances, and a quantitative presentation of VA may require a variety of tests.

**WHAT FOR?** The answer to this question depends on the “users” of the outcomes of VA assessment – their purposes, the targeted populations and individuals. For infants and children, monitoring of VA is required to ensure appropriate conditions for normal maturation of their visual system and appropriate vision care in kindergartens and schools, and later in colleges and universities. Outcomes of population-based surveys serve, among others, for providing optimal lighting conditions at work and in studying places. In individuals, monitoring of VA is required for detecting a visual impairment, its appropriate diagnosis and effective vision correction, for certification of vision loss, or ascertaining the requirements of safe performance in vision-demanding activities (e.g., driving).

**WHAT IS A PROPER REPRESENTATION, OR NOTATION, OF VA?** We assume that professionals have to comprehend clearly the VA notation preferred by them and that the VA notation they use meets well their professional objectives.

**HOW WILL THE VA DATA BE USED IN PRACTICE?** The information on VA measures is used for abundant purposes, to name just a few. Management of any institution or company has to take VA measures into account for designing optimal lighting workplace conditions to prevent visual discomfort, visual strain, or asthenopia. In production of textbooks, the letter font type and size are supposed to be aligned with the VA population data of potential readers of various ages. At schools, teachers are expected to consider individual students' VA for their optimal placing in a classroom. Specialist committees that certify visual impairment determine the degree of claimant's disability and eligibility for the corresponding social benefits. National offices of statistics analyze changes in population VA caused by various ecological and/or socioeconomic factors. In each case, the employed tool and condition of its administration determine the VA notation. However, to be optimal, the choice of the VA notation is to be guided by the aims and objectives of the professionals.

**The aim of the present overview** is to reflect on the principles and procedures underlying VA assessment within a general framework of information processing in the human visual system. Various factors are considered that affect outcomes of the VA measurement, as well as certain preferences in the choice of the tools for its measurement stipulated by everyday needs and the tasks on the agenda of various professional communities.

The impetus for this paper was a recent ardent debate that raised critique of certain VA notations. The dispute's pinnacle was the Editorial "The good (log-MAR), the bad (Snellen), and the ugly (BCVA, number of letters read) of visual acuity measurement" in *Ophthalmic and Physiological Optics* by David B. Elliott (2016), its Editor-in-Chief. Noteworthy, the Editorial compared various currently available charts that imply different ways of obtaining VA outcomes, i.e., focusing on the VA assessment tools, but leaving the matter of VA notations aside. The debate also revealed that some critics of certain VA notations committed the fallacy of confusing the following concepts: units of measurement; scales of measurement; specification of the reference levels in the measuring tools; scaling of the primary measurement outcomes, etc.

By approaching the matter of VA assessment from these different conceptual vantages, we strive to present arguments that for VA assessment, one can hardly arrive at a "gold standard" of VA measurement or the "best" VA notation. Moreover, the choice of the most appropriate VA notation could vary dependent on the theoretical approach, conditions, and objectives of the VA assessment, namely:

- screening, monitoring, functional correction, survey;
- target population (infants, teenagers, school and university students; adults; healthy observers or individuals with visual impairment);
- specific setting (clinic, field study, or laboratory);
- resources available for academic and clinical projects, field work, epidemiological studies, or routine practice (funds; experts, practitioners and supporting workforce; equipment; time constraints);
- social significance of the obtained VA data (for medical, educational or economic purposes).

An auxiliary aim of the present paper is to provide a tutorial to those at the beginning of their research on visual acuity. The definitions and conceptual clarifications are intended to help embedding the issues of VA assessment in a broader theoretical context.

## PRACTICAL ISSUES OF VISUAL ACUITY ASSESSMENT WITHIN A FRAMEWORK OF INFORMATION PROCESSING IN THE HUMAN VISUAL SYSTEM

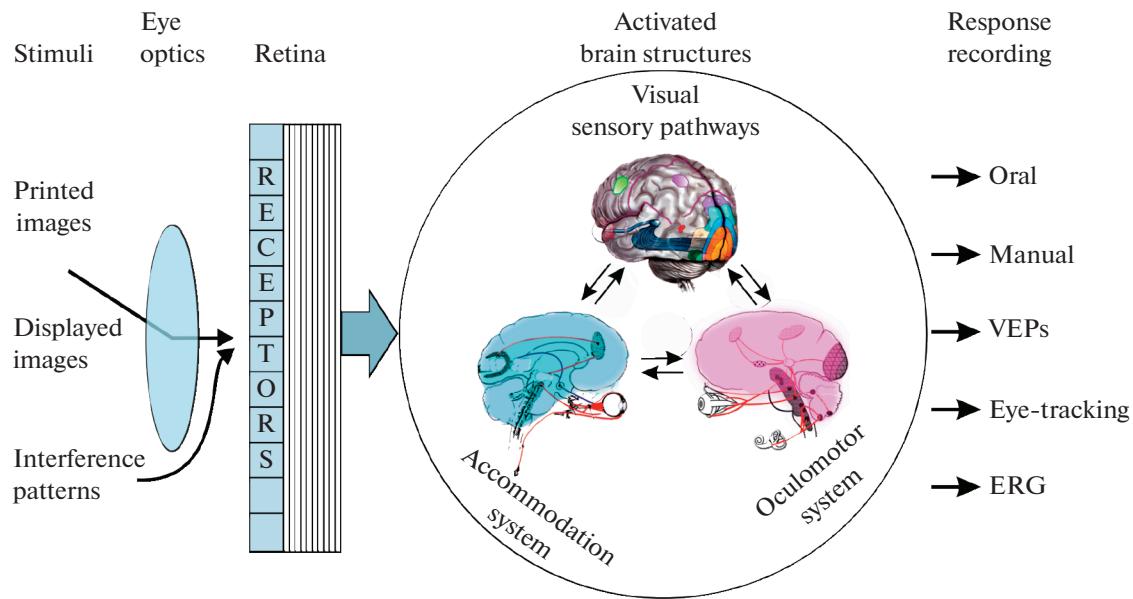
Innumerable studies of vision, that employed various neurophysiological, psychophysical and behavioral methods, accumulated evidence of the inherent basis of the functioning of the human visual system, namely, that it has multiple specific mechanisms underlying performance in different visual tasks. Moreover, in the functioning of the visual system, visual processing is interrelated with the functioning of the oculomotor and accommodation systems that optimize viewing conditions.

Generally, comprehensive characteristics of visual abilities of an individual would require a thorough and extended investigation by employing many tests that would result in dozens of critical parameters, indices, and scores – the multidimensional quantitative foundation of vision diagnostics.

In accordance with a general comprehension of VA as an ability to detect and recognize small objects, vision scientists define VA as the measure of accuracy of the spatial analysis. In practice, for VA assessment a great variety of test stimuli are employed (also termed targets, test images, symbols, signs, or optotypes) and different visual tasks are used, which inevitably give rise to the problem of converting between the obtained measures.

The other issue is that different visual stimuli and designs employed in the VA tests implicate processing by different modules of the visual system. VA assessment, an unfolding process, depends on many conditions. Impairment of vision and a VA measurement outcome could be stipulated by many factors affecting any of the stages of visual processing. The implicated stages are optical image formation in the eye; transformation of the optical image into a neural retinal image on the eye fundus (more precisely, a set of neural images with selected information about different features of the retinal optical image); its transmission to the brain upstream the visual system; subsequent transformation and processing of the neural images within different brain visual areas; formation of a percept; analysis and recognition of the viewed object.

Main stages of the visual processing are illustrated in Fig. 1. At the input end, a test stimulus passes through the eye optics and activates the photoreceptor layer of the retina, the two initial processing stages that are common for any manner of assessing VA (except a specific interference method). At the output end, the observer's responses to the test stimuli can be of two main types – either subjective (a verbal judgment, pointing, pressing a key etc.; e.g., Bach, 1996, 2007; Radner, 2017), or objective (ERG (Tehrani et al., 2015); recordings of eye movements (Wolin, Dillman, 1964; Hyon et al., 2010) or of brain activity (VEPs; e.g., Zheng et al., 2020), accompanying visual pro-



**Fig. 1.** An illustration of the information processing flow in the human visual system during VA assessment.

cessing with no necessarily explicit participant's action).

In relation to VA assessment, it is important to bear in mind that, between the input and the output, in cases of different test images quite different neural pathways are activated that underlie processing of different image features. Various neurophysiological and psychophysical studies, as well as clinical data of patients with certain visual brain injuries provide evidence that a localized impairment can have a very selective effect on stimulus recognition. For example, a left occipito-parietal ischemic infarction may cause literal alexia, i.e., the condition that affects patient's ability to recognize individual letters or numerals, while other visual recognition abilities remain unaffected.

The above implies that an outcome of **VA measurement using a given set of test stimuli would only characterize the functional state of the pathways activated by these stimuli**. In other words, the obtained results indicating, say, "normal" VA could testify only to normal functioning of a certain subnetwork of all visual pathways that subserves the certain type of test stimuli. In a general case, the structures forming this specific subnetwork could either belong to one functional unit – module – or include several modules. As indicated above, the eye optics and the photoreceptors in the retina are structures that are common for processing all types of the visual stimuli. This implies that impairment of the eye optics and/or a retinal pathology could affect VA measures regardless of the type of the test stimuli. One important exception is a specific case of grating stimuli created on the retina by means of the coherent interference technique bypassing the eye optics and, thus, enabling assessment of the so-called

"retinal VA". By comparing the outcomes of usual VA assessment by means of naturally projected stimuli with the values of retinal VA in patients, one can disentangle impacts of the optical and neuronal (retinal & post-retinal) impairments. To localize the structure (or structures) upstream in the visual system that cause VA loss is impossible without reaching for various other diagnostic methods of examination.

#### PARADIGMS AND PROCEDURES FOR THE VISUAL ACUITY ASSESSMENT

Accuracy required for VA assessment varies significantly depending on both the visual task and the assessment aim. The granularity of accuracy of VA assessment can be captured in terms of the measurement scales varying in power and precision: ordinal, interval or ratio (Stevens, 1946).

For everyday purposes, it is often sufficient to qualify VA using a coarse categorization in common terms: *low* (poor, weak), *normal* (good, fine), and *excellent* (supernormal, perfect). Such categorization is based on the individual's visual abilities compared to other people from a representative population. The implicated visual abilities comprise the individual's speed of visual search, dexterity (e.g., reaching and grasping, shooting, driving), and navigation accuracy, as well as presence/absence of asthenopia symptoms (visual discomfort, visual stress, visual fatigue, etc.). This type of VA notation, i.e., low–good–excellent, implies ranking the vision quality on an ordinal scale. It is apparent that such values are rather vague, fuzzy and subjective: for gauging the quality of vision, an assessor is guided by his/her (tacit) reference points and an (implicit) "yardstick".

**Table 1.** Characteristics of the typical examination paradigms for VA assessment

Examples of the test stimuli	Visual task	Instruction for the observer	Parameter to be measured
Single-point targets Crumbs Single thin lines	<b>Detection</b>	To detect and indicate or to grasp the smallest target on a uniform background	Threshold diameter Threshold line width (angular size)
Two-point targets Teller gratings Gabor patches	<b>Resolution</b>	To distinguish one-point and two-point targets. To indicate position of a periodic structure	Threshold angular distance between the points or critical spatial frequency
Tumbling E Lea symbols Letters	<b>Recognition</b>	To identify and name the presented optotype from a given set	Angular size of the critical parameter
Vernier stimuli	<b>Detection of a shift</b> of the two well seen stimulus halves <b>(Hyperacuity)</b>	To detect the shift and to determine its direction (left-right)	Threshold angular size of the shift

In clinical practice, however, precision of VA assessment is crucial for functional correction and monitoring of vision development, or for stimulation of the visual function during recovery from an injury: the higher the precision of the measurement, the earlier the anticipated VA changes (positive or negative) can be detected, which ensures sooner and more perceptible effects of a timely, patient-tailored treatment (therapeutic intervention). Clinical cases require accurate *measuring* of VA by means of standardized test procedures and presenting the results in a *quantitative form* (in conventional units cognized by professionals).

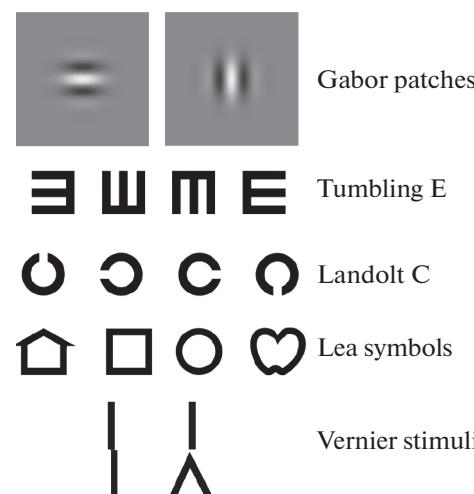
Quantitative clinical methods of VA assessment are based on estimating either the smallest angular size of the test stimuli detected/recognized by the testee, or the smallest test image components that can be *resolved*, i.e., can be seen as distinct. Different examination paradigms imply solving different visual tasks (Table 1). Some examples of the test stimuli are shown in Fig. 2.

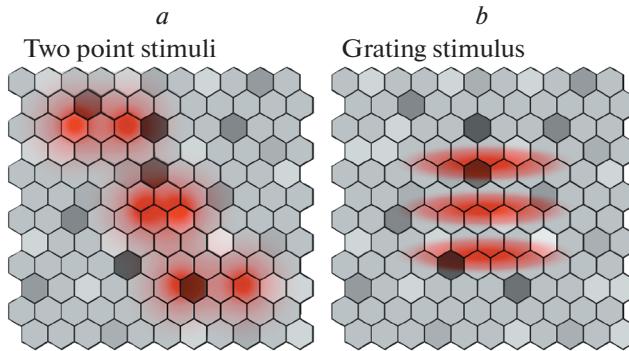
In this paper, we mainly consider clinical VA measures, conventionally termed *resolution VA* and *recognition VA*; detection acuity and Vernier acuity have lesser fields of application. We would also like to underscore that most paradigms of the VA assessment imply not just one visual task but a task composition (cf. Heinrich, Bach, 2013, on distinguishing the resolution and recognition in tests with modified Landolt C optotypes).

The notion of resolution is broadly used in physics and engineering. Originally, in physics, the term *resolution* was coined to characterize the quality of an optical device – its ability to produce clear image of the finest structural elements of the test pattern. Quantitative assessment of the quality of an optical device is accomplished by estimating its *resolution threshold*. The standard method for this is to estimate the *minimum angular distance*,  $\alpha_m$ , between two light test points,

which allows detecting (either visually or by a light-sensitive instrument) the two luminance maxima and a luminance minimum, corresponding to the test points that are sufficiently well separated in the image produced by the device in question.

Lord Rayleigh (1879) proposed the criterion for calculating the resolution threshold based on the analysis of the light distribution in the two-point images created by an optical device. It is well known that, due to diffraction, the optical image of each light point consists of the central light spot and a series of dark and light annuli. The pattern of diffraction depends on the aperture size and the wavelength of light (with shorter wavelengths being diffracted at a greater angle than longer ones). According to Rayleigh, the resolution threshold  $\beta$  corresponds to the distance between the two near test points that is equal to the radius of the

**Fig. 2.** Various types of the test stimuli for assessment of visual acuity and hyperacuity.



**Fig. 3.** Schematic illustration of the projection of two-point test stimuli (a) and a test grating (b) onto a hexagonal array of photosensors varying in sensitivity. Light distribution in the optical images is represented by red; differences in light sensitivity of photosensors are represented in a grayscale (where lighter hexagons indicate higher sensitivity).

central spot. This radius  $\beta$  (in angular minutes) is given by the equation  $\beta = 1.22\lambda/D$  (where  $\lambda$  is wavelength,  $D$  is the aperture size). For such a distance, in case of two light sources with incoherent radiation, it was found that, in the midpoint between the two maxima, the light level is equal to 3/4 of the maximum. The application of the Rayleigh criterion to the human eye seems problematic for several reasons: e.g., calculation of the effective aperture from the pupil diameter; the required correction for optical properties of the eye media, etc. However, for realistic sizes of the pupil one can assume certain coarse approximations of the resolution threshold  $\alpha_m$  as quite likely.

Another way of assessing the quality of optical devices is to estimate the highest frequency of the spatial grating that can be reproduced by the device in question, or the *critical spatial frequency*,  $F_c$ . Apparently, in real conditions it is much easier to measure  $F_c$  than  $\alpha_m$ . This is illustrated schematically in Fig. 3 inspired by representation of a regular hexagonal lattice of photoreceptors in the foveal region of the human retina by Curcio et al. (1990). Fig. 3 shows images of two-point tests (a) and a test grating (b) projected onto a regular mosaic of photosensors. Variation of light sensitivity of individual photosensors is represented by cell lightness (on grayscale). In any real case, sensitivity of each photosensor varies due to internal noise, as well as a result of exposure to preceding stimuli. In turn, the variability of the photosensor sensitivity stipulates that visibility of the light test significantly depends on the position of its projection onto the photosensor array. The effect of the photosensor variability is stronger for single light points than for gratings since visibility (detectability) of the latter gain an advantage from summation of responses from multiple photosensors.

It is worth noting that in physics, rather than the resolution threshold, the broadly accepted measure of the quality of an optical device is *resolving power*, the

value reciprocal to resolution threshold,  $1/\alpha_m$ . The term was introduced to reconcile the physics terms with an intuitive idea of a better quality, whereby a better (higher) optical quality implies a higher resolving power, but a lower (smaller) resolution threshold. Unfortunately, the term *resolution* is often used without this clarification, which leads to confusion. In many publications, the term “high resolution” implies, in fact, *high resolving power*. For instance, the term *resolution* is used as the synonym of *resolving power* in computer science. It reflects the potential of either generating fine-grained images on a given display screen that consist of discrete elements (pixels), or printing fine pictures by a printer that produces images from tiny dye dots. It seems reasonable that a display resolution is characterized by the pixel density, the number of pixels per inch (ppi) in a line. Similarly, a printer resolution is expressed in dots per inch (dpi), i.e., the number of dots that can be printed in a 1-inch long line segment.

In ophthalmology and vision science, both notations, resolution threshold ( $\alpha_m$ ) and its reciprocal value  $1/\alpha_m$ , are used. However, in these disciplines the term *resolving power* has not been brought into broad use, and the measure  $1/\alpha_m$  is termed *visual acuity* (VA). When using the term *visual resolution*, researchers quite often imply resolution threshold, unlike resolving power that is used in optics and computer science.

#### PRIMARY MEASURE OF VISUAL ACUITY AND HISTORICALLY FIRST VISUAL ACUITY NOTATION

Over time, the following procedures proved to be most feasible for the assessment of VA:

- (1) determining the minimum size of the test stimuli that can be recognized at the chosen viewing distance;

- (2) determining the maximum viewing distance at which certain test stimuli can be recognized.

In both cases, the outcome implies the identical spatial metrics and is expressed in the same measure — as the minimum angular size of the smallest elements comprising the test images that is sufficient for satisfactory recognition of the elements. In vision science, this angle is termed the *minimum angle of resolution*, MAR.

It so happened that the term MAR became used by ophthalmologists not only as the acronym but also as a *mathematical quantity* in algebraic formulae. For most algebraists, such denotation is a violation of the conventional rule to use one letter for one quantity, which sometimes causes misunderstanding and lengthening of formulae. (However, it is rather pointless to raise an objection against use of the MAR denotation, since currently software developers often use even much longer acronyms in their software codes.)

When periodic gratings are employed as the test stimuli, another critical value – the highest discernible spatial frequency (critical frequency,  $F_c$ ) – is determined instead of the MAR. It is apparent that, ideally,  $F_c$  can be calculated from  $\alpha_m$ : at the threshold, one period of the just resolvable grating has to span  $2\alpha_m$ , i.e., one dark and one light bar; hence, visual angle of  $1^\circ$  ( $60'$ ) corresponds to  $60'/2\alpha_m$  periods, or  $F_c = 30/\alpha_m$  cpd (cycles per degree).

In the history of development of VA metrics, professionals initially pursued the physics approaches by treating the human visual system as an optical device (Colenbrander, 2008). Also vision scientists of the 18th and 19th centuries applied approaches, notions and notations that had been developed in physics for optics research.

Frans Cornelis Donders, Dutch physiologist, was the first to introduce the notion of VA in 1861 and to develop the standard procedure and the measurement unit for VA assessment (Pfeiffer, 1936). Donders assumed that calculation of an observer's VA should be accomplished in the same way as calculation of resolving power of an optical device, i.e., as the reciprocal of resolution threshold. Adapting this approach to human vision, Donders introduced the notion of a "standard eye" which, at threshold, can recognize letters as small as  $5'$  height (without errors or with sufficiently high probability of a correct response). VA of the standard eye was proposed to be taken as the measurement baseline.

Donders' protocol of VA assessment implied comparison of the testee's threshold letter size with the threshold letter size of  $5'$  accepted for the standard eye. Thus, in the course of the measurement, the examiner determined linear magnification ( $M$ ) of the test letters (in relation to the  $5'$  size) required for the observer to provide the standard level of letter recognition. The VA value,  $v$ , was calculated as the inverse of the magnification value:  $v = 1/M$ .

Thus, Donders introduced the first correct quantitative VA notation and the first unit for the VA measurement that is compatible with the modern concepts in metrology: "Unit of measurement – real scalar quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a number." (International Vocabulary of Metrology. 2008. Basic and General Concepts and Associated terms, p. 6).

Donders based the proposed way of VA estimation on the premise that, for normal human vision, the resolution threshold is equal to 1 arcmin, i.e., for the standard eye  $\alpha_{st} = 1'$ . Considering that the Donders test letters were drawn within a  $5 \times 5$  matrix, and the stroke width was  $1/5$  of the letter size, the threshold size of the letters for the standard eye can be assumed as  $5' \times 5' = 5\alpha_{st} \times 5\alpha_{st}$ . If an individual observer's threshold letters appeared to be, say,  $M \cdot 5'$  high, it is

apparent that the corresponding value of MAR,  $\alpha_i$ , was  $M$  times larger than  $\alpha_{st}$ , i.e.,  $\alpha_i = M\alpha_{st}$ . Therefore, the magnification  $M$  will be equal to  $\alpha_i/\alpha_{st}$  and, for the VA value,  $v$ , one can use the following expressions:

$$v = 1/M = \alpha_{st}/\alpha_i = 1'/\alpha_i.$$

The expression  $v = 1'/\alpha_i$  for VA was introduced by Monoyer (1875) and coined *decimal notation*. In the present paper we denote it as  $V_d$ . If the examination outcome is  $\alpha_i = \alpha_{st} = 1'$ , the testee's VA is characterized as  $V_d = 1.0$ , i.e., equal to that of the standard eye.

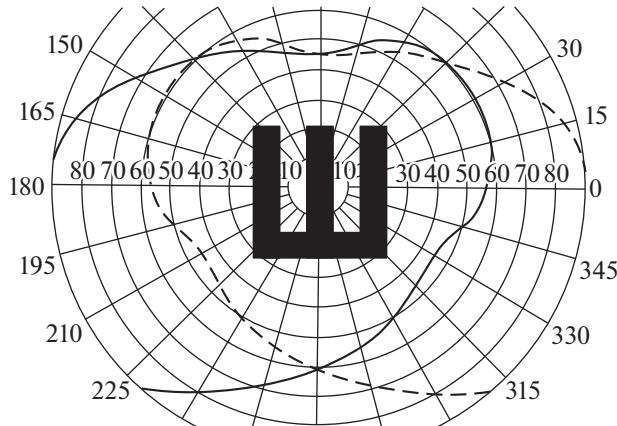
What are potential ranges of the MAR and corresponding VA values in humans? Apparently, these ranges are limited by the minimum and maximum sizes of the test stimuli that can be projected onto the retina and perceived as discernible visual images. The maximum test size is ultimately limited by the size of the visual field, while the minimum size is limited by the optics of the eye (diffraction, optical aberrations, intraocular scattering, etc.; for reviews, see Westheimer, 1970; 2001; 2010; Artal, 2014), by properties of the retina (the size and packing density of the photoreceptors; structure of the retinal neural networks), and by processing of retinal signals upstream the visual system.

To estimate a theoretical *lower* limit of VA, consider the central area of the binocular visual field, whereto the largest test image can be presented (Fig. 4). Its extents in various directions is somewhat varying, and spans about  $100^\circ$  in angular terms. However, a test image should occupy no more than one third of this area, i.e., about  $33^\circ$ , leaving a blank space around. A critical size of  $1/5$  of that test extent would be about  $400'$ , which corresponds to  $\text{MAR} = 400'$  and  $V_d = 1'/400'$ , or about 0.0025.

To find out an *upper* limit of the VA, one needs to consider the following factors (Westheimer, 2010):

- the transfer function of the eye's optics (optical transfer function, OTF:), which determines the quality of the image projected onto the retina and its highest spatial frequency;
- discrete structure of the retina limiting the accuracy of image sampling by the size and density of the retinal photoreceptors in the light sensory layer;
- neural processing of the retinal image upstream the visual system, up to the perceptual level.

With regard to the OTF of the typical human eye, the highest spatial frequency that can be reproduced unambiguously,  $F_c$ , is about 60 cpd, i.e., is 1 cycle per 1 arcmin (Campbell, Green, 1965). Thus, the minimum resolution threshold should be no less than  $0.5'$  (half a period of the spatial grating), and decimal  $V_d$ , accordingly, not higher than 2.0, if it is solely the OTF that stipulates the maximum VA. However, certain in-built factors of the visual system exert their inherent influences on the OTF-based VA, namely: the size of the retinal photoreceptors and their packing density



**Fig. 4.** Combined visual fields of the left and right eyes with the projected optotype (tumbling E) of the maximum size that can be presented binocularly for VA assessment.

can reduce upper limits of VA values, whereas neural processing at higher levels of the visual system and eye movements can improve the OTF-based VA (Duncan, Boynton, 2003).

In popular ophthalmology literature, one can find a rather simplified approach to visual resolution assessment, or the so-called “receptor theory” that implicates the limitation of the photoreceptor size in the fovea of the human eye, “grains” or “pixels”. The “receptor theory” proposed by Helmholtz (1867; English translation: 1924, v. 2, p. 33) and treated as quite plausible up to the middle of the past century (e.g., Hecht, Mintz, 1939; Polyak, 1941, p. 430) is based on a seemingly reasonable assumption that if two test points project on two adjacent cones, they cannot be distinguished from a larger single object projecting on those adjacent cones. For discerning the two points, at least one unstimulated or a less stimulated cone is required in-between. This simplified model of resolution threshold is often used as an illustration despite a vast amount of knowledge on visual optics, morphology and physiology providing evidence of the model flaws.

The real picture behind the visual processing is very complicated. Firstly, under natural viewing conditions it is impossible to stimulate two single photoreceptors in the configuration comprising two activated photo-sensors and one silent in-between: even in the case of an infinitely small test point (a Dirac pulse), the optical point image (the point spread function) is smeared over tens of photoreceptors due to diffraction and eye optics aberrations. Moreover, the stimulation pattern is permanently drifting over the retina because of the eye micromovements. A reliable stimulation of single photoreceptors became possible only in artificial conditions with the development of adaptive optics that compensates for the eye’s aberrations and counteracts eye micromovements (Roorda, Williams, 1999; Roorda et al., 2002; Porter et al., 2006; Artal, 2014).

Secondly, neural processing at the post-receptoral stages of the visual pathway was shown to provide significantly better resolution and higher VA exceeding the value of 1' indicated for the standard eye by Donders on the basis of using letters as the test stimuli (Artal et al., 2004). In this case the so-called Vernier acuity, or hyperacuity, is implicated, i.e., the well-known fact that it is possible to detect the misalignment of just few arcsec between two thin lines or fine segments if they are sufficiently long (Westheimer, 2010). It is apparent that such considerable upsurge of visual analysis of fine stimulus details is achieved due to the aggregate higher-level processing of visual inputs from many photoreceptors activated by stimuli extending beyond a single point (Westheimer, 1975; 2010).

With regard to the constraints posed by the discrete structure of the retina, the highest spatial frequency that can be successfully coded and reproduced in the visual pathway is limited by Shannon’s sampling theorem (Jerri, 1977). In the human fovea, the photoreceptor mosaic can be considered as resembling a regular hexagonal lattice (Curcio et al., 1990; Putnam et al., 2005). For such spatial geometrical organization, the highest spatial frequency that can be resolved unambiguously is given by the Nyquist frequency based on two-dimensional sampling theory (Snyder, Miller, 1977; Miller, Bernard, 1983):

$$F_N = \pi P / (180\sqrt{3}r),$$

where  $F_N$  is the spatial frequency in cpd,  $P$  is the posterior nodal distance (16.7 mm for a standard human eye), and  $r$  is the photoreceptor spacing. Taking into account the data of Polyak (1941) and Curcio et al. (1990), for the minimum size of the foveal receptors (about 1.5  $\mu$ ) the calculation results in  $F_N$  of about 100 cpd, that is larger than the theoretical OTF limit (60 cpd).

In (Rozhkova, Matveev, 2007, p. 145), the data are presented providing evidence that the VA values of 3.0 and somewhat higher (in decimal units), corresponding to  $F_N \geq 90$  cpd, were really encountered, although seldom, especially in 11–14-year-old children. In this investigation, the VA was assessed using tumbling E.

#### VISUAL ACUITY NOTATIONS SUGGESTED FOR VARIOUS PURPOSES

By designing the first letter chart for VA assessment (in 1862), Herman Snellen, Donders’ successor, proposed the VA notation that, in essence, was equivalent to the Donders notation but, in addition, included information on the viewing distance  $D_o$  at which VA was measured (Pfeiffer, 1936; Colenbrander, 2008; Cole, 2014). Based on the threshold letter height  $h_i$  that was estimated for the individual testee and the viewing distance  $D_o$ , Snellen suggested to express VA as the ratio of the viewing distance,  $D_o$ , to the conventional view-

ing distance,  $D_i$ , at which the testee can see the threshold letters of height  $h_i$  as well as the standard letter of height  $h_s$  is seen by the standard eye at the viewing distance  $D_o$ :  $V_s = D_o/D_i$ . (Fig. 5). One can easily see that  $D_o$  and  $D_i$  are the distances, where the threshold letters  $h_s$  and  $h_i$  for the standard and the testee's eye, respectively, are of equal angular size. Based on denotations in Fig. 5, where  $\tan \alpha_{ms} = h_s/D_o = h_i/D_i$  and  $\tan \alpha_{mi} = h_i/D_o$ , one obtains  $V_s = h_s/h_i = D_o/D_i = \alpha_{ms}/\alpha_{mi}$  (since  $\alpha_{ms}$  and  $\alpha_{mi}$  are small and  $\tan \alpha \approx \alpha$ ).

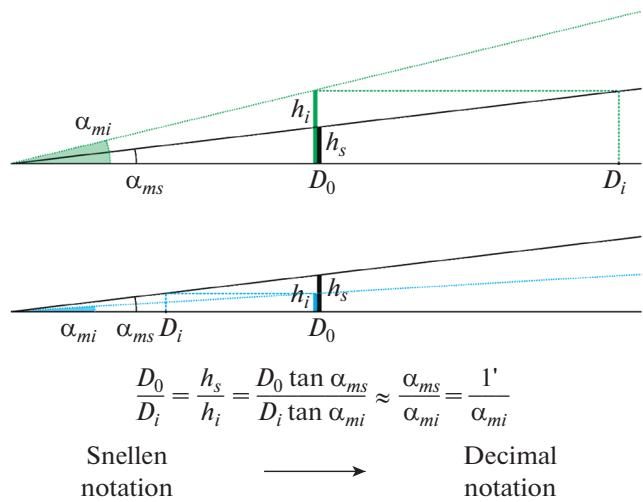
Historically, the numerator and denominator of the Snellen fraction are expressed either in feet or meters for far vision, and occasionally, for near vision, either in inches or centimeters. Thus, for normal vision of the standard eye ( $MAR = 1'$ ), the fractions 20/20 and 6/6 are different expressions corresponding to one and the same viewing distance expressed either in feet or meters (20 ft  $\approx$  6 m); the fraction 14/14 corresponds to normal near vision at the viewing distance expressed in inches (14 inch  $\approx$  35 cm).

When the Snellen fraction is converted to decimal form, one obtains the value identical to the VA in the Donders (decimal) notation, since both can be expressed as  $\alpha_{ms}/\alpha_{mi}$  (see Fig. 5). Thus, the Snellen VA notation is quantitatively equivalent to the Donders notation (and to the Monoyer decimal notation which is equivalent to it), but is more informative, since it also contains the indication of the viewing distance.

Knowing the viewing distance is essential, since, in general, VA depends on it (and not only in cases of refractive anomalies: cf. Heron et al., 1995; Rozhkova et al., 2004, 2005). However, in practice, the examiner regularly performs long series of measurements at a fixed, standard viewing distance (e.g., in examination of schoolchildren or in population surveys). It is therefore often reasonable to indicate the viewing distance only once for the whole assessment series and, in a dataset, to use corresponding decimal values of the Snellen fractions (or their denominators) in further statistical analysis.

For more than 150 years the Snellen notation of VA was prevalent, especially in Anglo-Saxon countries, and until recently remained broadly used for vision examination by clinicians who conducted routine measurements at standard far and near distances. However, instigated by significant progress in vision science and medical technology, in the 20<sup>th</sup> century there emerged the need to introduce other VA notations.

An extensive development, since 1960s, of the concept of spatial frequency and Fourier analysis in visual psychophysics, neurophysiology, as well as theoretical analysis of experimental findings on pattern recognition resulted in an increased use of the critical spatial frequency,  $F_c$ , of test gratings as the measure of VA (Campbell, Green, 1965; Campbell, Robson, 1968; Teller, 1979, 1997; Anderson, Thibos, 1999; a.o.). Actually, gratings had already been used for VA measure-



**Fig. 5.** The Snellen notation of VA. The top graph represents a below-standard visual acuity and the bottom graph an above-standard visual acuity. The distances  $D_0$  and  $D_i$  are the reference and the testee's threshold distances, respectively, providing equal angular sizes of  $1'$  for  $h_s$  at  $D_0$  and  $h_i$  at  $D_i$ .

ments more than a century ago: for instance, at the end of the 19<sup>th</sup> century, Wertheim (1894) designed small gratings (grids) from thin wire and estimated the largest distance at which the grating structure could be resolved. In effect, by this means Wertheim was measuring  $F_c$ .

The advancements of many recent investigations made the use of  $F_c$  as the VA measure more compelling than other VA measures and notations. One can outline at least following advantages of employing  $F_c$ .

- It is appealing that the  $F_c$  measure links the assessment of VA and contrast sensitivity function (CSF), since  $F_c$  corresponds to the highest spatial frequency discerned at the highest contrast level. Ideally,  $F_c$  can be estimated in the assessment of the full range of CSF, however, many researchers are interested only in the range of optimal spatial frequencies, thus limiting their measurements of CSF by the spatial frequencies which are essentially lower than  $F_c$ .

- In cross-sectional studies,  $F_c$  can be estimated in observers over the entire lifespan, from the neonatal age onwards (Teller, 1979; 1997; Vital-Durand et al., 1996; Woodhouse et al., 2007; Sturm et al., 2011), thus enabling investigation of age dynamics of VA in different populations. As the uniform and most unambiguous measure,  $F_c$  also provides an opportunity to monitor VA in individuals in longitudinal studies.

- The values of  $F_c$ , ranging from low to high, correspond to an intuitive interpretation of low and high vision quality.

- Graphic representation of the  $F_c$  scale for VA as the frequency-modulated pattern gives the most ade-

quate and direct representation of resolving power that increases with the increasing VA.

— Using the  $F_c$  scale for measuring VA provides an opportunity of a direct comparison of the human eye and optical devices with regard to their resolving power.

Significant advances in knowledge of the visual system functioning, on the one hand, and in technology, on the other hand, instigated attempts to facilitate association of VA scores with visual performance and life quality. The new forms of VA measures were also developed with the intention to make them more suitable for statistical analysis of VA data, and for comparison of different methods and procedures of VA assessment. For different purposes, various notations were proposed, discussed and modified, such as logMAR, Visual Efficiency (VE), Visual Acuity Score (VAS), Visual Acuity Rating (VAR), letter-by-letter scores of the Early Treatment of Diabetic Retinopathy Study Group (ETDRS) and others (Snell, Sterling, 1925; Ogle, 1953; Bailey, Lovie, 1976; Ferris et al., 1982; Holladay, 1997; Carkeet, 2001; Bourne et al., 2003; Kaido et al., 2007; Plainis et al., 2007; Gregory et al., 2010).

Each of these notations was developed in response to a certain need of VA assessment and was supposed to meet the requirements of a certain group of professionals, to achieve ease and efficiency of using VA data in the specific form. Each notation has its advantages, shortcoming, and is tailored to a specific area of application, being beneficial for some but not always for other purposes. Due to various purposes of assessing VA, one would anticipate that different VA notations will be used in parallel, depending on specific aims and objectives of individual investigations. Unexpectedly though, there arose an unwarranted competition between different notations, which culminated in the alleged “gold standard” among the tools for VA assessment and the claim of the corresponding “best notation” (Elliott, 2016). To elucidate foundations of VA assessment, it is reasonable to scrutinize different VA notations from various viewpoints.

**The VE notation**, visual efficiency, was introduced by Snell and Sterling in 1925. These researchers proceeded from the assumption that each increase of MAR by 1' would reduce the efficiency of the visual system functioning (in a broad sense) to 84% of the initial level. In literature, there are two different formulae approximating such relationship (Snell, Sterling, 1925; Westheimer, 1979):

$$VE = 0.2^{(MAR-1)/9}$$

$$VE = 100 \exp(-0.179[MAR - 1'])$$

where MAR is expressed in arcmin and VE in percent.

**The logMAR notation** was introduced by Bailey and Lovie (1976), Australian optometrists. The notation reflected the logMAR principles, implemented in the

Bailey-Lovie VA letter chart. Each chart line has five equi-legible<sup>1</sup> letters; from line to line, there is a geometric progression in letter size (constant ratio of size, 1.26 corresponding to 0.1 steps in logarithmic scale); the intervals between the letters and lines are proportional to the letter sizes, to hold constant contour interaction and letter-by-letter scoring, thus, ensuring a similar difficulty of the visual task. As Cole (2014, p. 1) graphically put it, the chart of Bailey and Lovie “took the ophthalmic world by storm. Their principles for the design of a visual acuity chart covered all the bases so that letter size is the only significant variable in measuring visual acuity”. Thus, the core idea behind the logMAR notation is to use the logarithmic function of MAR instead of the directly estimated MAR value. Remarkably, already in 1868 Green (1868; also, Green, 1905) developed a letter chart with a logarithmic scale; also the multiplier of the size progression was the same, as in the Bailey-Lovie chart, 1.26, if one advances from the smallest letters (Green indicated, although, the value  $0.79 = 1/1.26$ , since he proceeds from the largest letters at the top of his chart.) Unfortunately, at that time, the idea of logarithmic scaling was not yet appreciated.

The logMAR notation has been keenly promoted since its inception, to become widely used in the coming decades and dominating in both academic research and clinical practice. The main attractive feature of logMAR is that it transforms geometric progression of MAR values to arithmetic progression. LogMAR chart became very popular for VA assessment as it could provide equal accuracy in the whole range of measures. (However, it is an advantage of the chart design but not the logMAR notation.)

**The ETDRS letter scores** are based on the number of letters read correctly in the charts with logMAR design. They were proposed as the result of the experience obtained in the process of working with such charts to make the scores more easily collected, compared and interpreted. By designing the chart, the ETDRS Group adopted all of the principles of the Bailey-Lovie chart, although with some modifications (Ferris et al., 1982). The ETDRS chart has firmly enshrouded the logMAR principles. By now, various VA charts have been developed on the same principles, i.e., with a proportional design and the multiplier of 1.26 for letter sizes at the reference levels, which led to coining the term “logMAR design”.

**The VAR notation**, the visual acuity rating score, was proposed as an alternative way of denoting VA in logarithmic scale, while counteracting one of the main inconveniences of the logMAR notation – its negative values in a significant part of the VA range (Bailey, Lovie-Kitchin, 2013). Specifically, the logMAR zero value corresponds to MAR = 1.0' (standard vision); for

<sup>1</sup> Authors of the charts implementing the logMAR principle not infrequently refer to the employed letters as “equi-legible”; however, the criteria of equi-legibility vary between studies.

**Table 2.** Approximate correspondence of VA values in various VA notations and ranges of VA values reflecting the vision quality conventionally adopted in ophthalmology, optometry and vision science

Vision quality, normative levels	MAR, arcmin	logMAR	Donders and Monoyer (decimal value)	Snellen fraction (m/m) and its reduced value	$F_c$ , cpd	VE, %	VAR = VAS, %	ETDRS, number of letters read
Excellent-to-normal vision								
Standard eye	0.4	-0.4	2.5	6/2.4 = 2.5	75			
	0.5	-0.3	2.0	6/3.0 = 2.0	60	109.4	115	100
	0.63	-0.2	1.6	6/3.75 = 1.6	47.5	106.8	110	95
	0.79	-0.1	1.25	6/4.8 = 1.25	37.8	103.6	105	90
	<b>1.00</b>	<b>0.0</b>	<b>1.00</b>	<b>6/6.0 = 1.00</b>	<b>30</b>	<b>100.0</b>	<b>100</b>	<b>85</b>
	1.26	+0.1	0.8	6/7.5 = 0.8	23.8	95.6	95	80
Near-normal vision								
Driver license	1.58	+0.2	0.63	6/9.5 = 0.63	18.9	89.8	90	75
	<b>2.00</b>	<b>+0.3</b>	<b>0.50</b>	<b>6/12 = 0.5</b>	<b>15</b>	<b>83.6</b>	<b>85</b>	<b>70</b>
	2.51	+0.4	0.40	6/15 = 0.4	11.9	76.5	80	65
	3.16	+0.5	0.32	6/18.7 = 0.32	9.5	67.5	75	60
Moderate-to-low vision								
	3.98	+0.6	0.25	6/24 = 0.25	7.5	58.5	70	55
	5.01	+0.7	0.20	6/30 = 0.2	6.0	48.9	65	50
	6.31	+0.8	0.16	6/37.5 = 0.16	4.7	38.8	60	45
	7.94	+0.9	0.125	6/48 = 0.125	3.8	28.6	55	40
Severe low vision								
Legal blindness	<b>10.00↓</b>	<b>+1.00↓</b>	<b>0.100↓</b>	<b>6/60 = 0.1↓</b>	<b>3.0↓</b>	<b>20.0↓</b>	<b>50↓</b>	<b>35↓</b>

better-than-standard visual acuities ( $\text{MAR} < 1.0'$ ), logMAR values become negative, which is counterintuitive. As pointedly noted by Colenbrander (2008, p. 65), logMAR is a measure of *vision loss*, rather than a measure of visual acuity, so, an increase in MAR (or logMAR) means a decrease in vision.

The formula for calculating VAR is a simple transformation of the logMAR scale:

$$\text{VAR} = 100 - 50 \log \text{MAR}.$$

Basic information useful for comparison of the VA notations and VA notations discussed herein is presented in Table 2.

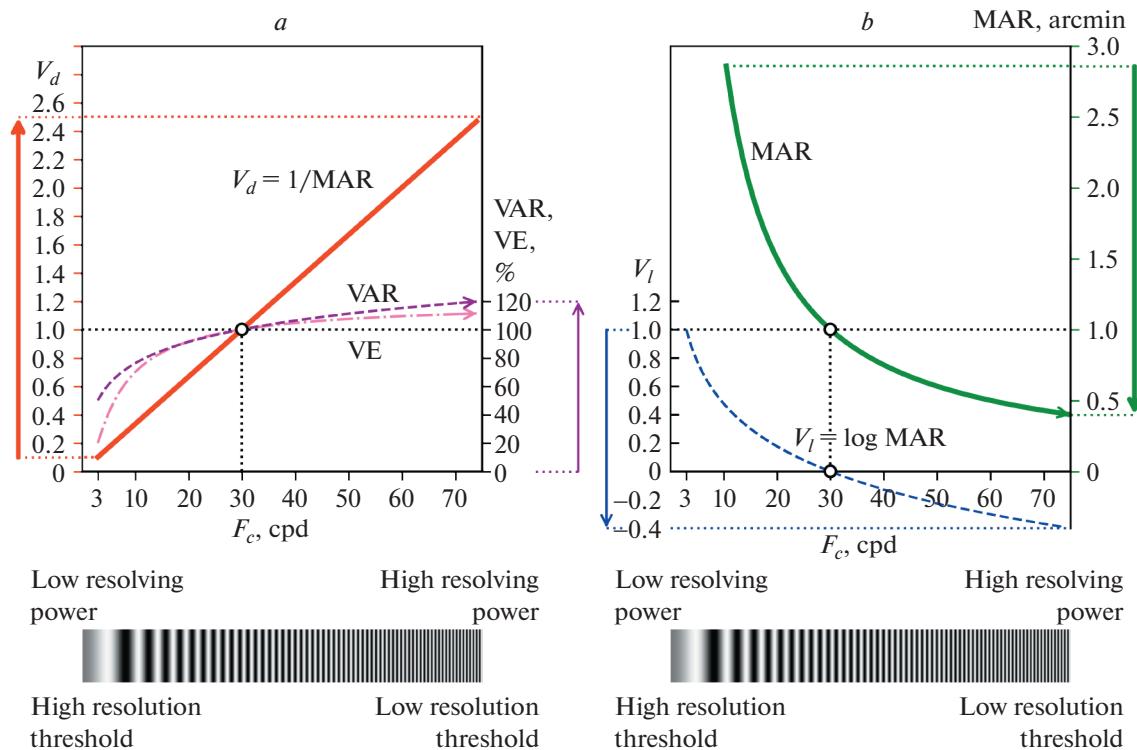
Table 2 presents the ranges of VA values, in different VA notations, aligned with the categorized quality of vision in accord with experts' conventions, as reflected in some handbooks (e.g., Somov, 1989; Bondarko et al., 1999; Shamshinova, Volkov, 1999; Colenbrander, 2001, 2008; Lennie, Van Hemel, 2002; Rozhkova, Matveev, 2007). Notably, the correspondence between the values presented in Table 2 can be considered only tentative. One need to bear in mind that the conversion of outcomes obtained by different tools, methods and under varying testing protocols is problematic, since for each method of measurement and

each VA notation, the outcome is stipulated by different physical parameters of the (threshold) test stimuli.

For example, Wesemann (2003) undertook comparison of VA estimates (in decimal notation) obtained by using three tests – the Landolt test, the Freiburg Visual Acuity & Contrast Test (FrACT), and the Bailey-Lovie chart. He found that VA values obtained by administering the Landolt test were lower by 0.5 units compared to the FrACT values, but the difference was within the German DIN tolerance limits; in comparison, VA values obtained using the Bailey-Lovie chart were on average 0.9 units lower than the Landolt test estimates. A consideration of VA scores based on grating and non-grating optotypes could be found in Thorn, Schwartz, 1990; Stiers, 2003; 2004; Strasburger et al., 2018, and a general discussion of dependence of VA on the employed measuring tools and assessment protocol one could see in (Cole, 2014).

We mention these findings in passing since the issue of comparability of the values (expressed in identical units) obtained by individual VA assessment tools is beyond the scope of the present paper.

We consider it more insightful to determine *relational functions* between the currently used VA measures. Crucially, this opens a new avenue in the quest to find the “grounding truth” measure of VA assess-



**Fig. 6.** Comparative characteristics of different VA measures as functions of the critical spatial frequency  $F_c$  ( $F_c = 30/\text{MAR}$ ) that signpost the quality of functional vision: *a* – VA measures based on resolving power (positively correlate with intuitive understanding of VA); *b* – VA measures based on resolution thresholds (negatively correlate with intuitive understanding of VA).

ment. In our view, the VA measure is supposed to reflect an intuitive understanding that higher VA values reflect better spatial vision, and that all VA values  $\geq 0$ , since even the smallest degree of spatial vision is certainly positive in comparison to blindness. Predicated by this understanding, the critical spatial frequency,  $F_c$ , was used as the benchmark for measuring VA.  $F_c$  is inversely proportional to MAR,  $F_c = 30/\text{MAR}$ ; so, unlike MAR values,  $F_c$  values increase for better visual acuities. In Fig. 6 we present the conventionally used VA measures as functions of  $F_c$ , signposting the quality of observers' functional vision. In addition, in the bottom of Fig. 6, vertical bars of increasing density illustrate the increasing frequency of spatial gratings (cpd), the measurement tool that reflects an increase in VA, or the quality of functional vision.

The two parts of Fig. 6 presents functions for two principally different types of the VA measures (and corresponding notations) in their relationship to  $F_c$ . Specifically, Fig. 6a shows ascending functions that correspond to the three VA measures reflecting resolving power – decimal (1/MAR), VE, and VAR. Each function shows an increase of the VA scores with the increasing quality of vision, as indicated by increasing  $F_c$ . The decimal measure,  $V_d = 1/\text{MAR}$ , that is directly proportional to  $F_c$ , has a linear positive relationship of VA scores with  $F_c$ . In comparison, the VAR and

VE functions rather resemble a logarithm relationship; although similar, they slightly differ in slopes at the lower and upper ends of the  $F_c$  range. The larger differences in the slopes of the VAR and VE functions, compared to 1/MAR (a linear function of  $F_c$ ), are correlated with an overestimation or underestimation of certain VA changes by practitioners guided by VAR and VE scores. In particular, with focus on visual impairments, practitioners are inclined to pay a greater attention to VA improvement within a range of poor vision, hence, might overestimate this improvement; conversely, they pay lesser attention to an improvement within a range of good or excellent vision, and might underestimate the improvement.

Fig. 6b, in contrast, presents two descending functions that are based on resolution thresholds and correspond to the MAR and logMAR measures. We reiterate that the thresholds are negatively related to the vision quality: ***the higher the threshold, the poorer the vision quality or the greater the vision loss***. The MAR and logMAR scores, apparently, are inverted in relation to the quality of spatial vision.

The inverted scores might be as good as the direct ones; it is essential, however, to take into account their properties. The shortcomings of the logMAR notation were discussed by Rozhkova (2017) in her paper in Russian and are briefly summarized here. These are of different kinds: some of them are related to metrolog-

ical properties of logMAR as the scale of measurement; the others concern an opaque relationship between logMAR values around its zero, characteristic for the majority of the population, and the VA values in other notations.

— The zero value of logMAR, which corresponds to the conventional normal vision (the standard eye with MAR = 1'), precludes using a ratio measurement scale, required by the principles of metrology, for comparison of the vision quality across various cases, since division by zero is impossible.

— Neither can other logMAR scale values be adequately compared via ratios, in particular, because it would be problematic to interpret negative ratio of outcomes in cases where pairs of negative and positive logMAR values are to be related.

— It is also difficult to fathom a meaningful relationship between logMAR values and VA values obtained using the other notations for observers with normal and near-normal vision (cf. Table 2). Specifically, in the decimal notation, VE and VAR, small changes in functional vision are reflected by commonsensical small changes of VA values over the whole VA range; whereas in the logMAR notation, around the logMAR zero, some of such small changes in functional vision may effect change of the sign of the VA value, prompting a spurious impression that these changes in the vision quality are more specific than those in other range.

— During preschool and school ages, most of children with normal and excellent vision quality have negative VA values on logMAR scale (see the blue parts of the histograms in the Fig. 7a–c). From psychological view, negative VA estimates, being associated with impairments, might appear deceptively disturbing for parents.

— Moreover, on the widespread logMAR charts (with 0.1 logMAR steps between the lines), the number of reference levels in the negative scale range are too few for timely revealing significant changes in a child's vision quality.

To illustrate the latter point, let us inspect Fig. 7. The top graphs (Fig. 7a–c) present histograms of VA values in decimal units, based on the data obtained without optical correction for binocular far vision in three age groups of primary schoolchildren using specially elaborated charts with tumbling E optotypes and the reference level steps of 0.2 decimal units (Rozhkova et al., 2001; Rozhkova, Matveev, 2007). The bottom graph, Fig. 7d, shows the logMAR function and, in the inset, repeated data for the 7-year-old children in a modified form (with the inverted order of bins to line up the direction of the vision quality changes with the logMAR scale). When the black bar, conventionally normal vision (decimal VA = 1), is aligned with logMAR = 0 (corresponding to MAR = 1'), it is apparent that, in the logMAR notation, more than 70% of the tested 7-year-old children have negative VA values

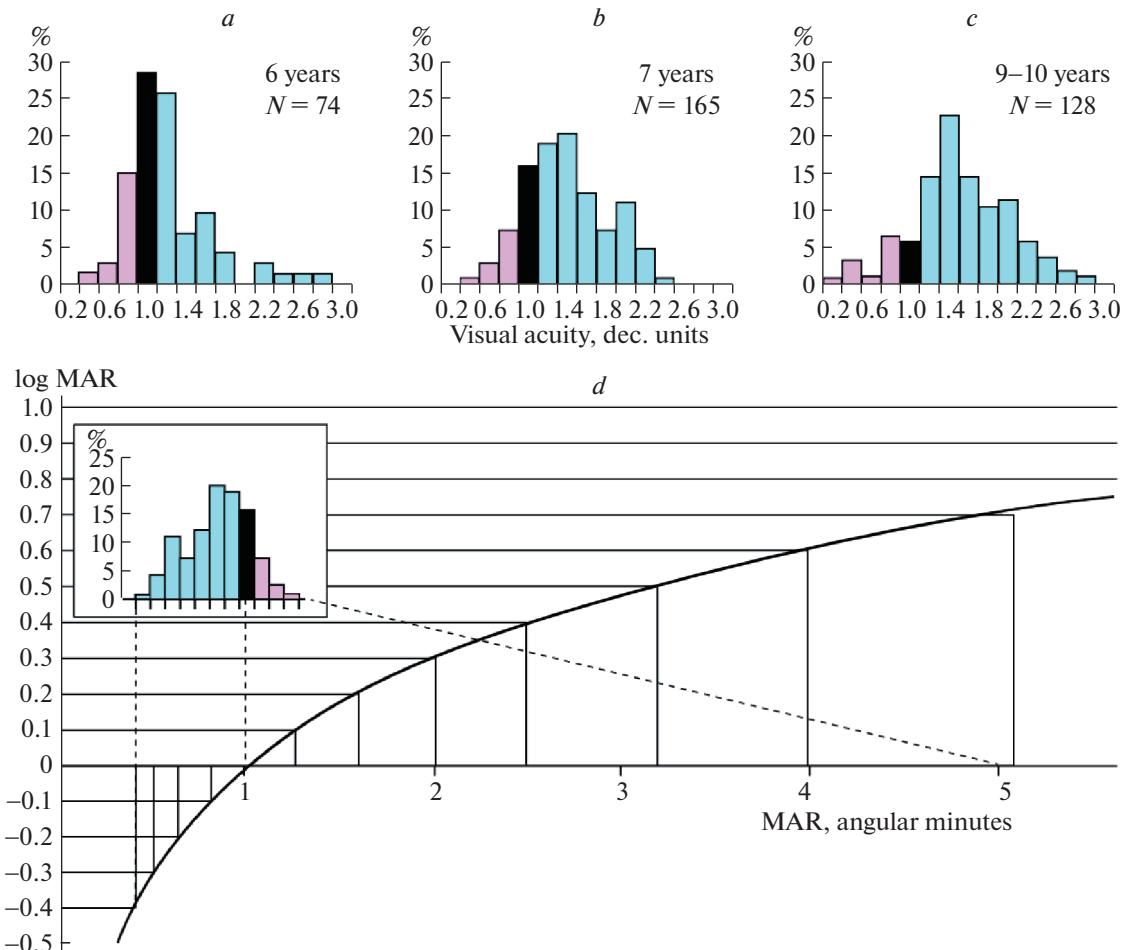
(blue bins). Note also that in this negative logMAR range, there are only four reference levels (with 0.1 logMAR steps). By comparing the three decimal VA histograms presented in Fig. 7abc, one could see that the proportion of children with negative logMAR values (corresponding to blue bins in the graphs) noticeably increases with age. It is problematic to expose this age dynamics properly using only four indicated logMAR reference levels presented in most VA charts based on logMAR notation.

Like many authors comparing various VA notations, we deliberately did not separate the VA notations *per se* from their embodiment in the measuring tools developed for practical use, although this would be required for a comprehensive analysis. However, this task is not viable in a brief overview. It seemed for us more important to provide researchers with a cautionary advice to abstain from conclusions inferred from a limited experience and/or uncritically rely on an authority's opinion.

Assessing the advantages of certain VA notation for practical use, the researcher considers and analyzes the experimental data obtained with the existing tools. However, ideally, it is necessary to investigate certain representative sample of tools including various optotypes, designs and notations. Sensitivity and specificity of the testing method, and reliability of its outcomes are dependent on many factors and conditions, in particular, on the parameters of the employed tool. Currently, the tools for VA assessment are predominantly charts. Regarding comparison the VA notations, it is important to bear in mind, that the design of the VA chart and the VA notation used for designating numeric outcomes are actually independent. This evident fact was clear to the author of the first chart with logarithmic (proportional) design but with decimal notation for designation of reference levels (Green, 1868). Unfortunately, in the XX century there appears and become spreading a groundless belief that the chart design implies the use of definite VA notation for outcomes of measurements.

However, we can indicate at least one opposite example – the chart elaborated by Kholina using proportional design and decimal notation for the reference levels (Kholina, 1930). Her chart contained 32 lines for the VA values varied from 0.1 to 2.0 with logarithmic steps of approximately 0.05 logMAR (the multiplier of the geometric progression for the optotype size was close to  $10^{1/24} \approx 1.1$ ), i.e. corresponding levels were: 0.1; 0.11; 0.12; 0.13; 0.15; ...; 0.8; 0.9; 1.0; 1.1; 1.2; 1.35; 1.50; 1.65; 1.8 and 2.0.

Fig. 8 shows real sets of the reference levels used in the described chart of Kholina with logMAR steps and decimal notation. The decimal notation is used in practically all Russian charts (Golovin, Sivtsev, 1926; Roslyakov, 2001; Gracheva et al., 2019; Stulova et al., 2019), on the one hand, and in the ETDRS charts with logMAR notation, predominantly employed in the



**Fig. 7.** An illustration of the relationship between the logMAR function and the decimal VA measures, presented as histograms of the data obtained for schoolchildren of three age groups (adopted from Rozhkova, 2017, Fig. 3).

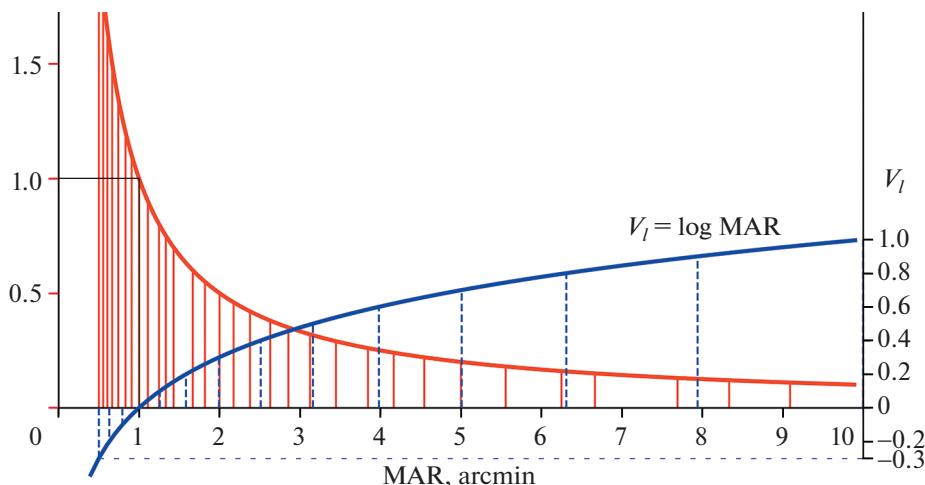
USA and Europe (Ferris et al., 1982; Plainis et al., 2007), on the other hand. As is well known, in ETDRS chart, the intervals between the neighboring reference levels are equal to 0.1 logMAR that is equivalent to multiplying the letter size by 1.26 (26%-steps).

Comparing the two presented distributions of the reference levels, one can conclude that, because of its greater density, the chart of Kholina could provide better accuracy of measurement in the whole range of the VA values.

It is important to emphasize that, in all cases, one could vary the sets of the reference levels independently of the VA notation. It seems puzzling that by now the set of reference levels used for a certain chart is considered to be linked to the VA notation. A trivial reason of linking the reference levels and certain notation is probably to simplify numerical values of outcomes by using “round” values for practical purposes. From such viewpoint, the easiest and widespread way (often implicitly used in practice and noted in some guidelines) is to indicate the reference line numbers

corresponding to testee’s thresholds: “He sees the 5th line”, or “She sees the 3rd line and 3 letters in the 4th line”. During an examination session, it is sufficient to make such recordings which contain information required for practical purposes, while leaving the task of data interpretation and analysis for future or to other researchers. Noteworthy, many optometrists often employ this scoring of VA by the line numbers since it is simple and understandable to every participant of the measuring procedure.

By close of this overview, we cannot but remark that, with time, the task of optimizing the sets of the reference levels in printed charts becomes less actual, since in the near(est) future these will surely be supplanted by computerized procedures for VA assessment, implying the shift to the task of developing optimal algorithms for varying stimulus sizes. Such algorithms were already emerging in the second half of the 20th century (Taylor, Creelman, 1967; Watson, Pelli, 1983) and continue to be improving at present. More than that: global computerization will enable both standardization and variegation of the VA procedure



**Fig. 8.** Reference levels used in the two charts with proportional designs but different VA notations: the chart of Kholina (1930) with decimal notation of VA (red:  $V_d = 1/MAR$ ) and the ETDRS chart with logMAR notation of VA (blue:  $V_l = \log MAR$ ).

stipulating quite another issues and approaches to its execution — novel paradigms, principles, optotypes, and protocols.

Another factor that probably will positively change the practice of VA assessment is the envisaged transition from sporadic VA screening to its frequent monitoring. In the latter case, short intervals between the measurement sessions would make employment of such tools as the VA charts problematic, since frequent presentation of the same chart during the test session may ensue memorizing of its elements/lines, which would result in biased outcomes of the retest.

In computer-aided methods, the notion of the reference test level actually becomes unnecessary since the test stimulus size can be voluntarily modified after each stimulus presentation. In these cases, the crucial feature is the display resolution as the limiting factor of the possible increments/decrements of the optotype sizes. This is illustrated in Fig. 9 that shows feasible sizes of the stimuli calculated on the basis on the screen resolution (pixel sizes) for three different display brands at the viewing distance of 5 m (Terekhin et al., 2015; Rozhkova, Malykh, 2017). The points on the graph inside the three gray bars correspond to varying stimulus size by adding pixel by pixel. It is clearly seen that such one-pixel steps do not limit essentially the number of the stimulus size gradations in the range of the lowest VA but could be critical for the range of the high VA values.

To develop proper and convenient computer-aided methods of the VA measurement, both scientific and technical components of the elaborated setup become equally important.

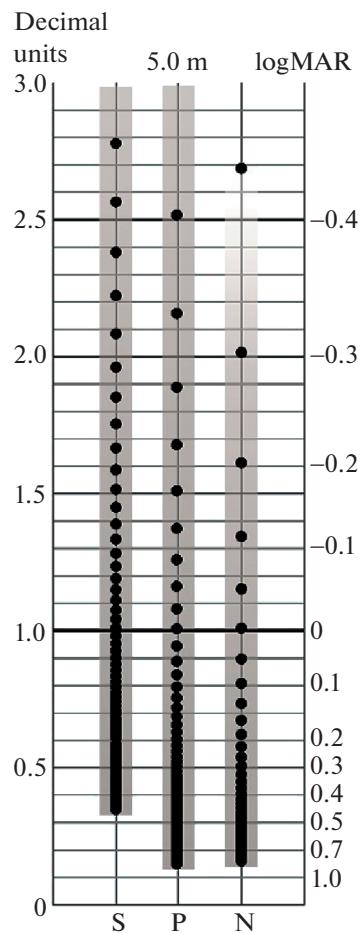
## CONCLUDING REMARKS

More than 40 years ago Westheimer (1979, p. 327) wrote in his paper on the VA measurement and scaling (italics added by the authors):

“We base the discussion on the premise that it is possible to arrive at *a single value for the visual acuity of an eye under a given set of conditions. It is a threshold,  $\alpha$ , the minimum angle of resolution (MAR), and is to be measured in minutes of arc. The method of determining the threshold is immaterial for this discussion, and so is the nature of the target.* The arguments apply equally to Snellen letter, checkerboard, and grating targets, and to evoked potentials, optokinetic, or forced-choice psychophysical responses. *What concerns us here is the scale in which to place a set of visual acuity measurements expressed in minutes of arc.*”

In the present paper we share this approach bar reservation about Westheimer's stance that the method of determining the threshold is “immaterial for this discussion”. In our opinion, the task of interrelating results of different measuring procedures that determine the MAR is not yet resolved and, in principle, cannot be solved in a uniform way. As was pointedly worded by Pirenne (1962, p. 175) in his frequently cited quotation, “[t]here are in fact as many different “visual acuities” as there are types of test objects”. This problem was, however, not addressed here and will be considered in more detail in the second part of our overview.

The undertaken analysis of different VA measures and notations shows that the advantages, shortcomings and applicability of a given notation should be assessed by considering multiple factors: the essence of the VA notation in question; the methods implicating its use in the direct form (i.e., without recalculations); intuitive comprehension of the notation; easiness of its understanding for both testees and the examiner; the



**Fig. 9.** The sets of feasible reference levels of VA (points) in decimal (left scale) and logMAR (right scale) notations calculated for the case of the tumbling-E optotype presentation on a smartphone (S), iPhone (P) and Notebook (N) from the observation distance of 5 meters. S – Samsung Galaxy S6; pixel size 0.044 mm; P – iPhone 5; pixel size 0.078 mm; N – Sony VAIO; pixel size 0.18 mm.

notation's applicability for assessing individuals of different age groups; insights from the population-based surveys, etc.

We hope that this paper provides convincing arguments in favor of a wider use of the maximal resolvable spatial frequency – the critical frequency,  $F_c$  – as a measure of VA. This metric is close to both the optical standards and ones' intuitive understanding of VA: the better VA, the higher  $F_c$ ; in addition, the relationship between  $F_c$  and VA can be compellingly presented graphically.

The Donders and Monoyer notations, as well as the Snellen fraction in its decimal form are all proportional to  $F_c$ , and, thus, are equally suited for expressing VA. Visual efficiency, VE, is similar to  $F_c$ , and the Donders and Monoyer notations in that it corresponds to one's intuitive concepts. The nonlinear dependence of VE on  $F_c$ , in the low-vision range, designates a greater impact of VA increase (or decrease), compared with the case of good and excellent vision.

There are some cautious signs that the logMAR appeal has passed its peak: recently, there appeared a tendency to return to Snellen and decimal notations of VA. As remarked by Tsou et al. (2017, p. 1): "...many ophthalmologists do not understand non-Snellen formats, such as logarithm of the Minimum Angle of Resolution (logMAR) or Early Treatment Diabetic Retinopathy Study (ETDRS) letter scores. As a result, some journals, since at least 2013, have instructed authors to provide approximate Snellen equivalents next to non-Snellen visual acuity values."

We conclude with a suggestion of an "inclusive practice": may each researcher and practitioner use the method of VA assessment and the VA notation that is most appropriate for addressing their research question or purpose in a clinical setting, for collecting and analyzing the data required for these – provided the researcher and the practitioner profoundly understand theoretical and methodological bases of the VA measurement and take into account the metric properties of the chosen notation.

#### ACKNOWLEDGMENTS

The authors are very grateful to Hans Strasburger for fruitful discussions and valuable comments, and to Alexander Belokopytov for technical assistance.

#### COMPETING INTERESTS

The authors have declared no competing interests.

#### REFERENCES

- Altman D.G., Gore S.M., Gardner M.J., Pocock S.J. Statistical guidelines for contributors to medical journals. *Annals of Clinical Biochemistry*. 1992. V. 29 (1). P. 1–8. <https://doi.org/10.1177/000456329202900101>
- Anderson R.S., Thibos L.N. Sampling limits and critical bandwidth for letter discrimination in peripheral vision. *Journal of the Optical Society of America A*. 1999. V. 16 (10). P. 2334–2342. <https://doi.org/10.1364/JOSAA.16.002334>
- Artal P. Optics of the eye and its impact in vision: a tutorial. *Advances in Optics and Photonics*. 2014. V. 6 (3). P. 340–367. <https://doi.org/10.1364/AOP.6.000340>
- Artal P., Chen L., Fernández E.J., Singer B., Manzanera S., Williams D.R. Neural compensation for the eye's optical aberrations. *Journal of Vision*. 2004. V. 4 (4). P. 281–287. <https://doi.org/10.1167/4.4.4>
- Bach M. The Freiburg Visual Acuity test –Automatic measurement of visual acuity. *Optometry and Vision Science*. 1996. V. 73 (1). P. 49–53. <https://doi.org/10.1097/00006324-199601000-00008>
- Bach M. The Freiburg Visual Acuity Test-Variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*. 2007. V. 245 (7). P. 965–971. <https://doi.org/10.1007/s00417-006-0474-4>

- Bailey I.L., Lovie J.E. New design principles for visual acuity letter charts. *American Journal of Optometry and Physiological Optics*. 1976. V. 53 (11). P. 740–745.  
<https://doi.org/10.1097/00006324-197611000-00006>
- Bailey I.L., Lovie-Kitchin J.E. Visual acuity testing. From the laboratory to the clinic. *Vision Research*. 2013. V. 90. P. 2–9.  
<https://doi.org/10.1016/j.visres.2013.05.004>
- Beck R.W., Moke P.S., Turpin A.H., Ferris III, F.L., San-Giovanni J.P., Johnson C.A., Birch E.E., Chandler D.L., Cox T., Blair C., Kraker, R.T. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *American Journal of Ophthalmology*. 2003. V. 135 (2). P. 194–205.  
[https://doi.org/10.1016/S0002-9394\(02\)01825-1](https://doi.org/10.1016/S0002-9394(02)01825-1)
- Bohigian G.M. An ancient eye test—using the stars. *Survey of Ophthalmology*. 2008. V. 53 (5). P. 536–539.  
<https://doi.org/10.1016/j.survophthal.2008.06.009>
- Bondarko V.M., Danilova M.V., Krasil'nikov N.N., Leushina L.I., Nevskaya A.A., Shelepin Yu.E. *Prostranstvennoe zrenie [Spatial vision]*. St. Petersburg: Nauka. 1999. 218 p. (in Russian).
- Bourne R.R.A., Rosser D.A., Sukudom P., Dineen B., Laidlaw D.A.H., Johnson G.J., Murdoch J.E. Evaluating a new logMAR chart designed to improve visual acuity assessment in population-based surveys. *Eye*. 2003. V. 17. P. 754–758.  
<https://doi.org/10.1038/sj.eye.6700500>
- Campbell F.W., Green D.G. Optical and retinal factors affecting visual resolution. *The Journal of Physiology*. 1965. V. 181 (3). P. 576–593.  
<https://doi.org/10.1113/jphysiol.1965.sp007784>
- Campbell F.W., Robson J.G. Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*. 1968. V. 197 (3). P. 551–566.  
<https://doi.org/10.1113/jphysiol.1968.sp008574>
- Carkeet A. Modeling logMAR visual acuity scores: effects of termination rules and alternative forced-choice options. *Optometry and Vision Science*. 2001. V. 78. P. 529–538.  
<https://doi.org/10.1097/00006324-200107000-00017>
- Cole B.L. Measuring visual acuity is not as simple as it seems. *Clinical and Experimental Optometry*. 2014. V. 97. P. 1–2.  
<https://doi.org/10.1111/cxo.12123>
- Colenbrander A. Measuring vision and vision loss. In Tasman W., Jaeger E.A. (Eds.), *Duane's clinical ophthalmology*. Vol. 5 (Ch. 51). Philadelphia, PA: Lippincott Williams & Wilkins. 2001. P. 2–42.
- Colenbrander A. The historical evolution of visual acuity measurement. *Visual Impairment Research*. 2008. V. 10 (2–3). P. 57–66.  
<https://doi.org/10.1080/1388235080263240>
- Curcio C.A., Sloan K.R., Kalina R.E., Hendrickson A.E. Human photoreceptor topography. *The Journal of Comparative Neurology*. 1990. V. 292 (4). P. 497–523.  
<https://doi.org/10.1002/cne.902920402>
- Duncan R.O., Boynton G.M. Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*. 2003. V. 38 (4). P. 659–671.  
[https://doi.org/10.1016/S0896-6273\(03\)00265-4](https://doi.org/10.1016/S0896-6273(03)00265-4)
- Elliott D.B. The good (logMAR), the bad (Snellen) and the ugly (BCVA, number of letters read) of visual acuity measurement. *Ophthalmic and Physiological Optics*. 2016. V. 36 (4). P. 355–358.  
<https://doi.org/10.1111/opo.12310>
- Elliott D.B., Sheridan M. The use of accurate visual acuity measurements in clinical anti-cataract formulation trials. *Ophthalmic and Physiological Optics*. 1988. V. 8 (4). P. 397–401.  
<https://doi.org/10.1111/j.1475-1313.1988.tb01176.x>
- Ferris F.L., Kassoff A., Bresnick G.H., Bailey I.L. New visual acuity charts for clinical research. *American Journal of Ophthalmology*. 1982. V. 94 (1). P. 91–96.  
[https://doi.org/10.1016/S0896-6273\(03\)00265-4](https://doi.org/10.1016/S0896-6273(03)00265-4)
- Golovin S.S., Sivtsev D.A. *Tablitsy dlya issledovaniya ostroty zreniya [Charts for measurement of visual acuity]*. Moscow: Gosidatel'stvo. 1926 (in Russian).
- Gracheva M.A., Kazakova A.A., Pokrovskiy D.F., Medvedev I.B. *Tablitsy dlya otsenki ostroty zreniya: analiticheskii obzor, osnovnye terminy [Visual acuity charts: analytical review, basic terms]*. *Annals of the Russian Academy of Medical Sciences*. 2019. V. 74 (3). P. 192–199 (in Russian)  
<https://doi.org/10.15690/vramn1142>
- Green J. On a new series of test-letters for determining the acuteness of vision. *Transactions of the American Ophthalmological Society*. 1868. V. 1 (4–5). P. 68–71.
- Green J. Notes on the clinical determination of the acuteness of vision, including the construction and graduation of optotypes, and on systems of notation. *Transactions of the American Ophthalmological Society*. 1905. V. 10 (3). P. 644–654.
- Gregory N.Z., Feuer W.F., Rosenfeld P.J. Novel method for analyzing Snellen visual acuity measurements. *Retina*. 2010. V. 30 (7). P. 1046–1050.  
<https://doi.org/10.1097/IAE.0b013e3181d87e04>
- Hecht S., Mintz E.U. The visibility of single lines at various illuminations and the retinal basis of visual resolution. *Journal of General Physiology*. 1939. V. 22 (5). P. 593–612.  
<https://doi.org/10.1085/jgp.22.5.593>
- Heinrich S.P., Bach M. Resolution acuity versus recognition acuity with Landolt-style optotypes. *Graefe's Archive for Clinical and Experimental Ophthalmology*. 2013. V. 251 (9). P. 2235–2241.  
<https://doi.org/10.1007/s00417-013-2404-6>
- Helmoltz H.L.F. *Handbuch der physiologischen Optik*. 1867. L. Voss: Leipzig (in German).
- Helmholtz H.L.F. *Helmholtz' treatise on physiological optics*. Trans. From third German ed., ed. by J.P.C. Southall, published by Optical Society of America. 3 vols. Menasha, WI: G. Banta Co. 1924–1925.
- Heron G., Furby H.P., Walker R.J., Lane C.S., Judge O.J.E. Relationship between visual acuity and observation distance. *Ophthalmic and Physiological Optics*. 1995. V. 15 (1). P. 23–30.  
<https://doi.org/10.1046/j.1475-1313.1995.9592788.x>
- Holladay J.T. Proper method for calculating average visual acuity. *Journal of Refractive Surgery*. 1997. V. 13 (4). P. 388–391.  
<https://doi.org/10.3928/1081-597X-19970701-16>

- Hyon J.Y., Yeo H.E., Seo J.M., Lee I.B., Lee J.H., Hwang J.M. Objective measurement of distance visual acuity determined by computerized optokinetic nystagmus test. *Investigative Ophthalmology & Visual Science*. 2010. V. 51 (2). P. 752–757.  
<https://doi.org/10.1167/iovs.09-4362>
- International Vocabulary of Metrology – Basic and General Concepts and Associated Terms* (3rd ed.), Joint Committee for Guides in Metrology. 2008. P. 6  
[http://www.bipm.org/utils/common/documents/jcgm/JCGM\\_200\\_2008.pdf](http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2008.pdf) (accessed on 08.01.2021)
- Jerri A.J. The Shannon sampling theorem – its various extensions and applications: A tutorial review. *Proceedings of the IEEE*. 1977. V. 65 (11). P. 1565–1596.  
<https://doi.org/10.1109/PROC.1977.10771>
- Kaido M., Dogru M., Ishida R., Tsubota K. Concept of functional visual acuity and its applications. *Cornea*. 2007. V. 26. P. S29–S35.  
<https://doi.org/10.1097/ICO.0b013e31812f6913>
- Kholina A. Novaya tablitsa dlya issledovaniya ostroty zreniya [A new chart for assessment of visual acuity]. *Russkii oftalmologicheskii zhurnal*. 1930. V. 11 (1). P. 42–47 (in Russian).
- Koskin S.A. *Sistema opredeleniya ostroty zreniya v tselyakh vrachebnoi ekspertizy* [A system of determining visual acuity for medical expertise]. MD thesis. St. Petersburg. 2009. 48 p. (in Russian).
- Lennie P., Van Hemel S.B. (Eds.). *Visual impairments: Determining eligibility for social security benefits*. Washington, DC: National Academies Press. 2002.  
<https://doi.org/10.17226/10320>
- Lovie-Kitchin J.E. Validity and reliability of visual acuity measurements. *Ophthalmic and Physiological Optics*. 1988. V. 8 (4). P. 363–370.  
<https://doi.org/10.1111/j.1475-1313.1988.tb01170.x>
- Lovie-Kitchin J.E., Brown B. Repeatability and intercorrelations of standard vision tests as a function of age. *Optometry and Vision Science*. 2000. V. 77 (8). P. 412–420.  
<https://doi.org/10.1097/00006324-200008000-00008>
- Miller W.H., Bernard G.D. Averaging over the foveal receptor aperture curtails aliasing. *Vision Research*. 1983. V. 23 (12). P. 1365–1369.  
[https://doi.org/10.1016/0042-6989\(83\)90147-5](https://doi.org/10.1016/0042-6989(83)90147-5)
- Monoyer F. Échelle typographique décimauex pour mesurer l'acuité visuelle. 1875. *Gazette Médicale de Paris*. V. 21. 258–259 (in French).
- Ogle K.N. On the problem of an international nomenclature for designating visual acuity. *American Journal of Ophthalmology*. 1953. 36 (7). P. 909–921.  
[https://doi.org/10.1016/0002-9394\(53\)92172-2](https://doi.org/10.1016/0002-9394(53)92172-2)
- Pfeiffer R.L. Frans Cornelis Donders Dutch physiologist and ophthalmologist. *Bulletin of the New York Academy of Medicine*. 1936. V. 12 (10). P. 566–581.
- Pirenne M.H. Visual acuity. In H. Davson (Ed.), *The eye. Vol. 2: The visual process*. New York/London: Academic Press. 1962. P. 175–195.  
<https://doi.org/10.1016/B978-1-4832-3089-4.50018-2>
- Plainis S., Tzatzala P., Orphanos Y., Tsilimbaris M.K. A modified ETDRS visual acuity chart for European-wide use. *Optometry and Vision Science*. 2007. V. 84 (7). P. 647–653.  
<https://doi.org/10.1097/OPX.0b013e3180dc9a60>
- Polyak S.L. *The retina: The anatomy and the histology of the retina in man, ape, and monkey, including the consideration of visual functions, the history of physiological optics, and the histological laboratory technique*. Chicago: University of Chicago Press. 1941. 607 p.
- Porter J., Queener H., Lin J., Thorn K., Awwal A.A.S. (Eds.). *Adaptive optics for vision science: principles, practices, design, and applications*. New York: Wiley. 2006. 595 p.
- Putnam N.M., Hofer H.J., Doble N., Chen L., Carroll J., Williams D.R. The locus of fixation and the foveal cone mosaic. *Journal of Vision*. 2005. V. 5 (7). P. 632–639.  
<https://doi.org/10.1167/5.7.3>
- Radner W. Reading charts in ophthalmology. *Graefe's Archive for Clinical and Experimental Ophthalmology*. 2017. V. 255 (8). P. 1465–1482.  
<https://doi.org/10.1007/s00417-017-3659-0>
- Rayleigh L. Investigations in optics, with special reference to the spectroscope. *Philosophical Magazine*. 1879. V. 8 (49): 261–274.  
<https://doi.org/10.1080/14786447908639684>
- Roorda A., Williams D.R. The arrangement of the three cone classes in the living human eye. *Nature*. 1999. V. 397 (6719). P. 520–522.  
<https://doi.org/10.1038/17383>
- Roorda A., Romero-Borja F., Donnelly W.J., Queener H., Hebert T.J., Campbell M.C. Adaptive optics scanning laser ophthalmoscopy. *Optics Express*. 2002. V. 10 (9). P. 405–412.  
<https://doi.org/10.1364/OE.10.000405>
- Roslyakov V.A. Novye tablitsy dlya izmereniya ostroty zreniya [New charts for visual acuity assessment]. *Rossiiskii oftalmologicheskii zhurnal*. 2001. (1). P. 36–38 (in Russian).
- Rosser D.A., Cousins S., Murdoch I.E., Fitzke F.W., Laidlaw D.A. How sensitive to clinical change are ETDRS logMAR visual acuity measurements? *Investigative Ophthalmology & Visual Science*. 2003. V. 44 (8). P. 3278–3281.  
<https://doi.org/10.1167/iovs.02-1100>
- Rozhkova G.I. LogMAR dlya ostroty zreniya khuzhe, chem loshadinya sila dlya moshchnosti elektricheskoi lampochki [LogMAR is worse for visual acuity than horsepower for electric lamp]. *Sensornye sistemy* [Sensory Systems]. 2017. V. 31 (1). P. 31–43 (in Russian).
- Rozhkova G.I. Est' li real'nye osnovaniya scitat' tablitsy ETDRS "zolotym standartom" dlja izmerenij ostroty zrenija? [Are there true reasons to consider ETDRS charts as a "golden standard" for measuring visual acuity?] *Russian Military Medical Academy Reports*. 2018. V. 37 (2). P. 120–123 (in Russian).
- Rozhkova G.I., Malykh T.B. Sovremennye aspekty standartizacii vizometrii [Current issues of standardization in visometry]. *Aviakosmicheskaya i ekologicheskaya meditsina*. 2017. V. 51 (6). P. 5–16 (in Russian).  
<https://doi.org/10.21687/0233-528X-2017-51-6-5-16>
- Rozhkova G.I., Matveev S.G. *Zrenie detei: problemy otsenki i funktsional'noi korrektssi* [Vision in children: Problems of the assessment and functional correction]. Moscow: Nauka, 2007. 315 p. (in Russian).

- Rozhkova G.I., Tokareva V.S., Vaschenko D.I., Vasiljeva N.N. Vozrastnaya dinamika ostroty zreniya u shkol'nikov. I. Binokulyarnaya ostrota zreniya dlya dali [Age dynamics of visual acuity in schoolchildren. I. Binocular visual acuity for far distance]. *Sensornye sistemy [Sensory Systems]*. 2001. V. 15 (1). P. 47–52 (in Russian).
- Rozhkova G.I., Tokareva V.S., Nikolaev D.P., Ognivov V.V. Osnovnye tipy zavisimosti ostroty zreniya ot rasstoyaniya u cheloveka v raznom vozraste po rezul'tatam diskriminantnogo analiza [Main types of dependence of visual acuity on the distance in individuals of different age based on discriminant analysis results]. *Sensornye sistemy [Sensory Systems]*. 2004. V. 18 (4). P. 330–338 (in Russian).
- Rozhkova G.I., Podugolnikova T.A., Vasiljeva N.N. Visual acuity in 5–7-year-old children: Individual variability and dependence on observation distance. *Ophthalmic and Physiological Optics*. 2005. V. 26. P. 66–80. <https://doi.org/10.1111/j.1475-1313.2004.00263.x>
- Shamshinova A.M., Volkov V.V. *Funktional'nye metody issledovaniya v oftal'mologii [Functional methods of investigation in ophthalmology]*. Moscow: Nauka. 1999. 416 p. (in Russian).
- Siderov J., Tiu A.L. Variability of measurements of visual acuity in a large eye clinic. *Acta Ophthalmologica Scandinavica*. 1999. V. 77. P. 673–676. <https://doi.org/10.1034/j.1600-0420.1999.770613.x>
- Sloan L.L. Needs for precise measures of acuity: Equipment to meet these needs. *Archives of Ophthalmology*. 1980. V. 98 (2). P. 286–290. <https://doi.org/10.1001/archopht.1980.01020030282008>
- Snell A.C., Sterling S. The percentage evaluation of macular vision. *Transactions of the American Ophthalmological Society*. 1925. V. 23. P. 204–227.
- Snellen H., Graham C.H. *Probebuchstaben zur Bestimmung der Sehschärfe [Test letters for determining visual acuity]*. Utrecht: Van de Weijer. 1862. 19 p. (in German).
- Snyder A.W., Miller W.H. Photoreceptor diameter and spacing for highest resolving power. *Journal of the Optical Society of America*. 1977. V. 67 (5). P. 696–698. <https://doi.org/10.1364/JOSA.67.000696>
- Somov E.E. *Metody oftal'moergonomiki [Methods of ophthalm ergonomics]*. Leningrad: Nauka, 1989. 157 p. (in Russian).
- Stiers P., Vanderkelen R., Vandebussche E. Optotype and grating visual acuity in preschool children. *Investigative Ophthalmology & Visual Science*. 2003. V. 44 (9). P. 4123–4130. <https://doi.org/10.1167/iovs.02-0739>
- Stiers P., Vanderkelen R., Vandebussche E. Optotype and grating visual acuity in patients with ocular and cerebral visual impairment. *Investigative Ophthalmology & Visual Science*. 2004. V. 45 (12). P. 4333–4339. <https://doi.org/10.1167/iovs.03-0822>
- Stevens S.S. On the theory of scales of measurement. *Science*. 1946. V. 103 (2684). P. 677–680. <https://doi.org/10.1126/science.103.2684.677>
- Strasburger H., Bach M., Heinrich S. P. Blur unblurred – A mini tutorial. *i-Perception*. 2018. V. 9 (2). P. 1–15. <https://doi.org/10.1177/2041669518765850>
- Stulova A.N., Semenova N.S., Akopyan V.S. Otsenka ostroty zreniya: vzglyad v proshloe i sovremennoe tendentsii [Visual acuity assessment: historical overview and current trends]. *Vestnik oftal'mologii*. 2019. V. 135 (6). P. 141–146 (in Russian). <https://doi.org/10.17116/oftalma2019135061141>
- Sturm V., Cassel D., Eisenman M. Objective estimation of visual acuity with preferential looking. *Investigative Ophthalmology & Visual Science*. 2011. V. 52 (2). P. 708–713. <https://doi.org/10.1167/iovs.09-4911>
- Taylor M.M., Creelman C.D. PEST: Efficient Estimates on Probability Functions. *The Journal of the Acoustical Society of America*. 1967. V. 41 (4A). P. 782–787. <https://doi.org/10.1121/1.1910407>
- Tehrani N.M., Riazi-Esfahani H., Jafarzadehpur E., Mirzajani A., Talebi H., Amini A., Mazloumi M., Roohipoor R., Riazi-Esfahani M. Multifocal electroretinogram in diabetic macular edema; correlation with visual acuity and optical coherence tomography. *Journal of Ophthalmic & Vision Research*. 2015. V. 10 (2). P. 165–171. <https://doi.org/10.4103/2008-322X.163773>
- Teller D. The forced-choice preferential looking procedure: A psychophysical technique for use in human infants. *Infant Behavior & Development*. 1979. V. 2. P. 135–153. [https://doi.org/10.1016/S0163-6383\(79\)80016-8](https://doi.org/10.1016/S0163-6383(79)80016-8)
- Teller D. The first glances: The vision of infants. *Investigative Ophthalmology & Visual Science*. 1997. V. 38 (11). P. 2183–2203.
- Terekhin A.P., Gracheva M.A., Rozhkova G.I., Lebedev D.S. *Svidetel'stvo 2015616714 Rossiiskaya Federatsiya. [Certificate of the state registration of a computer program. Interactive program for visual acuity assessment based on the accurate threshold measurement using three optotypes "Tip-Top"]*; the applicant and copyright holder: A.A. Kharkevich FGBU IPPI RAN (RU). – No.2014619697; submitted: 26.09.2014; published: 19.06.2015. (in Russian). [iitp.ru/ru/patents/1293.htm](http://iitp.ru/ru/patents/1293.htm)
- Thorn F., Schwartz F. Effects of dioptric blur on Snellen and grating acuity. *Optometry and Vision Science*. V. 67 (1). P. 3–7. <https://doi.org/10.1097/00006324-199001000-00002>
- Tsou B.C., Bressler N.M. Visual acuity reporting in clinical research publications. *JAMA Ophthalmology*. 2017. V. 135 (6). P. 651–653. <https://doi.org/10.1001/jamaophthalmol.2017.0932>
- Vital-Durand F., Ayzac L., Pinzar G. Acuity cards and the determination of risk factors in 6–8 months infants. In F. Vital-Durand, J. Atkinson, O.J. Braddick (Eds.), *Infant vision*. New York: Oxford University Press. 1996. P. 185–200.
- Watson A.B., Pelli D.G. QUEST: A general multidimensional bayesian adaptive psychometric method. *Perception & Psychophysics*. 1983. V. 33 (2). P. 113–120. <https://doi.org/10.1167/17.3.10>
- Wertheim T. Üeber die indirekte Sehschärfe [On indirect visual acuity]. *Zeitschrift für Psychologie & Physiologie der Sinnesorgane*. 1894. V. 7. P. 172–187 (in German).
- Wesemann W. Die Grenzen der Sehschärfe, Teil 6: Welche Sehschärfe erreicht der Mensch? [The limits of visual acuity, Part 6: What is the best visual acuity in possible in a human being?]. *Optometrie*. 2003. (2). P. 42–47 (in German).

- Westheimer G. Image quality in the human eye. *Optica Acta*. 1970. V. 17 (9). P. 641–658.  
<https://doi.org/10.1080/713818355>
- Westheimer G. Visual acuity and hyperacuity. *Investigative Ophthalmology & Visual Science*. 1975. V. 14 (8). P. 570–572.
- Westheimer G. Scaling of visual acuity measurements. *Archives of Ophthalmology*. 1979. V. 97 (2). P. 327–330.  
<https://doi.org/10.1001/archophth.1979.01020010173020>
- Westheimer G. Updating the classical approach to visual acuity. *Clinical and Experimental Optometry*. 2001. V. 84 (5). P. 258–263.  
<https://doi.org/10.1111/j.1444-0938.2001.tb05035.x>
- Westheimer G. Visual acuity and hyperacuity. In M. Bass, C. DeCusatis, J.M. Enoch, V. Lakshminarayanan, G. Li, C.A. MacDonald, V.N. Mahajan, E. Van Stryland (Eds.), *Handbook of optics*, 3<sup>rd</sup> ed. *Vision and vision optics*. 2010. V. 3. P. 41–417. New York: McGraw-Hill.
- Williams M.A., Moutray T.N., Jackson A.J. Uniformity of visual acuity measures in published studies. *Investigative Ophthalmology & Visual Science*. 2008. V. 49 (10). P. 4321–4327.  
<https://doi.org/10.1167/iovs.07-0511>
- Wolin L.R., Dillman A. Objective measurement of visual acuity: Using optokinetic nystagmus and electro-oculography. *Archives of Ophthalmology*. 1964. V. 71 (6). P. 822–826.  
<https://doi.org/10.1001/archopht.1964.00970010838008>
- Woodhouse J.M., Morjaria S.A., Adler P.M. Acuity measurements in adult subjects using a preferential looking test. *Ophthalmic and Physiological Optics*. 2007. V. 27 (1). P. 54–59.  
<https://doi.org/10.1111/j.1475-1313.2006.00454.x>
- Zheng X., Xu G., Zhang K., Liang R., Yan W., Tian P., Jia Y., Zhang S., Du C. Assessment of human visual acuity using visual evoked potential: A review. *Sensors (Switzerland)*. 2020. V. 20 (19). P. 1–26.  
<https://doi.org/10.3390/s20195542>

## ОЦЕНКА ОСТРОТЫ ЗРЕНИЯ.

### 1. ПЕРВИЧНЫЕ ИЗМЕРЕНИЯ И РАЗЛИЧНЫЕ МЕРЫ

Г. И. Рожкова<sup>a, #</sup>, М. А. Грачева<sup>a</sup>, Г. В. Парамей<sup>b</sup>

<sup>a</sup> Институт проблем передачи информации им. А.А. Харкевича РАН, 127051 Москва, Б. Калетный пер, 19/1, Россия

<sup>b</sup> Department of Psychology, Liverpool Hope University, Hope Park, L16 9JD Liverpool, UK

<sup>#</sup>E-mail: gir@iitp.ru

В статье приведены краткий обзор и анализ различных подходов к оценке остроты зрения (ОЗ), имеющий целью прояснить причины отсутствия среди специалистов однозначного мнения по поводу существующих способов представления ОЗ, их достоинств и недостатков. С течением времени количество методов и процедур, предложенных для оценки ОЗ, нарастало, они становились разнообразнее, увеличивалась их сложность. Такое разностороннее совершенствование шло параллельно с увеличением числа профессионалов, которые оценивали ОЗ в разных целях, что привело к определенному расхождению взглядов на интерпретацию результатов измерений и к введению разных систем их обозначения. Однако сейчас наличие консенсуса между специалистами становится критичным, поскольку многие проекты предусматривают совместную работу профессионалов из разных учреждений и требуют международного сотрудничества. В поисках общего основания для консенсуса и унификации терминологии логично начать с рассмотрения сути оценки ОЗ и причин значительного разнообразия современных точек зрения на решение этой задачи на практике. Обзор существующих подходов к оценке ОЗ показывает, что имеется только одна первичная мера ОЗ, которую можно получить непосредственным измерением и выразить в основных единицах пространственных метрик – минимальный угол разрешения (MAR,  $\alpha_m$  – minimum angle of resolution), определяемый минимальным угловым расстоянием, необходимым, чтобы две точки воспринимались раздельно. Имеется еще одна мера – критическая пространственная частота ( $F_c$ ), – которую можно оценить непосредственно, вместо MAR, если использовать набор тестовых решеток разной пространственной частоты и определить максимальную (критическую) воспринимаемую частоту, при превышении которой полосы решетки уже не различаются. Все остальные меры остроты зрения резонно считать вторичными, производными, поскольку они вычисляются как функции  $\alpha_m$ . Введение дроби Снеллена, десятичных единиц, logMAR, VE (visual efficiency), VAR (visual acuity rating) и других мер было связано с потребностью иметь для ОЗ альтернативные представления, более удобные и понятные, чем  $\alpha_m$ , для специалистов, имеющих дело с оценкой ОЗ на практике и использующих получаемые данные для разных целей – скрининга, мониторинга, диагностики, реабилитации, установления инвалидности, статистических оценок, разработки новых тестов и др. В связи с существенным различием требований и критериев у разных специалистов, по-видимому, невозможно найти для ОЗ представление, оптимальное со всех точек зрения, равно как и наилучший единый практический метод оценки ОЗ – “золотых” стандартов и методов не существует.

**Ключевые слова:** зрение человека, острота зрения, шкалирование, MAR, дробь Снеллена, десятичные единицы, logMAR, VAR

## КОМПЛЕКСНЫЙ АНАЛИЗ ФУНКЦИОНАЛЬНЫХ ПОКАЗАТЕЛЕЙ БИНОКУЛЯРНОЙ ЗРИТЕЛЬНОЙ СИСТЕМЫ У ШКОЛЬНИКОВ

© 2021 г. Н. Н. Васильева<sup>1,\*</sup>

<sup>1</sup> Институт проблем передачи информации им. А.А. Харкевича РАН  
127051 Москва, Большой Картеный переулок, д. 19, Россия

\*E-mail: nn\_vasiliyeva@mail.ru

Поступила в редакцию 29.03.2020 г.

После доработки 22.04.2021 г.

Принята к публикации 29.04.2021 г.

Представлен анализ результатов систематического изучения развития бинокулярного зрения (БЗ) у детей на протяжении школьного обучения. В исследовании приняли участие 750 школьников в возрасте от 6 до 18 лет. Оценивалась возрастная динамика функциональных показателей БЗ: остроты бинокулярного стереозрения, конвергентных и дивергентных фузионных резервов, скорости бинокулярной интеграции, баланса бинокулярных и монокулярных механизмов пространственного восприятия. Комплексный сравнительный анализ показателей БЗ в разном возрасте позволил уточнить сроки и темпы гетерохронного созревания отдельных механизмов, возраст достижения максимального уровня перед намечающимся ослаблением функций. Установлено, что наиболее быстрые темпы увеличения вклада бинокулярных механизмов в пространственное восприятие и улучшения различных показателей БЗ характерны для периода от 6–7 до 9–10 лет. По средним данным, фузионные резервы достигают максимальных значений в 12–13 лет, острота бинокулярного стереозрения и скорость бинокулярной интеграции – в 14–15 лет. Выявлена тенденция к снижению функциональных показателей БЗ в юношеском возрасте, что может быть связано с естественными возрастными процессами постепенной потери пластичности зрительных механизмов после их полного созревания, а также с влиянием различных неблагоприятных факторов. В ходе дискриминантного анализа данных определены показатели, имеющие наибольшее диагностическое значение для характеристики возрастных особенностей пространственного зрительного восприятия у детей: конвергентные фузионные резервы, острота стереоскопического зрения для пространственной частоты 0.7 цикл/град, бинокулярная острота зрения для близи, показатели успешности бинокулярной интеграции для времени экспозиции тестовых объектов 100 и 400 мс, соотношение данных стерео-кинетического теста в бинокулярных и монокулярных условиях наблюдения.

**Ключевые слова:** зрение детей, возрастная динамика, бинокулярная зрительная система, пространственное восприятие, острота стереозрения, фузионные резервы, бинокулярная интеграция

**DOI:** 10.31857/S0235009221030069

### ВВЕДЕНИЕ

В процессе обучения и осуществления многих видов деятельности для человека очень важны способности правильно оценивать размеры, форму объектов и их взаимное расположение в окружающем трехмерном пространстве, анализировать информацию о положении, ориентации и движении собственного тела, обеспечивая осуществление эффективной деятельности. Эти способности в значительной мере базируются на бинокулярных зрительных механизмах.

Бинокулярное зрение по сравнению с монокулярным обеспечивает более совершенное восприятие трехмерной формы объектов, более точную оценку их абсолютной и относительной удаленности, лучшее восприятие яркости и кон-

траста. При бинокулярном восприятии уменьшается утомляемость и растет вероятность распознавания объекта восприятия. Состояние бинокулярного зрения и зрелость бинокулярных механизмов являются определяющими для полноценного восприятия фото-, кино-, видеопродукции в 3D-формате. При этом бинокулярная зрительная подсистема представляет собой очень тонкий и подверженный внешним воздействиям механизм, функциональная незрелость и разлад которого в значительной мере отражаются на всем процессе зрительного восприятия.

Базовыми механизмами, обуславливающими преимущества бинокулярного пространственного восприятия по сравнению с монокулярным зрением, являются бинокулярный стереопсис, бинокулярная фузия и бинокулярная интеграция

(Кропман, 1966; Котлярский, 1982; Рожкова, 1992; Blake, Fox, 1973; Tyler, 1983; Howard, Rogers, 2012; Rogers, 2017 и др.).

На фоне большого числа публикаций о зрении детей объем работ, посвященных изучению возрастных аспектов бинокулярного зрения сравнительно невелик. Анализ работ, направленных на изучение онтогенеза зрительной системы и зрительного восприятия, показывает, что многие зрительные функции, в том числе бинокулярные, не заканчивают своего формирования к концу сенситивного периода и могут совершенствоваться в период школьного обучения (Базарный, 1991; Васильева, 2002; Фарбер, Бетелева, 2005; Рожкова, Матвеев, 2007; Морозова, 2016; Бондарко, 2018; Жукова и др., 2018; Николаева, Джалаева, 2020; Oduntan et al., 1998; Chen, Abidin, 2002; Irving et al., 2006; Dyea, Bavelierb, 2010; Namaeh et al., 2020 и др.). Усложнение деятельности зрительной системы, повышение скорости и точности зрительного анализа, совершенствование разнообразных зрительно-моторных актов в постнатальном онтогенезе связаны с морфофункциональным созреванием разных модулей зрительной сенсорной, аккомодационной и глязодвигательной систем.

Одной из основных закономерностей морфофункционального созревания растущего организма является гетерохронность развития. Гетерохронность проявляется в разной скорости созревания различных систем и функций и достижении ими полной зрелости на разных этапах индивидуального развития. Гетерохронность развития – специфический для онтогенеза механизм адаптации живой системы, благодаря которому в каждом возрасте наиболее эффективно используются те структуры и их функциональные возможности, которые к данному времени созрели настолько, чтобы отвечать запросам следующего этапа развития организма. Таким образом, каждый возрастной этап имеет характерную для него неповторимую морфофункциональную структуру физиологических систем, определяющую функциональные возможности данного возраста. Причем достижение адаптивного результата не исключает в последующем перестройки функциональных систем в связи с изменением круга задач и созреванием других более эффективных или более специализированных механизмов.

Результаты фундаментальных исследований указывают на необходимость проведения комплексного исследования бинокулярной зрительной системы детей и подростков, учитывая наличие в зрительной системе большого числа параллельных каналов переработки информации, различающихся между собой по принципам и механизмам анализа поступающей информации (Леушина, 1967; Супин, 1974; 1981; Шевелев,

1984; Плосконос, 1989; Глезер и др., 1992; Пигарев, 1992; Рожкова, 1992; Рожкова, Матвеев, 2007; Алексеенко, 2018; Hubel, Wiesel, 1962; Wolfe, 1986; Cogan, 1987; Hubel, Livingstone, 1987).

Использование комплексного подхода имеет принципиально важное значение, поскольку позволяет уточнить данные о сроках и темпах созревания различных механизмов, периодах интенсивного формирования функций, условиях достижения максимального уровня и причинах намечающегося ослабления функций. Эти ключевые вопросы имеют не только теоретический интерес, позволяя раскрыть закономерности развития зрительной системы в периоды высокой адаптационной нагрузки растущего организма, но и практическое значение при создании условий для своевременной реализации времяспецифических функций, обосновании технологий и методик обучения, разработки подходов к охране и укреплению здоровья детей и подростков.

В настоящей статье кратко суммированы результаты долгосрочного систематического комплексного изучения функциональных показателей бинокулярной зрительной системы у школьников 6–18-летнего возраста.

## МАТЕРИАЛЫ И МЕТОДЫ

**Испытуемые.** Исследования были проведены в период с 2003 по 2020 г. С учетом имеющихся возможностей в разных сериях экспериментов участвовало различное количество испытуемых, но оно всегда было достаточным для получения достоверных результатов. В целом в исследовании приняли участие 750 человек в возрасте от 6 до 18 лет, обучающиеся в образовательных учреждениях г. Москвы и г. Чебоксары.

Испытуемые участвовали в исследовании добровольно, от совершеннолетних было получено информированное согласие на проведение процедуры измерений. Родители (законные представители) несовершеннолетних были проинформированы о проводившемся исследовании и дали письменное согласие на участие в нем детей. Отбор испытуемых проводили на основе анализа анамнестических данных и медицинских карт. По результатам стандартного медицинского осмотра были отобраны дети, у которых аномалий бинокулярного зрения не выявлено. Условно можно считать, что эти испытуемые имели нормальное бинокулярное зрение, хотя не исключено, что у незначительной части имелись скрытые или устранившие ранее дефекты, которые не выявляются при стандартном медицинском исследовании зрения. У всех участников проводили оценку монокулярной и бинокулярной остроты дальнего (4 м) и ближнего (0.5 м) зрения и измеряли рефракцию.

Организация проведения исследования соответствовала принципам биомедицинской этики, представленным в Хельсинкской декларации 1964 г. и ее последующих обновлениях.

**Основные оцениваемые бинокулярные показатели.** На основе анализа и обобщения представлений о функциональной организации пространственного восприятия человека, современных нейрофизиологических данных о бинокулярных и монокулярных механизмах, а также многоканальности переработки информации в зрительной системе, нами были выделены четыре функциональных показателя, которые, на наш взгляд, позволяют достаточно разносторонне охарактеризовать состояние бинокулярного зрения:

- *острота бинокулярного стереозрения*, отражающая качество стереописа, т.е. способность определять малые сдвиги объектов по глубине на основе бинокулярной диспаратности;
- *фузионные резервы*, характеризующие совместную работу окуломоторного аппарата и сенсорных механизмов фузии, отражающие прочность, помехоустойчивость и гибкость механизмов фузии, обеспечивающих формирование единого трехмерного образа в результате совместной обработки изображений объекта на сетчатках двух глаз;
- *скорость бинокулярной интеграции*, отражающая временные характеристики бинокулярного восприятия, а именно – способность зрительной системы за короткое время сформировать целостный образ объекта на основе двух неполных изображений в правом и левом глазу;
- *バランス бинокулярных и монокулярных механизмов пространственного восприятия*, характеризующий особенности взаимодействия бинокулярных и монокулярных механизмов и вклад бинокулярных механизмов в пространственное восприятие.

**Обоснование выбора методик.** Острота бинокулярного стереозрения и фузионные резервы оцениваются в настоящее время в офтальмологической и профессионально-экспертной практике при необходимости уточнения и постановки диагноза или выявлении нарушений бинокулярного зрения. В то же время для решения поставленных в исследовании задач ограничиться уже известными и апробированными методами, широко применяемыми в офтальмологической практике, не представляло возможным в силу ряда обстоятельств.

Во-первых, для оценки остроты бинокулярного стереозрения доступные тесты Titmus test и Randot stereo test ограничены точностью 20" и 40" соответственно. Анализ опубликованных в настоящее время данных о количественных характеристиках стереописа у взрослых позволяет заключить, что стереопороги у школьников могут быть значительно ниже 20".

Во-вторых, синоптофор, применяемый для оценки фузионных резервов, не всегда позволяет точно и объективно оценить эти показатели, так как достоверность данных в значительной мере зависит от субъективных ответов обследуемого ребенка, а значит, от правильного понимания им своих зрительных ощущений и тех требований, которые к нему предъявляются в процессе обследования.

В-третьих, цветотест, применяемый для оценки способности зрительной системы формировать единый образ объекта, комбинируя фрагменты его проекций в правом и левом глазу, не позволяет количественно охарактеризовать скорость бинокулярной интеграции.

В-четвертых, чтобы обеспечить возможность проведения исследования по единой комплексной методике в широком возрастном интервале, важно было подобрать такой комплекс измерительных процедур, который был бы понятен и доступен испытуемым младших возрастных групп и в то же время не снизил интерес к исследованию в старших возрастных группах.

Отбор методов осуществляли на основе следующих принципов:

- максимальное исключение влияния субъективных факторов на результат;
- возможность количественного выражения результата;
- высокий уровень успешности тестирования: возможность довести процедуру до конца в подавляющем большинстве случаев;
- возможность реализации метода в игровой форме (важно для детей младших возрастных групп);
- возможность автоматизации и компьютеризации процедуры.

В итоге комплексное исследование развития механизмов бинокулярной зрительной системы школьников в данной работе включало оценку остроты стереоскопического зрения, конвергентных и дивергентных фузионных резервов, скорости бинокулярной интеграции, баланса бинокулярных и монокулярных механизмов пространственного восприятия.

#### *Описание тестовых процедур*

**Оценка остроты бинокулярного стереозрения.** Для количественной оценки остроты стереозрения измеряли пороги стереовосприятия методом пространственно-частотной стереовизометрии при помощи компьютерной программы "Стереопись", диагностические возможности которой были продемонстрированы при исследовании офтальмологически здоровых взрослых и лиц с различными заболеваниями зрительной системы (Белозе-

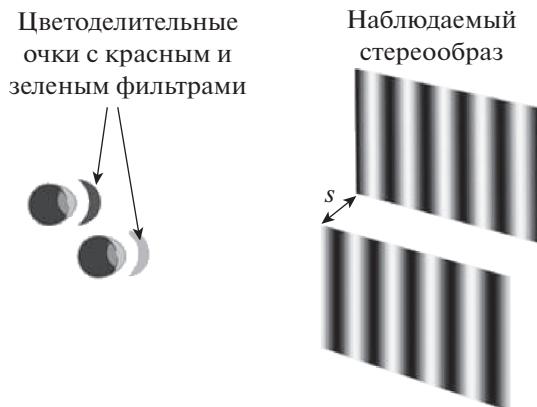


Рис. 1. Схема измерения стереопорогов.

ров, 2003). Достоинством программы является возможность оценки малых стереопорогов, величиной до 1". Процедура ее применения подробно описана в публикации (Васильева и др., 2010а).

Тестовым стимулом являлась стереопара, занимающая всю поверхность экрана монитора и изображающая две расположенные одна над другой вертикальные решетки с синусоидальным профилем яркости и высоким контрастом (рис. 1). В каждом стимуле решетки имели одинаковую пространственную частоту, но различную диспаритетность (которую варьировали), поэтому при разделении полей зрения они казались отстоящими друг от друга по глубине, если величина диспаритетности превосходила порог.

Измерения проводили на пространственных частотах: 0.4; 0.5; 0.7; 1.0; 1.4; 2.0; 2.8; 4.0 цикл/град в условиях дихоптического предъявления стимулов на расстоянии 2.27 м от экрана (расстояние от глаз наблюдателя до экрана автоматически задавалось программой, учитывающей размеры монитора, чтобы обеспечить постоянство углового размера экрана по горизонтали – 8.3°). Разделение полей зрения осуществляли методом анагlyphической гаплоскопии, тестирование проходило в красно-зеленых очках.

Задача испытуемого состояла в определении взаимного расположения воспринимаемых решеток, занимающих верхнюю и нижнюю часть экрана по отношению друг к другу, т.е. испытуемый должен был сказать, какая решетка находится ближе к нему, верхняя или нижняя.

За порог стереозрения принимали минимальную относительную диспаритетность верхней и нижней решеток, при которой испытуемый надежно определял их взаимное расположение.

**Оценка фузионных резервов.** Фузионные резервы измеряли при помощи интерактивной компьютерной программы “Фурор”, разработанной в ИППИ РАН. Данная программа использова-

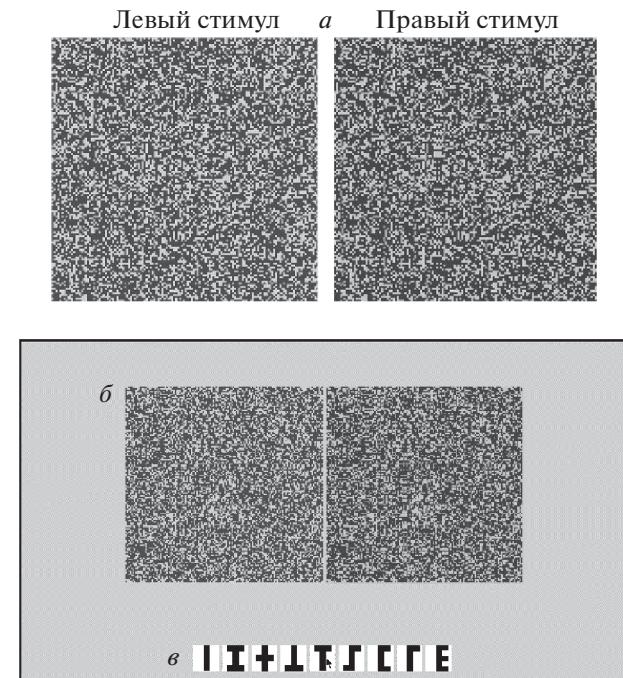


Рис. 2. Пояснения к методике: *а* – стереограмма, кодирующая основной ТО – вертикальную полоску; *б* – вид экрана после измерения с остановленной стереограммой, кодирующей ТО в виде буквы Т; *в* – набор тест-объектов.

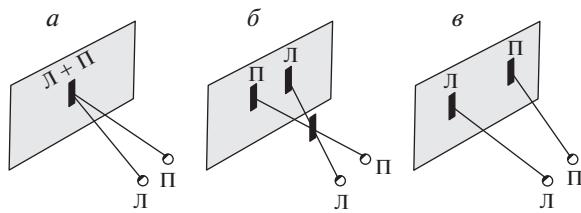
лась нами во многих исследованиях и подробно описана в публикациях (Васильева, Рожкова, 2009; Рожкова, Васильева, 2010).

Тестовым стимулом являлась случайно-точечная стереограмма (СТС). СТС на поле 140 × 140 мм кодировала тест-объект (ТО) в виде вертикальной полоски размером 114 × 26 мм, к которой по ходу процедуры измерения в случайном порядке добавлялись маркеры, меняющие вид ТО (рис. 2).

При измерении конвергентных фузионных резервов правый стимул двигался влево, а левый – вправо, что приводило к постепенному увеличению угла конвергенции. При измерении дивергентных фузионных резервов правый стимул двигался вправо, а левый – влево, что приводило к дивергенции зрительных осей (рис. 3).

Скорость движения стимулов по экрану составляла 50 мм/мин. Применение СТС позволяло контролировать сохранение у испытуемых способности к формированию полноценных стереообразов на основе бинокулярной диспаритетности на всем протяжении измерения до момента распада сфузированного образа при достижении критических углов конвергенции/дивергенции.

Измерения проводили на расстоянии 50 см от экрана в условиях дихоптического предъявления стимулов. Разделение полей зрения осуществляли методом анагlyphической гаплоскопии, тести-



**Рис. 3.** Схемы расположения левого (Л) и правого (П) стимулов в исходный момент (а) и при измерении конвергентных (б) и дивергентных (в) фузионных резервов.

рование проходило в красно-зеленых очках. В отличие от призм и синоптофора, используемых в клинической практике для оценки фузионных резервов, данная программа позволяла установить момент пропадания бинокулярного образа очень точно: как только испытуемый прекращал фузировать СТС, он переставал видеть закодированный ТО.

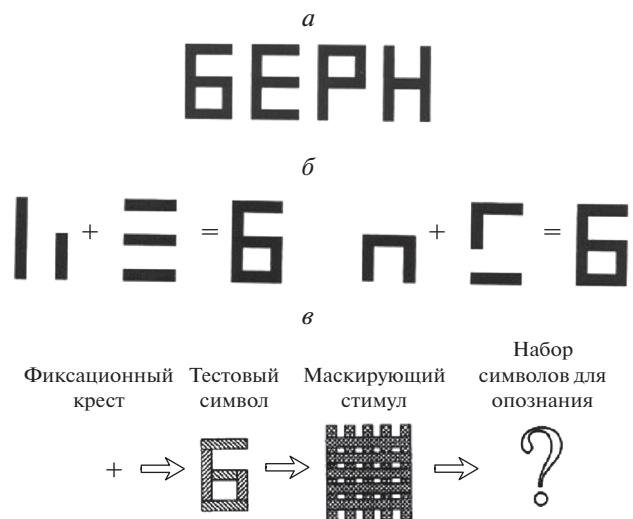
Перед испытуемым ставилась задача как можно дольше удерживать состояние фузии, следя за сменой тест-объектов и констатируя момент потери фузии и исчезновения бинокулярного образа.

Оценкой фузионных резервов служили те предельные углы конвергенции и дивергенции, превышение которых вызывало срыв фузии, исчезновение циклопического объекта и диплопию.

**Оценка скорости бинокулярной интеграции.** Для оценки скорости бинокулярной интеграции была использована интерактивная компьютерная программа, разработанная в ИППИ РАН. Процедура ее применения подробно описана в публикации (Рожкова и др., 2002).

В качестве тестовых объектов использовали схематические изображения букв, составленных из фрагментов, которые случайным образом разделялись на две части и предъявлялись правому и левому глазу на определенное время (рис. 4, а, б). Линейные размеры прямоугольной области, в которую вписывались тестовые символы, на экране составляли  $24 \times 37.5$  мм; толщина линий равнялась 5 мм. Время экспозиции варьировало в диапазоне от 50 до 800 мс. Измерения проводили на расстоянии 50 см от экрана в условиях дихотического предъявления стимулов с использованием цветоделительных очков. Для одного глаза фрагменты буквы были красными, для второго – синими.

В каждом испытании для опознания последовательно предъявляли 10 тестовых символов (букв). После каждого предъявления тестовой буквы на экран выводилась маскирующая решетка, перекрывающая все тестовое поле и содержащая горизонтальные и вертикальные элементы,



**Рис. 4.** Примеры тестовых объектов и схема процедуры эксперимента: а – образцы букв; б – два варианта комбинации буквы Б из левого и правого фрагментов; в – последовательность предъявления изображений в ходе тестирования.

из которых были составлены символы. Затем на экране возникал образец со всеми тестовыми знаками, из которых надо было выбрать только что увиденную букву (рис. 4, в).

В каждой серии фиксировали вероятность правильного ответа (отношение количества правильно названных букв к общему числу предъявленных тестовых символов). Об интегративных способностях судили по минимальному времени предъявления фрагментов букв, обеспечивающему 75% правильных ответов.

Для обеспечения единообразия в тестировании испытуемых из различных образовательных учреждений измерения проводили при помощи одной и той же переносной аппаратуры. Интерактивные компьютерные программы были установлены на ноутбуке ASUS, размер экрана составлял  $210 \times 330$  мм.

**Оценка баланса бинокулярных и монокулярных механизмов.** Для оценки баланса бинокулярных и монокулярных механизмов пространственного восприятия был использован стереокинетический тест. Данный тест применялся Л.Н. Могилевым при исследовании взрослых испытуемых и в наших предыдущих работах (Могилев, 1982; Рожкова, Васильева, 2001; Васильева, Рожкова, 2010б; Васильева, 2011). Мы усовершенствовали методику применения данного теста, сделав возможным количественную оценку баланса бинокулярных и монокулярных механизмов, на основе введенного нами коэффициента  $B$ , и адаптировали для детей, дополнив рядом эталонов.

Тестовым стимулом служил плоский диск диаметром 12 см с эксцентрическим кольцевым черно-белым паттерном. Диск вращался со скоростью 60 об/мин. Расстояние от глаз испытуемого до тест-объекта было равно 3 м. Когда стимул был неподвижен, он воспринимался как плоский или почти плоский объект. При вращении стимула возникала сильная иллюзия объемности: диск воспринимался как конический объект, вершина которого была направлена либо к испытуемому, либо от него. В данном тесте разные механизмы противопоставлены друг другу: монокулярные механизмы порождают сильную стереокинетическую иллюзию объемности плоского объекта, а бинокулярные механизмы должны противодействовать этой иллюзии.

Задача испытуемого заключалась в оценке иллюзорной глубины (высоты виртуального конуса) в монокулярных и бинокулярных условиях наблюдения.

Для характеристики баланса бинокулярных и монокулярных влияний на пространственное восприятие вычисляли введенный нами коэффициент вклада бинокулярных механизмов  $B$  по формуле:

$$B = (h_m - h_b)/h_m,$$

где  $h_m$  – средняя для двух глаз монокулярная оценка глубины,  $h_b$  – бинокулярная оценка.

В работе (Рожкова, Васильева, 2001) представлены расчеты результирующего коэффициента вклада бинокулярных механизмов с учетом представлений о сложной структуре бинокулярной зрительной системы и возможные его варианты в зависимости от функционирования у человека монокулярных и бинокулярных подсистем зрения. Положительные значения коэффициента  $B$  соответствуют случаям уменьшения иллюзорной глубины при переходе от монокулярных условий наблюдения к бинокулярным и должны регистрироваться у испытуемых с нормальным бинокулярным зрением. В идеале коэффициент, отражающий вклад бинокулярных механизмов, должен быть равен +0.3. При “полной победе” бинокулярных механизмов  $h_b = 0$  и  $B = 1$ . Нулевые значения  $B$  свидетельствуют об отсутствии влияния бинокулярных механизмов вследствие недостаточного их развития либо нарушения функционирования, а отрицательные значения  $B$  – о принципиальной перестройке взаимоотношений между монокулярными механизмами на фоне бездействия бинокулярных. Таким образом, возрастная и индивидуальная вариабельность этого коэффициента у испытуемых, по-видимому, определяется вариациями весовых коэффициентов, отражающими степень развития и баланс разных подсистем в процессе индивидуального жизненного опыта.

*Статистический анализ полученного цифрового материала проводили методами описательной и вариационной статистики с использованием программного пакета статистического анализа “StatSoft Statistica 6.0”. Весь цифровой материал до оценки значимости различий проверен на тип распределения при помощи критериев Колмогорова–Смирнова и Лиллифорса с построением гистограмм. Достоверность отличий определяли по t-критерию Стьюдента. В случаях, когда распределение данных не соответствовало нормальному, мы использовали непараметрический U-критерий Манна–Уитни. Для определения влияния разных факторов на изучаемые показатели проводили дисперсионный анализ. С целью уточнения классификации групп и выявления наиболее значимых функциональных показателей для характеристики формирования механизмов бинокулярного зрения был применен дискриминантный анализ.*

## РЕЗУЛЬТАТЫ

### *Возрастные показатели механизмов бинокулярного стереописца*

Величина порогов стереозрения во всех возрастных группах зависела от пространственной частоты стимула: при увеличении пространственной частоты от 0.4 до 0.7 цикл/град стереопороги понижались ( $p < 0.05$ ); в диапазоне частот от 1.0 до 2.0 цикл/град стереопороги были минимальными, а с повышением пространственной частоты от 2.8 до 4.0 цикл/град было зафиксировано их повышение ( $p < 0.05$ ). Во всех возрастных группах наименьшие пороги были зарегистрированы на частотах 1.0–2.0 цикл/град. При этом число испытуемых, имеющих порог ниже 10", неуклонно повышалось до 15 лет ( $p < 0.05$ ): 12% в возрастной группе 7–8 лет; 28% – в 9–10 лет; 51% – в 11–12 лет; 89% – в 14–15 лет; 82% – в 16–17 лет.

Особенности возрастной динамики стереопорогов, измеренных на пространственных частотах 1.0–2.0 цикл/град, показаны на рис. 5. Ход кривых свидетельствует о монотонном снижении порогов с возрастом.

Характер возрастных изменений стереопорогов заключался в статистически достоверном их снижении на возрастном интервале от 7 до 14–15 лет и увеличении остроты стереоскопического зрения на разных частотах в 3–4 раза ( $p < 0.05$ ). Это позволяет говорить о совершенствовании механизмов стереозрения на данном возрастном отрезке. В возрастной группе 16–18 лет выявлено небольшое, но достоверное ухудшение показателей стереоскопического зрения на частотах 1.4–4.0 цикл/град по сравнению с возрастной группой 14–15 лет.

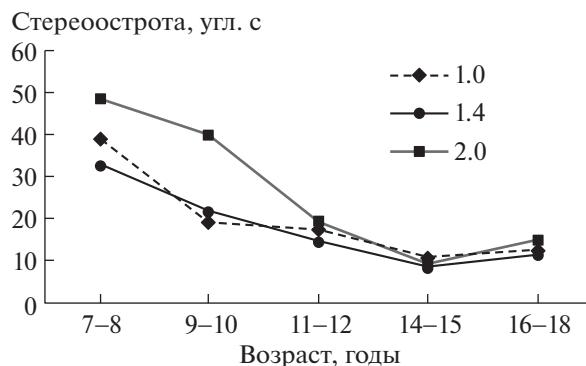


Рис. 5. Возрастная динамика стереоостроты.

#### Особенности возрастной динамики механизмов аккомодационно-окуломоторного обеспечения фузии

Проведенная оценка фузионных резервов показала, что наиболее быстрый темп увеличения конвергентных и дивергентных фузионных резервов на исследованном возрастном интервале характерен для детей младшего школьного возраста (рис. 6). В возрастной период от 7 до 11 лет конвергентные резервы нелинейно нарастают от  $14.0 \pm 3.1^\circ$  до  $18.2 \pm 3.8^\circ$  ( $p < 0.001$ ), а дивергентные от  $-6.2 \pm 1.5^\circ$  до  $-8.1 \pm 1.9^\circ$  ( $p < 0.01$ ). После некоторого пологого максимума, характерного для возрастного периода от 11 до 13 лет ( $18.2 \pm 3.6^\circ$  в 12 лет;  $18.6 \pm 3.9^\circ$  в 13 лет,  $p > 0.05$ ), отмечается постепенное снижение конвергентных резервов после 14 лет (от  $17.6 \pm 4.1^\circ$  в 14 лет до  $15.1 \pm 4.2^\circ$  в 18 лет,  $p < 0.05$ ). По средним данным дивергентные резервы, достигнув максимальных значений в 13 лет ( $-8.6 \pm 1.8^\circ$ ), с возрастом изменяются незначительно ( $p > 0.05$ ).

Результаты оценки фузионных резервов у учащихся 1, 4 и 10 классов при помощи синоптофора с использованием слайдов с СТС, представленные в работе (Вашенко, 2001), свидетельствуют о том, что наивысшие значения конвергентных фузионных резервов отмечены у обучающихся четвертого класса.

Опираясь на полученные в нашем исследовании данные и ориентировочные средние значения, полученные разными авторами на взрослых при измерениях на синоптофоре (Сомов, 1989; Рожков, Овсянникова, 2003), мы ввели следующие критерии для высоких и низких показателей фузионных резервов. Положительные резервы, не превышающие  $10^\circ$ , мы считали низкими, превосходящие  $20^\circ$  – высокими, находящиеся в диапазоне от  $10$  до  $20^\circ$  – средними; отрицательные резервы до  $-3^\circ$  – низкими, свыше  $-10^\circ$  – высокими, находящиеся в диапазоне от  $-3^\circ$  до  $-10^\circ$  – средними.

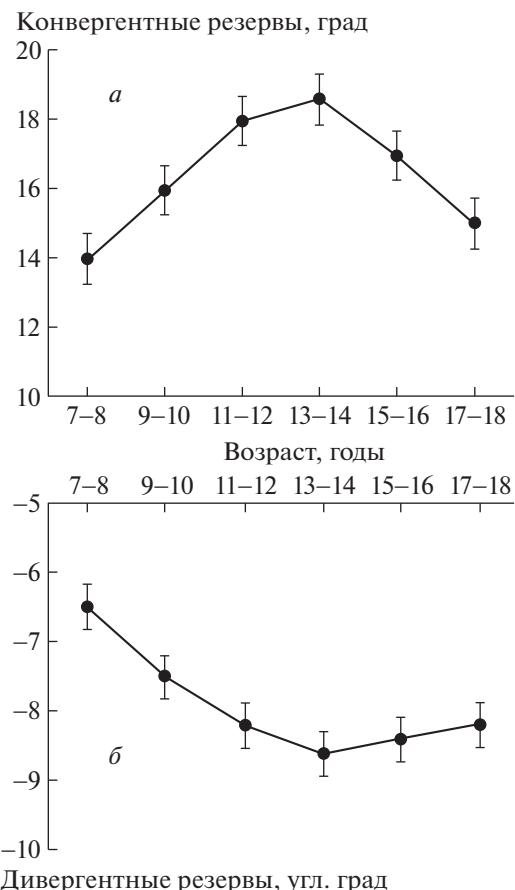
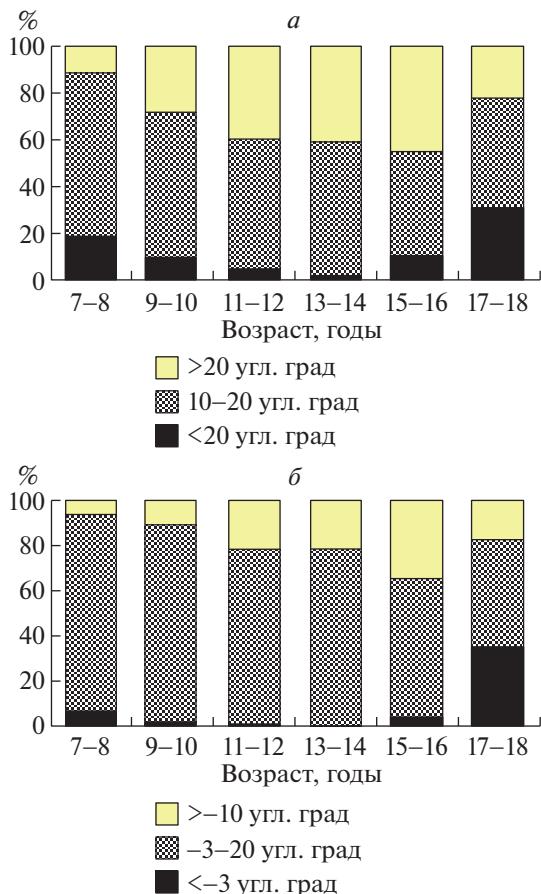


Рис. 6. Возрастная динамика конвергентных (а) и дивергентных (б) фузионных резервов.

Классифицируя полученные данные по выделенным критериям, можно видеть, что с увеличением возраста постепенно уменьшается доля детей с низкими показателями конвергентных и дивергентных фузионных резервов (рис. 7).

В возрастном интервале от 6 до 16 лет число детей с высокими конвергентными фузионными резервами (более  $20^\circ$ ) по мере взросления неуклонно нарастало (0% – в 6 лет; 28% – в 9–10 лет; 40% – в 13–14 лет; 45% – в 15–16 лет,  $p < 0.05$ ), при этом количество детей с низкими конвергентными резервами (менее  $10^\circ$ ) уменьшалось до 13–14 лет (34% – в 6 лет; 9% – в 9–10 лет; 2% – в 13–14 лет,  $p < 0.05$ ). В возрастной группе 17–18 лет наблюдалось уменьшение числа обучающихся с высокими конвергентными резервами (до 22%,  $p < 0.05$ ) и увеличение с низкими (до 30%,  $p < 0.05$ ). Иная закономерность выявлена в динамике высоких и низких значений дивергентных фузионных резервов. Установлено, что число детей с низкими дивергентными резервами в возрастном интервале от 6 до 16 лет существенно не менялось (1–3%), в то же время с высокими с возрастом неуклонно увеличивалось до 15–16 лет

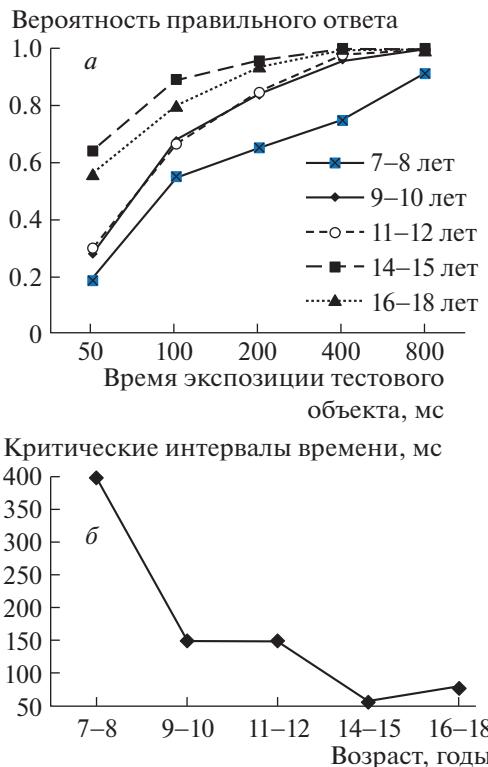


(2% – в 6 лет; 11% – в 9–10 лет; 21% – в 13–14 лет; 35% – в 15–16 лет,  $p < 0.05$ ).

#### Возрастные изменения скорости бинокулярной интеграции

В период от 7 до 15 лет происходит неуклонное повышение скорости бинокулярной интеграции при переходе к более старшему возрасту (рис. 8, *а*). Кривые зависимости средних значений вероятности правильного ответа от времени экспозиции для 9–10 и 11–12 лет практически совпадают. После 16 лет обнаружена тенденция к снижению способности к успешному комбинированию информации, содержащейся в левом и правом сетчаточных изображениях ( $0.05 < p < 0.1$ ).

Опираясь на средние данные, можно констатировать, что ускорение интегративных процессов происходит дважды: в младшем школьном до 9–10 лет и подростковом возрасте с 12 до 14–15 лет. При анализе минимального времени предъявления фрагментов тест-объектов, обеспечивающего 75% правильных ответов, установлено



**Рис. 8.** Возрастная динамика скорости бинокулярной интеграции: *а* – кривые зависимости средних значений вероятности правильного ответа от времени экспозиции; *б* – критические интервалы времени для осуществления бинокулярной интеграции у детей разных возрастных групп.

ны возрастные критические интервалы времени для осуществления бинокулярной интеграции (рис. 8, *б*): в 7–8 лет – 400 мс; в 9–10 лет и 11–12 лет – 150 мс; в 14–15 лет – 60 мс; в 16–18 лет – 80 мс.

#### Возрастная динамика баланса бинокулярных и монокулярных механизмов пространственного восприятия

В процессе анализа данных установлено, что во всех возрастных группах индивидуальный разброс показателя *B* был достаточно большим: от –1.3 до +1.0. В каждой группе в бинокулярных условиях наблюдения встречались случаи полного отсутствия иллюзии, т.е. врачащийся диск мог восприниматься как плоский объект. Число таких случаев было наибольшим в возрастной группе 9–10 лет (18% от общей численности группы). В других возрастных группах на долю испытуемых с полным отсутствием стереокинетической иллюзии при бинокулярном наблюдении приходилось от 1 до 8%.

Анализируя возрастную динамику средних показателей коэффициента *B* (рис. 9), можно за-

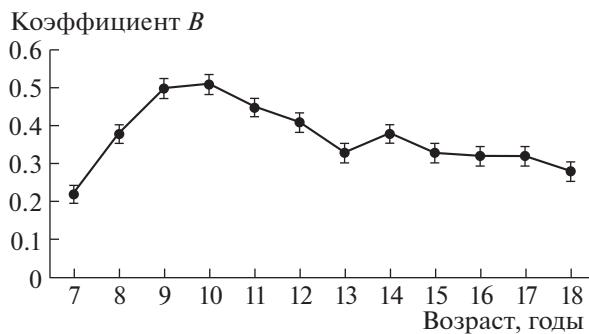


Рис. 9. Возрастная динамика баланса бинокулярных и монокулярных механизмов.

ключить, что в возрастном интервале от 7 до 9 лет имеет место быстрое развитие бинокулярных механизмов пространственного зрительного восприятия. Средние значения коэффициента  $B$  в этот возрастной период увеличиваются от  $0.22 \pm 0.27$  до  $0.50 \pm 0.28$  ( $p < 0.05$ ).

В возрасте от 10 до 12 лет коэффициент  $B$  в среднем остается на высоком уровне ( $0.51 \pm 0.25$  в 10 лет;  $0.48 \pm 0.20$  – в 11 лет;  $0.43 \pm 0.21$  – в 12 лет;  $p > 0.05$ ). В возрастном интервале от 13 до 17 лет значения  $B$  приближаются к 0.3.

Соотношение низких и высоких значений коэффициента  $B$  в разные возрастные периоды меняется: к подростковому возрасту заметно уменьшается количество испытуемых с отрицательными и нулевыми значениями и увеличивается число детей с  $B \geq 0.3$  (рис. 10).

#### Общий анализ массива полученных данных

Анализируя весь массив полученных данных, можно отметить, что в исследованных возрастных группах индивидуальный разброс показателей весьма велик. Наглядное представление о вариабельности индивидуальных показателей демонстрируют рис. 11 и 12, на которых приведен разброс индивидуальных значений стереопорогов, измеренных на разных пространственных частотах, и фузионных резервов соответственно.

Следует отметить, что неоднородность групп школьников наиболее выражена на начальном этапе обучения и в старших классах школы. Наименьший разброс индивидуальных показателей отмечался в возрастном интервале от 10 до 14 лет.

Анализ индивидуальных показателей сформированности разных механизмов у конкретного ребенка показал, что они могут различаться и быть независимыми от показателей остроты зрения. В качестве примера на рис. 13 приведены индивидуальные профили сформированности механизмов бинокулярного зрения в соответствии с

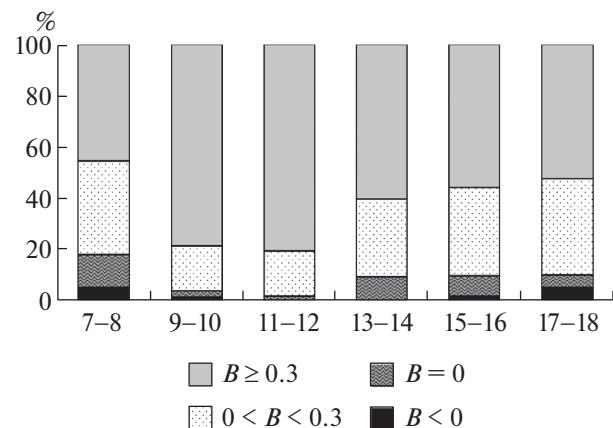


Рис. 10. Процентные соотношения детей с высокими и низкими значениями  $B$  в разных возрастных группах.

выделенными показателями у двух школьников 9 и 13 лет.

У обоих испытуемых монокулярные показатели остроты зрения левого и правого глаза находились в пределах нормативных значений для своей возрастной группы. Допуская наличие разных индивидуальных вариантов профилей сформированности бинокулярных зрительных функций,

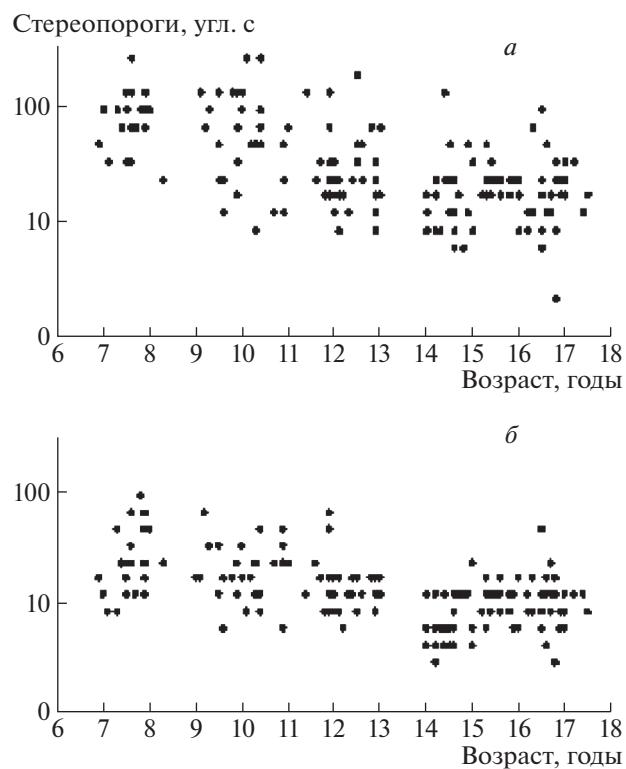
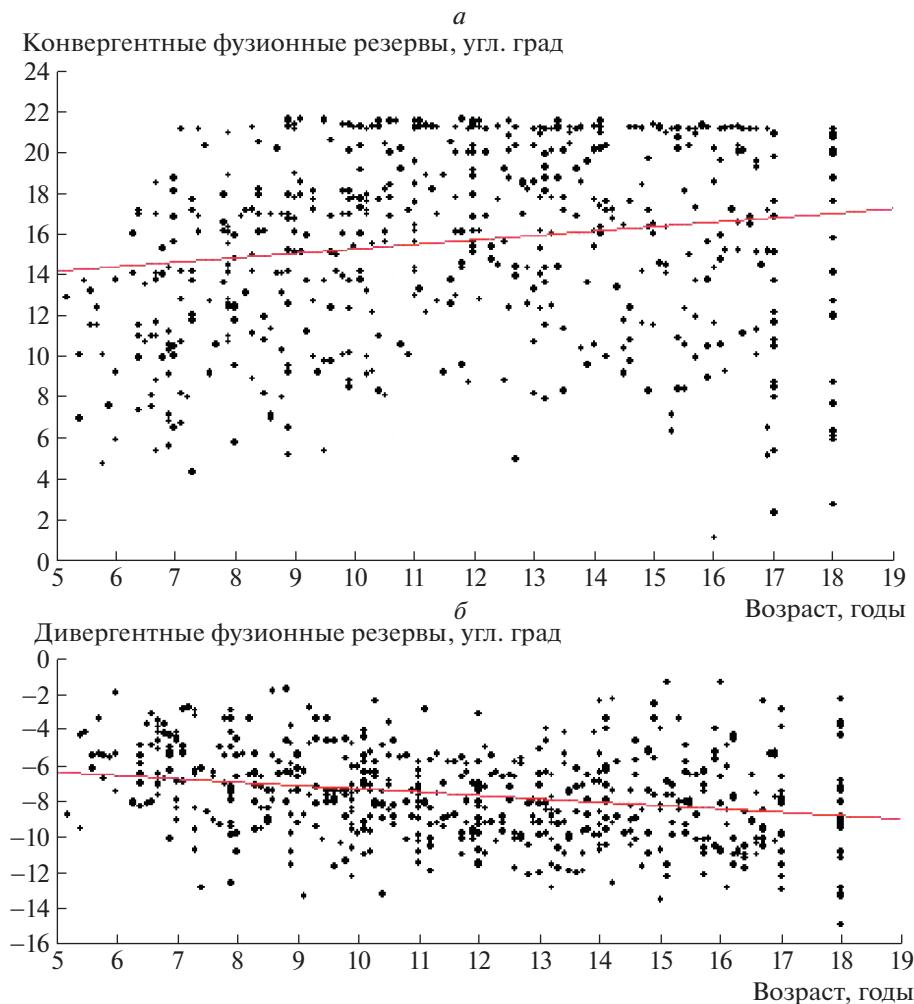


Рис. 11. Разброс индивидуальных значений стереопорогов для пространственных частот 0.5 (а) и 1.4 (б) цикл/град.



**Рис. 12.** Разброс индивидуальных значений конвергентных (*а*) и дивергентных (*б*) фузионных резервов.

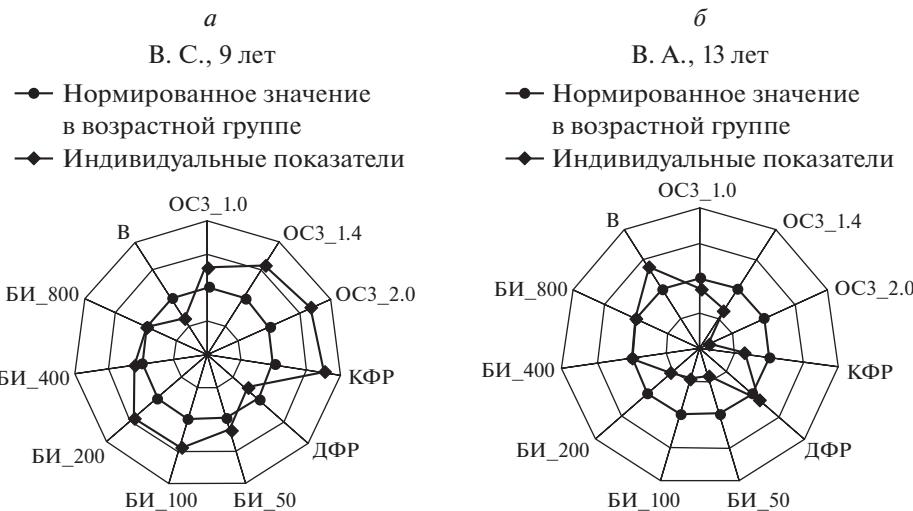
отметим, что приведенные на рис. 13 графики в некоторой степени отражают общую тенденцию: у многих детей в период роста и развития организма структурные преобразования в зрительной системе выражаются в том, что различные индивидуальные бинокулярные показатели могут быть как выше, так и ниже среднестатистических значений по своей возрастной группе.

**Таблица 1.** Информативность признаков, вошедших в дискриминантную модель

Показатели	Критерий F	Значение <i>p</i>	Толерантность
КФР	7.07	0.001	0.89
БО	8.51	0.0002	0.81
ОСЗ <sub>0.7</sub>	11.22	0.00002	0.73
БИ <sub>400</sub>	5.22	0.006	0.65
БОЗ <sub>6</sub>	3.92	0.021	0.90
БИ <sub>100</sub>	3.98	0.020	0.45

Для определения паттерна наиболее информативных функциональных показателей применительно к оценке развития бинокулярной зрительной системы у детей и построения дискриминантной модели был применен дискриминантный анализ. Группирующим признаком была выбрана принадлежность испытуемых к одной из исследуемых групп: 1 – 7–10 лет; 2 – 11–14 лет; 3 – 15–18 лет. Установлено, что самыми информативными показателями, определяющими принадлежность к одной из исследуемых групп, оказались шесть переменных, имеющих наибольшие отношения межгрупповой и внутригрупповой дисперсий (табл. 1):

КФР – конвергентные фузионные резервы;  
БО – бинокулярная оценка высоты виртуального конуса;  
ОСЗ<sub>0.7</sub> – острота стереоскопического зрения для пространственной частоты 0.7 цикл/град;  
БОЗ<sub>6</sub> – бинокулярная острота зрения для близи;



**Рис. 13.** Круговые диаграммы, демонстрирующие соотношение групповых и индивидуальных функциональных показателей бинокулярной зрительной системы (в относительных единицах): *а* – исп. В.С., 9 лет; *б* – исп. В.А., 13 лет  
ОСЗ – острота стереоскопического зрения для пространственных частот 1.0, 1.4, 2.0 цикл/град; КФР – конвергентные фузионные резервы; ДФР – дивергентные фузионные резервы; БИ – показатели скорости бинокулярной интеграции для времени экспозиции 50, 100, 200, 400 и 800 мс; В – баланс бинокулярных и монокулярных механизмов.

БИ<sub>100</sub>, БИ<sub>400</sub> – показатели скорости бинокулярной интеграции для времени экспозиции 100 и 400 мс.

В ходе дискриминантного анализа были получены следующие дискриминантные уравнения (канонические корни  $R_1 = 0.952372$ ,  $x^2 = 555.6767$ ,  $p < 0.000000$ ;  $R_2 = 0.423159$ ,  $x^2 = 42.6188$ ,  $p < 0.002295$ ):

$$F_1 \text{ группа (7–10 лет)} = 0.269\text{КФР} + 0.468\text{БО} + 0.094\text{ОСЗ}_{0.7} + 182.015\text{БИ}_{400} + 7.044 \text{БОЗ}_6 - 29.093\text{БИ}_{100}$$

$$F_2 \text{ группа (11–14 лет)} = 0.196\text{КФР} + 0.568\text{БО} + 0.141\text{ОСЗ}_{0.7} + 170.501\text{БИ}_{400} + 5.459 \text{БОЗ}_6 - 34.628\text{БИ}_{100}$$

$$F_3 \text{ группа (15–18 лет)} = -0.165\text{КФР} + 0.933\text{БО} + 0.161\text{ОСЗ}_{0.7} + 147.603\text{БИ}_{400} + 1.804 \text{БОЗ}_6 - 34.191\text{БИ}_{100}$$

Как следует из полученных уравнений, с изменением возраста закономерно наиболее связаны показатели, вошедшие в дискриминантную функцию.

## ОБСУЖДЕНИЕ

Использование в работе комплекса методов позволило получить значительный массив данных, иллюстрирующих закономерности формирования механизмов пространственного зрительного восприятия у детей в возрастном интервале от 6 до 18 лет, сроки и темпы гетерохронного созревания отдельных механизмов, возраст дости-

жения максимального уровня перед намечающимся ослаблением функций.

Данные проведенного исследования показывают, что развитие и совершенствование бинокулярных зрительных механизмов продолжается в течение длительного периода. Наиболее быстрые темпы увеличения вклада бинокулярных механизмов в пространственное восприятие и улучшения различных функциональных показателей бинокулярного зрения на исследованном возрастном интервале отмечаются в период от 6–7 до 9–10 лет. По средним показателям, фузионные резервы достигают максимальных значений в 12–13 лет, острота бинокулярного стереозрения и скорость бинокулярной интеграции – в 14–15 лет.

Важной теоретической базой, обеспечивающей понимание закономерностей развития зрительных механизмов в онтогенезе, является теория системогенеза (Анохин, 1975). Неодновременное и избирательное созревание отдельных элементов и постепенная консолидация функциональной системы позволяют достигать полезного приспособительного результата задолго до того, как все ее компоненты получат окончательное структурное оформление. Основу гетерохронности развития бинокулярных зрительных функций составляют разная скорость и направленность дифференцировочных процессов в клеточных популяциях. Известно, что участвующие в анализе трехмерной формы на основе бинокулярной диспаритности затылочные, теменные, височные, лобные области коры мозга (Gulyás et al., 1994; Nishida et al., 2001; Naganuma et al., 2005; Fischermeister, Bauer, 2006) имеют длительные перио-

ды моррофункционального созревания. Исследование корковых формаций большого мозга, участвующих в реализации зрительных функций, показало, что в течение всего восходящего онтогенеза от рождения до 20 лет при высокой индивидуальной вариабельности в коре больших полушарий нарастает объем внутрикорковых волокон, что свидетельствует о постепенном увеличении метаболического ресурса для наращивания новых межнейронных контактов и повышения эффективности обработки информации (Цехмистренко и др., 2017). В динамике постнатального развития микроструктуры зон неокортекса большого мозга, связанных со зрительно-пространственным восприятием, на исследованном нами возрастном интервале значимые количественные изменения кластерной организации коры происходят в период от 5–6 до 8–9 лет (Цехмистренко и др., 2019). В 11–12 лет наблюдается более зрелый тип вовлечения дорсолатеральной префронтальной коры, приводящий к повышению эффективности зрительного опознания по сравнению с детьми младшего возраста (Фарбер и др., 2015).

Высокая эффективность вклада бинокулярных механизмов в пространственное восприятие к концу младшего школьного возраста и постепенное ослабление в ходе индивидуального развития является отражением принципа биологической надежности организма, в соответствии с которым биологическая надежность обеспечивается не только высоким приспособительным эффектом гетерохронного формирования функциональных систем, но и такими свойствами живой системы, как избыточность элементов, их дублирование, взаимозамещаемость, быстрота возврата к относительному постоянству и динамичность отдельных звеньев системы.

Полученные нами результаты, свидетельствующие об улучшении бинокулярных функциональных показателей с возрастом, согласуются с имеющимися в литературе данными о совершенствовании других механизмов зрительной системы на этом возрастном этапе и положительной динамике целого ряда зрительных показателей – остроты зрения, точности зрительного восприятия сложных объектов, размеров поля зрения, окуломоторной активности, способности к различению цветов (Рыбалко, 1969; Рожкова и др., 2001; Медведев, Шошина, 2004; Жукова, Циркин, 2008; Безруких и др., 2019; Knoblauch et al., 2001). При этом в ряде исследований было установлено, что, как правило, показатели обнаруживают быстрый рост в начале школьного периода и достигают максимальных значений к подростковому возрасту, после чего сохраняются на достигнутом уровне, либо медленно снижаются.

Использованный в нашей работе комплексный подход позволил выявить степень относи-

тельной зрелости элементов бинокулярной зрительной системы, а также характер их взаимодействия в процессе возрастного индивидуального развития. Согласно данным дискриминантного анализа, нами было установлено, что наибольшее диагностическое значение для характеристики возрастных особенностей пространственного зрительного восприятия у детей имеют конвергентные фузионные резервы, данные стереокинетического теста в бинокулярных и монокулярных условиях наблюдения, острота стереоскопического зрения для пространственной частоты 0.7 цикл/град, бинокулярная острота зрения для близи, показатели успешности бинокулярной интеграции для времени экспозиции тестовых объектов в 100 и 400 мс.

Проведенное исследование показало, что развитие бинокулярной зрительной системы – это длительный процесс, с одной стороны, подчиненный определенным общим закономерностям и направленный на расширение функционального диапазона, и достижение высокой эффективности функционирования, а с другой – в значительной степени обусловленный влиянием наследственности и социально-экономическими факторами, в ходе которого складывается индивидуальное своеобразие.

Обучение в школе является одним из сложных периодов в жизни ребенка как в физиологическом, так и социо-психологическом плане. Адаптация к комплексу факторов, связанных с систематическим обучением и возрастающими умственными и зрительными нагрузками, сопровождается значительным напряжением во всех системах организма, в том числе и зрительной. Очевидно, что результирующую возрастную динамику исследованных зрительных функций определяют онтогенетические изменения, и условия функционирования зрительной системы. В связи с этим на восходящей фазе онтогенеза естественно наблюдать индивидуальный характер качественных преобразований созревающих функций и проявления сенситивных периодов и, соответственно, значительную вариабельность индивидуальных показателей бинокулярной зрительной системы.

Анализ индивидуальных вариантов функционирования бинокулярной зрительной системы у разных школьников подтвердил, что оптимальные режимы деятельности функциональных систем организма не универсальны (не едины для всех сверстников), они определяются непрерывными динамическими процессами, протекающими во времени в сложной сети коадаптированных регуляторных механизмов. Это требует комплексного подхода к анализу реальных функциональных возможностей системы пространственного зрительного восприятия конкретного ребенка не только на основе количественной оценки

отдельных бинокулярных показателей, отражающих зрелость отдельных механизмов, но и качественной характеристики особенностей их взаимодействия и функционирования. Если качественная специфика функционирования физиологических систем на отдельных этапах развития остается неучтеною, то понятие возрастной нормы перестает отражать реальные функциональные возможности организма в определенные возрастные периоды (Безруких, Фарбер, 2014). Сочетание возрастного и индивидуального подходов к изучению особенностей функционирования ребенка может обеспечить разработку адекватных коррекционных и педагогических мер, способствующих полноценному здоровью и нормальному развитию организма ребенка.

Установленная нами тенденция к снижению функциональных показателей бинокулярного зрения в юношеском возрасте может быть связана с естественными возрастными процессами постепенной потери пластичности зрительных механизмов после их полного созревания, а также с влиянием различных неблагоприятных факторов.

## ВЫВОДЫ

— Общий характер возрастных изменений порогов стереовосприятия в период от 7 до 14–15 лет свидетельствует о совершенствовании механизмов бинокулярного стереописца на данном возрастном интервале. На протяжении этого периода острота стереоскопического зрения на разных пространственных частотах увеличивается в 3–4 раза.

— На интервале от 6 до 16 лет с увеличением возраста постепенно повышается доля детей с высокими конвергентными фусионными резервами. Относительно быстрое развитие механизмов аккомодационно-окуломоторного обеспечения фузии осуществляется до 11 лет.

— Возрастные изменения показателей скорости бинокулярной интеграции в период от 7 до 18 лет свидетельствуют о том, что ускорение интегративных процессов формирования единого бинокулярного образа происходит дважды: в младшем школьном возрасте с 7 до 9–10 лет и в подростковом возрасте с 12 до 14–15 лет.

— По результатам стереокинетического теста выявлено постепенное нарастание с возрастом вклада бинокулярных механизмов в пространственное зрительное восприятие и достижение высокой эффективности их функционирования к концу младшего школьного возраста. В ходе дальнейшего развития наблюдается тенденция к ослаблению влияния бинокулярных механизмов, выраженность которой у разных школьников значительно варьирует.

— Выбранный комплекс методов для оценки функциональных показателей бинокулярного зрения можно использовать в целях дифференциальной диагностики нарушений бинокулярного зрения на основе полученных возрастных данных, контроля за динамикой индивидуального развития и разработки программ оптимизации и коррекции функционального состояния бинокулярной зрительной системы.

Автор выражает благодарность главному научному сотруднику ИППИ РАН Г.И. Рожковой за обсуждение работы и консультации при написании статьи, профессору кафедры медицинской биологии с курсом микробиологии и вирусологии ЧГУ им. И.Н. Ульянова В.А. Козлову за техническую помощь при проведении дискриминантного анализа данных.

## КОНФЛИКТ ИНТЕРЕСОВ

Автор данной статьи подтвердил отсутствие конфликта интересов, о котором необходимо сообщить.

## СПИСОК ЛИТЕРАТУРЫ

- Алексеенко С.В. Нейронные сети, обеспечивающие бинокулярное зрение. *Нейротехнологии*. Коллективная монография под ред. Ю.Е. Шелепина, В.Н. Чихмана. СПб.: ООО “Издательство ВВМ”, 2018. С. 186–205.
- Анохин П.К. *Очерки по физиологии функциональных систем*. М.: Медицина, 1975. 447 с.
- Базарный В.Ф. *Зрение у детей. Проблемы развития*. Новосибирск: Наука, 1991. 138 с.
- Безруких М.М., Иванов В.В., Голландцева А.И. Возрастные особенности окуломоторной активности при чтении текста с различных электронных устройств отображения информации у подростков. *Новые исследования*. 2019. № 4 (60). С. 4–29.
- Безруких М.М., Фарбер Д.А. Актуальные проблемы физиологии развития ребенка. *Новые исследования*. 2014. № 3 (40). С. 4–19.
- Белозеров А.Е. *Разработка и внедрение компьютерных функциональных методов в офтальмологии*. Дисс. докт. биол. наук. М. 2003. 285 с.
- Бондарко В.М. Острота зрения и краудинг-эффект у взрослых и детей различного возраста. *Нейротехнологии*. Коллективная монография под ред. Ю.Е. Шелепина, В.Н. Чихмана. СПб.: ООО “Издательство ВВМ”, 2018. С. 46–116.
- Васильева Н.Н. Особенности взаимодействия моно- и бинокулярных механизмов пространственного восприятия у слабовидящих детей. *Вестник офтальмологии*. 2011. Т. 127. № 1. С. 19–21.
- Васильева Н.Н. *Развитие бинокулярных зрительных функций у детей дошкольного и младшего школьного возраста*. Автореф. дисс. канд. биол. наук. М.: 2002. 22 с.
- Васильева Н.Н., Рожкова Г.И. Возрастная динамика фусионных резервов, измеренных при помощи

- циклических тест-объектов с маркерами. *Сенсорные системы*. 2009. Т. 23. № 1. С. 40–50.
- Васильева Н.Н., Рожкова Г.И. Возрастные изменения взаимодействия монокулярных и бинокулярных механизмов пространственного восприятия. *Сенсорные системы*. 2010а. Т. 24. № 1. С. 18–26.
- Васильева Н.Н., Рожкова Г.И., Белозеров А.Е. Возрастная динамика остроты стереозрения у школьников. *Сенсорные системы*. 2010б. Т. 24. № 3. С. 179–187.
- Вашенко Д.И. Возрастная динамика бинокулярных зрительных функций у детей в процессе школьного обучения. Автореф. дисс. канд. биол. наук. М. 2001. 23 с.
- Глазер В.Д., Праздникова Н.В., Леушина Л.И., Невская А.А., Павловская М.Б. Опознание зрительных образов. *Физиология зрения*. М.: Наука, 1992. С. 466–527.
- Жукова Е.А., Циркин В.И. Две тенденции в возрастной динамике остроты зрения мальчиков и девочек на протяжении обучения в средней школе. *Сенсорные системы*. 2008. Т. 22. № 3. Р. 241–247.
- Жукова Е.А., Шитова О.В., Коротаева К.В., Попова Л.И. Анализ индивидуальной динамики изменения остроты зрения вблизи и вдали у детей первых 4 лет обучения в школе. *Вятский медицинский вестник*. 2018. № 4 (60). С. 49–54.
- Котлянский А.М. Функционально-системный анализ бинокулярного зрения. Автореф. дисс. докт. биол. наук. М.: 1982. 66 с.
- Кропман И.Л. *Физиология бинокулярного зрения и расстройства его при содружественном косоглазии*. Л.: Медицина. 1966. 206 с.
- Леушина Л.И. О раздельности каналов для опознания формы и оценки размера изображения. *Проблемы физиологической оптики*. М., Л.: 1967. Т. 14. С. 128–134.
- Медведев Л.Н., Шошина И.И. Влияние возраста на проявление зрительной иллюзии Поггендорфа. *Сенсорные системы*. 2004. Т. 18. № 4. С. 326–330.
- Могилев Л.Н. *Механизмы пространственного зрения*. М.: Наука, 1982. 112 с.
- Морозова Л.В. Специфика психофизиологического обеспечения зрительного восприятия детей 7–8 лет с разным темпом его формирования. *Вестник Северного (Арктического) федерального университета. Серия: Медико-биологические науки*. 2016. № 1. С. 63–72.
- Николаева Е.И., Джалаева А.К. Сенсомоторная интеграция в подростковом возрасте. Обзор литературных источников. *Вестник психофизиологии*. 2020. № 3. С. 11–25.
- Пигарев И.Н. Экстраприарные зрительные зоны коры мозга. *Физиология зрения*. Под ред. Бызова А.Л. М.: Наука, 1992. С. 345–400.
- Плосконос Г.А. *Функциональные подсистемы бинокулярного зрения и их взаимодействие у детей*. Автореф. дисс. канд. мед. наук. М.: 1989. 21 с.
- Рожков С.Н., Овсянникова Н.А. *Стереоскопия в кино, фото-, видеотехнике. Терминологический словарь*. М.: Парадиз, 2003. 136 с.
- Рожкова Г.И. Бинокулярное зрение. *Физиология зрения*. Под ред. Бызова А.Л. М.: Наука, 1992. С. 586–664.
- Рожкова Г.И., Васильева Н.Н. Взаимодействие бинокулярного и стереокинетического механизмов восприятия глубины у детей с нормальным и нарушенным бинокулярным зрением. *Сенсорные системы*. 2001. Т. 15. № 1. С. 61–68.
- Рожкова Г.И., Васильева Н.Н. Компьютерный метод оценки фузионных резервов с объективным контролем нарушения фузии. *Физиология человека*. 2010. Т. 36. № 3. С. 135–137.
- Рожкова Г.И., Васильева Н.Н., Токарева В.С. Бинокулярная интеграция у детей дошкольного и младшего школьного возраста. *Сенсорные системы*. 2002. Т. 16. № 3. С. 221–229.
- Рожкова Г.И., Матвеев С.Г. *Зрение детей: проблемы оценки и функциональной коррекции*. М.: Наука, 2007. 315 с.
- Рожкова Г.И., Токарева В.С., Родионова Е.И., Вашенко Д.И., Васильева Н.Н. Возрастная динамика остроты зрения у школьников. III. Соотношение монокулярных и бинокулярных показателей. *Сенсорные системы*. 2001. Т. 15. № 3. С. 266–275.
- Рыбалко Е.Ф. *Возрастные особенности объема и структуры зрительного поля*. Л.: Изд-во ЛГУ, 1969. 135 с.
- Сомов Е.Е. *Методы офтальмогеномики*. Л.: Наука, 1989. 157 с.
- Супин А.Я. *Нейронные механизмы зрительного анализа*. М.: Наука, 1974. 192 с.
- Супин А.Я. *Нейрофизиология зрения млекопитающих*. М.: Наука, 1981. 252 с.
- Фарбер Д.А., Бетелева Т.Г. Формирование системы зрительного восприятия в онтогенезе. *Физиология человека*. 2005. Т. 31. № 5. С. 26–36.
- Фарбер Д.А., Курганский В.А., Петренко Н.Е. Мозговая организация преднастройки к зрительному опознанию у детей предподросткового возраста. *Физиология человека*. 2015. Т. 41. № 5. С. 5–15.
- Цехмистренко Т.А., Васильева В.А., Шумейко Н.С. Структурные преобразования коры большого мозга у детей и подростков. *Физиология человека*. 2017. Т. 43. № 2. С. 5–14.
- Цехмистренко Т.А., Обухов Д.К., Васильева В.А., Мазлоев А.Б., Шумейко Н.С. Структурные возрастные преобразования нейронных группировок коры большого мозга и мозжечка у детей. *Журнал анатомии и гистопатологии*. 2019. № 8 (4). С. 42–48.
- Шевелев И.А. *Нейроны зрительной коры. Адаптивность и динамика рецептивных полей*. М.: Наука, 1984. 232 с.
- Blake R., Fox R. The psychophysical inquiry into binocular summation. *Ibid.* 1973. V. 14. P. 161–185.
- Chen A.H., Abidin A.H.Z. Vergence and accommodation system in Malay primary school children. *Malaysian J. Med. Sci.* 2002. V. 9 (1). P. 9–15.
- Cogan A.I. Human binocular interaction: towards a neural model. *Vision Res.* 1987. V. 27 (12). P. 2125–2139. [https://doi.org/10.1016/0042-6989\(87\)90127-1](https://doi.org/10.1016/0042-6989(87)90127-1)

- Dyea M.W.G., Bavelier D. Differential development of visual attention skills in school-age children. *Vision Res.* 2010. V. 50 (4). P. 452–459.  
<https://doi.org/10.1016/j.visres.2009.10.010>
- Fischmeister F.P.S., Bauer H. Neural correlates of monocular and binocular depth cues based on natural images: A LORETA analysis. *Vision Res.* 2006. V. 46. P. 3373–3380.  
<https://doi.org/10.1016/j.visres.2006.04.026>
- Gulyás B., Roland P., Heywood C., Popplewell D., Cowey A. Visual form discrimination from luminance or disparity cues: functional anatomy by PET. *Neuroreport*. 1994. V. 5 (17). P. 2367–2371.  
<https://doi.org/10.1097/00001756-199411000-00039>
- Howard I.P., Rogers B.J. *Perceiving in depth. V. 2. Stereoscopic vision*. Oxford. Oxford University Press, 2012. 635 p.
- Hubel D.H., Livingston M.C. Segregation of form, color and stereopsis in primate area 18. *J. Neurosci.* 1987. V. 7 (1). P. 3378–3415.  
<https://doi.org/10.1523/JNEUROSCI.07-11-03378.1987>
- Hubel D.H., Wiesel T.N. Receptive fields, binocular interaction, and functional architecture in the cat visual cortex. *Ibid.* 1962. V. 160. P. 105–256.  
<https://doi.org/10.1113/jphysiol.1962.sp006837>
- Irving E.L., Steinbach M.J., Lillakas L., Babu R.J., Hutchings N. Horizontal Saccade Dynamics across the Human Life Span. *Investigative Ophthalmology & Visual Science*. 2006. V. 47 (6). P. 2478–2484.  
<https://doi.org/10.1167/iovs.05-1311>
- Knoublauch K., Vital-Durand F., Barbur J.L. Variation of chromatic sensitivity across the life span. *Vision Res.* 2001. V. 41. № 1. P. 23–36.  
[https://doi.org/10.1016/s0042-6989\(00\)00205-4](https://doi.org/10.1016/s0042-6989(00)00205-4)
- Naganuma T., Nose I., Inoue K., Takemoto A., Katayama N., Taira M. Information processing of geometrical features of a surface based on binocular disparity cues: an fMRI study. *Neuroscience Research*. 2005. V. 51 (2). P. 147–155.  
<https://doi.org/10.1016/j.neures.2004.10.009>
- Namaeh M., Scheiman M.M., Yaramothu Ch., Alvarez T.L. A Normative Study of Objective Measures of Disparity Vergence and Saccades in Children 9 to 17 Years Old. *Optometry and Vision Science*. 2020. V. 97 (6). P. 416–423.  
<https://doi.org/10.1097/OPX.0000000000001515>
- Nishida Y., Hayashi O., Iwami T., Kimura M., Kani K., Ito R., Shiino A., Suzuki M. Stereopsis-processing regions in the human parieto-occipital cortex. *Neuroreport*. 2001. V. 12 (10). P. 2259–2263.  
<https://doi.org/10.1097/00001756-200107200-00043>
- Oduntan A., Al-Chamdi M., Al-Dosari H. Randot stereoaucuity norms in a population of Saudi Arabian children. *Clinical and Experimental Optometry*. 1998. № 5. P. 193–197.  
<https://doi.org/10.1111/j.1444-0938.1998.tb06734.x>
- Rogers B. *Perception: a very short introduction*. OUP Oxford. University press, 2017. 162 p.
- Tyler C.W. Sensory processing of binocular disparity. *Vergence eye movements: Basic and clinical aspects*. Boston. Butterworths, 1983. P. 199–295.
- Wolfe J.M. Stereopsis and binocular rivalry. *Psychol. Rev.* 1986. V. 93 (3). P. 262–282.

## Analysis of the functions of binocular visual systems in schoolchildren

N. N. Vasilyeva<sup>a, #</sup>

<sup>a</sup> Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences  
 127051 Moscow, B. Karetny per. 19, Russia

#E-mail: nn\_vasilyeva@mail.ru

Presented herein is the analysis of the results of a systematic study of the development of binocular vision (BV) in schoolchildren. The study involved 750 schoolchildren aged 6 to 18 years. The age-related dynamics of functional parameters of the BV were assessed: the acuity of binocular stereovision, convergent and divergent fusional reserves, the speed of binocular integration, the balance of binocular and monocular mechanisms of spatial perception. Comprehensive comparative analysis of BV indicators at different ages made it possible to clarify the timing and rates of heterochronous maturation of individual mechanisms and the age of reaching the maximum level before the outlined weakening of functions. It has been established that the fastest rates of increase in the contribution of binocular mechanisms to spatial perception and improvements towards various indicators of biological health are characteristic for the period from 6–7 to 9–10 years. According to average data, fusional reserves reach maximum values at 12–13 years old, binocular stereovision acuity and binocular integration rate at 14–15 years old. A tendency towards decrease in the functional parameters of BV in adolescence was revealed, which may be associated with the natural age-related processes of the gradual loss of plasticity of visual mechanisms after their full maturation, as well as with the influence of various unfavorable factors. In the course of discriminant analysis of the data, the indicators that have the greatest diagnostic value for characterizing the age-related characteristics of spatial visual perception in children were determined: convergent fusional reserves, stereoscopic visual acuity for a spatial frequency of 0.7 cpd, binocular near-sighted visual acuity, indicators of success in binocular integration within exposure time of test objects between 100 and 400 ms, and the ratio of the stereokinetic test data in binocular and monocular observation conditions.

**Key words:** vision of children, age dynamics, binocular visual system, spatial perception, stereoacuity, fusional reserves, binocular integration

## REFERENCES

- Alekseenko S.V. Neironnye seti, obespechivayushchie binokulyarnoe zrenie [Neural networks that provide binocular vision]. *Neirotekhnologii* [Neurotechnology]. Kollektivnaya monografiya pod red. Yu.E. Shelepinia, V.N. Chikhmana. St. Petersburg. OOO "Izdatel'stvo VVM", 2018. P. 186–205 (in Russian).
- Anokhin P.K. Essays on the physiology of functional systems. Moscow: Meditsina, 1975. 447 p (in Russian).
- Bazarnyi V.F. Vision in children. Development problems. Novosibirsk: Nauka, 1991. 138 p (in Russian).
- Bezrukikh M.M., Ivanov V.V., Gollandtseva A.I. Vozrastnye osobennosti okulomotornoi aktivnosti pri chtenii teksta s razlichnykh elektronnykh ustroistv otoobrazheniya informatsii u podrostkov [Age peculiarities of eye movements when reading texts from various electronic devices in adolescents]. *Novye issledovaniya* [New research]. 2019. V. 4 (60). P. 4–29 (in Russian).
- Bezrukikh M.M., Farber D.A. Aktual'nye problemy fiziologii razvitiya rebenka. [Actual problems of the physiology of child development]. *Novye issledovaniya*. [New research]. 2014. V. 3 (40). P. 4–19 (in Russian).
- Belozerov A.E. Razrabotka i vnedrenie komp'yuternykh funktsional'nykh metodov v oftal'mologii. [Development and implementation of computer functional methods in ophthalmology Dr. biol. sci. diss.]. Moscow. 2003. 285 p (in Russian).
- Bondarko V.M. Ostrota zreniya i krauding-effekt u vzroslykh i detei razlichnogo vozrasta [Visual acuity and crowding effect in adults and children of different ages]. *Neirotekhnologii* [Neurotechnology]. Kollektivnaya monografiya pod red. Yu.E. Shelepinia, V.N. Chikhmana. SPb.: OOO "Izdatel'stvo VVM". 2018. P. 46–116 (in Russian).
- Vasilyeva N.N. Osobennosti vzaimodeistviya monokulyarnykh i binokulyarnykh mekhanizmov prostranstvennogo vospriyatiya u slabovidyashchikh detei [Features of the interaction of monocular and binocular mechanisms of spatial perception in visually impaired children]. *Vestnik oftal'mologii* [Annals of Ophthalmology]. 2011. V. 127 (1). P. 19–21 (in Russian).
- Vasilyeva N.N. Razvitiye binokulyarnykh zritel'nykh funktsii u detei doshkol'nogo i mlashego shkol'nogo vozrasta [The development of binocular visual functions in preschool and primary school children. Dr. biol. sci. diss.]. Moscow. 2002. 22 p (in Russian).
- Vasilyeva N.N., Rozhkova G.I. Vozrastnaya dinamika fuzionnykh rezervov, izmerennykh pri pomoshchi tsiklopicheskikh test-ob'ektorov s markerami [Age dynamics of fusional reserves assessed by means of cyclopean stimuli with markers]. *Sensornye sistemy* [Sensory systems]. 2009. V. 23 (1). P. 40–50 (in Russian).
- Vasilyeva N.N., Rozhkova G.I. Vozrastnye izmeneniya vzaimodeistviya monokulyarnykh i binokulyarnykh mekhanizmov prostranstvennogo vospriyatiya [Age dynamics of interaction between monocular and binocular mechanisms of spatial perception]. *Sensornye sistemy* [Sensory systems]. 2010. V. 24 (1). P. 18–26 (in Russian).
- Vasilyeva N.N., Rozhkova G.I., Belozerov A.E. Vozrastnaya dinamika ostroty stereozreniya u shkol'nikov [Age dynamics of stereoacuity in school children]. *Sensornye sistemy* [Sensory systems]. 2010. V. 24 (3). P. 179–187.
- Vashchenko D.I. Vozrastnaya dinamika binokulyarnykh zritel'nykh funktsii u detei v protsesse shkol'nogo obucheniya [Age-related dynamics of binocular visual functions in children in the process of schooling. Dr. biol. sci. diss.]. Moscow. 2001. 23 p (in Russian).
- Glezer V.D., Prazdnikova N.V., Leushina L.I., Nevskaya A.A., Pavlovskaya M.B. Opoznanie zritel'nykh obrazov [Recognition of visual images]. *Fiziologiya zreniya* [Vision physiology]. Moscow. Nauka, 1992. P. 466–527 (in Russian).
- Zhukova E.A., Tsirkin V.I. Dve tendentsii v vozrastnoi dinamike ostroty zreniya mal'chikov i devochek na protyazhenii obucheniya v srednei shkole [Two Trends in Age Dynamics of Visual Acuity of Boys and Girls During School Years]. *Sensornye sistemy* [Sensory systems]. 2008. V. 22 (3). P. 241–247 (in Russian).
- Zhukova E.A., Shitova O.V., Korotaeva K.V., Popova L.I. Analiz individual'noi dinamiki izmeneniya ostroty zreniya v blizi i vdal' u detei pervykh 4 let obucheniya v shkole [Analysis of individual dynamics of changing acute of vision near and far in children of the first 4 years of training at school]. *Vyatskii meditsinskii vestnik* [Medical Newsletter of Vyatka]. 2018. V. 4 (60). P. 49–54 (in Russian).
- Kotlyarskii A.M. Funktsional'no-sistemnyi analiz binokulyarnogo zreniya [Functional-system analysis of binocular vision. Dr. biol. sci. diss.]. Moscow. 1982. 66 p (in Russian).
- Kropman I.L. Fiziologiya binokulyarnogo zreniya i rasstroistva ego pri sodruzhestvennom kosoglavii [Physiology of binocular vision and its disorder in concomitant strabismus]. Leningrad. Meditsina. 1966, 206 p (in Russian).
- Leushina L.I. O razdel'nosti kanalov dlya opoznaniya formy i otsenki razmera izobrazheniya [Separation of channels for shape recognition and image size estimation]. *Problemy fiziologicheskoi optiki* [Physiological optics problems]. Moscow. Leningrad. 1967. V. 14. P. 128–134 (in Russian).
- Medvedev L.N., Shoshina I.I. Vliyanie vozrasta na proyavlenie zritel'noi illyuzii Poggendorffa [Influence of Age on Display of Optical Poggendorff Illusion]. *Sensornye sistemy* [Sensory systems]. 2004. V. 18. (4). P. 326–330 (in Russian).
- Mogilev L.N. Mekhanizmy prostranstvennogo zreniya [Spatial vision mechanisms]. Moscow. Nauka, 1982. 112 p (in Russian).
- Morozova L.V. Spetsifika psikhofiziologicheskogo obespecheniya zritel'nogo vospriyatiya detei 7–8 let s raznym tempom ego formirovaniya [Specificity of psychophysiological support of visual perception in children of 7–8 years old with a different rate of its formation]. *Vestnik Severnogo (Arkticheskogo) federal'nogo universiteta*. Seriya: Mediko-biologicheskie nauki [Arctic Environmental Research]. 2016. V 1. P. 63–72 (in Russian).

- Nikolaeva E.I., Dzhalaeva A.K. Sensomotornaya integratsiya v podrostkovom vozraste. Obzor literaturnykh i stochnikov [Sensomotor integration in adolescence. Review of literature sources]. *Vestnik psikhofiziologii* [Psychophysiology News]. 2020. V. 3. P. 11–25 (in Russian).
- Pigarev I.N. Ekstrastiarnye zritel'nye zony kory mozga [Extrastriate visual areas of the cerebral cortex]. *Fiziologiya zreniya*. [Vision physiology] Moscow. Nauka, 1992. P. 345–400 (in Russian).
- Ploskonos G.A. Funktsional'nye podsistemy binokulyarnogo zreniya i ikh vzaimodeistvie u detei [Functional subsystems of binocular vision and their interaction in children. Dr. med. sci. diss.]. Moscow. 1989. 21 p (in Russian).
- Rozhkov S.N., Ovsyannikova N.A. Stereoskopiya v kino-, foto-, videotekhnike. Terminologicheskii slovar' [Stereoscopy in film, photo, video equipment. Terminological dictionary]. Moscow. Paradiz. 2003. 136 p (in Russian).
- Rozhкова G.I. Binokulyarnoe zrenie [Binocular vision]. *Fiziologiya zreniya*. [Vision physiology] Moscow. Nauka, 1992. P. 586–664 (in Russian).
- Rozhкова G.I., Vasilyeva N.N. Vzaimodeistvie binokulyarnogo i stereokineticeskogo mekhanizmov vospriyatiya glubiny u detei s normal'nym i narushennym binokulyarnym zreniem [Interaction of binocular and stereokinetic depth mechanisms in children with normal and impaired binocular vision]. *Sensornye sistemy*. [Sensory systems]. 2001. V. 15 (1). P. 61–68 (in Russian).
- Rozhкова G.I., Vasilyeva N.N. Komp'yuternyi metod otsenki fuzionnykh rezervov s ob"ektivnym kontrolem narusheniya fuzii [Computer-Aided Method for the Evaluation of Fusional Reserves with Objective Control of Fusion Break]. *Fiziologiya cheloveka* [Human Physiology]. 2010. V. 36 (3). P. 135–137 (in Russian).
- Rozhкова G.I., Vasilyeva N.N., Tokareva V.S. Binokulyarnaya integratsiya u detei doshkol'nogo i mlashego shkol'nogo vozrasta [Binocular integration in children of preschool and junior school age]. *Sensornye sistemy*. [Sensory systems]. 2002. V. 16 (3). P. 221–229 (in Russian).
- Rozhкова G.I., Matveev S.G. Zrenie detei: problemy otsenki i funktsional'noi korrektssi [Children's vision: problems of assessment and functional correction]. Moscow. Nauka, 2007. 315 p (in Russian).
- Rozhкова G.I. Tokareva V.S., Rodionova E.I., Vashchenko D.I., Vasilyeva N.N. Vozrastnaya dinamika ostroty zreniya u shkol'nikov. III. Sootnoshenie monokulyarnykh i binokulyarnykh pokazatelei [Age dynamics of visual acuity in schoolchildren. III. Relationship of monocular and binocular estimates]. *Sensornye sistemy*. [Sensory systems]. 2001. V. 15 (3). P. 266–275 (in Russian).
- Rybalko E.F. Vozrastnye osobennosti ob"ema i struktury zritel'nogo polya [Age features of the volume and structure of the visual field]. Leningrad. Izd-vo LGU, 1969. 135 p (in Russian).
- Somov E.E. Metody oftal'moergonomiki [Ophthalm ergonomics methods]. Leningrad. Nauka, 1989. 157 p (in Russian).
- Supin A.Ya. Neironnye mekhanizmy zritel'nogo analiza [Neural mechanisms of visual analysis]. Moscow. Nauka, 1974. 192 p (in Russian).
- Supin A.Ya. Neirofiziologiya zreniya mlekopitayushchikh [Neurophysiology of mammalian vision]. Moscow. Nauka, 1981. 252 p (in Russian).
- Farber D.A., Beteleva T.G. Formirovanie sistemy zritel'nogo vospriyatiya v ontogeneze [Formation of the system of visual perception in ontogeny]. *Fiziologiya cheloveka* [Human Physiology]. 2005. V. 31 (5). P. 26–36 (in Russian).
- Farber D.A., Kurganskii V.A., Petrenko N.E. Mozgovaya organizatsiya prednastroiki k zritel'nomu opoznaniyu u detei predpodrostkovogo vozrasta [Brain organization of the preparation to visual recognition in preadolescent children]. *Fiziologiya cheloveka* [Human Physiology]. 2015. V. 41 (5). P. 5–15 (in Russian).
- Tsekhmistrenko T.A., Vasil'eva V.A., Shumeiko N.S. Strukturnye preobrazovaniya kory bol'shogo mozga u detei i podrostkov [Structural Transformations in the Cerebral Cortex in Children and Adolescents]. *Fiziologiya cheloveka* [Human Physiology]. 2017. V. 43 (2). P. 5–14 (in Russian).
- Tsekhmistrenko T.A., Obukhov D.K., Vasil'eva V.A., Mazzloev A.B., Shumeiko N.S. Strukturnye vozrastnye preobrazovaniya neironnykh gruppakov kory bol'shogo mozga i mozzhechka u detei [Structural Transformations in Neural Clusters of the Cerebral and Cerebellar Cortex in Children]. *Zhurnal anatomii i histopatologii* [Journal of anatomy and histopathology]. 2019. V. 8 (4). P. 42–48 (in Russian).
- Shevelev I.A. Neirony zritel'noi kory. Adaptivnost' i dinamika retseptivnykh polei [Neurons of the visual cortex. Adaptability and dynamics of receptive fields]. Moscow. Nauka, 1984. 232 p (in Russian).
- Blake R., Fox R. The psychophysical inquiry into binocular summation. *Jbid.* 1973. V. 14. P. 161–185.
- Chen A.H., Abidin A.H.Z. Vergence and accommodation system in Malay primary school children. *Malaysian J. Med. Sci.* 2002. V. 9 (1). P. 9–15.
- Cogan A.I. Human binocular interaction: towards a neural model. *Vision Res.* 1987. V. 27 (12). P. 2125–2139. [https://doi.org/10.1016/0042-6989\(87\)90127-1](https://doi.org/10.1016/0042-6989(87)90127-1)
- Dyea M.W.G., Bavelier D. Differential development of visual attention skills in school-age children. *Vision Res.* 2010. V. 50 (4). P. 452–459. <https://doi.org/10.1016/j.visres.2009.10.010>
- Fischmeister F.P.S., Bauer H. Neural correlates of monocular and binocular depth cues based on natural images: A LORETA analysis. *Vision Res.* 2006. V. 46. P. 3373–3380. <https://doi.org/10.1016/j.visres.2006.04.026>
- Gulyás B., Roland P., Heywood C., Popplewell D., Cowey A. Visual form discrimination from luminance or disparity cues: functional anatomy by PET. *Neuroreport*. 1994. V. 5 (17). P. 2367–2371. <https://doi.org/10.1097/00001756-199411000-00039>
- Howard I.P., Rogers B.J. Perceiving in depth. V. 2. Stereoscopic vision. Oxford. Oxford University Press, 2012. 635 p.
- Hubel D.H., Livingston M.C. Segregation of form, color and stereopsis in primate area 18. *J. Neurosci.* 1987. V. 7

- (1). P. 3378–3415.  
<https://doi.org/10.1523/JNEUROSCI.07-11-03378.1987>
- Hubel D.H., Wiesel T.N. Receptive fields, binocular interaction, and functional architecture in the cat visual cortex. *Ibid.* 1962. V. 160. P. 105–256.  
<https://doi.org/10.1113/jphysiol.1962.sp006837>
- Irving E.L., Steinbach M.J., Lillakas L., Babu R.J., Hutchings N. Horizontal Saccade Dynamics across the Human Life Span. *Investigative Ophthalmology & Visual Science.* 2006. V. 47 (6). P. 2478–2484.  
<https://doi.org/10.1167/iovs.05-1311>
- Knoublauch K., Vital-Durand F., Barbur J.L. Variation of chromatic sensitivity across the life span. *Vision Res.* 2001. V. 41. № 1. P. 23–36.  
[https://doi.org/10.1016/s0042-6989\(00\)00205-4](https://doi.org/10.1016/s0042-6989(00)00205-4)
- Naganuma T., Nose I., Inoue K., Takemoto A., Katayama N., Taira M. Information processing of geometrical features of a surface based on binocular disparity cues: an fMRI study. *Neuroscience Research.* 2005. V. 51 (2). P. 147–155.  
<https://doi.org/10.1016/j.neures.2004.10.009>
- Namaeh M., Scheiman M.M. Yaramothu Ch., Alvarez T.L. A Normative Study of Objective Measures of Disparity Vergence and Saccades in Children 9 to 17 Years Old. *Optometry and Vision Science.* 2020. V. 97 (6). P. 416–423.  
<https://doi.org/10.1097/OPX.0000000000001515>
- Nishida Y., Hayashi O., Iwami T., Kimura M., Kani K., Ito R., Shiino A., Suzuki M. Stereopsis-peocessing regions in the human parieto-occipital cortex. *Neuroreport.* 2001. V. 12 (10). P. 2259–2263.  
<https://doi.org/10.1097/00001756-200107200-00043>
- Oduntan A., Al-Chamdi M., Al-Dosari H. Randot stereoaucuity norms in a population of Saudi Arabian children. *Clinical and Experimental Optometry.* 1998. № 5. P. 193–197.  
<https://doi.org/10.1111/j.1444-0938.1998.tb06734.x>
- Rogers B. Perception: a very short introduction. OUP Oxford. University?press, 2017. 162 p.
- Tyler C.W. Sensory processing of binocular disparity. *Vergence eye movements: Basic and clinical aspects.* Boston. Butterworths, 1983. P. 199–295.
- Wolfe J.M. Stereopsis and binocular rivalry. *Psychol. Rev.* 1986. V. 93 (3). P. 262–282.

## ОКУЛОМОТОРНАЯ АКТИВНОСТЬ ПРИ РЕШЕНИИ ЗРИТЕЛЬНЫХ КОГНИТИВНЫХ ЗАДАЧ В РАЗЛИЧНЫХ ВРЕМЕННЫХ УСЛОВИЯХ

© 2021 г. А. И. Талеева<sup>1,\*</sup>, Н. В. Звягина<sup>1</sup>

<sup>1</sup>ФГАОУ ВО “Северный (Арктический) федеральный университет им. М.В. Ломоносова”  
163002 г. Архангельск, набережная Северной Двины, д. 17, Россия

\*E-mail: a.taleeva@narfu.ru

Поступила в редакцию 12.01.2021 г.

После доработки 04.03.2021 г.

Принята к публикации 04.03.2021 г.

Окуломоторная активность является физиологическим маркером зрительной когнитивной деятельности. Параметры зрительно-моторных реакций отражают мозговые процессы восприятия, обработки, осмыслиения полученной информации. В современном обществе человеку приходится жить в условиях быстрых социально-экономических и технологических перемен, высоких информационных нагрузок, ускоренного ритма жизни и все чаще возникающего цейтнота. Деятельность в условиях острой нехватки времени или работы, выполняемая в рамках строго ограниченного времени, характеризуется использованием дополнительных физиологических резервов, нарушением психологического и психофизиологического баланса, изменениями в активности мозговой деятельности, что может привести к развитию стресс-реакции. У представителей с разным типом вегетативной регуляции функций реакции на стресс имеют специфические проявления, которые будут прослеживаться в функционировании наиболее реактивных систем организма, влиять на мозговую активность и, соответственно, при реализации зрительной когнитивной деятельности отражаться на параметрах окуломоторной активности. Обследованы 70 студентов Северного (Арктического) федерального университета им. М.В. Ломоносова. Изучали окуломоторные реакции студентов с разным вегетативным статусом при решении когнитивных задач в произвольном темпе и в условиях ограниченного времени. Выявлены наиболее чувствительные параметры окуломоторных реакций при чтении в условиях лимита времени: частота фиксаций, средняя продолжительность фиксаций, дисперсия фиксаций, амплитуда и скорость саккад. При этом показано, что у представителей с различными индивидуально-типологическими особенностями вегетативного статуса эти маркеры могут варьировать. У представителей с разным статусом вегетативной нервной системы в процессе когнитивной работы в разных временных условиях зафиксирована разная эффективность выполнения заданий. Наиболее эффективная и скоростная обработка зрительной информации в условиях временного лимита, сопровождаемая увеличением параметров частоты фиксаций и снижением длительности фиксаций, характерна для представителей с нормотоническим вегетативным статусом.

**Ключевые слова:** окуломоторная активность, саккады, фиксации, когнитивная деятельность, вегетативный статус, физиологическая адаптация

**DOI:** 10.31857/S0235009221030045

### ВВЕДЕНИЕ

Современный мир характеризуется динамичным развитием, активным дополнением имеющейся и появлением новой информации из разных областей знаний, а значит, и необходимостью освоения большого объема информации за ограниченный промежуток времени. Деятельность в условиях цейтнота может повлиять на качество выполненной работы, сопровождаться развитием стресс-реакций, нарушением психофизиологического и психоэмоционального баланса человека (Канеман, 2006). Так, показано, что нехватка времени негативно влияет на адекватность выбора товаров в онлайн магазинах, бо-

лее того, временной лимит искажает восприятие качества и ценности покупки (Mou, Shin, 2018).

Существуют множество физиологических маркеров реакции организма на стресс-агенты, в том числе на лимитирующие временные факторы. Известно, что стресс-реакция сопровождается усилением работы центральной и вегетативной нервных систем, увеличением кровоснабжения мозга и изменениями мозговой активности. При реализации зрительной когнитивной деятельности отражением мозговой активности являются глазодвигательные реакции, которые выполняют различные функции – поиск информации, ее сканирование, построение и опознавание зри-

тельного образа. Движения глаз во время зрительного восприятия информации носят циклический характер и состоят из последовательных саккад и фиксаций (Ярбус, 1965).

Несомненно, в условиях временного ограничения параметры мозговой активности будут меняться, что повлечет и изменения параметров движений глаз (Candelieri et al., 2011). Исследования подтверждают, что в стрессовой ситуации происходят изменения параметров окуломоторной активности (Меркулова, Калинина, 2017). При выполнении зрительной когнитивной нагрузки в условиях стресса наблюдаются удлинение продолжительности фиксаций и резкое снижение числа саккад, которые увеличиваются по амплитуде (Филин, 2002). Эти данные согласуются с результатами исследования изменений в движениях глаз при напряженной работе на симуляторе вождения DriveSafety (Reyes, Lee, 2008). В исследованиях доказано, что выполнение когнитивной задачи в условиях ограничения времени приводит к изменению диаметра зрачка и увеличению умственного напряжения (Hertzum, Holmegaard, 2013). В работе (Stern, 1980) было обнаружено удлинение интервалов между морганиями при высокой когнитивной сосредоточенности.

Общеизвестно, что деятельность в условиях стресса сопровождается изменениями активности и других систем организма, прежде всего сердечно-сосудистой – как наиболее реактивной по отношению к стресс-факторам. Так, было обнаружено, что показатели вариабельности сердечного ритма (ВСР) значимо меняются с изменением напряженности когнитивной деятельности, эти изменения ВСР оказались сильно коррелированными с параметрами движения глаз (Ryu, Myung, 2005).

Изучая изменения деятельности организма и его систем под влиянием стресс-агентов, необходимо учитывать индивидуальные функциональные особенности. Многочисленные исследования свидетельствуют о том, что в условиях психоэмоционального напряжения обнаружаются отчетливые индивидуальные различия в устойчивости людей к стрессу (Александров, Судаков, 2005).

Ограничение времени при выполнении любой работы может оптимизировать все процессы в организме и привести к повышению эффективности, а может иметь и абсолютно противоположный эффект (Бабаева и др., 2012; Базаров, Туманян, 2012). В исследованиях (Кольцова, 2003) упоминается, что ограничение времени, как стресс-фактор, действует на организм обследуемых по-разному, в зависимости от их индивидуальных физиологических и психофизиологических особенностей. Имеются данные о влиянии личностных особенностей на регуляцию сердеч-

ного ритма при выполнении когнитивных нагрузок в условиях стресса. Сопоставление индивидуальных сдвигов показателей функционального состояния в ситуации мобилизационной готовности и при выполнении когнитивных заданий с комфортной и максимальной скоростью в условиях дефицита времени позволило выявить два основных типа вегетативного реагирования на нагрузку у детей 12–13 лет. Первый тип – реагирование преимущественно по симпатическому типу, второй – реагирование по парасимпатическому типу (Криволапчук и др., 2015). Также доказано, что параметры фиксаций и саккад зависят не только от вида и условий выполнения зрительной когнитивной деятельности, но и от индивидуально-типологических особенностей организма (Фокин, 2014).

Таким образом, в ряде работ снижение резерва времени оценивается как стресс-фактор, получены данные о реакциях отдельных систем организма на деятельность в условиях ограничения по времени. Однако имеющиеся данные носят разрозненный и усредненный характер, без учета индивидуальных физиологических особенностей обследованных.

Цель нашего исследования – выявить особенности окуломоторной активности в процессе зрительной когнитивной деятельности без/с ограничения по времени у студентов с разным вегетативным статусом.

## МАТЕРИАЛЫ И МЕТОДИКА

В исследовании приняли участие 70 человек (студенты Северного (Арктического) федерального университета им. М.В. Ломоносова, 35 юношей и 35 девушек). Средний возраст обследованных составил  $19 \pm 1.5$  года.

Участники исследования были соматически здоровы, не имели острых или хронических заболеваний сердечно-сосудистой системы, не имели патологий зрительной системы. Исследование проводили с добровольного согласия участников, вне периода сессии в первой половине дня с соблюдением норм биомедицинской этики (Declaration of Helsinki and European Community directives, 8/609 EC), в изолированном помещении, где действие различных отвлекающих факторов было сведено к минимуму.

Для выявления вегетативного статуса обследованных студентов и деления их на группы использовали запись кардиоинтервалографии при помощи аппаратно-программного комплекса “ВНС-спектр” (“Нейрософт”, Иваново), в первом стандартном отведении (левая рука и правая рука) с применением прижимных электродов (прищепки). Запись показателей электрокардиограммы проводили в состоянии спокойного бодрствования

<i>a</i>					<i>b</i>				
д_ля	х_лм	ш_аб	к_уб	ч_рт	у_ас	н_та	л_на	в_но	ч_до
н_чъ	з_ея	п_эт	т_ло	с_ед	г_аз	п_юс	з_ук	и_ея	м_ст
п_ро	б_да	к_ан	с_ет	п_ля	к_ша	т_мп	р_ма	д_ша	м_шь
р_ка	с_ок	д_ло	щ_ль	о_ыт	з_ма	б_ат	л_цо	с_ыд	к_ай
в_ус	в_ач	с_ло	с_ор	м_зг	у_од	н_бч	с_ль	г_уз	б_ак
г_иб	с_ег	в_ра	у_ок	р_на	б_ро	т_нь	б_нк	б_рт	с_ук
у_ар	к_но	р_за	ц_ет	с_ул	х_еб	к_юч	ю_ор	г_ра	с_ех
у_ро	д_ор	к_са	м_на	с_ол	т_уд	п_ле	н_ша	ж_ра	н_га

Рис. 1. – Пример когнитивной пробы для чтения.

*a* – без ограничения по времени, *b* – в условиях ограничения времени.

ния (положение сидя, при ровном дыхании) и после ортостатической пробы по методике А.М. Вейна (Вейн, 1998). Продолжительность каждой записи для обследуемого составляла 2 мин. Регистрировали временные характеристики сердечного ритма и некоторые показатели вариационной пульсометрии по Р.М. Баевскому (Баевский, 2004).

Вегетативный баланс определяли по значению индекса напряжения (ИН) в состоянии спокойного бодрствования, однако изменения данного показателя значительно варьируют в ортостатической пробе даже у лиц с ваготоническим типом. В связи с этим вегетативный статус определяли не по исходному значению ИН, а по его динамике (реактивности) в ортостатической пробе: если ИН в ортопробе увеличивался более чем на 10% относительно фонового значения, респондентов относили к группе с симпатотоническим вариантом реактивности (группа 1 – 48% обследованных), если ИН уменьшался, то таких лиц относили к группе с ваготоническим вариантом реактивности (группа 2 – 32% обследованных). В случае если колебания ИН оставались в пределах  $\pm 10\%$ , респондентов относили к группе с нормотоническим вариантом (группа 3 – 20% обследованных).

Экспериментальная часть исследования была основана на бинокулярной регистрации окуломоторной активности в процессе решения когнитивных задач. Запись осуществляли при помощи стационарной системы бинокулярного трекинга глаз iView X™ RED, 500 Гц.

Стимульный материал представлял собой два набора по 40 слов из четырех букв с пропущенной второй буквой. В каждом стимуле было восемь строк по пять слов (рис. 1). Слова были взяты из словаря О.Н. Ляшевская, С.А. Шаров “Частотный словарь современного русского языка”. Из 20 тысяч слов было выбрано 2731 существитель-

ное, из них были отобраны слова из четырех букв (296 слов). Рандомно из 296 слов были выбраны 80 слов и далее случайным выбором были разделены на две группы. Слова были напечатаны 24 кеглем шрифтом Times New Roman, черным цветом, с расстоянием между соседними словами и строками 1 см и предъявлялись на экране монитора. Все слова являлись существительными в именительном падеже в единственном числе. В качестве верных ответов считались варианты слов, соответствующие условию эксперимента, например, сало-село, угар-удар. Имена собственные не считались правильным ответом.

Исследование проводили индивидуально с каждым студентом. Для записи параметров окуломоторной активности обследуемый располагался в регулируемом по высоте кресле перед монитором 22” на расстоянии 65–70 см от системы удаленной регистрации движения глаз. Перед началом исследования обследуемому сообщалось, что на экране монитора будут представлены слова с одной пропущенной буквой, и его задача заключается в том, чтобы прочитать их по порядку слева направо, вслух, вставляя пропущенные буквы. Чтение набора слов (когнитивная нагрузка) осуществляли в двух временных режимах: в произвольном темпе без ограничений по времени; с ограничением времени на выполнение задания и сопутствующей инструкцией о необходимости прочитать как можно больше слов за установленный промежуток времени (60 с), при этом в правом верхнем углу монитора на секундомере демонстрировался отсчет оставшегося времени.

Перед началом исследования осуществляли калибровку (запись проводилась с погрешностью не более  $0.1^\circ$ ). После успешной калибровки в центре экрана монитора высвечивалась фиксационная точка, по мере готовности участника к решению заданий на ее месте предъявляли последовательно стимульный материал для работы в двух

временных режимах. Полученные данные трекинга глаз анализировали с применением программного пакета SMI BeGaze. Для проведения статистического анализа были выбраны основные параметры трекинга глаз: частота саккад, частота морганий, средняя длительность саккады, частота фиксаций, средняя длительность фиксации, средний разброс фиксаций, средняя амплитуда саккад.

Во время регистрации окуломоторной активности в процессе решения когнитивных задач одновременно фиксировали верные ответы и время выполнения задания для расчета эффективности когнитивной деятельности и времени обработки одного слова. Эффективность выполнения задания определяли как процент верно названных слов от общего количества слов в стимуле. Время, затрачиваемое на обработку одного слова, рассчитали, деля общее время обработки стимула (в секундах) на количество обработанных слов (ошибки и верные слова). Средняя продолжительность исследования одного студента составила 10–15 мин.

Статистический анализ изучаемых параметров проводили с применением набора компьютерных программ Statistical Package for the Social Sciences (SPSS) for Windows v.23.0. В качестве описательного анализа был выбран метод непараметрической обработки данных, основанный на определении медиан ( $Me$ ) и квартилей ( $Q1-Q3$ ). Для исследования структуры взаимосвязей изучаемых переменных применяли корреляционный анализ с вычислением рангового коэффициента Спирмена. Оценку достоверности различий устанавливали при помощи непараметрического критерия знаковых рангов Уилкоксона. Непараметрические критерии использовали, поскольку распределение выборки не является нормальным. Различия параметров окуломоторной активности между группами определяли на основании U-критерия Манн–Уитни.

## РЕЗУЛЬТАТЫ

Наиболее чувствительными параметрами окуломоторной деятельности при решении когнитивных задач в разных временных условиях у обследованных студентов являются частота и длительность фиксаций, амплитуда и скорость саккад. Именно по этим параметрам были выявлены статистически значимые различия (табл. 1). У представителей всех трех групп при решении когнитивных задач в условиях ограничения времени значительно увеличивается частота фиксаций в секунду ( $p \leq 0.05$ ). У студентов с симпатотоническими и нормотоническими статусами ВНС средняя продолжительность фиксаций снижается ( $p \leq 0.05$ ), у парасимпатотоников – снижается на уровне тенденции. Кроме того, у представите-

лей с симпатотоническим вегетативным статусом обнаружено значимое снижение разброса (дисперсии) фиксаций (при  $p = 0.041$ ). Анализ параметров саккад у обследованных в разных временных условиях работы выявил изменения по величине амплитуды и скорости саккад. Значимый характер изменений этих параметров зафиксирован только в группе с парасимпатическим вегетативным статусом: при ограничении временного ресурса наблюдалось достоверное увеличение указанных параметров. У нормотоников и симпатотоников подобные изменения зарегистрированы на уровне тенденции (табл. 1).

Анализируя особенности окуломоторной активности при зрительной когнитивной деятельности у представителей с разным вегетативным статусом в одинаковых временных условиях, мы выявили наиболее реактивные параметры: частота фиксаций в секунду, частота саккад в секунду и задержка саккады. При чтении в свободном временном режиме у студентов с симпатическим вегетативным статусом частота фиксаций была значимо ниже, чем у представителей с нормотоническим вегетативным статусом ( $p \leq 0.05$ ). При деятельности в условиях ограничения времени частота фиксаций у симпатотоников возрасала, но была значимо ниже, чем у нормо- и парасимпатотоников ( $p \leq 0.05$ ). Также у студентов с симпатотоническим вегетативным статусом в сравнении с нормотониками выявлены статистически значимые различия по параметрам частота саккад в секунду и задержка саккад при деятельности в условиях ограничения времени. У представителей этой группы отмечались достоверно повышенные параметры задержки саккад и достоверно сниженные значения частоты саккад ( $p \leq 0.05$ ) (табл. 1).

Согласно результатам статистического анализа, у представителей с одинаковым вегетативным статусом при сравнении параметров эффективности решения когнитивных задач и времени обработки одного слова в разных временных условиях были установлены достоверные отличия (табл. 2). Ограничение времени достоверно снижает эффективность выполнения когнитивной задачи у симпатотоников ( $p = 0.005$ ) и парасимпатотоников ( $p = 0.018$ ). У представителей с нормотоническим вегетативным статусом эффективность выполнения когнитивной задачи увеличивается на уровне тенденции ( $p = 0.395$ ). У обследованных всех трех групп время на обработку одного слова значительно снижается ( $p \leq 0.05$ ).

Для выявления статистической взаимосвязи между параметрами окуломоторной активности и показателями эффективности деятельности у представителей с разным вегетативным статусом с учетом временных условий выполнения когнитивных задач был проведен корреляционный анализ (рис. 2).

**Таблица 1.** Показатели окуломоторной активности у представителей с разным вегетативным статусом при чтении слов в различных временных условиях

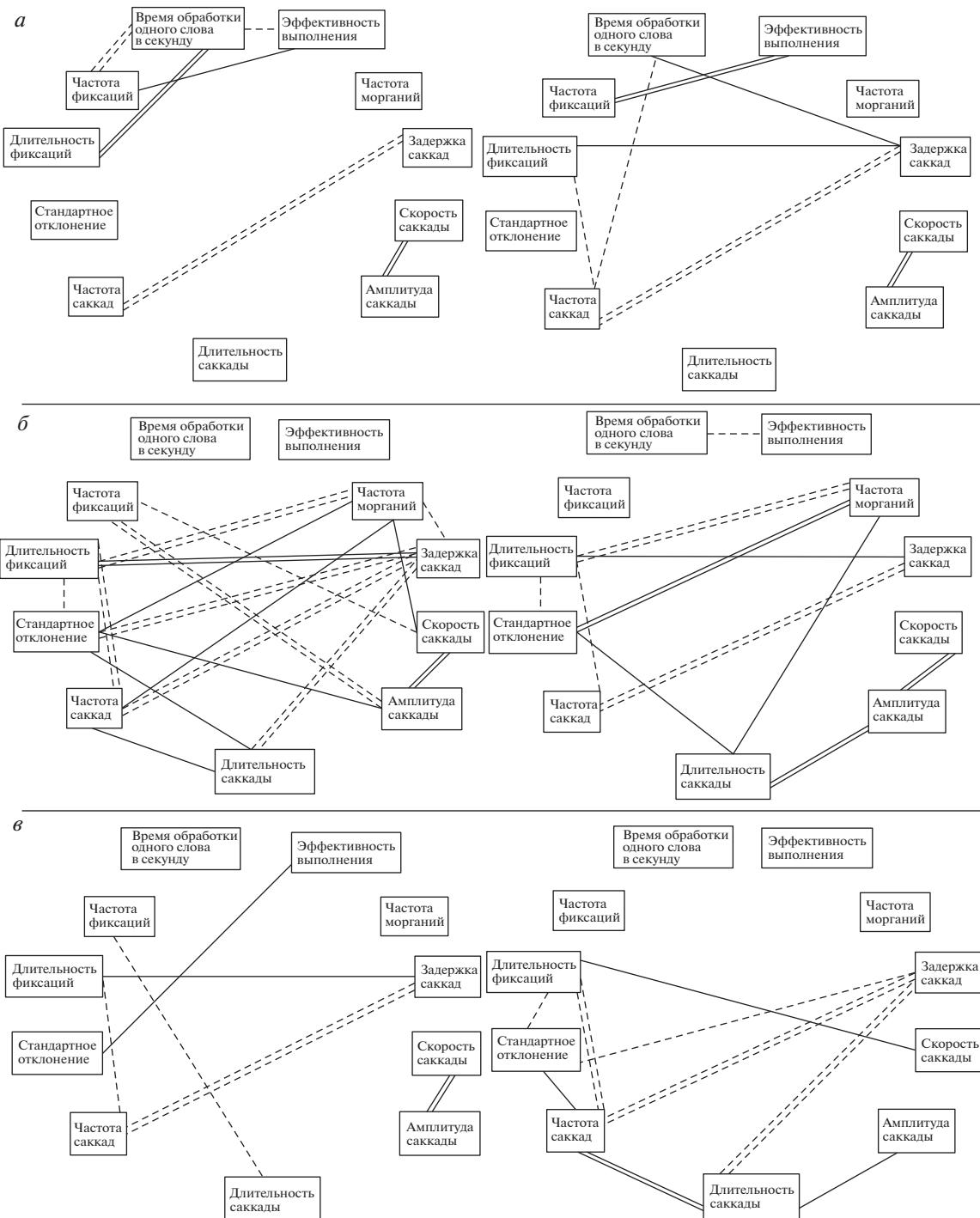
Показатель	Тип	Чтение без лимита Ме (Q1;Q3)	Чтение с лимитом Ме (Q1;Q3)	<i>p</i>
Частота фиксаций /с	Н	2.7 (2.1; 3.3)	3.0 (2.3; 3.6)	0.041*
	С	2.2 (1.9; 2.5) #	2.4 (2.1; 2.8) # ~	0.003*
	П	2.4 (2; 2.7)	2.9 (2.3; 3)	0.006*
Длительность фиксаций, мс	Н	237.2 (206.7; 340.2)	222.9 (164.5; 301.1)	0.028*
	С	296.3 (226.7; 361.5)	264.3 (203.7; 321.7)	0.004*
	П	257 (229.1; 321.9)	253.2 (211.4; 291.6)	0.249
Стандартное отклонение (дисперсия фиксаций), мс	Н	72.4 (66.5; 81.9)	76.5 (57.2; 126.5)	0.398
	С	78.4 (59.7; 91.4)	67.8 (57.8; 92.3)	0.041*
	П	75.2 (64.1; 88.5)	71.7 (57.5; 86.3)	0.294
Частота саккад /с	Н	2.8 (2.3; 4.5)	3.9 (2.8; 4.8)	0.176
	С	2.2 (2; 3.8)	2.7 (2.2; 2.9) #	0.334
	П	3.2 (1.9; 4.4)	3.1 (2.2; 4.3)	0.875
Длительность саккады, мс	Н	36.8 (34.5; 46.8)	46.5 (37.2; 49.7)	0.128
	С	40.7 (36.4; 42.4)	39.2 (37; 43.8)	0.895
	П	39.2 (35.6; 44.1)	39.5 (36.1; 45.8)	0.638
Амплитуда саккады, град	Н	3.3 (3; 7.2)	4.1 (3.3; 7)	0.236
	С	4 (3; 5)	3.9 (3.6; 5.3)	0.859
	П	3.9 (2.9; 4.7)	4.7 (4.1; 5.2)	0.033*
Скорость саккады, град/с	Н	76.2 (68; 113.9)	82.7 (59.7; 167.7)	1.0
	С	85 (66.3; 109.6)	84.5 (73.7; 99.7)	0.923
	П	75.8 (65.9; 98.1)	89.5 (85.2; 106.9)	0.033*
Задержка саккады, мс	Н	304.6 (184.4; 401.3)	215.8 (145.9; 301.6)	0.128
	С	409.6 (222.1; 471.3)	333.4 (308.3; 410.9) #	0.337
	П	260.1 (190.7; 465.3)	287 (187.9; 400.3)	0.422
Частота морганий /с	Н	0.5 (0.1; 1)	0.5 (0.1; 0.8)	0.236
	С	0.5 (0.3; 0.8)	0.4 (0.2; 0.7)	0.174
	П	0.4 (0.2; 0.5)	0.3 (0.1; 0.6)	0.314

Примечание: отличия показателей окуломоторной активности значимы при  $p < 0.05$ : \* с учетом временных условий деятельности; # при сравнении представителей с симпатотоническим и нормотоническим вегетативным статусом; ~ при сравнении представителей с симпатотоническим и парасимпатотоническим статусом; Ме – медиана; (Q1–Q3) – квартили; Н – нормотоники; С – симпатотоники; П – парасимпатотоники.

**Таблица 2.** Показатели успешности выполнения когнитивной задачи в различных временных условиях у представителей с разным вегетативным статусом

Показатель	Тип	Чтение без лимита Ме (Q1;Q3)	Чтение с лимитом Ме (Q1;Q3)	<i>p</i>
Эффективность выполнения	Н	0.82 (0.70; 0.92)	0.85 (0.77; 0.90)	0.395
	С	0.90 (0.80; 0.92)	0.82 (0.75; 0.92)	0.005*
	П	0.95 (0.88; 0.95)	0.85 (0.78; 0.92)	0.018*
Время обработки одного слова (в секундах)	Н	2.07 (1.42; 2.55)	1.10 (1.03; 1.25)	0.019*
	С	2.22 (1.62; 2.82)	1.35 (1.20; 1.53)	0.001*
	П	2.33 (1.34; 2.60)	1.35 (1.04; 1.50)	0.005*

Примечание: \* значения, отличия которых значимы на уровне  $p < 0.05$ ; Ме – медиана; (Q1–Q3) – квартили; Н – нормотоники; С – симпатотоники; П – парасимпатотоники.



**Рис. 2 – Статистическое взаимодействие параметров окуломоторной активности и успешности выполнения зрительных когнитивных задач у испытуемых с разным вегетативным статусом.**

*a* – нормотоники; *b* – симпатотоники; *c* – парасимпатотоники.

Левый столбец корреляционных матриц – чтение в свободном временном режиме.

Правый столбец корреляционных матриц – чтение в условиях временного лимита.

— прямые корреляции, статистически значимые на уровне 0.05,

— обратные корреляции, статистически значимые на уровне 0.05,

— прямые корреляции, статистически значимые на уровне 0.01,

— обратные корреляции, статистически значимые на уровне 0.01.

У представителей с нормотоническим вегетативным статусом при деятельности в свободном временном режиме выявляются прямые статистические связи между параметрами фиксации и эффективностью деятельности, скоростью и амплитудой саккад. При этом выявляются сильные обратные связи между частотой саккад и латентным временем саккад, и с частотой фиксаций и временем обработки одного слова в секунду, а также между параметрами эффективности деятельности и временем обработки одного слова в секунду (рис. 2, *a*). При выполнении когнитивной задачи в условиях лимита времени у представителей этой группы наблюдается усиление прямого взаимодействия между эффективностью деятельности и частотой фиксации. Временной параметр эффективности напрямую зависит от задержки саккады и снижается при увеличении частоты фиксации.

Для представителей с симпатотоническим вегетативным статусом при чтении без временного ограничения в корреляционной плеяде представлены разнообразные достоверные прямые и обратные связи, количество и сила которых снижаются при введении лимитирующего временного фактора (рис. 2, *b*). Очевидно, такие изменения статистического взаимодействия при влиянии лимита времени являются маркерами снижения эффективности когнитивной деятельности.

У представителей с парасимпатотоническим вегетативным статусом при когнитивной деятельности в произвольном временном режиме наблюдается минимальное статистическое взаимодействие изучаемых параметров. Определяющими показателями для данной деятельности являются скорость и амплитуда саккад, латентное время саккад, а также параметры фиксаций (рис. 2, *c*). При этом эффективность деятельности зависит от дисперсии фиксаций. Изменения статистического взаимодействия изучаемых параметров у представителей этой группы обследованных при влиянии лимитирующего временного фактора исключают корреляции с параметрами эффективности.

## ОБСУЖДЕНИЕ

У обследованных с разным вегетативным статусом реализация зрительного восприятия и когнитивных процессов в разных временных условиях сопровождается специфическими изменениями параметров движений глаз, что подтверждается данными зарубежных исследователей (Rendon-Velez et al., 2016).

При сравнении изучаемых параметров в группах с одинаковым вегетативным статусом и с учетом разных временных условий работы были выявлены наиболее реактивные параметры дви-

жений глаз: частота фиксаций, средняя продолжительность фиксаций, дисперсия фиксаций. При действии лимитирующего временного фактора во всех трех группах частота фиксаций увеличивалась, а длительность — снижалась. Некоторые исследователи отмечают, что увеличение частоты фиксаций и уменьшение длительности фиксаций при зрительной когнитивной деятельности могут быть связаны с более сложными условиями работы, с влиянием стресс-фактора, в данном случае таковым является снижение резерва времени (Барabanщиков, Жегало, 2014). При работе в условиях временного ограничения студенты успевают обработать большее количество слов, совершая при этом большее количество фиксаций меньшей продолжительности, т.е. зрительное восприятие в данных условиях характеризуется увеличением аналитической компоненты при меньших временных затратах. При этом дисперсия фиксаций снижается, что свидетельствует о том, что движения глаз становятся более стереотипными.

Были установлены достоверные отличия параметров саккад у представителей с одинаковым вегетативным статусом, но в разных временных условиях работы. По результатам исследований (Барabanщиков, Жегало, 2014) известно, что увеличение амплитуды и скорости саккад свидетельствует об увеличении сложности решаемой задачи и является следствием активации когнитивной деятельности. Следует отметить, что в естественных условиях длительность саккад пропорциональна их амплитуде и скорости (при увеличении длительности саккады амплитуда и скорость также возрастают). Вероятно, такое соотношение параметров саккад является наиболее оптимальным. В условиях работы с ограничением временного ресурса у представителей с парасимпатическим вегетативным статусом значимо увеличиваются амплитуда и скорость саккады при неизменной временной длительности саккады.

Как показали наши исследования, реализация зрительной деятельности в обычных временных условиях и в условиях снижения резерва времени у представителей с разным вегетативным статусом осуществляется по своей специфической программе. Маркерами специфичности являются частота фиксаций, частота саккад и задержка саккад.

Некоторая специфичность была выявлена и при оценке эффективности выполнения задания. У представителей с нормотоническим вегетативным статусом при выполнении когнитивных задач в условиях ограничения времени наблюдаются увеличение эффективности и снижение времени на обработку одного слова. Такая динамика свидетельствует о стимулирующем влиянии временного ограничения на успешность выполне-

ния когнитивной задачи. У представителей с симпатотоническим и парасимпатотоническим вегетативным статусами зафиксировано снижение эффективности выполнения зрительной когнитивной задачи в условиях лимита времени.

Сравнивая особенности статистического взаимодействия изучаемых параметров у представителей с разным вегетативным статусом, можно выделить те изменения корреляций, которые обеспечивают положительную динамику эффективности деятельности в условиях ограничения времени (рис. 2).

В группе нормотоников, где наблюдался рост эффективности деятельности, формируются статистические связи между минимальным количеством изучаемых параметров. Для нормотоников при работе в произвольном ритме определяющими параметрами восприятия зрительной информации являются амплитуда и скорость саккад, длительность и частота фиксаций, при этом эффективность деятельности напрямую зависит от частоты фиксаций, а время на обработку одного слова будет снижаться при увеличении частоты фиксаций и снижении длительности фиксаций. При выполнении когнитивной задачи в условиях лимита времени у представителей этой группы наблюдается усиление статистического прямого взаимодействия, выявленного при деятельности без временных ограничений. Поскольку по результатам оценки эффективности деятельности у нормотоников наблюдается положительная динамика по сравнению с другими группами обследуемых, то, вероятно, именно такие изменения статистического взаимодействия являются наиболее эффективными.

Для представителей с симпатотоническим вегетативным статусом корреляционная плеяда, отражающая взаимодействие параметров окуломоторной активности и эффективности деятельности при чтении слов без временного ограничения, характеризуется разнообразными достоверными прямыми и обратными связями, количество и сила которых незначительно снижаются при введении лимитирующего временного фактора. Однако эти изменения статистического взаимодействия приводят к снижению эффективности когнитивной деятельности.

У представителей с парасимпатотоническим вегетативным статусом при когнитивной деятельности в произвольном временном режиме корреляционная плеяда объединяет скорость и амплитуду саккад, латентное время саккад, а также параметры фиксаций. При этом эффективность деятельности зависит от дисперсии фиксаций. При действии лимитирующего временного фактора статистические перестройки, связанные с потерей взаимосвязей между параметрами эф-

ективности деятельности и показателями окуломоторной активности, приводят к значимому снижению успешности выполнения когнитивной задачи.

## ЗАКЛЮЧЕНИЕ

Как показали наши исследования, реализация зрительного восприятия у представителей с разным вегетативным статусом и в разных временных условиях имеет специфические особенности. Известно, что вегетативный статус организма, а значит, и тип реакции на стресс, определяются преобладанием активации симпатического или парасимпатического отделов ВНС, в связи с чем деление всех участников нашего исследования на группы по типу реакции ВНС обосновано.

Острый недостаток времени является мощным лимитирующим фактором, который может запустить специфические реакции ВНС, повлиять на эффективность и качество работы, энергетические и физиологические затраты. Несомненно, лимит времени следует рассматривать как важный фактор, влияющий на реализацию когнитивных функций.

Поскольку в индивидуальной стратегии адаптации человека к лимитирующему факторам, создающим стрессовые условия, механизмы ЦНС, ВНС и висцеральных функций взаимосвязаны, то параметры окуломоторных реакций в процессе выполнения когнитивной задачи в условиях временного ограничения могут служить маркерами функциональных изменений в организме и успешности деятельности. В результате наших исследований выявлены наиболее чувствительные параметры окуломоторной активности при когнитивной деятельности в условиях лимита времени – частота фиксаций, средняя продолжительность фиксаций, дисперсия фиксаций, амплитуда и скорость саккад. При этом необходимо учитывать, что у представителей с различными индивидуально-типологическими особенностями вегетативного статуса эти маркеры могут варьировать: у нормотоников наиболее чувствительными к лимитирующему временному фактору являются частота фиксаций, средняя продолжительность фиксаций; у симпатотоников – частота фиксаций, средняя продолжительность фиксаций, дисперсия фиксаций; у парасимпатотоников – частота фиксаций, амплитуда и скорость саккад. У представителей выделенных нами групп в процессе когнитивной работы в разных временных условиях зафиксированы разная эффективность выполнения заданий и скорость обработки одного слова. В условиях ограничения времени у симпатотоников и парасимпатотоников снижается эффективность, а скорость выпол-

нения задачи увеличивается, однако, выраженность этих изменений разная. Наиболее эффективная и скоростная обработка зрительной информации в условиях временного лимита, обеспечивающаяся увеличением частоты фиксаций и снижением длительности фиксаций, характерна для представителей с нормотоническим вегетативным статусом.

### ИСТОЧНИК ФИНАНСИРОВАНИЯ

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-313-90062.

### КОНФЛИКТ ИНТЕРЕСОВ

Авторы данной статьи подтвердили отсутствие конфликта интересов, о котором необходимо сообщить.

### СПИСОК ЛИТЕРАТУРЫ

- Александров Е.А., Судаков К.В. Информационные свойства функциональных систем и их математическое моделирование. *Информационные модели функциональных систем*. М.: Фонд “Новое тысячелетие”, 2004. С. 7–32.
- Бабаева Ю.Д., Ротова Н.А., Сабадош П.А. Детерминанты выполнения теста интеллекта в условиях ограничения времени. *Психологические исследования: электронный научный журнал*. 2012.
- Баевский Р.М. Анализ вариабельности сердечного ритма: история и философия, теория и практика. *Клиническая информатика и телемедицина*. 2004. № 1. С. 54–64.
- Базаров Т.Ю., Туманян Д.Г. Влияние дефицита времени на решение творческих задач. *Национальный психологический журнал*. 2012. № 2 (8). С. 116–123.
- Барабанчиков В.А., Жегало А.В. *Айтреинг: методы регистрации движения глаз в психологических исследованиях и практике*. М.: Когито-Центр, 2014. 128 с.
- Вейн А.М. *Вегетативные расстройства. Клиника, диагностика, лечение* / Под ред. А.М. Вейна. М.: Медицина, 1998. 740 с.
- Канеман Д. *Внимание и усилие*. М.: Смысл, 2006. 287 с.
- Кольцова М.М. *Медлительные дети*. СПб.: Речь, 2003. 96 с.
- Криволапчук И.А., Чернова М.Б., Савушкина Е.В. Функциональное состояние детей 12–13 лет при выполнении когнитивных заданий. *Новые исследования*. 2015. № 4 (45). С. 24–32.
- Меркулова А.Г., Калинина С.А. Распределение зрительного внимания при подготовке пилотов-курсантов к летной деятельности. *Гигиена и санитария*. 2017. № 96 (8). С. 752–755.  
<https://doi.org/10.1882/0016-9900-2017-96-8-752-755>
- Филин В.А. *Автоматия саккад*. М.: Изд-во МГУ, 2002. 240 с.
- Фокин В.Ф., Пономарева Н.В., Клопов В.И., Танашян М.М., Лагода О.В. Сосудистая реактивность, вызванная когнитивной нагрузкой, у больных дисциркуляторной энцефалопатией. *Асимметрия*. 2014. № 3. С. 4–22.
- Ярбус А.Л. *Роль движения глаз в процессе зрения*. М: Наука, 1965. 173 с.
- Candelieri A., Riganello F., Cortese D., Sannita W.G. Functional status and the eye-tracking response: A data mining classification study in the vegetative and minimally conscious states. *Proceedings of the International Conference on Health Informatics*. 2011. P. 138–141.  
<https://doi.org/10.5220/0003128201380141>
- Hertzum M., Holmegaard K.D. Perceived Time as a Measure of Mental Workload: Effects of Time Constraints and Task Success. *Intern. J. Human-Computer Interaction*. 2013. V. 29 (1). P. 26–39.  
<https://doi.org/10.1080/10447318.2012.676538>
- Mou J., Shin D. Effects of social popularity and time scarcity on online consumer behaviour regarding smart healthcare products: An eye-tracking approach. *Computers in Human Behavior*. 2018. V. 78. P. 74–89.  
<http://doi.org/10.1016/j.chb.2017.08.049>
- Rendon-Velez E., van Leeuwen P. M., Happee R., Horváth I., van der Vegte W.F., de Winter J.C. F. The effects of time pressure on driver performance and physiological activity: A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2016. V. 41. P. 150–169.  
<https://doi.org/10.1016/j.trf.2016.06.013>
- Reyes M.L., Lee J.D. Effects of cognitive load presence and duration on driver eye movements and event detection performance. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2008. V. 11 (6). P. 391–402.  
<https://doi.org/10.1016/j.trf.2008.03.004>
- Ryu K., Myung R. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*. 2005. V. 35. P. 991–1009.  
<http://doi.org/10.1016/j.ergon.2005.04.005>
- Stern J.A., Aspects of Visual Search Activity Related to Attentional Processes and Skill Development (Final Report, Contract F49620-79-C0089). *Air Force Office of Scientific Research*. Washington, DC. 1980.

## Oculomotor activity in solving visual cognitive tasks under different time conditions

A. I. Taleeva<sup>a,\*</sup> and N. V. Zvyagina<sup>a</sup>

<sup>a</sup> Northern (Arctic) Federal University 163002 Arkhangelsk, Severnaya Dvina Emb. 17, Russia

\*E-mail: a.taleeva@narfu.ru

Oculomotor activity is a physiological marker of cognitive activity. The parameters of visual-motor reactions reflect the brain processes of perception, processing, and comprehension of received information. In con-

temporary society, a person lives in conditions of rapid socio-economic and technological changes, high information loads, accelerated pace of life and ever-increasing time pressure. Activity in time pressure conditions or work performed within strictly limited time are characterized by the use of additional physiological reserves, psychological and psychophysiological disbalance, changes in the brain activity, which can lead to the development of a stress reaction. In representatives with different types of autonomic regulation of functions, stress reactions have specific manifestations that will be traced in the functioning of the most reactive systems of the body, affect brain activity and, accordingly, during the implementation of visual cognitive activity, affect the parameters of oculomotor activity. Seventy students of the Northern (Arctic) Federal University named after M.V. Lomonosov have participated in the research. We studied oculomotor reactions of students with different vegetative status when solving cognitive tasks at an arbitrary pace and under time pressure. The most sensitive parameters of oculomotor reactions when reading under time limit conditions were identified: fixations frequency, average duration of fixations, dispersion of fixations, amplitude and speed of saccades. At the same time, it was shown that in representatives with different individual-typological features of the vegetative status, these markers can vary. In representatives with different status of the autonomic nervous system in the process of cognitive work in different time conditions, different performance of tasks was recorded. The most effective and high-speed processing of visual information in the conditions of the time limit, accompanied by an increase in the parameters of the frequency of fixations and a decrease in the duration of fixations, is characteristic of representatives with a normotonic vegetative status.

**Key words:** Oculomotor activity, saccades, fixations, cognitive task, vegetative status, physiological adaptation

## REFERENCES

- Aleksandrov Ye.A., Sudakov K.V. *Informatsionnye svoystva funktsional'nykh sistem i ikh matematicheskoye modelirovaniye* [Informational properties of functional systems and their mathematical modeling]. *Informatsionnye modeli funktsional'nykh system* [Information models of functional systems]. Moscow. Fond "Novoye tsvyacheletiye", 2004. S. 7–32.
- Babayeva Yu.D., Rotova N.A., Sabadosh P.A. Determinanty vypolneniya testa intellekta v usloviyah ograniceniya vremeni [Determinants of intellectual test performance under time pressure] *Psikhologicheskiye issledovaniya: elektronnyy nauchnyy zhurnal*. [Psychological research: electronic scientific journal]. 2012.
- Bayevskiy R.M. *Analiz variabel'nosti serdechnogo ritma: istoriya i filosofiya, teoriya i praktika* [The analysis of heart rate variability: history and philosophy, theory and practice]. *Klinicheskaya informatika i telemeditsina*. [Clinical informatics and telemedicine]. 2004. V. 1. P. 54–64.
- Bazarov T.Yu., Tumanyan D.G. *Vliyaniye defitsita vremeni na resheniye tvorcheskikh zadach* [The effect of time pressure on creative problem solving]. *Natsional'nyy psikhologicheskiy zhurnal* [National Psychological Journal]. 2012. V. 2(8). P. 116–123.
- Barabanschikov V.A., Zhegallo A.V. *Aitreking: Metody registratsii dvizhenii glaz v psikhologicheskikh issledovaniyakh i praktike* [Eyetracking: Eye movements registration methods in psychological research and practice]. Moscow. Kogito-tsentr. 2014. 128 p.
- Veyn A.M. *Vegetativnyye rasstroystva. Klinika, diagnostika, lecheniye* [Vegetative disorders. Clinical picture, diagnosis, treatment]. Pod red. A.M. Veyna. Moscow. Meditsina. 1998. 740 p.
- Kaneman D. *Vnimaniye i usiliye* [Attention and Effort]. Moscow. Smysl. 2006. 287 p.
- Kol'tsova M.M. *Medlitel'nyye deti* [Slow children]. Sankt-Peterburg: Rech'. 2003. 96 p.
- Krivolapchuk I.A., Chernova M.B., Savushkina Ye.V. *Funktsional'noye sostoyaniye detey 12–13 let pri vypolnenii kognitivnykh zadaniy* [The functional state of children 12–13 years old when performing cognitive tasks]. *Novyye issledovaniya*. [New research]. 2015. V. 4 (45). P. 24–32.
- Merkulova A.G., Kalinina S.A. *Raspredeleniye zritel'nogo vnimaniya pri podgotovke pilotov-kursantov k lotnoy deyatel'nosti* [The distribution of the visual attention in the training of student-pilots for the flight activity]. *Gigiena i sanitariya*. [Gigiena i Sanitaria]. 2017. № 96 (8). P. 752–755. <http://doi.org/10.1882/0016-9900-2017-96-8-752-755>
- Filin V.A. *Avtomatiya sakkad* [Automatic saccades]. Moscow: Izdatel'stvo MGU. 2002. 240 p.
- Fokin V.F., Ponomareva N.V., Klopov V.I., Tanashian M.M., Lagoda O.V. *Sosudistaya reaktivnost', vyzvannaya kognitivnoy nagruzkoj, u bol'nykh distsirkulyatornoy entselopatiyey* [Vascular reactivity induced by cognitive task in patients with vascular encephalopathy]. Asimmetriya. [Asymmetry]. 2014. № 3. P. 4–22.
- Yarbus A.L. *Rol' dvizheniya glaz v protsesse zreniya* [Eye Movements and Vision]. M: Nauka, 1965. 173 p.
- Candelieri A., Riganello F., Cortese D., Sannita W.G. Functional status and the eye-tracking response: A data mining classification study in the vegetative and minimally conscious states. Proceedings of the International Conference on Health Informatics. 2011. P. 138–141. <https://doi.org/10.5220/0003128201380141>
- Hertzum M., Holmegaard K.D. Perceived Time as a Measure of Mental Workload: Effects of Time Constraints and Task Success. Intern. J. Human-Computer Interaction. 2013. V. 29 (1). P. 26–39. <https://doi.org/10.1080/10447318.2012.676538>
- Mou J., Shin D. Effects of social popularity and time scarcity on online consumer behaviour regarding smart healthcare products: An eye-tracking approach. Computers in Human Behavior. 2018. V. 78. P. 74–89. <http://doi.org/10.1016/j.chb.2017.08.049>

- Rendon-Velez E., van Leeuwen P.M., Happee R., Horváth I., van der Vegt W.F., de Winter J.C. F. The effects of time pressure on driver performance and physiological activity: A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2016. V. 41. P. 150–169.  
<https://doi.org/10.1016/j.trf.2016.06.013>
- Reyes M.L., Lee J.D. Effects of cognitive load presence and duration on driver eye movements and event detection performance. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2008. V. 11 (6). P. 391–402.  
<https://doi.org/10.1016/j.trf.2008.03.004>
- Ryu K., Myung R. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*. 2005. V. 35. P. 991–1009.  
<http://doi.org/10.1016/j.ergon.2005.04.005>
- Stern J.A. Aspects of Visual Search Activity Related to Attentional Processes and Skill Development (Final Report, Contract F49620-79-C0089). Air Force Office of Scientific Research. Washington, DC. 1980.

## РАЗЛИЧЕНИЕ ЗВУКОВЫХ СИГНАЛОВ С ГРЕБЕНЧАТЫМИ СПЕКТРАМИ В ПРИСУТСТВИИ ДОПОЛНИТЕЛЬНЫХ СИГНАЛОВ

© 2021 г. М. С. Томозова<sup>1</sup>, Д. И. Нечаев<sup>1</sup>, О. Н. Милехина<sup>1</sup>, А. Я. Супин<sup>1,\*</sup>

<sup>1</sup> Федеральное Государственное бюджетное учреждение науки

Институт проблем экологии и эволюции им. А.Н. Северцова РАН 119071 Москва, Ленинский просп., 33, Россия

\*E-mail: alex\_supin@mail.ru

Поступила в редакцию 29.03.2021 г.

После доработки 15.04.2021 г.

Принята к публикации 28.04.2021 г.

Исследовали различение звуковых сигналов (тестового и референтного) с гребенчатой структурой спектра в условиях, когда на тестовый и референтный сигналы накладывался дополнительный сигнал. В тестовом сигнале каждые 400 мс происходила реверсия фазы гребней. Референтный сигнал имел либо гребенчатый спектр (та же плотность гребней, что в тестовом сигнале, но без изменения фазы гребней), либо спектр с неразличимой гребенчатой структурой. Дополнительный сигнал имел ту же спектральную полосу и тот же уровень, что тестовый и референтный сигналы, и имел спектр либо плоский, либо гребенчатый с плотностью гребней от 2 до 7 цикл/окт. При гребенчатом референтном сигнале разрешение плотности гребней составило 8.7 цикл/окт в контроле (без наложения дополнительного сигнала), 8.5 цикл/окт при плоском дополнительном сигнале и от 8.6 до 9.2 цикл/окт при дополнительном сигнале с плотностью гребней от 2 до 7 цикл/окт; отличия этих значений от контроля не были статистически достоверны. При референтном сигнале с неразличимой спектральной структурой разрешение плотности гребней в тест-сигнале составило 51.0 цикл/окт в контроле, 16.7 цикл/окт при плоском дополнительном сигнале и от 11.0 до 11.1 цикл/окт при дополнительном сигнале с плотностью гребней от 2 до 7 цикл/окт; отличия всех этих значений от контроля были статистически достоверны. Из данных следует, что наложение дополнительного сигнала вызывает эффекты, связанные со спектральным рисунком дополнительного сигнала, причем эти эффекты различны для спектрального и временного механизмов частотного анализа, участвующих в различении сигналов.

**Ключевые слова:** слух, гребенчатые спектры, частотный анализ

**DOI:** 10.31857/S0235009221030057

### ВВЕДЕНИЕ

Звуковые сигналы с гребенчатыми частотными спектрами активно применяются для тестирования частотной разрешающей способности (ЧРС) слуха. В гребенчатых спектрах чередуются максимумы и минимумы спектральной плотности (гребни), создавая таким образом спектральную “решетку”. Максимальная плотность гребней, при которой слуховая система способна отличить тест-сигнал от сигнала с другим расположением спектральных максимумов-минимумов на частотной шкале или от сигнала, спектр которого не имеет гребенчатой структуры (“плоский” спектр) – информативный индикатор ЧРС. Плотность гребней может быть представлена как количество гребней либо на линейную единицу частоты (цикл/кГц), либо на логарифмическую единицу частоты (цикл/окт). Для спектров с равномерным распределением гребней на частотной шкале адекватной является размерность

цикл/кГц, для спектров с логарифмическим распределением – цикл/окт.

Установлено, что ЧРС слуха человека зависит от ряда факторов. В частности, существенно разные оценки ЧРС были получены в зависимости от того, какие сигналы использовались в задаче различения. Если референтный сигнал имеет гребенчатый спектр с той же плотностью гребней, что и тест-сигнал, но с другим расположением гребней на частотной шкале (различие по фазе гребней), то эти два сигнала могут различаться при плотности гребней около 10 цикл/окт (Supin et al., 1994, 1997, 1998, 1999; Anderson et al., 2011; Nечаев, Supin, 2013). При этом возможны варианты: если тест-сигнал содержит несколько сегментов с противоположной фазой гребней, а референтный сигнал характеризуется постоянной фазой гребней, то порог различения составляет 9–10 цикл/окт. Если же и тест-сигнал, и референтный сигнал характеризуются постоянной

фазой гребней, но фаза различна для этих двух сигналов, то порог различения ниже: около 8 цикл/окт (Anderson et al., 2012; Milekhina et al., 2018) или даже 5 цикл/окт (Henry et al., 2005). Различие объяснялось тем, что в первом случае для решения задачи различия кратковременная память не участвует (в тест-сигнале звуки с противоположными фазами гребней спектра непосредственно сменяют друг друга), а во втором случае необходимо участие кратковременной памяти, поскольку тестовый и референтный сигналы предъявляются последовательно, и один из сигналов сравнивается с сохраненным в памяти образом другого.

Принципиально другие результаты получены в экспериментах, в которых требовалось различить тестовый сигнал с гребенчатым спектром и референтный сигнал с плоским спектром. В этом случае различие оказывалось возможным при плотности гребней в тестовом сигнале до 60 цикл/окт (Anderson et al., 2011, 2012) или по крайней мере более 20 цикл/окт (Nechaev et al., 2019). Различие в оценках ЧРС, зависящее от характера референтного сигнала, объясняли тем, что в разных задачах различия слуховая система использует разные механизмы частотного анализа: в одном случае – анализ частотного спектра на уровне слуховой улитки (условно обозначаемый как частотный анализ), в другом – анализ временной структуры сигнала в слуховых нервных центрах (условно обозначаемый как временной анализ). Термины “частотный” и “временной” в данном контексте нужно рассматривать как условные, поскольку любой сигнал и его преобразования могут быть одинаково полно описаны как в частотном, так и во временном представлении, но применительно к анализу сигналов в слуховой системе эти термины обозначают разные физиологические механизмы.

Участие временного механизма в анализе гребенчатых сигналов возможно потому, что они имеют скрытую периодичность: рисунок звуковых колебаний повторяется через временные интервалы, обратные частотным интервалам между гребнями спектра. Эта периодичность проявляется в автокорреляционной функции (АКФ) сигнала как наличие дополнительного сегмента с задержкой  $\tau = 1/\delta f$ , где  $\tau$  – задержка и  $\delta f$  – частотный интервал между гребнями. Для широкополосных сигналов с постоянными частотными интервалами между гребнями этот сегмент редуцируется до узкого пика. Участие временного механизма в анализе гребенчатых спектров первоначально предполагалось в работах (Yost et al., 1996; Patterson et al., 1976). Этот же механизм может быть ответственным за различные оценки ЧРС в экспериментах с гребенчатым и с плоским референтным сигналом. Если и тестовый, и референтный сигналы имеют гребенчатую структуру спектров с

одинаковыми интервалами между гребнями, то положение задержанного сегмента в их АКФ одинаково, и различие этих сигналов при помощи временного механизма невозможно; в различении участвует только частотный механизм. Если же референтный сигнал имеет плоский спектр, то задержанный сегмент в его АКФ отсутствует, т.е. сигнал не имеет скрытой периодичности, и по этому признаку тестовый и референтный сигналы могут различаться при помощи временного механизма.

Все приведенные выше данные о значениях ЧРС для обеих задач различия (с применением гребенчатых и плоских референтных сигналов) получены в условиях, когда тестовый и референтный сигналы предъявлялись испытуемому на “чистом” фоне, не содержащем других звуковых сигналов. Такая идеализированная ситуация не соответствует большинству реальных акустических сцен, в которых “целевой” (т.е. подлежащий обнаружению и анализу) сигнал возникает наряду с множеством других звуковых сигналов, которые по отношению к целевому могут рассматриваться как фоновый шум. Поэтому важно знать, каким образом звуковой фон влияет на ЧРС слуха, и как это влияние зависит от характера фоновых звуков.

Можно ожидать, что добавочный сигнал ухудшит различие гребенчатой структуры спектра тестового сигнала ввиду хорошо известного эффекта энергетической маскировки. Применительно к различию гребенчатых спектров этот эффект будет проявляться в том, что добавочный сигнал уменьшит глубину модуляции (перепад между спектральными максимумами и минимумами) тестового сигнала, а уменьшение глубины модуляции приводит к снижению ЧРС. Однако нетривиальный вопрос состоит в том, сводится ли эффект дополнительного сигнала только к энергетической маскировке, или он зависит еще и от характеристик дополнительного сигнала, в частности, от того, насколько структурирован его частотный спектр.

В настоящем исследовании для решения этого вопроса применена следующая экспериментальная модель. Измеряли ЧРС в условиях, когда тестовый и референтный сигналы суммируются с дополнительным сигналом, который имеет либо гребенчатый, либо плоский спектр. Влияние дополнительного сигнала на ЧРС исследовали при двух задачах различия: между двумя гребенчатыми сигналами (тестовым и референтным) и между гребенчатым тестовым и референтным сигналом с неразличимой спектральной структурой.

## МЕТОДИКА

*Испытуемые.* В исследовании принимали участие девять испытуемых (пять мужчин и четыре

женщины) в возрасте от 25 до 54 лет. Перед экспериментами для всех испытуемых была проведена стандартная аудиометрия, которая показала пороги слышимости не выше 10 дБ на частоте 2 кГц, на которой были центрированы все сигналы. Все испытуемые подписали письменное согласие на участие в экспериментах, предполагающих прослушивание звуковых сигналов с уровнем 70 дБ УЗД.

Во время измерения испытуемый находился в звукоизолированной кабине MINI 350, IAC (Германия), обеспечивающей заглушение внешних звуков не менее чем на 40 дБ.

Программа экспериментов была одобрена Комиссией по биоэтике Института проблем экологии и эволюции РАН.

**Стимулы.** Программа измерений включала предъявление двух видов стимулов: тестовые и референтные.

Тестовый стимул (рис. 1, а) имел спектральную полосу, ограниченную косинусоидальной огибающей с периодом 2 окт. Полоса была центрирована на частоте 2 кГц, т.е. вся полоса включала частоты от 1 до 4 кГц. В пределах огибающей спектр имел гребенчатую структуру, т.е. чередующиеся максимумы и минимумы спектральной амплитуды. Форма гребней описывалась косинусоидальной функцией от логарифма частоты, т.е. гребни спектра были частотно-пропорциональными. Плотность гребней задавалась в цикл/окт и могла принимать следующие значения: 2, 3, 5, 7, 10, 15, 20, 30, 50, 70, 100 цикл/окт (квазилогарифмическая шкала с шестью ступенями на 10-кратный интервал). Во время действия тестового стимула каждые 400 мс фаза спектральных гребней инвертировала (положения спектральных максимумов и минимумов на частотной шкале менялись на противоположные). Стимул содержал шесть сегментов с противоположными фазами гребней спектра, так что общая длительность стимула составляла 2400 мс.

Референтный стимул имел либо различимую гребенчатую, либо неразличимую структуру спектра. Оба варианта референтного стимула имели такую же огибающую спектра и такую же длительность (2400 мс), как тестовый сигнал.

Референтный стимул с различимой гребенчатой структурой (рис. 1, б) имел такую же плотность спектральных гребней, как и тестовый стимул. Отличие от тестового стимула состояло в том, что фаза гребней оставалась постоянной в течение всего времени действия стимула.

Референтный стимул с неразличимой структурой спектра (рис. 1, в) имел спектр, в котором максимумы и минимумы амплитуды чередовались с цикличностью  $2^9 = 512$  цикл/кГц, что соответствовало 362 цикл/окт на нижнем крае частотной полосы (1 кГц) и 90 цикл/окт на верхнем крае

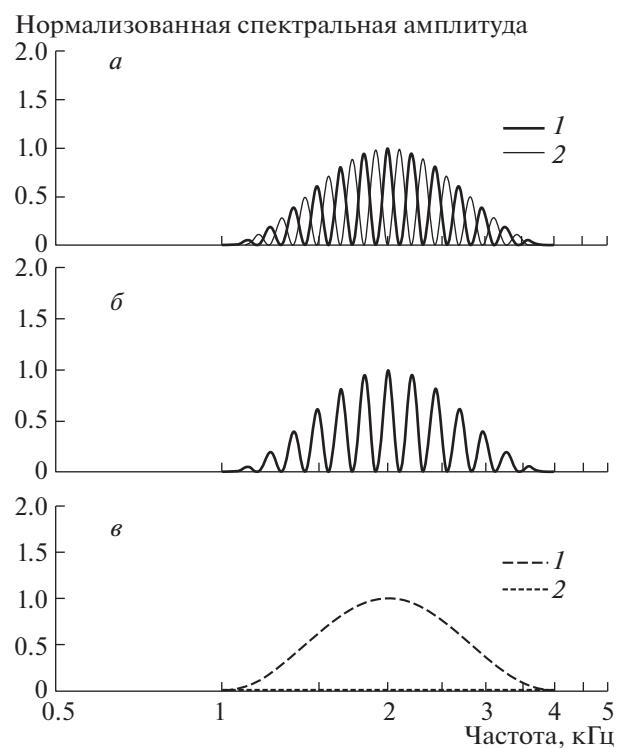


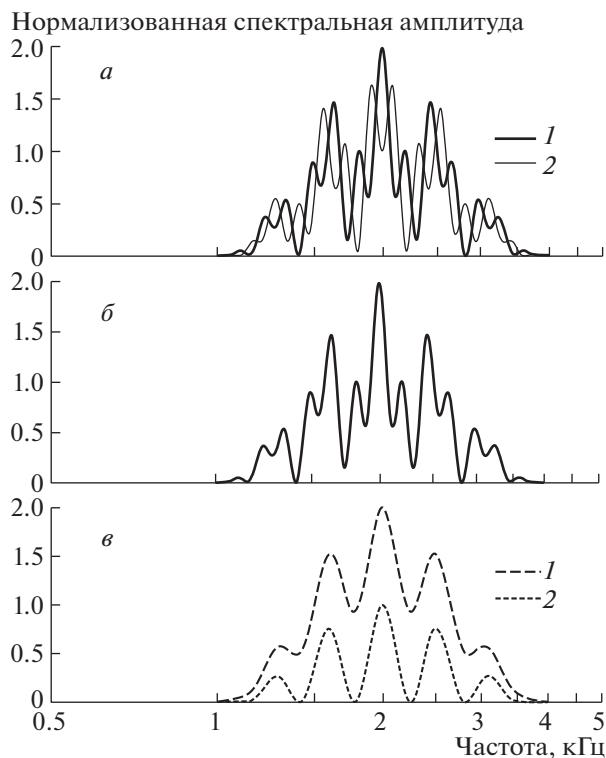
Рис. 1. Частотные характеристики фильтров, используемых для синтеза сигналов (они же – спектры длительных отрезков синтезированных сигналов без наложения дополнительного сигнала).

а – тест-сигнал с плотностью гребней 7 цикл/окт; 1 и 2 – спектры, сменяющие друг друга в течение действия стимула; б – референтный сигнал с гребенчатым спектром; в – референтный сигнал с неразличимой структурой спектра; гребни высокой плотности не могут быть воспроизведены на рисунке, поэтому показаны только верхняя (1) и нижняя (2) границы спектральных амплитуд.

(4 кГц). Эти значения намного превосходят известные данные о разрешении плотности гребней при любой форме эксперимента. Поэтому мы принимали, что структура спектра в таких стимулах заведомо неразличима, т.е. они могут рассматриваться как имитация стимула с плоским спектром. В отличие от стимулов с реально плоским спектром, они имели такой же перепад между минимумом и максимумом спектральных амплитуд, как у тестовых стимулов, что было принципиально важным при сравнении стимулов, на которые накладывался дополнительный сигнал.

Все стимулы, как тестовые, так и референтные имели одинаковые времена нарастания-спада 10 мс и УЗД 70 дБ.

Тестовый и референтный стимулы предъявлялись либо в чистом виде, без наложения дополнительных сигналов (рис. 1), либо на них накладывался дополнительный сигнал. Дополнительный сигнал имел либо плоский, либо гребенчатый



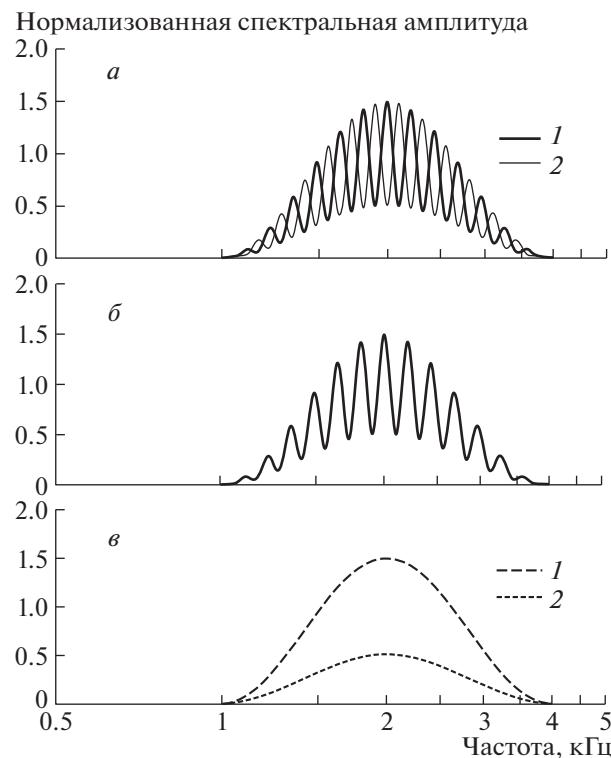
**Рис. 2.** Спектры тестового и референтного сигналов с наложением дополнительного сигнала с гребенчатым спектром 3 цикл/окт.

а – тест-сигнал с плотностью гребней 7 цикл/окт; 1 и 2 – спектры, сменяющие друг друга в течение действия стимула. б – референтный сигнал с гребенчатым спектром; в – референтный сигнал с неразличимой структурой спектра; показаны только верхняя и нижняя границы спектральных амплитуд.

спектр с такой же полосой, как у тестового и референтного сигналов. Соотношение спектральных амплитуд тестового или референтного сигнала и дополнительного сигнала составляло 1 : 1.

В случае гребенчатого дополнительного сигнала суммарные сигналы (тестовый + дополнительный и референтный + дополнительный) имели сложные спектры, описываемые суммой двух косинусоид (рис. 2). Плотность гребней в дополнительном сигнале варьировали ступенчато; она могла принимать значения 2, 3, 5 или 7 цикл/окт. Максимальная плотность 7 цикл/окт была выбрана потому, что она не превышала значений ЧРС, ожидаемых, исходя из литературных данных (Введение). Таким образом, обеспечивалось постоянное условие, состоящее в том, что интервал повторения рисунка в дополнительном сигнале не превышал соответствующий интервал в тестовом и референтном стимулах на пороге различия.

В случае плоского дополнительного сигнала его наложение приводило к уменьшению глуби-



**Рис. 3.** То же, что рис. 2, для дополнительного сигнала с плоским спектром.

ны модуляции спектра тестового и референтного сигналов вдвое по сравнению с “чистым” сигналом (рис. 3).

*Генерация стимулов.* Все стимулы синтезировали цифровым способом с использованием оригинальной программы, составленной в среде LabVIEW (National Instruments, США). Частота стробирования при синтезе сигналов составляла  $2^5 = 32$  кГц, и число точек стробирования составляло  $2^{14} = 16\,384$ , что давало дробность цифрового описания спектра  $2^9 = 512$  точек/кГц. Программа синтеза включала следующие шаги.

- Создавалось цифровое описание спектров тестового, референтного и дополнительного сигналов. Для тестового сигнала создавали цифровое описание двух спектров с одинаковой плотностью гребней, но с противоположными фазами гребней (противоположным положением максимумов и минимумов спектральной амплитуды на шкале частот).

- Если предполагалось применение дополнительного сигнала, то амплитудно-частотные спектры тестового и референтного сигналов суммировали с амплитудно-частотным спектром дополнительного сигнала.

- Обратным преобразованием Фурье получали временные передаточные функции тестового сигнала (с наложенным дополнительным сигна-

лом или без него, в зависимости от задачи эксперимента) и референтного сигнала (также с наложенным дополнительным сигналом или без него).

- Создавали цифровое описание белого шума как гауссовой последовательности чисел.

- Выполняли фильтрацию белого шума соответствующим цифровым фильтром. Фильтрацию проводили путем свертки белого шума с временной передаточной функцией фильтра. Для получения тестового сигнала каждые 400 мс белый шум перенаправляли со входа одного из двух фильтров с противоположными фазами спектральных гребней на вход другого и обратно; выходы фильтров суммировали. Шесть чередующихся сегментов с противоположными фазами спектральных гребней составляли цифровое описание тестового сигнала длительностью 2400 мс. Для получения референтного сигнала использовали один фильтр и отрезок белого шума длительностью 2400 мс.

- К полученному сигналу длительностью 2400 мс применяли амплитудную модуляцию, которая создавала линейное нарастание амплитуды в течение 10 мс и линейный спад в течение такого же времени.

- Амплитуду сигнала устанавливали на таком уровне, чтобы с учетом характеристик головных телефонов уровень выходного сигнала составил 70 дБ УЗД.

Полученные цифровые описания сигналов преобразовывали в аналоговую формуцифроаналоговым преобразователем карты сбора данных NI DAQ-6512 и воспроизводили диотически (идентично на оба уха) через головные телефоны Sennheiser HD580 (Германия).

**Процедура измерения.** ЧРС измеряли психофизическим методом с использованием трехальтернативной процедуры с принудительным выбором. В каждой пробе испытуемому предъявляли последовательность из трех стимулов (каждый длительностью 2400 мс) с паузами 400 мс между ними: один тестовый и два референтных стимула. Последовательность стимулов (тестовый стимул первый, второй или третий в последовательности) меняли от пробы к пробе в случайном порядке. Задача испытуемого состояла в том, чтобы указать, какой из трех стимулов отличается по звучанию от двух других, т.е. является тестовым.

От пробы к пробе плотность гребней в тестовом стимуле меняли по адаптивной процедуре, т.е. в зависимости от предыдущего ответа испытуемого. В случае правильных ответов в двух последовательных пробах плотность спектральных гребней в следующей пробе повышали на один шаг. В случае ошибки плотность спектральных гребней в следующей пробе понижали на один шаг. Такая процедура приводила плотность гребней к значению, которое обеспечивало 71% пра-

вильных обнаружений тестового стимула (Levitt, 1971), что близко к среднему значению (67%) между 100% при безошибочном определении тестового стимула и 33% при случайном угадывании одного стимула из трех. Поэтому значение плотности гребней, к которому приводила адаптивная процедура, принимали за разрешение плотности гребней структуры спектра, т.е. ЧРС. Если использовали референтный сигнал с гребенчатым спектром, то плотность его спектральных гребней меняли от пробы к пробе так же, как в тестовом сигнале. Плотность гребней в дополнительном сигнале при этом не меняли.

Предъявления стимулов с плотностью спектральных гребней, варьирующей от пробы к пробе, продолжали до получения 10 точек реверсии, т.е. перехода от повышения плотности спектральных гребней к понижению и обратно. Среднее геометрическое значение плотности спектральных гребней в этих 10 точках принимали за значение ЧРС в данной измерительной сессии, т.е. для комбинации параметров референтного и дополнительного сигналов, использованной в этой сессии.

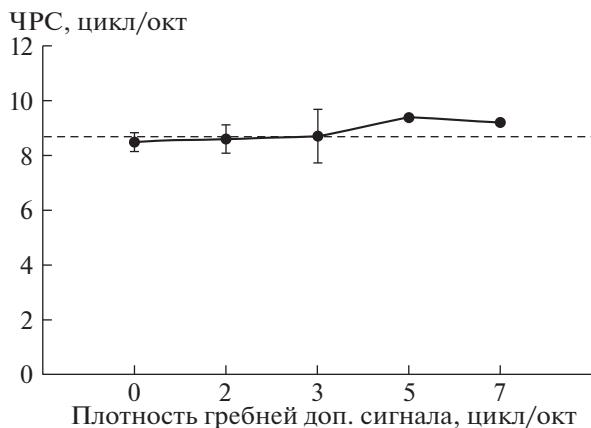
От сессии к сессии параметры референтного и дополнительного сигналов меняли. Для каждой комбинации параметров измерения повторяли трижды на каждом испытуемом. Средние значения по всем измерениям принимали за окончательную оценку ЧРС при данной комбинации параметров.

## РЕЗУЛЬТАТЫ

**Гребенчатый референтный сигнал.** В контроле (в отсутствие дополнительного сигнала) максимальная плотность гребней, при которой тестовый сигнал отличался от референтного, составила в среднем по всем измерениям 8.7 цикл/окт при стандартном отклонении 0.75 цикл/окт (рис. 4).

При наложении плоского (без гребенчатой структуры) дополнительного сигнала разрешение плотности гребней составило в среднем  $8.5 \pm 0.6$  цикл/окт. Отличие от контрольного значения в 0.2 цикл/окт не было статистически достоверным ( $p = 0.34$  по двухвыборочному  $t$ -тесту).

При наложении дополнительного сигнала с гребенчатой структурой разрешение плотности гребней в тест-сигнале также мало отличалось от контрольного: от  $8.6 \pm 0.4$  цикл/окт при плотности гребней в дополнительном сигнале 2 цикл/окт до  $9.2 \pm 1.3$  цикл/окт при плотности гребней в дополнительном сигнале 7 цикл/окт. Ни одно из этих значений не отличалось статистически достоверно от контрольного значения: от  $p = 0.52$  при плотности гребней в дополнительном сигнале 2 цикл/окт до  $p = 0.14$  при плотности гребней в дополнительном сигнале 7 цикл/окт.



**Рис. 4.** Зависимость оценки ЧРС от плотности гребней в спектре дополнительного сигнала в экспериментах с гребенчатым референтным сигналом.

Нулевая плотность гребней – плоский (без гребенчатой структуры) спектр дополнительного сигнала. Горизонтальная пунктирная линия – контрольное (без дополнительного сигнала) значение ЧРС. Планки погрешностей – стандартные отклонения по всем измерениям.

*Референтный сигнал с неразличимой структурой спектра.* В контроле (в отсутствие дополнительного сигнала) максимальная плотность гребней, при которой тестовый сигнал отличался от референтного, составила в среднем по всем измерениям 51.0 цикл/окт при стандартном отклонении 30.3 цикл/окт (рис. 5).

При наложении плоского (без гребенчатой структуры) дополнительного сигнала разрешение плотности гребней значительно снижалось по сравнению с контролем (рис. 4): оно составило в среднем  $16.7 \pm 5.9$  цикл/окт. Отличие от контрольного значения в 34.3 цикл/окт было статистически достоверно ( $P < 0.001$ ).

При наложении дополнительного сигнала с гребенчатой структурой спектра разрешение плотности гребней в тест-сигнале также существенно снижалось по сравнению с контролем, причем было еще ниже, чем при дополнительном сигнале с неразличимой структурой спектра (рис. 5): от  $11.0 \pm 3.5$  цикл/окт при дополнительном сигнале 2 цикл/окт до  $11.1 \pm 2.0$  цикл/окт при дополнительном сигнале 7 цикл/окт. При всех значениях плотности гребней в спектре дополнительного сигнала отличие показателей от таковых для дополнительного сигнала с неразличимой структурой спектра было статистически достоверно ( $p = 0.02$  для всех значений плотности гребней в дополнительном сигнале). С еще более высокой достоверностью эти значения отличались от контрольных (без дополнительного сигнала) значений ( $p < 0.001$  для всех значений плотности гребней в дополнительном сигнале).



**Рис. 5.** То же, что рис. 4, для экспериментов с референтным сигналом с неразличимой плотностью гребней. Плотность “0” дополнительного сигнала – плоский спектр.

## ОБСУЖДЕНИЕ

Присутствие дополнительного сигнала уменьшает глубину модуляции спектра тестового и референтного сигналов, что делает их менее различимыми. Однако приведенные выше данные показывают, что действие дополнительного сигнала не может быть сведено к этому тривиальному эффекту. Действие дополнительного сигнала оказалось принципиально различным при разных задачах различения: между гребенчатым тест-сигналом и гребенчатым референтным сигналом (рис. 4) или между гребенчатым тест-сигналом и негребенчатым референтным сигналом (рис. 5).

*Различение сигналов с гребенчатыми спектрами.* На основании ранее полученных данных (Введение) можно принять, что различие двух сигналов с гребенчатыми спектрами преимущественно происходит с участием спектрального механизма частотного анализа.

В данном исследовании парадоксальным выглядит тот факт, что при добавлении дополнительного сигнала разрешение плотности гребней тест-сигнала заметно не снижалось по сравнению с контролем. Ранее было показано, что с увеличением плотности гребней повышается порог спектрального контраста: при низкой плотности гребней он составляет всего 8.5%, но при повышении плотности гребней сверх 7 цикл/окт порог спектрального контраста приближается к максимально возможной величине 100%, что обуславливает порог плотности гребней 8.9 цикл/окт (Supin et al., 2019). Отсюда следует и обратная зависимость: чем меньше спектральный контраст, тем ниже должна быть разрешаемая плотность гребней. Однако в экспериментах, описанных выше, при наложении как гребенчатого, так и негребенчато-

го дополнительного сигнала порог плотности гребней не снижался по сравнению с контролем.

На самом деле, нет противоречия между данными настоящего исследования и ранее полученными данными. При наложении добавочного сигнала, который имеет плоский спектр и уровень которого равен уровню тестового и референтного сигналов, спектральный контраст снижается вдвое. Если спектральный контраст “чистых” тестового и референтного сигналов составлял 100%, то с наложением дополнительного сигнала он понижается до 50%. Пороги спектрального контраста 50 и 100% могут соответствовать лишь небольшой разнице в плотности спектральных гребней, так что понижение спектрального контраста со 100 до 50% в результате наложения “плоского” дополнительного сигнала может вызывать лишь небольшое снижение разрешения плотности гребней, которое не выявляется на фоне разброса данных.

Что касается дополнительных сигналов с гребенчатым спектром, их эффект может быть еще менее заметен, чем эффект плоского дополнительного сигнала. Наложение спектральных максимумов и минимумов двух сигналов – тестового или референтного и дополнительного – создает спектральный рисунок, в котором присутствуют как пики максимально возможной величины, так и провалы почти до нуля, что хорошо видно на рис. 2, а и б. Спектры с таким контрастом могут обеспечивать практически такое же разрешение плотности гребней, как и спектры с 100%-ной глубиной модуляции в отсутствие дополнительного сигнала.

*Различие между тестовым сигналом с гребенчатым спектром и референтным сигналом с неразличаемой структурой спектра.* В этом случае различие тестового и референтного сигналов происходит преимущественно за счет временного механизма частотного анализа (Anderson et al., 2011; Nechaev et al., 2019). В отсутствие дополнительного сигнала разрешение плотности гребней в несколько раз выше, чем при применении гребенчатого референтного сигнала: 51.0 против 8.7 цикл/окт.

При высоком спектральном контрасте сигнала временной механизм может обеспечить разрешение более высокой плотности гребней, чем спектральный механизм. Однако порог спектрального контраста для временного механизма выше 50% даже при низкой плотности гребней. Поэтому наложение дополнительного сигнала, снижающее спектральный контраст до 50%, делает временной механизм неэффективным. В результате разрешение плотности гребней должно падать до величины, характерной для спектрального механизма, что и наблюдалось в данном исследовании: разрешение снижалось с 51 цикл/окт в контроле

до 10–11 цикл/окт при наложении дополнительного сигнала с любой плотностью гребней, в том числе негребенчатого дополнительного сигнала.

Отдельного рассмотрения заслуживает тот факт, что при наложении негребенчатого дополнительного сигнала разрешение плотности гребней оказывалось хотя и ниже, чем в контроле, но выше, чем при наложении гребенчатого сигнала с любой плотностью гребней, т.е. негативный эффект дополнительного сигнала без гребенчатой структуры спектра был слабее, чем сигнала с гребенчатой структурой. Этот факт не может быть объяснен влиянием сниженного спектрального контраста. Можно допустить, что гребенчатый дополнительный сигнал вызывает срабатывание нейронных структур, которые включают линии задержки, соответствующие частотному интервалу между гребнями спектра дополнительного сигнала. Возможно, активация таких структур подавляет (по типу латерального торможения) активность аналогичных структур, “настроенных” на частотные интервалы между гребнями тестового и референтного сигнала, что затрудняет различение последних.

Таким образом, помимо маскировки наложение дополнительного сигнала на сигналы, подлежащие различию, вызывает специфические эффекты, зависящие от спектрального рисунка дополнительного сигнала. Эти эффекты различны для спектрального и временного механизмов частотного анализа, участвующих в различении сигналов.

Работа поддержана РФФИ (грант № 20-015-00054).

## REFERENCES

- Anderson E.S., Nelson D.A., Kreft H., Nelson P.B., Oxenham A.J. Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. *J. Acoust. Soc. Am.* 2011. V. 130. P. 364–375.
- Anderson E.S., Oxenham A.J., Nelson P.B., Nelson D.A. Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination on cochlear-implant users. *J. Acoust. Soc. Am.* 2012. V. 132. P. 3925–3934.
- Henry B.A., Turner C.W., Behrens A. Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *J. Acoust. Soc. Am.* 2005. V. 118. P. 1111–1121.
- Levitt H. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 1971. V. 49. P. 467–477.
- Milekhina O.N., Nechaev D.I., Supin A.Ya. Estimation of frequency resolving power of human hearing by different methods: Roles of sensory and cognitive factors. *Human Physiol.* 2018. V. 44. P. 123–130.
- Nechaev D.I., Milekhina O.N., Supin A.Ya. Estimates of ripple-density resolution based on the discrimination

- from rippled and nonrippled reference signals. *Trends Hearing*. 2019. V. 23. P. 1–9.
- Nechaev D.I., Supin A.Ya. Hearing sensitivity to shifts of rippled-spectrum patterns. *J. Acoust. Soc. Am.* 2013. V. 134. P. 2913–2922.
- Patterson R.D. Auditory filter shapes derived with noise stimuli. *J. Acoust. Soc. Am.* 1976. V. 59. P. 640–654.
- Supin A.Ya., Milekhina O.N., Nечаев Д.И. Rippled depth thresholds: Estimates obtained by discrimination from rippled and nonrippled reference signals. *Acta Acust. United Acust.* 2019. V. 105. P. 1198–1205.
- Supin A.Y., Popov V.V., Milekhina O.N., Tarakanov M.B. Frequency resolving power measured by rippled noise. *Hearing Res.* 1994. V. 78. P. 31–40.
- Supin A.Y., Popov V.V., Milekhina O.N., Tarakanov M.B. Frequency-temporal resolution of hearing measured by rippled noise. *Hearing Res.* 1997. V. 108. P. 17–27.
- Supin A.Y., Popov V.V., Milekhina O.N., Tarakanov M.B. Ripple density resolution for various rippled-noise patterns. *J. Acoust. Soc. Am.* 1998. V. 103. P. 2042–2050.
- Supin A.Y., Popov V.V., Milekhina O.N., Tarakanov M.B. Ripple depth and density resolution in rippled noise. *J. Acoust. Soc. Am.* 1999. V. 106. P. 2800–2804.
- Yost W.A., Patterson R.D., Sheft S. A time domain description for the pitch strength of iterated rippled noise. *J. Acoust. Soc. Am.* 1996. V. 99. P. 1066–1078.

## Discrimination of rippled sound signals with overlapping additional signals

**M. S. Tomozova<sup>a</sup>, D. I. Nечаев<sup>a</sup>, O. N. Milekhina<sup>a</sup>, and A. Ya. Supin<sup>a,\*</sup>**

<sup>a</sup> Institute of Ecology and Evolution RAS 119071 Moscow, Leninsky Prospect, 33, Russia

\*E-mail: alex\_supin@mail.ru

Discrimination between test and reference rippled sound signals with overlapping additional signals was measured. In the test signal, ripple phase inverted every 400 ms. The reference signal had either a rippled spectrum with the same ripple density as the test or a spectrum with unresolvable ripple pattern. The additional signal had the same spectrum band as the test and reference signals, either flat or rippled 2 to 7 ripples/oct. With a rippled reference signal, the ripple-density resolution was 8.7 ripples/oct in control (no additional signal), 8.5 ripples/oct with a flat additional signal, and 8.6 to 9.2 ripples/oct with a rippled additional signal) of 2 to 7 ripples/oct (all insignificantly differing from control). With a noinrippled reference signal, the ripple density resolution was 51.0 ripple/oct in control, 16.7 ripples/oct with a flat additional signal, and 11.0 to 11.1 ripples/oct with a rippled additional signal 2 to 7 ripples/oct (all significantly differing from control). The conclusion: overlapping additional signals result in specific effects depending on the spectrum pattern of the additional signal, with these effects differing for spectral and temporal mechanisms of frequency analysis.

*Key words:* hearing, rippled spectra, frequency analysis

## THE ROLE OF PROJECTIVE TRANSFORMATIONS IN IMAGE NORMALIZATION

© 2021 г. I. A. Konovalenko<sup>1,\*</sup> and P. P. Nikolaev<sup>1,2</sup>

<sup>1</sup> Institute for Information Transmission Problems of RAS (Kharkevich Institute),  
127051 Moscow, Bolshoy Karetny pereulok, 19, Russia

<sup>2</sup> Moscow Institute of Physics and Technology (National Research University),  
141701 Moscow Region, Dolgoprudny, Institutsky pereulok 9, Russia

\*E-mail: konovalenko@iitp.ru

Received March 25, 2021; Revised April 12, 2021; Accepted April 25, 2021

The analysis of an image captured under arbitrary conditions requires preliminary normalization: a conversion to such a form as if the image was captured under normal, i.e. convenient for the further analysis, conditions. This paper presents a review of modern methods, accuracy criteria, and applications of various normalization approaches. The main stages of the problem development are described. For the first time, the two most important special cases of normalization, conventionally considered independently in the literature, are examined in a unified way. The first special case covers only geometric issues, and the second one is concerned with color aspects exclusively. We demonstrate that the normalization procedure fundamentally involves two- and three-dimensional projective transformations within a general analytical framework, without regard to its color and geometric interpretation for practical problems. This implies the advantage of the suggested unified approach.

**Key words:** geometric and color normalization, projective transformation, homography matrix, root mean square and maximum coordinate discrepancies, normalization accuracy criteria, region of interest

**DOI:** 10.31857/S0235009221030021

### IMAGE NORMALIZATION

Images of the same object of some visual scene differ significantly under different capturing conditions. It is obvious, for example, that images can differ radically when using image-forming optical systems (hereinafter referred to as *cameras*) of different spectral ranges (see Fig. 1).

Note that even if the spectral ranges of cameras are perfectly matching, the captured images can differ due to the divergence in their sensitivity spectra (see Fig. 2).

Images can also vary greatly when using the same camera with different camera settings (focus, aperture, color-correction). When using the same camera with fixed settings, the lighting of the scene (see Fig. 3 and 4), the angle of shooting the object (see Fig. 5) and the optical properties of the environment (see Fig. 6) have a fundamental influence on the resulting image.

The dependence of images on the capturing conditions significantly complicates the analysis of the former. Therefore, when capturing conditions can be controlled, they, as a rule, are chosen to be convenient for the subsequent analysis of the resulting images (for example, a scanner is usually used for documents image-forming) – such conditions are usually called *normal*. For a flat object, an important aspect of normal imaging conditions is usually the orthogonality of the

camera optical axis to the plane of the object (Rodríguez-Piñeiro et al., 2011; Kholopov, 2017). However, the imaging conditions control may be technically difficult (see examples in (Nikolaev et al., 2016; Gladkov et al., 2017)), or impossible. In such cases, it is necessary to solve the *normalization* problem, i.e., to transform the image in such a way as if it was obtained under normal imaging conditions (see examples in Fig. 5 and 6). The initial (*input*) image transformed in this way is called a *normalized image* (see formulas (2) and (3)) (Murygin, 2010), and the imaginary camera which could have captured this normalized image is called a *virtual camera* (Kholopov, 2017).

In the literature, usually one of two special cases of image normalization is considered. In order to describe them, we will consider the image  $I$ , as a function:

$$I: \mathbb{D} \rightarrow \mathbb{V}, \quad (1)$$

where  $\mathbb{D} \subset \mathbb{R}^2$  is the domain of image,  $\mathbb{V}$  is codomain of image. Then, in the first case, the normalization is performed by transformation in the independent variables domain of images (see Fig. 5):

$$I_{\text{norm}}(H_g(\mathbf{r})) = I_{\text{input}}(\mathbf{r}), \quad \mathbf{r} \in \mathbb{D}_{\text{input}}, \quad (2)$$



**Fig. 1.** Images of the same area of the Earth surface taken in the radio (left) and optical (right) spectral bands. The images are reproduced from (Abulkhanov et al., 2018).



**Fig. 2.** Images of the same document captured by different cameras.

and in the second case, in images values domain (see Fig. 6):

$$I_{\text{norm}}(\mathbf{r}) = H_c(I_{\text{input}}(\mathbf{r})), \quad \mathbf{r} \in \mathbb{D}_{\text{norm}}. \quad (3)$$

We will call these types of normalizations *geometric* (Chen et al., 2002; Chekhlov and Ablameiko, 2004; Singh et al., 2008; Zeynalov et al., 2009), and *color normalizations* (Finlayson et al., 1995; Iyatomi et al., 2010; Gong et al., 2019; Kordecki, 2019) respectively. Note that formulas (2) and (3), in general, may not be sufficient to specify  $I_{\text{norm}}$  at each point  $\mathbb{D}_{\text{norm}}$ . For example, in Fig. 5(d), the undefined region of the image  $I_{\text{norm}}$  by formula (2) is shown by squares.

Active research on the topic of geometric normalization of images began with the work (Huttenlocher

et al., 1993) by D. Huttenlocher published in 1993. The term normalization in the described sense was used by E. Blake for the first time in 1994 (Sinclair and Blake, 1994), but it is not widely used in English-language literature. In Russian works on image analysis, the term normalization was introduced by V.A. Gorokhovatsky in 1997 (Triputen' and Gorokhovatskii, 1997) and it is now generally accepted (Putyatin et al., 1998; Lyubchenko and Putyatin, 2002; Chekhlov and Ablameiko, 2004; Vanichev, 2007; Kozlov et al., 2009; Bolotova et al., 2017). The term color image normalization was introduced by G. Finlayson (Finlayson et al., 1998). However, research on this topic began much earlier, for example, in the fundamental work of G. Healey (Healey, 1989) published in 1989.



**Fig. 3.** Images of the same color table taken with the Canon 5D Mark III camera with fixed settings, but under different illumination conditions. This visually demonstrates the color metameristic phenomenon: the colors in different areas of the color table match or differ depending on the lighting. The images are reproduced from the MLSDCR (Multiple Light Source Dataset for Colour Research) (Smagina et al., 2020).



**Fig. 4.** Images of the same banknote under visible (left) and ultraviolet (right) illumination, where a bright fluorescent area stands out.

If the normal imaging conditions are exclusive, then normalization is an *idempotent operation*: applying normalization twice to an image has the same result as applying it once (Finlayson et al., 1998; Nikolaidis, 2011). If, however, there are multiple sets of imaging conditions that can be considered normal (Murygin, 2010; Nikolayev, 2016), it is reasonable to require idempotence additionally. Then normalization will not change the image originally captured under normal conditions. The idempotence combines the operation of normalization of images with the operation of normalization of vectors (reducing them to a single unit length).

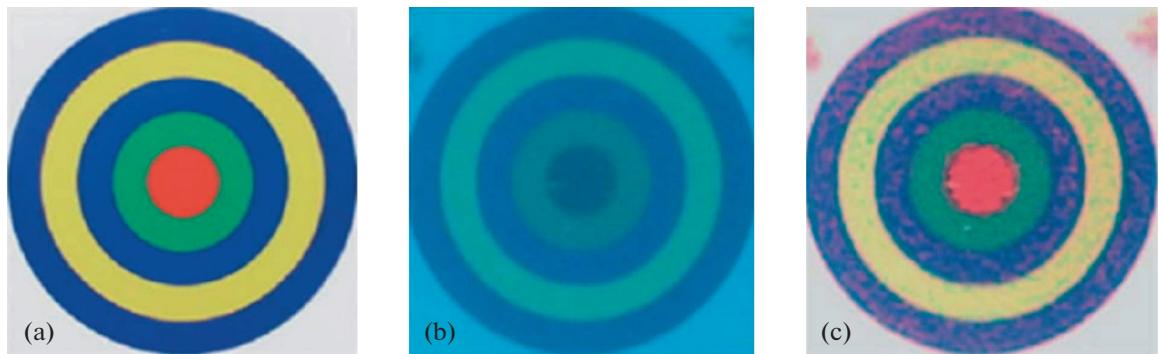
Images obtained under arbitrary imaging conditions are often considered as a result of the distortion of “imaginary” (hypothetical) images obtained under normal imaging conditions. Normalization in such a case is understood as *removal (correction, compensation)* of these distortions (Khlopov, 2017; Baltzopoulos, 1995; Calore et al., 2012; Tsviatkou, 2014).

Normalization of images is usually based on some model of the target object of the scene. For example, such a model can be a so-called *reference image* (Murygin, 2010; Nikolayev, 2010; Vanichev, 2007) – an image of a similar object obtained under normal conditions. In the described case, the geometric normalization of the image can be considered as an operation of its *alignment (image registration)* (Goshtasby, 2005) with the reference image.

Algorithms for the geometric normalization of images are usually based on the assumption that the target object of the scene is sufficiently rich in detail, well known *a priori*, and depicted informatively. Otherwise, most approaches do not yield satisfactory results. To eliminate this problem, a theory has been proposed. This theory provides methods for projective geometric normalization for the extremely complex cases where the information characterizing the object is minimal in the number of its normalizing (projectively invariant) features. An example of such an ap-



**Fig. 5.** Images of the same document. Images (a), (b), and (c) are captured from different angles. Image (d) is the result of geometric normalization of image (b) assuming that the capturing angle of image (c) is chosen as the normal imaging condition.



**Fig. 6.** Images of the same object. Image (a) was obtained in the air environment, (b) – in the water environment, (c) – the result of color normalization of image (b) assuming that the air environment is chosen as the normal imaging condition. Reproduced from (Shepelev et al., 2020).

proach to geometrical normalization is known for flat smooth shapes given by the family of ovals ( $\mathbf{o}$ ). Normalization is carried out by the procedure of its re-description, invariant with respect to the 2D homography transformation, i.e. the projective transformation of the plane  $\mathbf{o}$  in Cartesian 3D space. The algorithms of such processing are particularly simple in the case of symmetry of three kinds (radial, axial or rotational),

while the detection of symmetry elements (axes and/or centers) performed via fast universal procedures (Nikolaev, 2016). It is sufficient enough to obtain an invariant description of  $\mathbf{o}$  as a projection onto a “reference” 4-vertex shape or by calculating a closed curve of 2D wurf-mapping (Nikolaev, 2010). Let us add that the family  $\mathbf{o}$  (their subfamily are the Lamé curves, also called “superellipses” having at least two

axes plus the center of symmetry) rightfully belongs to the objects which do not include the “standard” varieties of projectively invariant points: inflection, fracture, second-order contact and etc. In the model, when the optical registration of  $\sigma$  is adequately described by a flat central projection given by a notional pinhole camera, such an invariant representation of  $\sigma$  (if technically necessary, for some problem of  $\sigma$  recognition) can be transformed (up to a scale accuracy) to a normalized version of representation as if the camera was capturing  $\sigma$  orthogonally. In the context of this approach, the cases when the target object is modeled, for example, by the composition of two ovals (Nikolaev et al., 2018), a composite oval (Nikolaev, 2010), an oval with hidden symmetries (Nikolaev, 2014; Nikolaev, 2017), an oval with an inner point (Savchik and Nikolaev, 2016), an outer straight line (Balitskiy et al., 2017), with two marked points (Savchik and Nikolaev, 2018) were studied.

## THE NORMALIZATION PROBLEM AND ITS PRACTICAL APPLICATIONS

The optical systems of practical interest have geometric aberrations – deviations from the pinhole camera model. One of such aberrations is radial distortion, which disturbs the collinear correspondence between the image and the subject and is typical in inexpensive optical systems equipped with wide-angle lenses and designed for wide-scale use. Geometric aberrations are related to imaging conditions, so their elimination is a special case of normalization. In (Kunina et al., 2016), a single-image blind radial distortion compensation algorithm was proposed.

The same cameras have different radial distortions when capturing images in air and underwater conditions. The paper (Sheshkus et al., 2020) provides an analytical description of visual geometric distortions occurring when capturing underwater objects. A normalization transformation was introduced to compensate for underwater distortion without an underwater calibration procedure. In (Titov et al., 2019), a method for the normalization of underwater color images is proposed.

It is reasonable to consider the algorithms of “automatic white balance” as the color normalization of images. Conventionally, white balance is performed before converting the color coordinates of the camera to the coordinate system of a standard observer for subsequent image finalization and rendering (Karaimer, Brown, 2016), which allows for significantly improved picture quality using simple linear models (Karaimer, Brown, 2018). At the moment, there are many different white balance algorithms. There are two reviews on the subject (Gijsenij et al., 2011; Das et al., 2018), new datasets are being created (Ershov et al., 2020), and the original, more complex formulations were proposed (Savchik et al., 2019).

Another example of color normalization is the reconstruction of multispectral images. The visual non-optimality of the latter is due to various deviations of imaging conditions from normal (Kober, Karnaukhov, 2016a; Kober, Karnaukhov, 2016b; Kober, Karnaukhov, 2015; Chochia, 2016).

Normalization is used for preprocessing of images in many image analysis tasks, some examples of both types of normalization will be discussed further.

## Document Image Analysis

It is common to apply geometric (Rodríguez-Piñeiro et al., 2011), (Skoryukina et al., 2020), and color (Polevoy et al., 2021) normalization to images of documents for the subsequent optical recognition of the latter. At the same time, inaccuracies in the normalization can lead to recognition errors. In (Bulatov et al., 2020), a dataset of video recordings of documents obtained from a variety of camera angles, and in (Smagina et al., 2020), a dataset of images obtained under different light conditions are presented.

In the field of automatic document analysis, the standard task is to remove the slope of letters and significant characters to be recognized, i.e. to perform geometric normalization of the image. There are many papers demonstrating the influence of tilt compensation on all subsequent stages of document processing. The main problem, in this case, is to determine this tilt angle. One of the standard approaches is the Hough image analysis of the document image. The paper (Bezmaternykh, Nikolaev, 2020) investigates the quality of tilt angle detection using the Hough image obtained using the fast Hough transform algorithm.

As part of the task of understanding a document, often the intermediate step is the recognition of certain attributes of the document. For this purpose, the localized attributes are sent as the input of the OCR module in the form of images of text fragments. Most of these modules are designed for the task of recognition of the standard font fragments, but italic and handwritten texts are quite common and also should be recognized. Often, this leads to complications of the standard fragment processing scheme, in particular, algorithms for segmentation of the string image into rasters of individual characters (Chernov et al., 2016). Correction of slanted text fragmentation is one of the classical stages of image normalization in OCR modules. Nowadays, there are many methods for determining the angle of slope of text fragments. Many of them are based on the application of fast Hough transform (Limonova et al., 2017; Bezmaternykh et al., 2018). However, the slant of the characters can occur not only because of the slanted font but also because of the inaccurate normalization of the imaging angle. In (Konovalenko et al., 2020b), an analytical expression for the maximum target direction normalization error for the document was proposed.

A classic step in automatic document processing is *image segmentation*. Its special case is the binarization task: all pixels of the source image are divided into two classes, which are usually referred to as object and background. Application of such segmentation is widely used in systems of document recognition, archival storage. This technique is also employed to improve the visual quality of document images, which is a special case of the problem of image normalization. The task of document binarization attracts the close attention of developers of automatic recognition procedures. There is even a special DIBCO competition to monitor the situation in this area, which is regularly held within ICDAR. Recently, the highest accuracy of binarization is demonstrated by the solutions which employ artificial neural networks. The most popular architecture is U-Net, based on of which new solutions are constantly being proposed. One such solution, for example, won a competition in 2017 (Bezmaternykh et al., 2019). The U-Net architecture has been the subject of in-depth research by scientists, both in finding optimal neuronal activation functions for training (Gayer et al., 2021) and in reducing the number of trainable coefficients in the model (Limonova et al., 2021). However, in some cases, the use of neural network solutions is not feasible, for example, due to limited resources on the device. One of the standard tools in such a case may be the Otsu method or its various generalization (Ershov et al., 2021).

Another relevant task is the recognition of matrix barcodes scanned or captured under conditions unknown *a priori*. Usually, the process of code recognition is divided into several stages. First, the code is localized in the image, then its corners are precisely determined and the normalization of the image is carried out. Then the code is divided into separate modules (matrix code cells), and the original message is extracted. However, sometimes it is possible to apply a generative approach to recognition, instead of individual modules extraction (Bezmaternykh et al., 2010).

In (Kunina et al., 2020), a method for color normalization of document illumination in a full-page scanner without moving parts was proposed, and in (Karnaughov and Kober, 2017), an adaptive method for eliminating shadows in the document image was proposed.

Normalization is also used for the automatic classification of document types (Awal et al., 2017).

Algorithms for geometric normalization of document images are often based on vanishing points (Shemiakina et al., 2020; Abramov et al., 2020). In (Konovalenko et al., 2020c), a method of vanishing point detection based on the principle of maximum likelihood was introduced, and in (Sheshkus et al., 2020), a method based on neural network involvement was demonstrated.

### *Traffic situation recognition*

In the recognition of the traffic situation by automated devices, the features that are localized in the plane of the road are of greatest interest: the markings of the roadway, its boundaries. Recognition of such features is greatly simplified if the image is received from above the plane of the roadway (the so-called birds-eye view) since in this case the parallel lines on the road are projected as the parallel lines on the image. In practice, cameras are mounted behind the windshield of cars. In order to switch from the camera images to the “birds-eye view on the road” view, a projective transformation can be applied that performs a virtual rotation of the camera. Such normalization is the first step of many well-known algorithms for the recognition of road boundaries and road markings (Panfilova et al., 2021; Shipitko et al., 2019; Shipitko et al., 2021; Prun et al., 2017), which is also used for navigation of unmanned vehicles (Abramov et al., 2019; Shipitko, Grigoryev, 2018). Image normalization is also applied to license plate recognition (Murygin, 2010; Povolotskiy et al., 2019; Povolotskiy et al., 2018).

### *Computed tomography scan*

In classical computed tomography, the probing radiation is considered to be monochromatic and the tomographic reconstruction is reduced to the problem of reversal of the Radon transform. However, modern tomographs use X-ray tube with polychromatic radiation as a radiation source. Applying this technique to the projections registered in the polychromatic model obtained by classical reconstruction algorithms leads to distortions in the reconstructed image. To obtain a correct reconstructed image, it is necessary to bring the registered projections to monochromatic, i.e. to solve the normalization problem. It is impossible to solve the normalization problem exactly by any known mathematical transformation. There are many approaches to find an approximate solution. The works (Chukalina et al., 2017; Ingacheva, Chukalina, 2019) suggest numerically simulating monochromatic projections — by applying a one-parameter correction function with automatically found correction parameter-based on the measured in polychromatic mode.

### *Other image normalization applications*

In addition to the above, normalization is used to search for similar images in databases (Orrite and Herero, 2004), to identify television broadcasts from a TV screen image (Skoryukina et al., 2017), to compare space images of Earth with electronic maps (Kozlov et al., 2009), and to analyze medical images (Baltzopoulos, 1995). Brightness normalization is used to improve the accuracy of neural network methods for face verification (Ilyuhin et al., 2019a; Ilyuhin et al., 2019b). In (Nikolaev et al., 2015), normalization was

applied to the detection of diamonds in the ore stream. In addition, G. Legge showed that normalization of images can be applied to facilitate the visual perception by humans (Legge et al., 1985).

## IMAGE NORMALIZATION ACCURACY CRITERIA

Let us consider formally described in the literature interfaces of geometric image normalization algorithms. The image to be normalized is input  $I_{\text{input}}$ . In addition to it, some form of a priori information about the target object and information about what imaging conditions are considered normal may be fed to the input. Some normalization algorithms use independent data about the imaging conditions, such as illumination and perspective (Kholopov, 2017; Calore et al., 2012; Arvind et al., 2018; Karpenko et al., 2015). The normalization algorithm returns an *algorithmically normalizing transformation*  $\hat{H}$  (or its parameters), a transformation of pixel coordinates, application of which to the input image  $I_{\text{input}}$  yields an *algorithmically normalized image*  $I_{\text{alg}}$ . In this case, the image transformation itself is either not performed at all (when it is enough to know the transformation parameters) or, due to its non-triviality, is given for implementation to algorithms specially designed for this purpose. In addition, the estimation of certain imaging conditions can be returned.

Normalization algorithms are usually not required to perform precisely. In order to formalize an estimate of the accuracy of an algorithmically normalizing transformation  $\hat{H}$ , an *ideal normalizing transformation*  $H$  (ground truth) is usually expertly specified. The transform  $\hat{H}$  may then be regarded as an estimate of the ideal transform  $H$ . Let us call the image obtained by applying  $H$  to the input image  $I_{\text{input}}$ , an *ideally normalized image*  $I_{\text{ideal}}$  (see example in Fig. 7). Naturally, the more accurate the normalization is, the closer the transformation  $\hat{H}$  is to the transformation  $H$  or the closer the image  $I_{\text{alg}}$  is to the image  $I_{\text{ideal}}$ . However, there are many reasonable non-equivalent ways to formalize this closeness – the accuracy of geometric normalization inherently has many criteria. A large number of normalization accuracy criteria have been proposed in the literature. Before considering them, let us introduce the necessary notations.

Let us denote by  $\mathbf{r} \stackrel{\text{def}}{=} [x \ y]^T$  the Cartesian coordinates of pixels on the image plane  $I_{\text{ideal}}$ , and by  $\mathbf{q}$  the Cartesian coordinates of pixels in the image plane  $I_{\text{alg}}$  and let us determine the *residual distortion* as

$$V \stackrel{\text{def}}{=} \hat{H}H^{-1}, \quad (4)$$

for each visible point of the target object, the coordinates  $\mathbf{r}$  which translate its image in the image  $I_{\text{ideal}}$  into the coordinates  $\mathbf{q}$  of its image in the image  $I_{\text{alg}}$ :

$$\mathbf{q} = V(\mathbf{r}). \quad (5)$$

If the normalization algorithm works accurately, the residual distortion  $V$  is an identity transformation. Let us also introduce the residual distortion *coordinate discrepancy* (Kunina et al., 2016) (see example in Fig. 8)

$$d(\mathbf{r}) \stackrel{\text{def}}{=} \|\mathbf{r} - V(\mathbf{r})\|_2, \quad (6)$$

for each visible point of the target object. This expression shows the distance by which the image of the point in the image  $I_{\text{alg}}$  is shifted compared to the image of the same point in the image  $I_{\text{ideal}}$ .

In some cases, the normal imaging conditions themselves determine which area of the image  $I_{\text{ideal}}$  is of interest (for example, contains an image of the target object). Then we call it the *region of interest* and denote by

$$R \subset \mathbb{D}_{\text{ideal}}, \quad (7)$$

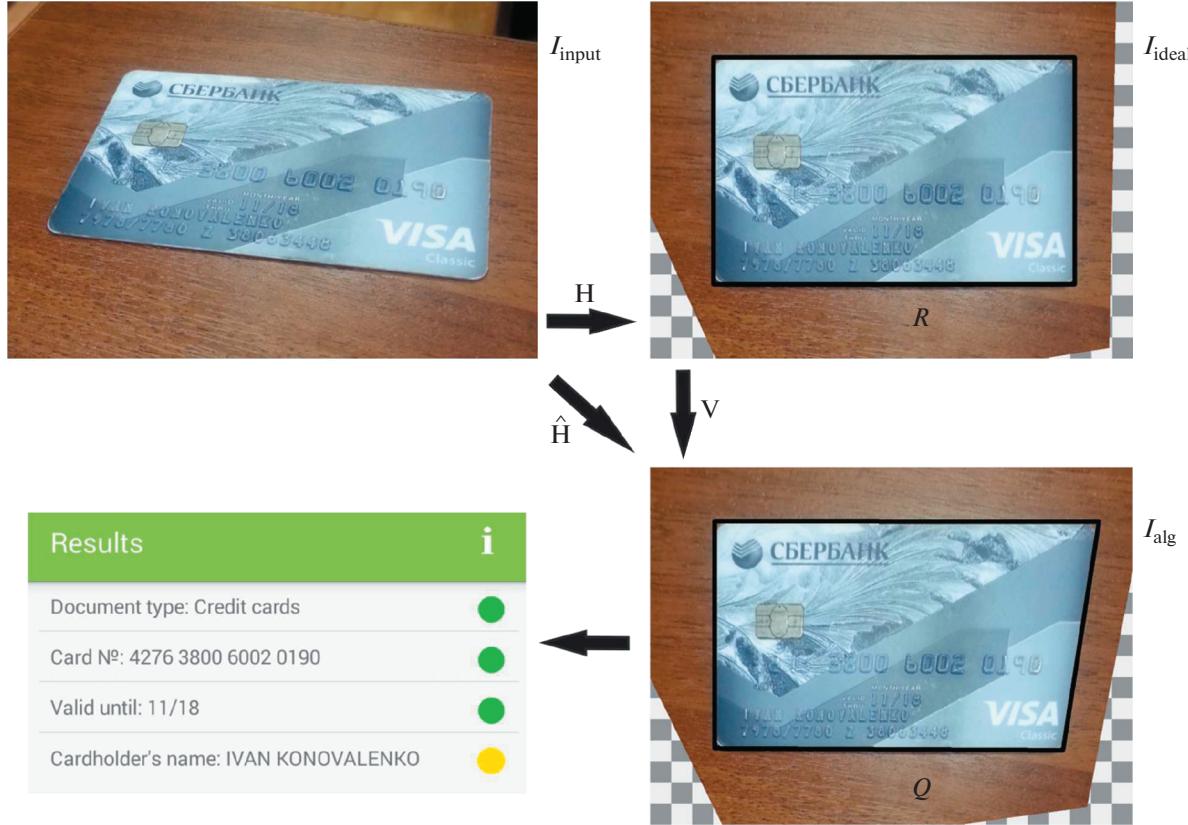
where  $\mathbb{D}_{\text{ideal}}$  is the domain of the image  $I_{\text{ideal}}$  (see (1)). Otherwise, we take  $R = \mathbb{D}_{\text{ideal}}$ . The region of interest  $R$  shows exactly where in the image  $I_{\text{ideal}}$  plane the normalization  $\hat{H}$  is required to be accurate. We will not consider the case of an empty region of interest. Since the image domain  $\mathbb{D}_{\text{ideal}}$  is always bounded, the region of interest  $R$  is also bounded. The sets  $R$  and  $\mathbb{D}_{\text{ideal}}$  will be considered closed, since in this case there is no practical reason not to, and it is mathematically convenient. We will not require the convexity and connectivity of the region of interest  $R$ , since these restrictions are insufficient for practical needs. The image  $Q$  of the region of interest  $R$  in the image  $I_{\text{alg}}$  plane is set by  $Q \stackrel{\text{def}}{=} V[R] \stackrel{\text{def}}{=} \{V(\mathbf{r}): \mathbf{r} \in R\}$ . Ideally,  $Q = R$ .

Hereafter, for brevity, we will use the notation as follows:

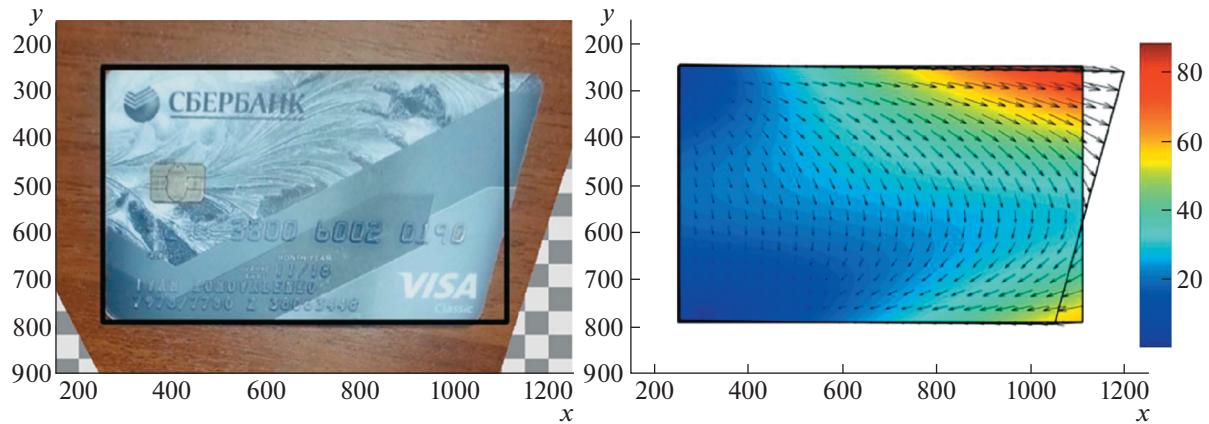
$$\begin{aligned} \max_X f &\stackrel{\text{def}}{=} \max_{\mathbf{x} \in X} f(\mathbf{x}), \\ \sup_X f &\stackrel{\text{def}}{=} \sup_{\mathbf{x} \in X} f(\mathbf{x}) \stackrel{\text{def}}{=} \sup \{f(\mathbf{x}): \mathbf{x} \in X\}, \end{aligned} \quad (8)$$

and we will call  $\sup_X f$  the supremum of the function  $f$  on the set  $X$ .

Now let us proceed to the review of the criteria proposed in the literature for the accuracy of the geometric normalization. The works (Clark et al., 2008; Singh et al., 2008; Zeynalov et al., 2009) suggest evaluating the accuracy of normalization visually. Formal criteria can be divided into three groups: intrasystem, color, and geometric. Further, we will consider these three groups separately.



**Fig. 7.** Example of image normalization and its subsequent analysis.  $H$  – ideal normalization transformation,  $\hat{H}$  – algorithmic normalization transformation,  $V$  – residual distortion,  $I_{\text{input}}$  – image to be normalized,  $I_{\text{ideal}}$  – algorithmically normalized image,  $I_{\text{alg}}$  – algorithmically normalized image, bottom left – result (protocol) of image  $I_{\text{alg}}$  analysis,  $R$  – region of interest,  $Q$  – image of the region of interest  $R$  in the image  $I_{\text{alg}}$  plane.



**Fig. 8.** Example of coordinate discrepancy  $d$ . Left: algorithmically normalized image  $I_{\text{alg}}$ , the area of interest  $R$  is marked by a black frame. Right: vector field of residual distortion displacements  $V(\mathbf{r}) - \mathbf{r}$ , color shows coordinate discrepancy values  $d(\mathbf{r}) = \|V(\mathbf{r}) - \mathbf{r}\|_2$ .

#### Intrasytem criteria

We have already shown that the normalization is applied as a stage of preprocessing of images for various image analysis problems. Thus, there is an ap-

proach where the accuracy of the normalization is defined as the quality of the solution of the problem in which it is applied. For example, in (Merino-Gracia et al., 2013; Lu et al., 2005; Zhang et al., 2008; Tong, Zhang, 2010; Takezawa et al., 2016) the criterion for

the accuracy of document image normalization is chosen as the quality of text recognition on an algorithmically normalized image  $I_{\text{alg}}$ ; in (Awal et al., 2017; Finlayson et al., 1998) – the quality of scene object recognition; and in (Skoryukina et al., 2017) – the proportion of correctly identified TV broadcasts. Since such accuracy criteria are defined exclusively within the framework of some image analysis system and its testing system, we will call them *intrasytem criteria*. They depend only on a single image or a set of images  $I_{\text{alg}}$ .

Intrasytem criteria of accuracy are undoubtedly useful, as improvement of normalization algorithm in the sense of intrasytem criterion by its definition means the improvement of quality of solution of the final image analysis problem, and exactly in the sense in which this quality is specified. However, the criteria proposed in the literature are not limited to intrasytem criteria. The reason for this lies in the violation of the principle of software modularity, which requires that the development (and hence testing) of modules should be carried out independently. This violation results in the following problems. The system of image analysis and/or system of its testing by the time of the introduction of the normalization algorithm may not be fully developed. If both of these systems already exist, they are usually in the process of constant change, so the intrasytem criteria for normalization accuracy are also changing. In addition, they are not mathematically formalized, difficult to analyze, make the software debugging process difficult, and do not imply the universality of normalization algorithms. Zeynalov et al. (2009) describe intrasytem criteria as incorrect for these reasons. There is an opposite approach, in which it is believed that the image analysis system must be such that the quality of its performance on normalized images correlated well with some simple fixed criterion for the accuracy of normalization of these images.

### Color criteria

By *color criteria* for the accuracy of geometric normalization, we will denote the criteria that necessarily depend on both images  $I_{\text{alg}}$  and  $I_{\text{ideal}}$ , and may depend on the region of interest  $R$  and its image  $Q$ . For example, the articles (Szeliski, 1996; Sawhney and Kumar, 1999; Calderon and Romero, 2007; Goshin et al., 2014) use the root mean square pointwise difference of these images throughout the region of interest as the normalization accuracy of single-channel images (see Fig. 9):

$$I_2(I_{\text{alg}}, I_{\text{ideal}}; R) \stackrel{\text{def}}{=} \sqrt{\frac{1}{S(R)} \int_R (I_{\text{alg}}(\mathbf{r}) - I_{\text{ideal}}(\mathbf{r}))^2 d\mathbf{r}}, \quad (9)$$

where  $S(R)$  is the area of the region of interest  $R$ ; in (Tsviatkou, 2014; Gong et al., 2019), peak signal-to-noise ratio (PSNR):

$$\begin{aligned} PSNR(I_{\text{alg}}, I_{\text{ideal}}; R) &\stackrel{\text{def}}{=} \\ &\stackrel{\text{def}}{=} 20 \log_{10} [I_{\text{max}} / I_2(I_{\text{alg}}, I_{\text{ideal}}; R)], \end{aligned} \quad (10)$$

where  $I_{\text{max}}$  is the maximum possible image value (usually  $I_{\text{max}} = 255$ ); and in (Gong et al., 2019), the structural similarity index:

$$\begin{aligned} SSIM(I_{\text{alg}}, I_{\text{ideal}}; R) &\stackrel{\text{def}}{=} \\ &= \frac{(2\mu_a\mu_i + c_1)(2\sigma_{ai} + c_2)}{(\mu_a^2 + \mu_i^2 + c_1)(\sigma_a^2 + \sigma_i^2 + c_2)}, \end{aligned} \quad (11)$$

where

$$\begin{aligned} \mu_a &\stackrel{\text{def}}{=} \frac{1}{S(R)} \int_R I_{\text{alg}}(\mathbf{r}) d\mathbf{r}, \\ \sigma_a^2 &\stackrel{\text{def}}{=} \frac{1}{S(R)} \int_R (I_{\text{alg}}(\mathbf{r}) - \mu_a)^2 d\mathbf{r}, \\ \mu_i &\stackrel{\text{def}}{=} \frac{1}{S(R)} \int_R I_{\text{ideal}}(\mathbf{r}) d\mathbf{r}, \\ \sigma_i^2 &\stackrel{\text{def}}{=} \frac{1}{S(R)} \int_R (I_{\text{ideal}}(\mathbf{r}) - \mu_i)^2 d\mathbf{r}, \end{aligned} \quad (12)$$

$$\sigma_{ai} \stackrel{\text{def}}{=} \frac{1}{S(R)} \int_R (I_{\text{alg}}(\mathbf{r}) - \mu_a)(I_{\text{ideal}}(\mathbf{r}) - \mu_i) d\mathbf{r}, \quad (13)$$

$$\begin{aligned} c_1 &= (k_1 I_{\text{max}})^2, & c_2 &= (k_2 I_{\text{max}})^2, \\ k_1 &= 0.01, & k_2 &= 0.03. \end{aligned} \quad (14)$$

The Wasserstein metric (Haker et al., 2004; Schmitzer and Schnörr, 2015; Su et al., 2015), informally called the “earth mover’s distance”, is also used for single-channel images:

$$W_p(I_{\text{alg}}, I_{\text{ideal}}; Q, R) \stackrel{\text{def}}{=} \inf_{\gamma \in \Gamma(\mu, \nu)} \left[ \int_{Q \times R} \|\mathbf{q} - \mathbf{r}\|_2^p d\gamma(\mathbf{q}, \mathbf{r}) \right]^{1/p}, \quad (15)$$

where

$$\begin{aligned} \mu(X) &= \frac{\int_X I_{\text{alg}}(\mathbf{q}) d\mathbf{q}}{\int_Q I_{\text{alg}}(\mathbf{q}) d\mathbf{q}}, & X \subseteq Q, \\ \nu(X) &= \frac{\int_X I_{\text{ideal}}(\mathbf{r}) d\mathbf{r}}{\int_R I_{\text{ideal}}(\mathbf{r}) d\mathbf{r}}, & X \subseteq R \end{aligned} \quad (16)$$

are the distributions of values of images  $I_{\text{alg}}$  and  $I_{\text{ideal}}$ , and  $\Gamma(\mu, \nu)$  is the set of all measures on  $Q \times R$  with marginal measures  $\mu$  and  $\nu$ . The interpretation of  $W_p$  in this case can be detailed as follows. If the measures  $\mu$  and  $\nu$  are understood as “piles of dirt”, then



**Fig. 9.** Single-channel images  $I_{\text{ideal}}$ ,  $I_{\text{alg}}$  and their difference module  $|I_{\text{alg}} - I_{\text{ideal}}|$ . The images are reproduced from (Goshin et al., 2014).

the Wasserstein metric  $W_p$  defines the minimal “cost” of turning one pile into the other, the cost is assumed to be proportional to the amount of dirt and the distance (raised to the power  $p$ ) by which it has to be moved.

Radial distortion compensation is an important special case of geometric normalization. To describe its accuracy, criteria detecting the presence of straight lines in the image  $I_{\text{alg}}$  are used (Kunina et al., 2016).

The main property of color criteria for geometric normalization accuracy is as follows: even for fixed transformations  $\hat{H}$ ,  $H$  and the region of interest  $R$ , i.e. when in the geometric sense normalization is made equally accurately, the value of any color criterion changes depending on images  $I_{\text{alg}}$  and  $I_{\text{ideal}}$ .

#### Geometric criteria

By *geometric criteria* of the normalization accuracy, we mean the criteria that depend only on transforma-

tions  $H$  and  $\hat{H}$ , or on their parameters, and on the region of interest  $R$ , and on the derived objects:  $Q$ ,  $V$ ,  $d$ . They do not depend on the images ( $I_{\text{input}}$ ,  $I_{\text{alg}}$  and  $I_{\text{ideal}}$ ). Geometric accuracy criteria are natural for geometric normalization and are more frequently met in the literature. Let us list them below.

1. The closeness of the parameters specifying the transformations  $\hat{H}$  and  $H$ . For example, in (Triputen' and Gorokhovatskii, 1997), the accuracy of normalization by an affine transformation given by the matrix  $\hat{A} = (\hat{a}_{ij}) \in \mathbb{R}^{2 \times 3}$ , was calculated as follows:

$$E_{\text{affin}}(\hat{A}, A) = \sum_{i=1}^2 \sum_{j=1}^3 \frac{|\hat{a}_{ij} - a_{ij}|}{|a_{ij}|}, \quad (17)$$

and in (Calderon and Romero, 2007), where normalization was performed by projective transformation, the accuracy was introduced by the homography matrix  $\hat{H} = (\hat{h}_{ij}) \in \mathbb{R}^{3 \times 3}$ :

$$E_{\text{proj}}(\hat{H}, H) = \|\hat{H} - H\|_F, \quad (18)$$

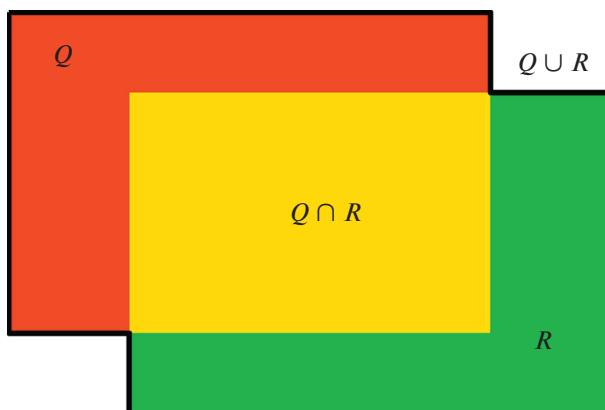
for the following homogeneity normalization:  $\hat{h}_{33} = h_{33} = 1$ , where  $\|\cdot\|_F$  is the Frobenius norm.

2. Jaccard's coefficient (Jaccard, 1901), equal to the area of intersection of sets  $Q$  and  $R$ , divided by the area of their union (see Fig. 10):

$$K_{\text{Jaccard}}(Q, R) \stackrel{\text{def}}{=} \frac{S(Q \cap R)}{S(Q \cup R)}. \quad (19)$$

It was used, for example, at the “Smartphone document capture competition” of the ICDAR conference (Zhukovskiy et al., 2018). The paper (Rezatofighi et al., 2019) suggests its more suitable modification for optimization.

3. Hausdorff metric, the greatest distance from the points of one set to their corresponding nearest points of the second set (see Fig. 11):



**Fig. 10.** Illustration of Jaccard's coefficient definition: intersection and union of sets  $Q$  and  $R$ .

$$d_H(Q, R) \stackrel{\text{def}}{=} \max \left\{ \sup_{q \in Q} \inf_{r \in R} \|q - r\|_2, \sup_{r \in R} \inf_{q \in Q} \|r - q\|_2 \right\}. \quad (20)$$

The Hausdorff metric has been used for arbitrary object detection (Sim et al., 1999), for alignment of partially occluded contours (Orrite and Herrero, 2004), for robust face detection (Jesorsky et al., 2001), and for calculation of closeness between two images (Huttenlocher et al., 1993). In the works (Dubuisson and Jain, 1994; Efimov and Novikov, 2016), its modifications were proposed.

4. In the case when instead of sets  $Q$  and  $R$ , two continuous curves are considered:  $Q, R: [0,1] \rightarrow \mathbb{R}^2$ , the Frechette distance, related to the Hausdorff metric, is used:

$$F(Q, R) \stackrel{\text{def}}{=} \inf_{a,b} \max_{t \in [0,1]} \|Q(a(t)), R(b(t))\|_2, \quad (21)$$

where  $a, b: [0,1] \rightarrow [0,1]$  are continuous non-decreasing surjections (reparametrization). This criterion was used to specify morphing accuracy (Har-Peled, 2002) and closeness of two contours after projective alignment (Pritula et al., 2015).

5. Mean square coordinate discrepancy  $d$ :

$$L_2(V; R) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{S(R)} \int_R d^2(r) dr & \text{for } 0 < S(R) < \infty \\ \sqrt{\frac{1}{|R|} \sum_{r \in R} d^2(r)} & \text{for } 0 < |R| < \infty \end{cases}, \quad (22)$$

has been used as a criterion of normalization accuracy in radial distortion elimination (Stein, 1997), panorama creation (Sawhney and Kumar, 1999; Hsu and Sawhney, 1998; Chen et al., 2002), space image matching (Kozlov et al., 2009; Katamanov, 2007), medical image analysis (Baltzopoulos, 1995), text recognition (Dance, 2001).

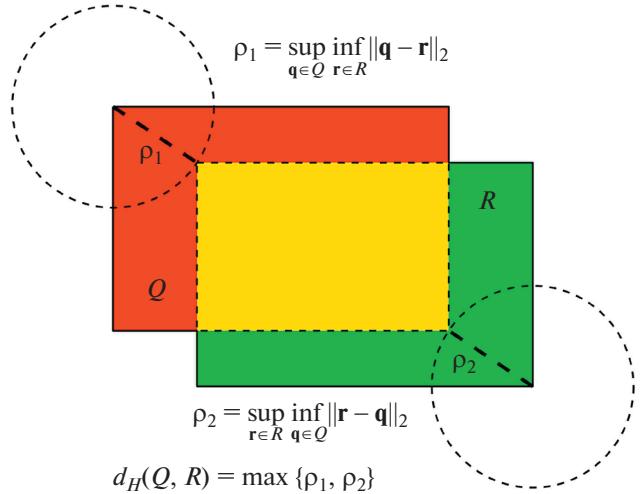
6. The mean coordinate discrepancy  $d$  was also employed (Kunina et al., 2016; Shemiakina et al., 2017):

$$L_1(V; R) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{S(R)} \int_R d(r) dr & \text{for } 0 < S(R) < \infty, \\ \frac{1}{|R|} \sum_{r \in R} d(r) & \text{for } 0 < |R| < \infty. \end{cases} \quad (23)$$

7. Finally, the maximum coordinate discrepancy  $d$  (minimax criterion) was also proposed as an accuracy criterion. In the general case, it is defined as the following supremum:

$$L_\infty(V; R) \stackrel{\text{def}}{=} \sup_{r \in R} d(r). \quad (24)$$

The maximum coordinate discrepancy was used as a criterion for the normalization accuracy in the tasks of space image georeferencing (Katamanov, 2007), face detection (Jesorsky et al., 2001), and text recogni-



**Fig. 11.** The Hausdorff metric  $d_H(Q, R)$  between the sets  $Q$  and  $R$ .

tion (Shemiakina et al., 2017; Skoryukina et al., 2018). In the case of projective normalization, some works (Shemiakina et al., 2017; Skoryukina et al., 2018) applied this criterion in a non-standard way. Instead of the entire region of interest  $R$ , it used only the *extreme points* of the region's convex hull  $E(\text{Conv}(R))$ :

$$\hat{L}_\infty(V; R) \stackrel{\text{def}}{=} \sup_{E(\text{Conv}(R))} d. \quad (25)$$

The authors assumed that equality  $\hat{L}_\infty(V; R) = L_\infty(V; R)$  is true. However, this assumption is not always correct. Let us consider a counterexample to the declared equality (see Fig. 12). Let

$$V(r) = \frac{1}{-2x + 10} \begin{bmatrix} 2x \\ -x + 2y + 4 \end{bmatrix}, \quad (26)$$

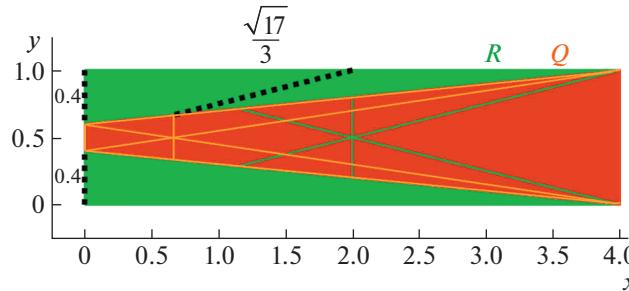
$$d(r) = \|r - V(r)\|_2,$$

and the region of interest is a rectangle  $R = [0, 4] \times [0, 1]$ , then:  $E(\text{Conv}(R)) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ , therefore:

$$\begin{aligned} \hat{L}_\infty(V; R) &= \sup_{E(\text{Conv}(R))} d = 0.4 < \frac{\sqrt{17}}{3} = d \left( \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in R \right) \leq \\ &\leq \sup_R d = L_\infty(V; R) \Rightarrow \hat{L}_\infty(V; R) < L_\infty(V; R). \end{aligned} \quad (27)$$

It is interesting to note that all of the listed geometric accuracy criteria depend on  $\hat{H}$  and  $H$  only through the residual distortion  $V = \hat{H}H^{-1}$  (4) and the coordinate discrepancy  $d$  derived from it (6). Therefore, the accuracy of the transformation  $H$  estimate  $\hat{H}$  can be understood as the closeness of the residual distortion  $V$  to the identity transformation.

Let us now examine the listed geometric criteria. The criteria based on proximity of parameters that define transformations  $\hat{H}$  and  $H$  are not suitable to de-



**Fig. 12.** A counterexample to the statement that the supremum of coordinate discrepancy of the projective transformation on a rectangle is achieved on its vertices. The projective transformation of a rectangle into a trapezoid is shown. The lengths of the dashed lines correspond to the coordinate discrepancies. The property is visually demonstrated: the coordinate discrepancy at vertices of the rectangle is smaller than at the point on its edge.

scribe the normalization accuracy, since they do not depend on the region of interest  $R$ , which shows where exactly in the image  $I_{\text{ideal}}$  plane the normalization should be accurate. The Jaccard coefficient, the Hausdorff metric, and the Frechet distance do not have this drawback, but, being utilized as criteria for normalization accuracy, they have another general drawback. They specify only the similarity of sets  $Q$  and  $R$ , while arbitrary distortions within a set are not accounted for. For example, in Fig. 13 we can see two examples of document image normalization, correct and incorrect, which have an ideal accuracy in terms of each of the three above-mentioned criteria.

Mean, root mean square and maximum coordinate discrepancies have no obvious disadvantages compared to the previous criteria.

## PROJECTIVE IMAGE NORMALIZATION AND ACCURACY CRITERIA

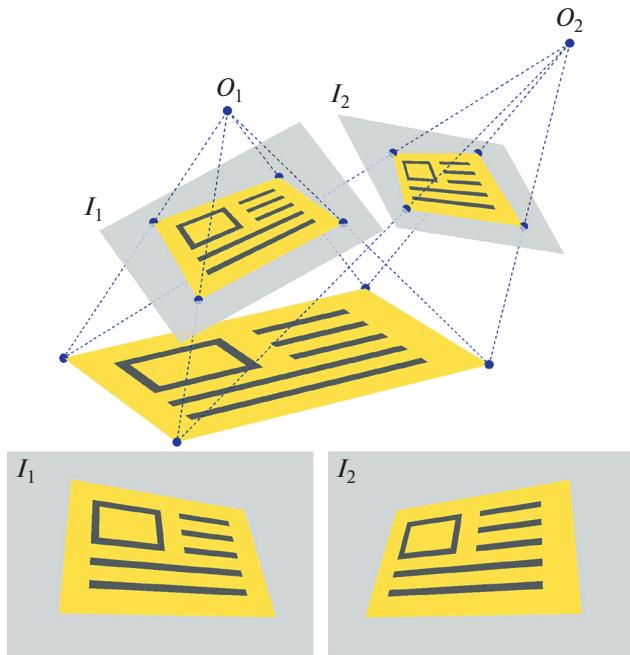
### Theoretical justification of projective image normalization

The following classes of transformations are used to implement geometric normalization of images: isometric (Murygin, 2010; Huttenlocher et al., 1993; Bolotova et al., 2017), affine (Triputen' and Gorokhovatskii, 1997; Putyatkin et al., 1998; Nikolaidis, 2011), polynomial (Kozlov et al., 2009), fractional-polynomial (Singh et al., 2008), radial-polynomial (to compensate for radial distortion) (Kunina et al., 2016), central-projective (Rodríguez-Piñeiro et al., 2011; Zhang and He, 2007; Kholopov, 2017), projective (Safari et al., 1997; Iwamura et al., 2007; Merino-Gracia et al., 2013; Shemiakina et al., 2017; Xie et al., 2018), and arbitrary (Jesorsky et al., 2001; Zeynalov et al., 2009).

The shape of scene objects is very often modeled by a polyhedron (approximation of their 3D shape is called *polyhedral*), formally approximating the optical system of its registration by a *pinhole camera* (the laws of geometric optics are approximated in this case by the *flat central projection*) (Forsyth, Ponce, 2002). Under such assumptions, the images of the same object face captured from arbitrary angles are connected by a two-dimensional *projective transformation* (Shemiakina, 2017; Hartley, Zisserman, 2003) (see Fig. 14 and 15). Thus, the ideal normalizing transformation  $H$  of this face image is projective, and the algorithmically normalizing transformation  $\hat{H}$  is chosen to be projective. Since the transformations  $H$  and  $\hat{H}$  are projective, the residual distortion  $V = \hat{H}H^{-1}$  is also projective. Such geometric normalization will be referred to as *projective normalization*. The projective transformation preserves straight lines. In Fig. 14 and 15, this funda-



**Fig. 13.** Two examples of document image normalization (left – correct, right – incorrect), which are perfectly accurate in terms of Jaccard coefficient, Hausdorff metric, and Frechette distance.



**Fig. 14.** Formation of images  $I_1$  and  $I_2$  of a flat rectangular object in a pinhole camera with optical centers  $O_1$  and  $O_2$ . The object and its images are pairwise connected by projective transformations.

mental property is expressed by the fact that the images of straight lines of the scene also remain straight.

The projective transformation is used not only for geometric normalization of images, but also for color normalization. In 1931, the International Commission on Illumination (CIE) introduced the standard CIE XYZ color space and formalized the concept of *chromaticity* (color without luminance) as an ordered pair of *chromatic coordinates* ( $x, y$ ) of the CIE xy color space (Smith et al., 1931), with the relationship between color and chromaticity specified as a projective mapping. Projective CIE xy chromaticity transformations have been considered since MacAdam's work (1937) published in 1937. However, due to their projective definition, we can assume that the projective chromaticity transformations have been considered since their introduction in 1931. Nevertheless, a formal theoretical justification for projective color normalization was published only in 2016 by G. Finlayson (Finlayson et al., 2016). It is worth noting several works where all three color coordinates are projectively transformed. The first appears to be a paper (Wallace et al., 2003) by a Princeton University Computer Science Department team, published in 2003. A three-dimensional projective transformation was used to compare the color bodies of different projection frameworks. Later, the same approach was applied to photorealistic color palette transfer between images (Gong et al., 2019) (see Fig. 16 for examples). In both cases, it is a mutual calibration of the two images,



**Fig. 15.** The real image obtained by the camera, which can be accurately simulated by a pinhole camera model. The flat faces of the scene are related to their images by projective transformations.

rather than a transition to a color space with the required properties. In the article (Smagina et al., 2019), the projective transformation was suggested to be used precisely for the transition to space with a simple, but meaningful, metric. In this work, it was shown that a fixed three-dimensional projective transformation of color coordinates can improve the results of color segmentation algorithms. In 2020, another paper (Kim et al., 2020) demonstrated the use of a three-dimensional projective transformation for color normalization (calibration) of micro-LED displays.

#### Accuracy criteria for projective image normalization

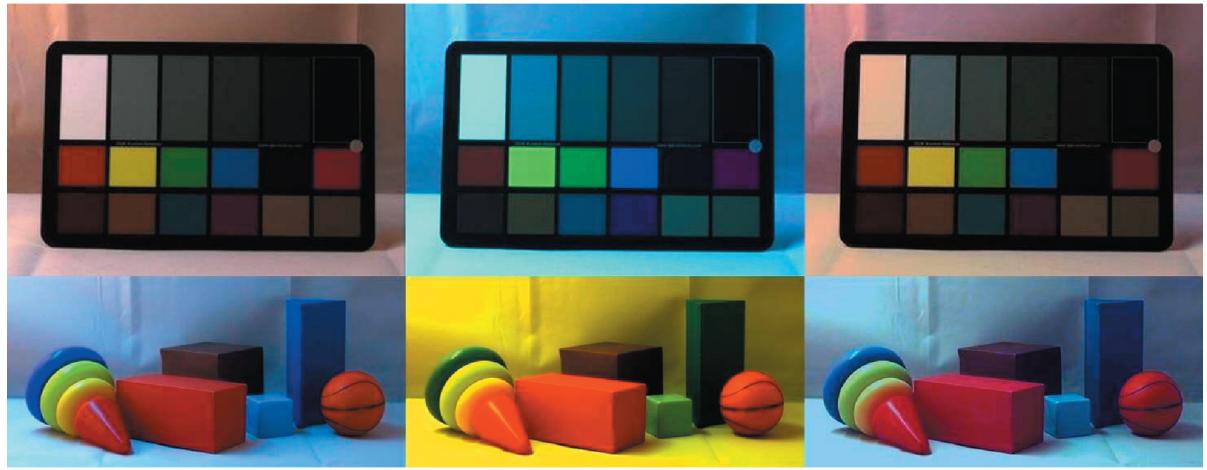
All the criteria described above are universal, and therefore applicable to the description of the projective normalization accuracy. However, for the case of projective normalization, special criteria have been proposed in the literature. They all assume that the region of interest  $R$  is a rectangle. Then its image  $Q$  is a quadrilateral. For example, in (Calore et al., 2012), the angle between the left and right sides of the quadrilateral  $Q$  was proposed as an accuracy criterion:  $\alpha(Q)$  (see Fig. 17), and in (Khlopov, 2017), the ratio between the minimum and maximum angles of the quadrilateral  $Q$  (see Fig. 18):

$$E(Q) = \frac{\alpha_{\min}(Q)}{\alpha_{\max}(Q)}, \quad (28)$$

in (Takezawa et al., 2016), the total relative proximity of the lengths of the opposite sides of  $Q$  was chosen:

$$D(Q) = \frac{|a - c|}{a + c} + \frac{|b - d|}{b + d}. \quad (29)$$

The articles (Rodríguez-Piñeiro et al., 2011; Zhang and He, 2007) define normalization accuracy as the accuracy of the normalization algorithm's estimate of the aspect ratio of a rectangle  $R$ . All the criteria listed here are invariant to the similarity transformation.



**Fig. 16.** Projective color normalization. The left images were obtained under the illumination chosen as normal, the center images were obtained under some other types of illumination. On the right, the results of projective color normalization of the central images are shown.

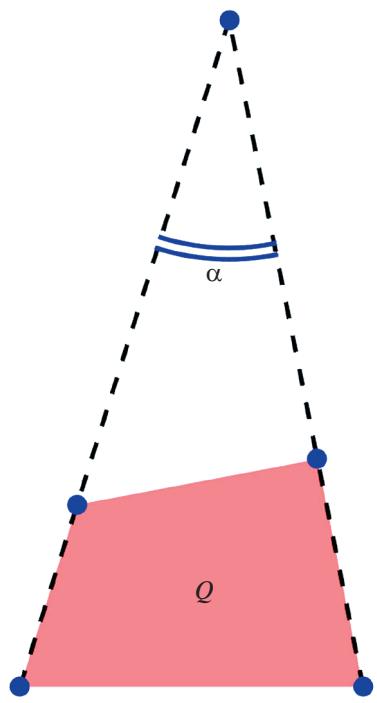
#### THEORETICAL JUSTIFICATION OF THE NORMALIZATION ACCURACY CRITERIA

It was shown above that a large number of normalization accuracy criteria have been already proposed, including those introduced specifically for projective normalization. Thus, the question of the proposed criteria applicability to various image analysis problems is relevant. The introduction of problem-oriented criteria allows for a theoretically justified approach to the

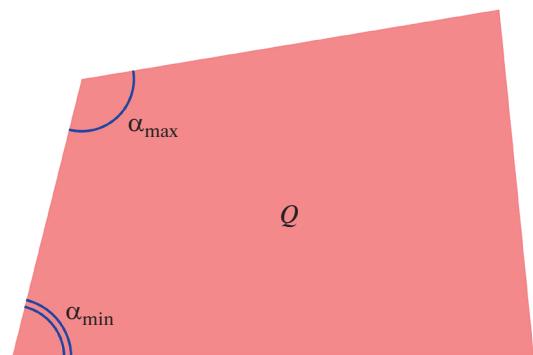
choice of optimal algorithm among existing algorithms of geometric normalization, and to the development of new algorithms.

In (Konovalenko et al., 2020d), for the case of a document with a fixed structure (Povolotskiy and Tropin, 2019), a normal probabilistic recognition model was introduced, according to which the probability of correct recognition of a symbol jumps to zero as the coordinate discrepancy of this symbol increases. For this model, it was proved that the accuracy criterion of image normalization, expressed as the maximum (by text fields of the document) coordinate discrepancy, is monotonically related to the probability of correct recognition of the entire document.

In (Konovalenko et al., 2020a), another model of recognition was introduced, according to which the probability of correct recognition of a symbol decays according to Gaussian with the growth of the coordinate discrepancy of this symbol. For this model, it was proved that the accuracy criterion of image normalization, equal to the mean square of the coordinate mis-



**Fig. 17.** The angle between the left and right sides of a quadrilateral  $Q$ .



**Fig. 18.** Minimum and maximum angles of a quadrilateral  $Q$ .



**Fig. 19.** An example of projective normalization in which the Jacobian of residual distortion  $V$  equals to either positive or negative values in the region of interest  $R$ . Above – ideally normalized image  $I_{\text{ideal}}$ . Bottom – algorithmically normalized image  $I_{\text{alg}}$ , obtained as a result of splitting the image  $I_{\text{ideal}}$  by residual projective distortion into two parts: for the right side the Jacobian of residual distortion is positive, and for the left side it is negative, which corresponds to “reflecting” the image.

alignment in the text fields of the document, is monotonically related to the probability of true recognition of the entire document.

In (Konovalenko and Shemiakina, 2018; Konovalenko et al., 2020a), the maximum and root mean square coordinate discrepancies were expressed analytically for the case of projective normalization.

#### LIMITS OF APPLICABILITY OF NORMALIZATION ACCURACY CRITERIA

##### *The case of arbitrary normalization*

All the above mentioned criteria of images normalization accuracy, except intrasystem, are not intended for the case when some part of the region of interest image is not mapped on the image  $I_{\text{alg}}$ . In this regard, it is necessary that the algorithmically normalized image completely contains the image of the region of interest:

$$Q = V[R] \subseteq \text{dom } I_{\text{alg}}, \quad (30)$$

and that the input image completely contains the pre-image of the region of interest:

$$H^{-1}[R] \subseteq \text{dom } I_{\text{input}}. \quad (31)$$

In addition, a residual distortion that is too “bad” at the region of interest also makes all of the above non-intrasystem accuracy criteria meaningless. Therefore, let us require that at each point the transformation has a Jacobian matrix and a positive Jacobian

$$\mathbf{r} \in R \Rightarrow \det(J(\mathbf{r})) > 0, \quad (32)$$

because a zero Jacobian usually leads to loss of information, and a negative one means “reflection” of the image and corresponds to imitation of the view of the object surface “from the back” (see Fig. 19).

##### *The case of projective normalization*

Let us consider in detail what these constraints mean for the case of projective normalization, when the camera is modeled by a pinhole camera and the object surfaces are modeled by planes. Let us introduce the projective transformation  $P$ , the inverse of  $H$ :

$$P = H^{-1} \quad (33)$$

– it translates the points of the image  $I_{\text{ideal}}$  into the points of the image  $I_{\text{input}}$ . The projective transformation  $V$  is parameterized by a homogeneous matrix (*homography matrix*)  $V \stackrel{\text{def}}{=} (v_{ij}) \in \mathbb{R}^{3 \times 3}$  in the following standard way:

$$V(\mathbf{r}) \stackrel{\text{def}}{=} \begin{bmatrix} v_{11}x + v_{12}y + v_{13} \\ v_{21}x + v_{22}y + v_{23} \\ v_{31}x + v_{32}y + v_{33} \end{bmatrix}. \quad (34)$$

Similarly, the projective transformations  $\hat{H}$ ,  $H$  and  $P$  are parameterized by the matrices  $\hat{H}$ ,  $H$  and  $p$  respectively. Let

$$V = \hat{H}H^{-1}, \quad P = H^{-1} \Rightarrow V = \hat{H}P. \quad (35)$$

Let us introduce the function

$$Z(\mathbf{r}) \stackrel{\text{def}}{=} v_{31}x + v_{32}y + v_{33}. \quad (36)$$

From the constraints (30), (31), and (32) it follows that the function  $Z$  has a constant sign on the set  $R$ :

$$\begin{cases} \mathbf{r} \in R \Rightarrow Z(\mathbf{r}) < 0, \\ \mathbf{r} \in R \Rightarrow Z(\mathbf{r}) > 0. \end{cases} \quad (37)$$

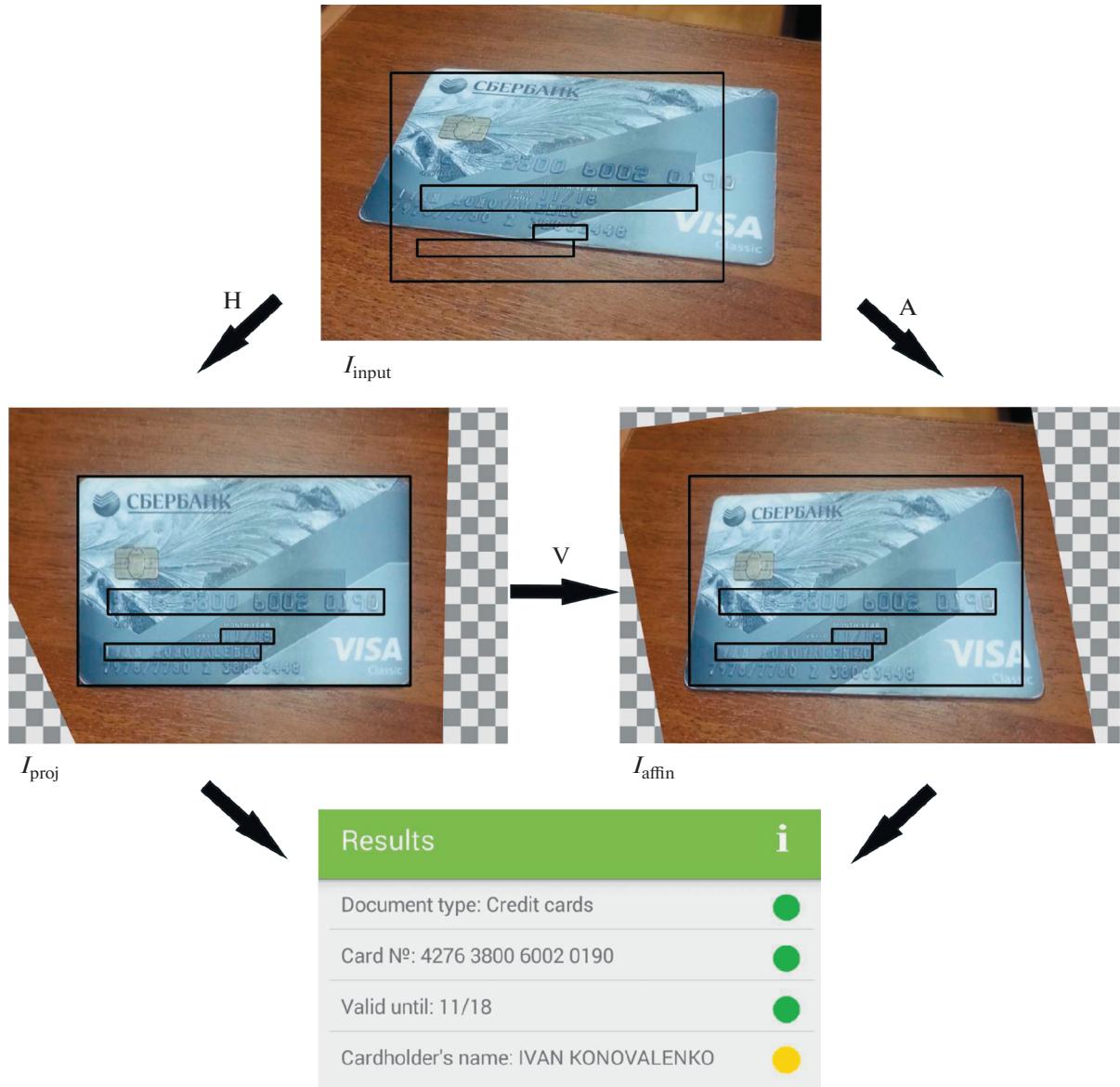
Let us prove this for the general case when the set  $R$  is not necessarily connected. Indeed, if for a point  $\mathbf{r} \in R$  is satisfied  $Z(\mathbf{r}) = 0$ , then the corresponding point  $V(\mathbf{r})$  in the image  $I_{\text{alg}}$  plane is infinitely distant (see (34)) and therefore cannot belong to this image, which violates the condition (30). Moreover, it follows from (34) and (36) that:

$$\det(J(\mathbf{r})) = \frac{V_{11}V_{22}V_{33} - V_{11}V_{23}V_{32} - V_{12}V_{21}V_{33} + V_{12}V_{23}V_{31} + V_{13}V_{21}V_{32} - V_{13}V_{22}V_{31}}{Z^3(\mathbf{r})}. \quad (38)$$

I.e. the sign  $Z(\mathbf{r})$  either everywhere coincides with the sign of the Jacobian of the projective transformation  $V$ , or differs from it everywhere, which means that to satisfy condition (32) it has to be either uniformly on  $R$  negative or positive, what was required to prove.

If condition (37) is ignored, the coordinate discrepancy (6) is predetermined naturally:

$$d(\mathbf{r}) \stackrel{\text{def}}{=} \begin{cases} \|\mathbf{r} - V(\mathbf{r})\|_2 & \text{for } Z(\mathbf{r}) \neq 0, \\ +\infty & \text{for } Z(\mathbf{r}) = 0. \end{cases} \quad (39)$$



**Рис. 20.** Projective (left) and affine (right) geometric normalizations of the document image  $I_{\text{input}}$  (top) and the result of its recognition (bottom). H – projective transformation, A – affine transformation,  $V = AH^{-1}$  – residual projective distortion. The black boxes show the ideal localization of a document and its text fields. Even though the camera's optical from normal, the text fields of the document were accurately normalized via affine normalization.

## AFFINE APPROXIMATION OF PROJECTIVE IMAGE NORMALIZATION

Due to the growing technical capabilities of mobile devices in recent years, autonomous image analysis on mobile devices without the involvement of the server has become relevant. A significant contribution to the development of this approach was made by V.V. Arlazarov. The computational power of modern mobile devices is such that the time required for projective image transformation turns out to be a critical factor (Trusov, Limonova, 2020). A competitive approach that increases the speed of image processing could be the affine transformation (Putyatkin et al., 1998; Wolberg,

1990). Note that a typical variant of the orientation of the camera optical axis with respect to the plane of the target object can be represented by an *orthogonal view model*. In this approximation, the camera-object system is naturally described by an *affine projection model* (Forsyth and Ponce, 2002), and the generally required projective normalization is replaced by frequently used *affine normalization* without significant loss of accuracy (Triputen' and Gorokhovatskii, 1997; Putyatkin et al., 1998; Nikolaidis, 2011) (see Fig. 20). Such a transition to a less computationally expensive model can provide the required acceleration of the normalization step.

The notion that in practice the projective transformation can be replaced by the affine transformation was suggested earlier in (Gruen, 1985). This property was used in (Ohta et al., 1981) to simplify further mathematical constructions. The affine approximation is widely used in image augmentation (Pavić et al., 2006) and in rendering (Wolberg, 1990; Heckbert, 1989; Lorenz and Döllner, 2009). In (Huang et al., 2015), the projective transformation is replaced by a simpler affine transformation in order to eliminate overtraining. A similar idea is used in the weak perspective camera model (Alter, 1992; Kutulakos and Vallino, 1996; Aradhye and Myers, 2010), where within each scene object the equidistance of its parts from the camera is assumed. Developing affine invariant methods instead of the much more complex projective invariant ones is common in the popular technology of special points (Mikolajczyk and Schmid, 2002; Mikolajczyk and Schmid, 2004; Morel and Yu, 2009), and in the related problem of significant regions detection (Kadir et al., 2004), although both approaches are practically invariant to the capturing perspective. The division into affine and projective methods also exists in the field of stereoreconstruction (Faugeras, 1995). Replacing a projective transformation with an affine one for rendering and image normalization purposes leads to the loss of accuracy (Putyatin et al., 1998; Zwicker et al., 2004).

In (Konovalenko et al., 2019), the maximum and root mean square coordinate discrepancies were proposed as accuracy criteria for affine approximation of projective normalization. Based on these criteria, the problems of the search for optimal affine approximations are formulated. The convexity of the obtained optimization problems is proved. A method for the employment of optimal affine approximations to save computational resources during image transformation is proposed. In (Konovalenko et al., 2021), the problem of finding an affine approximation optimal according to the criterion of coordinate mean square discrepancy was solved analytically.

## CONCLUSION

In the analytical part of this review paper, the authors demonstrated the following original results:

1. Among the known criteria for the accuracy of image normalization, several were chosen based on the following attributes. Ones that do not violate the software modularity principle, do not depend on the values of images, do take into account the area of interest, and do not require an ideal accuracy in the case of loose normalization. Such criteria include mean, root mean square, and maximum coordinate discrepancies (see section Theoretical justification of the normalization accuracy criteria).

2. For the case of projective normalization, an analytical expression for the maximum coordinate dis-

crepancy was previously described in the literature. This expression, however, was refuted by the authors: it is shown that the supremum of the coordinate discrepancy of 2D projective transformation on a closed bounded set is not necessarily reached at the extreme points of its convex hull (see item 7 in Geometrical Criteria section).

3. Limits of applicability of normalization accuracy criteria were introduced.

4. We proposed mean square and maximum coordinate discrepancies as accuracy criteria of affine approximation of projective normalization. It significantly reduces computational complexity when choosing optimal affine approximations, and the problem of affine approximation search has been studied analytically (Konovalenko et al., 2021).

## FUNDING

The work was partially financially supported by RFBR in the context of scientific projects №№ 18-29-26035 and 19-29-09075.

## CONFLICT OF INTEREST

All other authors report no conflicts of interest relevant to this article.

## REFERENCES

- Abramov M.P., Shipitko O.C., Grigoryev A.S., Ershov E.I. Poisk tochki skhoda dlya dinamicheskoi kalibrovki vneshnikh parametrov monokulyarnoi kamery pri uslovii pryamolineinogo dvizheniya [Vanishing point detection for monocular camera extrinsic calibration under translation movement] *Sensornye sistemy* [Sensory systems]. 2020. V. 34 (1). P. 32–43 (in Russian). <https://doi.org/10.31857/S0235009220010023>
- Abramov M.P., Shipitko O.S., Lukyanov A.S., Panfilova E.I., Kunina I.A., Grigoryev A.S. Sistema pozitsionirovaniya vnutri zdanii mobil'noi robototekhnicheskoi platformy na osnove detektsii kraev [Edge detection based mobile robot indoor localization] *Sensornye sistemy* [Sensory systems]. 2019. V. 33 (1). P. 30–43 (in Russian). <https://doi.org/10.1134/S0235009219010025>
- Abulkhanov D.A., Sidorchuk D.S., Konovalenko I.A. Obuchenie neirosetevykh deskriptorov osobykh tochek dlya sopostavleniya radiolokatsionnykh i opticheskikh izobrazhenii [Neural network-based feature point descriptors for registration of optical and SAR images] *Sensornye sistemy* [Sensory Systems]. 2018. V. 32 (3). P. 222–229 (in Russian). <https://doi.org/10.1134/S0235009218030034>
- Alter T.D. *3d pose from 3 corresponding points under weak-perspective projection*. Technical report. Massachusetts Inst Of Technology Artificial Intelligence Lab. 1992.
- Aradhye H., Myers G.K. *Method and apparatus for recognition of symbols in images of three-dimensional scenes*. US Patent. No. 7.738.706. 2010.

- Arvind C.S., Mishra R., Vishal K., Gundimeda V. Vision based speed breaker detection for autonomous vehicle. *Tenth International Conference on Machine Vision (ICMV 2017). International Society for Optics and Photonics.* 2018. V. 106960E. P. 1–9.
- Awal A.M., Ghanmi N., Sicre R., Furon T. Complex document classification and localization application on identity document images. *IAPR 2017-International Conference on Document Analysis and Recognition.* 2017. P. 427–431.  
<https://doi.org/10.1109/ICDAR.2017.77>
- Balitskiy A.M., Savchik A.V., Gafarov R.F., Konovalenko I.A. O proektivno invariantnykh tochkakh ovala s vydelennoy vneshey pryamoy [On projectively invariant points of an oval with a distinguished exterior line] *Problemy peredachi informatsii* [Problems of Information Transmission]. 2017. V. 53 (3). P. 84–89 (in Russian).
- Baltzopoulos V. A video fluoroscopy method for optical distortion correction and measurement of knee-joint kinematics. *Clinical Biomechanics.* 1995. V. 10 (2). P. 85–92.
- Bezmaternykh P.V., Ilin D.A., Nikolaev D.P. U-Net-bin: hacking the document image binarization contest. *Computer Optics.* 2019. V. 43 (5). P. 825–832.  
<https://doi.org/10.18287/2412-6179-2019-43-5-825-832>
- Bezmaternykh P.V., Nikolaev D.P. A document skew detection method using fast Hough transform. *Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019).* 2020. V. 11433. P. 114330J.  
<https://doi.org/10.1117/12.2559069>
- Bezmaternykh P.V., Nikolaev D.P., Arlazarov V.L. Textual Blocks Rectification Method Based on Fast Hough Transform Analysis in Identity Documents Recognition. *Tenth International Conference on Machine Vision (ICMV 2017). International Society for Optics and Photonics.* 2018. V. 10696. P. 1069606.  
<https://doi.org/10.1117/12.2310162>
- Bezmaternykh P.V., Vylegzhannin D.V., Gladilin S.A., Nikolaev D.P. Generativnoe raspoznavanie dvumernykh shtrikhkodov. *Iskusstvennyi intellekt i priyatiye reshenii* [Scientific and Technical Information Processing]. 2010. V. 2010 (4). P. 63–69 (in Russian).
- Bolotova Yu.A., Spitsyn V.G., Osina P.M. Obzor algoritmov detektsii tekstovykh oblastei na izobrazheniyakh i videozapisyakh [Review Of Algorithms For Text Detection In Images And Videos] *Komp'yuternaya optika* [Computer Optics]. 2017. V. 41 (3). P. 441–452 (in Russian).
- Bulatov K., Matalov D., Arlazarov V.V. MIDV-2019: Challenges of the Modern Mobile-Based Document OCR. *Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019).* 2020. V. 11433. P. 114332N.  
<https://doi.org/10.1117/12.2558438>
- Calderon F., Romero L. An accurate image registration method using a projective transformation model. *Eighth Mexican International Conference on Current Trends in Computer Science (ENC2007).* IEEE. 2007 P. 58–64.
- Calore E., Pedersini F., Frosio I. Accelerometer based horizon and keystone perspective correction. *Instrumentation and Measurement Technology Conference (I2MTC). 2012 IEEE International.* IEEE. 2012. P. 205–209.
- Cheklov D.O., Ablameiko S.V. Normalizatsiya izobrazhenii otnositel'no perspektivnogo preobrazovaniya na osnove geometricheskikh parametrov [Normalization of images relating to perspective transformation based on geometric options] *Informatika* [Informatics]. 2004. V. (3). P. 67–76 (in Russian).
- Chen H., Sukthankar R., Wallace G., Li K. Scalable alignment of large-format multi-projector displays using camera homography trees. *Proc. Conf. Visualization'02. IEEE Computer Society.* 2002. P. 339–346.
- Chernov T.S., Ilin D.A., Bezmaternykh P.V., Faradzhev I.A., Karpenko S.M. Research of Segmentation Methods for Images of Document Textual Blocks Based on the Structural Analysis and Machine Learning. *Vestnik RFFI.* 2016. I. 4. P. 55–71. (in Russian).  
<https://doi.org/10.22204/2410-4639-2016-092-04-55-71>
- Chochia P.A. Vosstanovlenie amplitudnykh kharakteristik monokhromnykh i mul'tispektral'nykh izobrazhenii, ispol'zuya funktsiyu gradientov [Recovering of the Amplitude Characteristics of Monochrome and Multispectral Images Using the Function of Gradients] *Informatsionnye protsessy.* 2016. V. 16 (2). P. 112–120. (in Russian).
- Chukalina M., Ingacheva A., Buzmakov A., Polyakov I., Gladkov A., Yakimchuk I., Nikolaev D. Automatic beam hardening correction for CT reconstruction. *Proc. ECMS 2017, European Council for Modeling and Simulation 2017.* 2017. P. 270–275.  
<https://doi.org/10.7148/2017-0270>
- Clark A.J., Green R.D., Grant R.N. Perspective correction for improved visual registration using natural features. *Image and Vision Computing New Zealand. 2008. IVCNZ 2008. 23rd International Conference.* 2008. P. 1–6.
- Dance C.R. Perspective estimation for document images. *Document Recognition and Retrieval IX. International Society for Optics and Photonics.* 2001. V. 4670. P. 244–255.
- Das P., Baslamisli A.S., Liu Y., Karaoglu S., Gevers T. Color constancy by GANs: an experimental survey. 2018. arXiv preprint arXiv:1812.03085.
- Dubuisson M.P., Jain A.K. A modified hausdorff distance for object matching. *Proc. of 12th international conference on patternrecognition.* IEEE. 1994. V. 1. P. 566–568.
- Efimov A.I., Novikov A.I. Algoritm poetapnogo utocheniya proektivnogo preobrazovaniya dlya sovmeshcheniya izobrazhenii [An algorithm for multistage projective transformation adjustment for image superimposition] *Komp'yuternaya Optika* [Computer Optics]. 2016. V. 40 (2). P. 258–265 (in Russian).  
<https://doi.org/10.18287/2412-6179-2016-40-2-258-265>
- Ershov E., Savchik A., Semenkov I., Banic N., Belokopytov A., Senshina D., Koscevic K., Subasi M., Loncaric S. The Cube++ Illumination Estimation Dataset. *IEEE Access.* 2020. V. 8. P. 227511–227527.  
<https://doi.org/10.1109/ACCESS.2020.3045066>
- Ershov E.I., Korchagin S.A., Kokhan V.V., Bezmaternykh P.V. A generalization of Otsu method for linear separation of two unbalanced classes in document image binarization. *Computer Optics.* 2021. V. 45 (1).

- P. 66–76.  
<https://doi.org/10.18287/2412-6179-CO-752>
- Faugeras O.D. What can be seen in three dimensions with an uncalibrated stereo rig? *European conference on computer vision. Springer.* 1992. P. 563–578.
- Finlayson G.D., Funt B.V., Barnard K. Color constancy under varying illumination. *Proc. of IEEE International Conference on Computer Vision.* 1995. P. 720–725.
- Finlayson G.D., Gong H., Fisher R.B. Color homography color correction. *Color and Imaging Conference.* 2016. V. 2016 (1). P. 310–314.
- Finlayson G.D., Schiele B., Crowley J.L. Comprehensive colour image normalization. *European conference on computer vision.* 1998. P. 475–490.
- Forsyth D.A., Ponce J. *Computer vision: a modern approach.* Prentice Hall Professional Technical Reference, 2002. 720 p.
- Gayer A.V., Sheshkus A.V., Nikolaev D.P., Arlazarov V.V. Improvement of U-Net Architecture for Image Binarization with Activation Functions Replacement. *Thirteenth International Conference on Machine Vision (ICMV 2020).* 2021. V. 11605. P. 116050Y.  
<https://doi.org/10.1117/12.2587027>
- Gijsenij A., Gevers T., Van De Weijer J. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing.* 2011. V. 20 (9). P. 2475–2489.
- Gladkov A.P., Kuznetsova E.G., Gladilin S.A., Gracheva M.A. Adaptivnaya stabilizatsiya yarkosti izobrazheniya v tekhnicheskoi sisteme raspoznavaniya krupnykh dvizhushchikhsya ob'ektorov [Adaptive image brightness stabilization for the industrial system of large moving object recognition]. *Sensornye sistemy* [Sensory systems]. 2017. V. 31 (3). P. 247–260 (in Russian).  
<https://doi.org/10.31857/S0235009220010047>
- Gong H., Finlayson G.D., Fisher R.B., Fang F. 3D color homography model for photo-realistic color transfer re-coding. *The Visual Computer.* 2019. V. 35 (3). P. 323–333.
- Goshin Y.V., Kotov A.P., Fursov V.A. Dvukhetapnoe formirovanie prostranstvennogo preobrazovaniya dlya sovmeshcheniya izobrazhenii [Two-stage formation of a spatial transformation for image matching]. *Kompyuternaya optika* [Computer Optics]. 2014. V. 38 (4). P. 886–891 (in Russian).
- Goshtasby A.A. *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications.* John Wiley & Sons, 2005. 280 p.
- Gruen A. Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography.* 1985. V. 14 (3). P. 175–187.
- Haker S., Zhu L., Tannenbaum A., Angenent S. Optimal mass transport for registration and warping. *International Journal of computer vision.* 2004. V. 60 (3). P. 225–240.
- Har-Peled S. New similarity measures between polylines with applications to morphing and polygon sweeping. *Discrete & Computational Geometry.* 2002. V. 28 (4). P. 535–569.
- Hartley R., Zisserman A. *Multiple view geometry in computer vision.* Cambridge, England, Cambridge university press. 2003. 655 p.
- Healey G. Using color for geometry-insensitive segmentation. *JOSAA.* 1989. V. 6 (6). P. 920–937.
- Heckbert P.S. Fundamentals of texture mapping and image warping. *University of California, Berkeley.* 1989. V. 2 (3). P. 1–86.
- Hsu S.C., Sawhney H.S. Influence of global constraints and lens distortion on pose and appearance recovery from a purely rotating camera. *Applications of Computer Vision. 1998. WACV'98. Proc. of the Fourth IEEE Workshop on.* IEEE. 1998. P. 154–159.
- Huang J.B., Singh A., Ahuja N. Single image super-resolution from transformed self-exemplars. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition.* 2015. P. 5197–5206.
- Huttenlocher D.P., Klanderman G.A., Rucklidge W.J. Comparing images using the hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence.* 1993. V. 15 (9). P. 850–863.
- Ilyuhin S.A., Chernov T.S., Polevoy D.V., Fedorenko F.A. A method for spatially weighted image brightness normalization for face verification. *Proc. SPIE 11041. Eleventh International Conference on Machine Vision (ICMV 2018).* 2019a. V. 11041. P. 1104118.  
<https://doi.org/10.1117/12.2522922>
- Ilyukhin S.A., Chernov T.S., Polevoy D.V. Povyshenie tochnosti neirosetevykh metodov verifikatsii lits za schet prostranstvenno-vzveshennoi normalizatsii yarkosti izobrazheniya [Improving the Accuracy of Neural Network Methods of Verification of Persons by Spatial-Weighted Normalization of Brightness Image]. *Informatsionnye tekhnologii i vychislitel'nye sistemy* [Journal Of Information Technologies And Computing Systems]. 2019b. V. 2019 (4). P. 12–20 (in Russian).  
<https://doi.org/10.14357/20718632190402>
- Ingacheva A.S., Chukalina M.V. Polychromatic CT Data Improvement with One-Parameter Power Correction. *Mathematical Problems in Engineering.* 2019. V. 2019. P. 1405365.  
<https://doi.org/10.1155/2019/1405365>
- Iwamura M., Niwa R., Kise K., Uchida S., Omachi S. Rectifying perspective distortion into affine distortion using variants and invariants. *Proc. of the Second International Workshop on Camera-Based Document Analysis and Recognition.* 2007. P. 138–145.
- Iyatomi H., Celebi M.E., Schaefer G., Tanaka M. Automated color normalization for dermoscopy images. *2010 IEEE International Conference on Image Processing.* IEEE. 2010. P. 4357–4360.
- Jaccard P. Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. *Bull Soc Vaudoise Sci Nat.* 1901. V. 37. P. 241–272.
- Jesorsky O., Kirchberg K.J., Frischholz R.W. Robust face detection using the hausdorff distance. *International Conference on Audio-and Video-Based Biometric Person Authentication.* Springer. 2001. P. 90–95.
- Kadir T., Zisserman A., Brady M. An affine invariantsalient region detector. *European conference on computer vision.* Springer. 2004. P. 228–241.

- Karaimer H.C., Brown M.S. A software platform for manipulating the camera imaging pipeline. *European Conference on Computer Vision*. 2016. P. 429–444.  
[https://doi.org/10.1007/978-3-319-46448-0\\_26](https://doi.org/10.1007/978-3-319-46448-0_26)
- Karaimer H.C., Brown M.S. Improving color reproduction accuracy on cameras. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018. P. 6440–6449.
- Karnaikhov V.N., Kober V.I. A locally adaptive algorithm for shadow correction in color images. *Applications of Digital Image Processing XL, International Society for Optics and Photonics, Nov. 2017*. 2017. V. 10396. P. 10396–23.  
<https://doi.org/10.1117/12.2272692>
- Karpenko S., Konovalenko I., Miller A., Miller B., Nikolaev D. Uav control on the basis of 3d landmark bearing-only observations. *Sensors*. 2015. V. 15 (12). P. 29802–29820.  
<https://doi.org/10.3390/s151229768>
- Katamanov S.N. Avtomaticheskaya privyazka izobrazhenii geostatsionarnogo sputnika mtsat-1r. *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa* [Current Problems In Remote Sensing Of The Earth From Space]. 2007. V. 1 (4). P. 63–68 (in Russian).
- Kholopov I.S. Algoritm korrektii proektivnykh iskazhenii pri malovysotnoi s'emke [Projective distortion correction algorithm at low altitude photographing] *Komp'yuternaya optika* [Computer Optics]. 2017. V. 41 (2). P. 284–290 (in Russian).
- Kim K., Lim T., Kim C., Park S., Park C., Keum C. High-precision color uniformity based on 4D transformation for micro-LED. *Proc. SPIE 11302, Light-Emitting Devices, Materials, and Applications XXIV*. 2020. V. 11302. P. 113021U.  
<https://doi.org/10.1117/12.2542728>
- Kober V.I., Karnaikhov V.N. Adaptivnaya korrektsiya neravnomernogo osveshcheniya na tsifrovyykh multispektral'nykh izobrazheniyakh [Adaptive correction of non-uniform illumination of multispectral digital images] *Informatsionnye protsessy*. 2016a. V. 19 (2). P. 152–161 (in Russian).
- Kober V.I., Karnaikhov V.N. Vosstanovlenie multispektral'nykh izobrazhenii, iskazhennykh prostranstvenno-neodnorodnym dvizheniem kamery [Restoration of multispectral images degraded by non-uniform camera motion] *Informatsionnye protsessy*. 2015. V. 15 (2). P. 269–277 (in Russian).
- Kober V.I., Karnaikhov V.N. Adaptive correction of non-uniform illumination of multispectral digital images. *JCTE*. 2016b. V. 61 (12). P. 1419–1425.  
<https://doi.org/10.1134/S1064226916120123>
- Konovalenko I.A. Srednekvadratichnaya nevyazka koordinat kak kriterii tochnosti normalizatsii izobrazhenii pri opticheskem raspoznavanii dokumentov [RMS coordinate discrepancy as accuracy criterion of images normalization at optical document recognition] *Informatsionnye protsessy*. 2020a. V. 20 (3). P. 215–230 (in Russian).
- Konovalenko I.A., Shemiakina J.A. Error values analysis for inaccurate projective transformation of a quadrangle. *Journal of Physics: Conference Series*. 2018. V. 1096 (1). P. 012038.  
<https://doi.org/10.1088/1742-6596/1096/1/012038>
- Konovalenko I.A., Kokhan V.V., Nikolaev D.P. Optimal'naya affinnaya approksimatsiya proektivnogo preobrazovaniya izobrazhenii [Optimal affine approximation of image projective transformation] *Sensornye sistemy* [Sensory systems]. 2019. V. 33 (1). P. 7–14 (in Russian).  
<https://doi.org/10.1134/S0235009219010062>
- Konovalenko I.A., Kokhan V.V., Nikolaev D.P. Optimal affine image normalization approach for optical character recognition. *Computer Optics*. 2021. V. 45 (1). P. 90–100.  
<https://doi.org/10.18287/2412-6179-CO-759>
- Konovalenko I.A., Polevoy D.V., Nikolaev D.P. Maksimal'naya nevyazka napravlenii kak kriterii tochnosti proektivnoi normalizatsii izobrazheniya pri opticheskom raspoznavanii teksta [Maximal directions discrepancy as accuracy criterion of images projective normalization for optical text recognition] *Sensornye sistemy* [Sensory systems]. 2020b. V. 34 (2). P. 131–146 (in Russian).  
<https://doi.org/10.31857/S0235009220020079>
- Konovalenko I.A., Shemyakina Y.A., Faradzhev I.A. Otseňka tochki skhoda otrezkov metodom maksimal'nogo pravdopodobiya [Calculation of a vanishing point by the Maximum likelihood estimation method] *Vestnik YuUrGU MMP* [Bulletin of the South Ural State University, Series: Mathematical Modelling, Programming and Computer Software]. 2020c. V. 13 (1). P. 107–117 (in Russian).  
<https://doi.org/10.14529/mmp200108>
- Konovalenko I.A., Kokhan V.V., Nikolaev D.P. Maximal coordinate discrepancy as accuracy criterion of image projective normalization for optical recognition of documents. *Bulletin of the South Ural State University, Series: Mathematical Modelling, Programming and Computer Software*. 2020d. V. 13 (3). P. 43–58.  
<https://doi.org/10.14529/mmp200304>
- Kordecki A. Practical testing of irradiance-independent camera color calibration. *Proc. SPIE 11041. Eleventh International Conference on Machine Vision (ICMV 2018)*. 2019. V. 11041. P. 340–345.
- Kozlov E.P., Egoshkin N.A., Eremeev V.V. Normalizatsiya kosmicheskikh izobrazhenii Zemli na osnove ikh sopostavleniya s elektronnymi kartami. *Tsifrovaya obrabotka signalov*. [Digital Signal Processing]. 2009. V. (3). P. 21–26 (in Russian).
- Kunina I.A., Aliev M.A., Arlazarov N.V., Polevoy D.V. A method of fluorescent fibers detection on identity documents under ultraviolet light. *Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019)*. 2020. V. 11433. P. 1–8.  
<https://doi.org/10.1117/12.2558080>
- Kunina I.A., Gladilin S.A., Nikolaev D.P. Blind radial distortion compensation in a single image using fast Hough transform. *Computer Optics*. 2016. V. 40 (3). P. 395–403.  
<https://doi.org/10.18287/2412-6179-2016-40-3-395-403>
- Kutulakos K.N., Vallino J. Affine object representations for calibration-free augmented reality. *Virtual Reality An-*

- nual International Symposium. Proc. of the IEEE 1996. IEEE. 1996. P. 25–36.
- Legge G.E., Pelli D.G., Rubin G.S., Schleske M.M. Psychophysics of reading—I. Normal vision. *Vision research*. 1985. V. 25 (2). P. 239–252.
- Limonova E., Bezmaternykh P.V., Nikolaev D., Arlazarov V. Slant Rectification in Russian Passport OCR System Using Fast Hough Transform. *Ninth International Conference on Machine Vision (ICMV 2016)*. 2017. V. 10341. P. 103410P.  
<https://doi.org/10.1117/12.2268725>
- Limonova E., Nikolaev D., Arlazarov V.V. Bipolar Morphological U-Net for Document Binarization. *Thirteenth International Conference on Machine Vision (ICMV 2020)*. 2021. V. 11605. P. 116050P.  
<https://doi.org/10.1117/12.2587174>
- Lorenz H., Döllner J. Real-time piecewise perspective projections. *GRAPP*. 2009. P. 147–155.
- Lu S., Chen B.M., Ko C.C. Perspective rectification of document images using fuzzy set and morphological operations. *Image and Vision Computing*. 2005. V. 23 (5). P. 541–553.
- Lyubchenko V.A., Putyatin E.P. Matematicheskie modeli razlozheniya proektivnykh preobrazovanii v zadachakh normalizatsii. *Radioelektronika i informatika*. 2002. V. 2 (19). P. 57–59 (in Russian).
- MacAdam D.L. Projective Transformations of I. C. I. Color Specification. *J. Opt. Soc. Am.* 1937. V. 27 (8). P. 294–299.  
<https://doi.org/10.1364/JOSA.27.000294>
- Merino-Gracia C., Mirmehdi M., Sigut J., González-Mora J.L. Fast perspective recovery of text in natural scenes. *Image and Vision Computing*. 2013. V. 31 (10). P. 714–724.
- Mikolajczyk K., Schmid C. An affine invariant interest point detector. *European conference on computer vision. Springer*. 2002. P. 128–142.
- Mikolajczyk K., Schmid C. Scale & affine invariant interest point detectors. *International journal of computer vision*. 2004. V. 60 (1). P. 63–86.
- Morel J.M., Yu G. ASIFT: A new framework for fully affine invariant image comparison. *SIAM journal on imaging sciences*. 2009. V. 2 (2). P. 438–469.
- Murygin K.V. Normalizatsiya izobrazheniya avtomobil'nogo nomera i segmentatsiya simvolov dlya posleduyushchego raspoznaniya [Normalization of the Image of a Car Plate and Segmentation of Symbols for the Subsequent Recognition] *Shtuchnyjintelekt* [Artificial Intelligence]. 2010. V. (3). P. 364–369 (in Russian).
- Nikolaev D.P., Gladkov A., Chernov T., Bulatov K. Diamond Recognition Algorithm using Two-Channel X-ray Radiographic Separator. *Seventh International Conference on Machine Vision (ICMV 2014). International Society for Optics and Photonics*. 2015. V. 9445. P. 944507.  
<https://doi.org/10.1117/12.2181204>
- Nikolaev D.P., Grigoryev A.S., Gladkov A.P. *Programma avtomaticheskogo soglasovaniya chuvstvit'nostei kamer stereopary*. Patent RF. No. RU 2016617966. 2016.
- Nikolaev P.P. Metod proektivno invariantnogo opisaniya ovalov s osevoi libo tsentral'noi simmetriei. *Informatiionnye tekhnologii i vychislitel'nye sistemy* [Journal Of Information Technologies And Computing Systems]. 2014. V. (2). P. 46–59 (in Russian).
- Nikolaev P.P. Proektivno invariantnoe raspoznavanie sostavnykh ovalov. *Informatiionnye tekhnologii i vychislitel'nye sistemy* [Journal Of Information Technologies And Computing Systems]. 2010. V. (4). P. 3–15 (in Russian).
- Nikolaev P.P. Raspoznavanie proektivno preobrazovannykh ploskikh figur. X. Metody poiska okteta invariantnykh tochek kontura ovala – itog vklyucheniya razvityoi teorii v skhemy ego opisaniya [Recognition of projectively transformed planar figures. X. Methods for finding an octet of invariant points of an oval contour – the result of introducing a developed theory into the schemes of oval description] *Sensornye sistemy* [Sensory systems]. 2017. V. 31 (3). P. 202–226 (in Russian).
- Nikolaev P.P., Savchik A.V., Konovalenko I.A. Proektivno invariantnoe predstavlenie kompozitsii dvukh ovalov [A Projectively Invariant Representation of a Composition of Two Ovals] *Informatiionnye protsessy*. 2018. V. 18 (4). P. 304–321 (in Russian).
- Nikolayev P.P. Proektivno invariantnoe opisanie ovalov s simmetriami trekh rodov [A Projective Invariant Description of Ovals with Three Possible Symmetry Genera] *Vestnik RFFI* [Vestnik RFFI]. 2016. V. (4). P. 38–54 (in Russian).  
<https://doi.org/10.22204/2410-4639-2016-092-04-38-54>
- Nikolaidis A. Affine transformation invariant image watermarking using moment normalization and radial symmetry transform. *18<sup>th</sup> IEEE International Conference on Image Processing. IEEE*. 2011. P. 2729–2732.
- Ohta T.I., Maenobu K., Sakai T. Obtaining surface orientation from texels under perspective projection. *IJCAI*. 1981. P. 746–751.
- Orrite C., Herrero J.E. Shape matching of partially occluded curves invariant under projective transformation. *Computer Vision and Image Understanding*. 2004. V. 93 (1). P. 34–64.
- Panfilova E.I., Shipitko O.S., Kunina I.A. Fast Hough Transform-Based Road Markings Detection For Autonomous Vehicle. *Thirteenth International Conference on Machine Vision (ICMV 2020)*. 2021 V. 11605. P. 116052B.  
<https://doi.org/10.1117/12.2587615>
- Pavić D., Schönenfeld V., Kobbelt L. Interactive image completion with perspective correction. *The Visual Computer*. 2006. V. 22 (9–11). P. 671–681.
- Polevoy D.V., Panfilova E.I., Ershov E.I., Nikolaev D.P. Color correction of the document owner's photograph image during recognition on mobile device. *Thirteenth International Conference on Machine Vision (ICMV 2020)*. 2021. V. 11605. P. 1160510.  
<https://doi.org/10.1117/12.2587627>
- Povolotskiy M.A., Kuznetsova E.G., Utkin N.V., Nikolaev D.P. Segmentatsiya registratsionnykh nomerov avtomobilei s primeneniem algoritma dinamicheskoi transformatsii vremennoi osi [Segmentation of vehicle registration plates based on dynamic time warping] *Sensornye sistemy* [Sensory systems]. 2018. V. 32 (1). P. 50–59 (in Russian).  
<https://doi.org/10.7868/S0235009218010080>

- Povolotskiy M.A., Tropin D.V. Dynamic Programming Approach to Template-based OCR. *Proc. SPIE 11041. Eleventh International Conference on Machine Vision (ICMV 2018)*. 2019. V. 11041. P. 110411T. <https://doi.org/10.1117/12.2522974>
- Povolotskiy M.A., Tropin D.V., Chernov T.S., Savel'ev B.I. Metod segmentatsii strukturirovannykh tekstovykh ob'ektov na izobrazhenii s pomoshch'yu dinamicheskogo programmirovaniya [Dynamic programming approach to textual structured objects segmentation in images] *Informatsionnye tekhnologii i vychislitel'nye sistemy* [Journal Of Information Technologies And Computing Systems]. 2019. V. 69 (3). P. 66–78 (in Russian). <https://doi.org/10.14357/20718632190306>
- Pritula N., Nikolaev D.P., Sheshkus A., Pritula M., Nikolaev P.P. Comparison of two algorithms modifications of projective-invariant recognition of the plane boundaries with the one concavity. *Seventh International Conference on Machine Vision (ICMV 2014)*. International Society for Optics and Photonics. ICMV 2014. 2015. V. 944508. P. 1–5. <https://doi.org/10.1117/12.2181215>
- Prun V.E., Polevoy D.V., Postnikov V.V. Forward Rectification – Spatial Image Normalization for a Video from a Forward Facing Vehicle Camera. *Ninth International Conference on Machine Vision (ICMV 2016)*. 2017. V. 10341. P. 103410W. <https://doi.org/10.1117/12.2268605>
- Putyatin E.P., Prokopenko D.O., Pechenaya E.M. Voprosy normalizatsii izobrazhenii pri proektivnykh preobrazovaniyakh. *Radioelektronika i informatika* [Radioelectronics & Informatics]. 1998. V. 2 (3). P. 82–86 (in Russian).
- Rezatofighi H., Tsoi N., Gwak J., Sadeghian A., Reid I., Savarese S. Generalized intersection over union: A metric and a loss for bounding box regression. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019. P. 658–666.
- Rodríguez-Piñeiro J., Comesáñ-Alfaro P., Pérez-González F., Malvido-García A. A new method for perspective correction of document images. *Document Recognition and Retrieval XVIII*. International Society for Optics and Photonics. 2011. V. 787410. P. 1–12.
- Safari R., Narasimhamurthi N., Shridhar M., Ahmadi M. Document registration using projective geometry. *IEEE transactions on image processing*. 1997. V. 6 (9). P. 1337–1341.
- Savchik A., Ershov E., Karpenko S. Color Cerberus. *ISPA 2019*. 2019. P. 355–359. <https://doi.org/10.1109/ISPA.2019.8868425>
- Savchik A.V., Nikolaev P.P. Metod proektivnogo sopostavleniya dlya ovalov s dvumya otmechenymi tochkami. *Informatsionnye tekhnologii i vychislitel'nye sistemy* [Journal Of Information Technologies And Computing Systems]. 2018. V. (1). P. 60–67 (in Russian).
- Savchik A.V., Nikolaev P.P. Teorema o pereschenii T-i H-polyar [The Theorem of T- and H- Polars Intersections Count]. *Informatsionnye protsessy*. 2016. V. 16 (4). P. 430–443 (in Russian).
- Sawhney H.S., Kumar R. True multi-image align mentand its application to mosaicing and lens distortion correc-
- tion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1999. V. 21 (3). P. 235–243.
- Schmitzer B., Schnörr C. Globally optimal joint image segmentation and shape matching based on Wasserstein modes. *Journal of Mathematical Imaging and Vision*. 2015. V. 52 (3). P. 436–458.
- Shemiakina J., Konovalenko I., Tropin D., Faradjev I. Fast projective image rectification for planar objects with Manhattan structure. *Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019)*. 2020. V. 11433. P. 114331N. <https://doi.org/10.1117/12.2559630>
- Shemiakina J.A., Zhukovsky A.E., Faradjev I.A. Issledovanie algoritmov vychisleniya proektivnogo preobrazovaniya v zadache navedeniya na planarnyi ob'ekt po osobyml tochkam [The research of the algorithms of a projective transformation calculation in the problem of planar object targeting by feature points] *Iskusstvennyi intellekt i prinyatie reshenii* [Artificial Intelligence And Decision Making]. 2017. V. 2017 (1). P. 43–49 (in Russian).
- Shemiakina Yu.A. Ispol'zovanie tochek i pryamykh dlya vychisleniya proektivnogo preobrazovaniya po dvum izobrazheniyam ploskogo ob'ekta [The usage of points and lines for the calculation of projective transformation by two images of one plane object] *Informatsionnye tekhnologii i vychislitel'nye sistemy* [Journal Of Information Technologies And Computing Systems]. 2017. V. 2017 (3). P. 79–91 (in Russian).
- Shepelev D.A., Bozhkova V.P., Ershov E.I., Nikolaev D.P. Modelirovanie drobovogo shuma tsvetnykh podvodnykh izobrazhenii [Simulating shot noise of color underwater images] *Komp'yuternaya optika* [Computer Optics]. 2020. V. 44 (4). P. 671–679 (in Russian). <https://doi.org/10.18287/2412-6179-CO-754>
- Sheshkus A., Ingacheva A., Arlazarov V., Nikolaev D. HoughNet: neural network architecture for vanishing points detection. *ICDAR 2019, IEEE*. 2020. V. 8978201. P. 844–849. <https://doi.org/10.1109/ICDAR.2019.00140>
- Shipitko O., Grigoryev A. Ground Vehicle Localization With Particle Filter Based On Simulated Road Marking Image. *ECMS 2018*. 2018. P. 341–347. <https://doi.org/10.7148/2018-0341>
- Shipitko O., Kibalov V., Abramov M. Linear Features Observation Model for Autonomous Vehicle Localization. *ICARCV 2020, Institute of Electrical and Electronics Engineers Inc.* 2021. V. 9305434. P. 1360–1365. <https://doi.org/10.1109/ICARCV50220.2020.9305434>
- Shipitko O.S., Abramov M.P., Lukyanov A.S., Panfilova E.I., Kunina I.A., Grigoryev A.S. Edge detection based mobile robot indoor localization system. *Proc. SPIE 11041. Eleventh International Conference on Machine Vision (ICMV 2018)*. 2019. V. 11041. P. 110412V. <https://doi.org/10.1117/12.2522788>
- Sim D.G., Kwon O.K., Park R.H. Object matching algorithms using robust hausdorff distance measures. *IEEE Transactions on image processing*. 1999. V. 8 (3). P. 425–429.
- Sinclair D., Blake A. Isoperimetric normalization of planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1994. V. 16 (8). P. 769–777.

- Singh S.K., Naidu S.D., Srinivasan T.P., Krishnaa B.G., Srivastava P.K. Rational polynomial modelling for cartosat-1 data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2008. V. 37. P. 885–888.
- Skoryukina N., Arlazarov V.V., Nikolaev D.P. Fast method of ID documents location and type identification for mobile and server application. *ICDAR 2019*. 2020. P. 850–857.  
<https://doi.org/10.1109/ICDAR.2019.00141>
- Skoryukina N., Chernov T., Bulatov K., Nikolaev D.P., Arlazarov V. Snapscreen: Tv-stream frame search with projectively distorted and noisy query. *Ninth International Conference on Machine Vision (ICMV 2016)*. 2017. V. 103410Y. P. 1–5. Bellingham.  
<https://doi.org/10.1117/12.2268735>
- Skoryukina N., Shemiakina J., Arlazarov V. L., Faradjev I. Document localization algorithms based on feature points and straight lines. *Tenth International Conference on Machine Vision (ICMV 2017)*. International Society for Optics and Photonics. 2018. V. 106961H. P. 1–8.  
<https://doi.org/10.1117/12.2311478>
- Smagina A., Bozhkova V.P., Gladilin S., Nikolaev D. Linear colour segmentation revisited. *Proc. SPIE 11041. Eleventh International Conference on Machine Vision (ICMV 2018)*. 2019. V. 11041. P. 110410F.  
<https://doi.org/10.1117/12.2523007>
- Smagina A., Ershov E., Grigoryev A. Multiple Light Source Dataset for Colour Research. *Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019)*. 2020. V. 11433. P. 114332C.  
<https://doi.org/10.1117/12.2559491>
- Smith T., Guild J. The C.I.E. colorimetric standards and their use. *Transactions of the Optical Society*. 1931. V. 33 (3). P. 73–134.  
<https://doi.org/10.1088/1475-4878/33/3/301>
- Stein G.P. Lens distortion calibration using point correspondences. *Computer Vision and Pattern Recognition, 1997. Proc. of the IEEE Computer Society Conference*. 1997. P. 602–608.
- Su Z., Zeng W., Wang Y., Lu Z.L., Gu X. Shape classification using Wasserstein distance for brain morphometry analysis. *International Conference on Information Processing in Medical Imaging*. Springer. 2015. P. 411–423.
- Szeliski R. Video mosaics for virtual environments. *IEEE computer Graphics and Applications*. 1996. P. 16 (2). P. 22–30.
- Takezawa Y., Hasegawa M., Tabbone S. Camera-captured document image perspective distortion correction using vanishing point detection based on Radon transform. *Pattern Recognition (ICPR). 2016 23rd International Conference on*. IEEE. 2016. P. 3968–3974.
- Titov V., Shepelev D., Nikolaev D. Opredelenie parametrov pogloscheniya i rasseyaniya na osnove bystrogo preobrazovaniya Khafa. *Sbornik trudov 43-i mezhdisciplinarnoi shkoly-konferentsii IPPI RAN "Informatsionnye tekhnologii i sistemy 2019" (ITiS 2019)*. [Proc. of the ITiS 2019]. 2020. P. 495–500 (in Russian).
- Tong L., Zhang Y. Correction of perspective text image based on gradient method. *Information Networking and Automation (ICINA). International Conference on*. IEEE. 2010. V. 2. P. 312–316.
- Triputen' V.V., Gorokhovatskii V.A. Algoritm parallel'noi normalizatsii affinnykh preobrazovanii dlya tsvetnykh izobrazhenii. *Radioelektronika i informatika* [Radioelectronics & Informatics]. 1997. V. (1). P. 97–98. (in Russian).
- Trusov A., Limonova E. The analysis of projective transformation algorithms for image recognition on mobile devices. *Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019)*. Wolfgang Osten and Dmitry P. Nikolaev editors. 2020. V. 11433. P. 250–257.
- Tsviatkou V.Yu. Geometricheskie modeli mnogorakursnykh izobrazhenii i proektivnaya kompensatsiya dvizheniya kamery [Geometric models of multi-angle images and projective compensation of camera motion] *Doklady Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki* [Doklady BGUIR]. 2014. V. 86 (8). P. 41–47 (in Russian).
- Vanichev A.Yu. Normalizatsiya siluetov ob"ektorov v sistemakh tekhnicheskogo zreniya. *Programmnye produkty i sistemy*. [Software & Systems]. 2007. V. (3). P. 86–88 (in Russian).
- Wallace G., Chen H., Li K. Color gamut matching for tiled display walls. *EGVE '03: Proc. of the workshop on Virtual environments 2003*. 2003. P. 293–302.  
<https://doi.org/10.1145/769953.769988>
- Wolberg G. *Digital Image Warping*. IEEE Computer Society Press, LosAlamitos, CA, 1990. 318 p.
- Xie Y., Tang G., Hoff W. Geometry-based populated chessboard recognition. *Tenth International Conference on Machine Vision (ICMV 2017)*. International Society for Optics and Photonics. 2018. V. 1069603. P. 1–5.
- Zeynalov R., Velizhev A., Konushin A. Vosstanovlenie formy stranitsy teksta dlya korrektii geometricheskikh iskazhenii [Document images geometrical distortions correction using text lines shape extraction]. *Proc. of the 19 International Conference GraphiCon-2009*. 2009. P. 125–128 (in Russian).
- Zhang W., Li X., Ma X. Perspective correction method for Chinese document images. *Intelligent Information Technology Application Workshops. 2008. IITAW'08. International Symposium on*. IEEE. 2008. P. 467–470.
- Zhang Z., He L.W. Whiteboard scanning and image enhancement. *Digital signal processing*. 2007. V. 17 (2). P. 414–432.
- Zhukovskiy A.E., Nikolaev D.P., Arlazarov V.V., Postnikov V.V., Polevoy D.V., Skoryukina N.S., Chernov T.S., Shemjakina Yu.A., Mukovozov A.A., Konovalenko I.A., Povolotskiy M.A. Segments graph-based approach for document capture in a smartphone video stream. *ICDAR2017*. IEEE Computer Society. 2018. V. 1. P. 337–342.  
<https://doi.org/10.1109/ICDAR.2017.63>
- Zwicker M., Rasanen J., Botsch M., Dachsbaecher C., Pauly M. Perspective accurate splatting. *Proc. of the Graphics interface 2004. Canadian Human-Computer Communications Society*. 2004. P. 247–254.

## Роль проективных преобразований в нормализации изображений

И. А. Коноваленко<sup>a, #</sup>, П. П. Николаев<sup>a, b</sup>

<sup>a</sup> Институт проблем передачи информации им. А.А. Харкевича РАН,  
127051 Москва, Большой Картеный переулок, д. 19, Россия

<sup>b</sup> Московский физико-технический институт,  
141701 г. Долгопрудный, Институтский переулок, д. 9, Россия

#E-mail: konovalenko@iitp.ru

Анализ сделанного с произвольного ракурса изображения требует его предварительной нормализации – преобразования к такому виду, как если бы оно было получено с удобного для анализа ракурса. Данная работа представляет собой обзор современных методов, критериев точности и приложений нормализации разнообразных типов, а также описывает основные этапы становления этой проблематики. Впервые единообразно рассматриваются два важнейших частных случая нормализации, в литературе традиционно рассматриваемые независимо: первый включает вопросы только геометрического характера, второй рассматривает исключительно цветовые аспекты. Плодотворность этого объединяющего подхода выражается еще и в том, что процедура нормализации оказывается фундаментальным образом привлекающей двумерные и трехмерные проективные преобразования с общим аналитическим аппаратом, безотносительно к цветовому и геометрическому ее истолкованию для практических задач.

**Ключевые слова:** геометрическая и цветовая нормализация, проективное преобразование, матрица гомографии, среднеквадратичная и максимальная невязки координат, критерии точности нормализации, область интереса

## ПРАВИЛА ДЛЯ АВТОРОВ ЖУРНАЛА “СЕНСОРНЫЕ СИСТЕМЫ”

DOI: 10.31857/S023500922101011X

Все рукописи подаются в Редакцию журнала через редакционно-издательскую систему (РИС). Для подачи рукописи автор-корреспондент ( тот, кто подает рукопись) должен зарегистрироваться в РИС. Данные, которые требуется указать автору при регистрации, необходимы для автоматического заполнения договоров с издателями. Регистрация пользователей в системе производится на сайте <https://publish.sciencejournals.ru/login>. После регистрации возможна подача рукописи через личный кабинет: загрузка всех необходимых файлов (текст, рисунки, при необходимости – сопроводительные материалы), включая договор о передаче авторского права. В дальнейшем все этапы прохождения рукописи, в том числе рецензии и решение Редколлегии, могут быть отслежены через личный кабинет автора. При подаче материалов обязательно дублировать статьи по почте editor@sensorysystems.ru.

Журнал “Сенсорные системы” публикует оригинальные статьи, обзоры, краткие сообщения. Назначение журнала – освещать физические, физиологические, морфологические и информационные аспекты структуры и функции биологических и технических сенсорных систем, принципы представления и переработки информации из окружающего мира и реконструкции его содержания, алгоритмы работы систем автоматического анализа сенсорной информации для управления аппаратами, заменяющими человека в разных сферах деятельности. Журнал принимает только те статьи, которые соответствуют тематике журнала.

Журнал придерживается правил международного Комитета по этике публикаций (Committee on Publication Ethics – COPE, подробнее можно ознакомиться на сайте: <https://sensorysystems.ru/ru/ethics.html>).

К публикации принимаются завершенные экспериментальные и теоретические работы, ранее нигде не публиковавшиеся и не представленные для публикации в другом издании. Поступающие рукописи проверяются на наличие плагиата.

Шаблон рукописи в формате doc можно скачать по ссылке: [http://sensorysystems.ru/ru/article\\_submission.html](http://sensorysystems.ru/ru/article_submission.html).

**К рукописи обязательно должны быть приложены:**

- подписанный лицензионный договор (бланк договора можно скачать по адресу: [http://sensorysystems.ru/ru/article\\_submission.html/](http://sensorysystems.ru/ru/article_submission.html/));
- номер телефона и адрес электронной почты контактного лица (на отдельном листе или в тексте письма).

### Прохождение рукописей

Все рукописи, поступившие в редакцию, **рецензируются** не менее чем двумя рецензентами “слепым” рецензированием. В роли рецензентов могут выступать внешние эксперты и члены редколлегии. Решение о публикации принимается Редколлегией журнала на основе полученных рецензий. Автор может рекомендовать двух-трех рецензентов для своей работы, но право выбора рецензентов остается за Редколлегией. Редколлегия может поставить условием публикации доработку рукописи в соответствии с рекомендациями рецензентов. Редколлегия оставляет за собой право отклонить рукопись, если она не соответствует профилю журнала, неудовлетворительна по научному содержанию или по техническому исполнению. Рецензии направляются авторам вместе с мотивированным заключением (“принять”, “доработать”, “отклонить”). В случае заключения “доработать”, авторы должны представить доработанный вариант рукописи в течение трех месяцев. При предоставлении измененной рукописи в более поздний срок, она рассматривается как новая. Рецензии и заключения Редколлегии хранятся в редакции и Издательстве в течение 5 лет.

### Авторские экземпляры

Авторам высыпается PDF-файл опубликованной статьи.

### Технические требования к рукописи

Рукописи представляются в электронном виде на русском или английском языках. Электронный файл должен быть в формате DOC. Один электронный файл должен содержать полный набор информации, т.е. основной текст, подрисуночные подписи, таблицы и рисунки.

Рукопись должна быть набрана шрифтом Times New Roman 12 pt, отформатирована через 1.5 интервала, на листах А4 с полями 3-5 см с левой стороны, 1 см с правой стороны, не менее 3 см сверху и снизу. Каждый абзац начинается с красной строки, межстрочный отступ должен быть равен межстрочному интервалу. Все страницы рукописи нумеруются по порядку.

При первом введении аbbревиатур дается их полная расшифровка. Другие сокращения слов, кроме единиц измерения и сокращения “т.е.”, не допускаются.

При наборе правильно используйте прописные и строчные буквы, буквы русского и латинского алфавитов и другие символы сходных начертаний. Помните, что в компьютерных текстах все эти символы имеют разные коды.

**Рисунки** располагаются в конце рукописи, а не в тексте. Каждый рисунок выполняется на отдельной странице с указанием номера рисунка. Разрешение рисунков должно быть не менее 300 dpi. Рекомендуется, чтобы ширина рисунка на странице составляла 8-8.5 см (на одну колонку) или 17–17.5 см (на две колонки). При этом с рукописью обязательно должны быть сданы исходные файлы иллюстраций.

Векторные иллюстрации должны быть представлены в стандартном формате файлов графического редактора, в котором они были подготовлены, также принимается формат EPS. Остальные иллюстрации принимаются в любых стандартных графических форматах, предпочтительно – TIFF.

Если рисунок содержит несколько фрагментов, они обозначаются по порядку курсивными строчными буквами русского алфавита: *a*, *b*, *v* и т.д.

Цветные рисунки допускаются только по предварительному согласованию с редакцией.

Если авторы используют в своей рукописи иллюстрации или таблицы из других публикаций (в том числе своих собственных), то им необходимо запросить у Издателей этих публикаций разрешение на перепечатку или использование материалов. Подробнее: <https://www.pleiades.online/ru/authors/permission/>

**Подписи к рисункам** даются в конце текста статьи на отдельной странице (все на одной) и должны быть лаконичными и не дублировать основной текст рукописи, однако все условные обозначения и символы должны быть расшифрованы, значения координатных осей указаны.

**Таблицы** набираются при помощи редактора таблиц. Каждая таблица выполняется на отдельной странице после текста рукописи.

#### **Дополнительные материалы**

Для более полного описания исследования, к статье могут прилагаться дополнительные мате-

риалы (аудио- и видеофайлы, презентации, дополнительные таблицы и рисунки и пр.) при условии, если автор является правообладателем прилагаемых материалов, и автором ранее не были переданы авторские права на их использование иным (кроме издателя) лицам, либо автор имеет письменное разрешение правообладателя на их использование в целях опубликования и распространения в журнале. Дополнительные материалы публикуются только в электронной версии на сайте <https://elibrary.ru>.

#### **Структура и оформление рукописи**

На первой странице рукописи приводятся:

- УДК (индекс Универсальной десятичной классификации).
- Заглавие. Заголовок статьи должен быть максимально конкретным, желательно не более 1.5-2 строк (лучше в пределах одной строки).
- Инициалы и фамилия (фамилии) автора (авторов).
- Полное название и адрес организации, где выполнена работа. Если авторы публикации – из разных учреждений, то аффилиация каждого отмечается надстрочной цифрой (например, И.И. Иванов<sup>1</sup>). В английском резюме аффилиации отмечаются латинскими буквами (I.I. Ivanov<sup>a</sup>). Контактный адрес e-mail одного из авторов. Контактный автор в тексте отмечается звездочкой (например, И.И. Иванов<sup>1,\*</sup>). В английском резюме контактный автор отмечается надстрочным знаком # (I.I. Ivanov<sup>#</sup>).

– Слова “Поступила в редакцию...”, После доработки..., Принята к публикации...” (даты будут вписаны редакцией).

– Аннотация на русском языке объемом до 3/4 страницы. Аннотация должна давать представление о предмете исследования, использованных методах и основных результатах.

– Ключевые слова.

– DOI: (DOI будет вписан редакцией).

Пример оформления шапки русскоязычной части статьи:

#### **НАЗВАНИЕ СТАТЬИ**

© 2018 г. И. И. Иванов<sup>1,\*</sup>, П. П. Петров<sup>1</sup>, С. С. Сидоров<sup>2</sup>

<sup>1</sup> Место работы первого автора, 127000 Москва, Один Переулок, д. 19, Россия

<sup>2</sup> Место работы второго автора, 127000 Москва, Другой Переулок, д. 19, Россия

\*E-mail: author1@mail.ru

Поступила в редакцию ...

После доработки ...

Принята к публикации ...

Последующие разделы начинаются со второй страницы. Стандартная структура рукописи, описывающая экспериментальное исследование, включает следующие разделы.

– **Введение** должно содержать изложение проблемы, уровень ее изученности на данный момент и вытекающие отсюда задачи исследования.

– **Описание методики** должно показать соответствие технологии и методов задачам работы и при необходимости обеспечить воспроизводимость результатов другими исследователями. Необходимо указать вид и число наблюдений, дозы веществ, параметры стимуляции, особенности наркоза и пр. При выполнении экспериментов на людях или животных, обязательно описание тех особенностей методики, которые обеспечивают соблюдение норм безвредности экспериментов для людей и гуманного обращения с животными.

– **Результаты исследования** должны содержать описание всех оригинальных данных, представляющих научную ценность и используемых в дальнейшем для обсуждения.

– **Обсуждение** должно касаться рассмотрения и оценки важнейших результатов. Привлекайте для обсуждения только свои оригинальные данные, изложенные в разделе “Результаты”, и данные цитируемых литературных источников. Не включайте в раздел информацию, если не можете точно указать ее источник. При необходимости может привлекаться любая опубликованная информация, но желательно избегать ссылок на нерецензируемые издания, не гарантирующие достоверность сведений (например, краткие тезисы конференций, научно-популярные публикации).

#### – Заключение или Выводы.

– **Финансирование.** Следует указать, каким фондом и грантом поддержано данное исследование и каждая часть работы в отдельности, если источники финансирования разные.

– **Благодарности.** Раздел не является обязательным, заполняется по желанию авторов.

– **Конфликт интересов.** Раздел является обязательным. В случае отсутствия конфликта интересов, авторы также указывают “Авторы данной статьи подтвердили отсутствие конфликта интересов, о котором необходимо сообщить”.

– **Соблюдение этических норм.** Раздел является обязательным. Если исследования проводились на животных, то в данном разделе указывается: “Все процедуры, выполненные в исследованиях с участием животных, соответствовали этическим стандартам учреждения, в котором проводились исследования, и утвержденным правовым актам РФ и международных организаций”. Если исследования проводились с участием людей, то в разделе “Соблюдение этических норм” указывается: “Все процедуры, выполненные в исследованиях с

участием людей, соответствуют этическим стандартам национального комитета по исследовательской этике и Хельсинкской декларации 1964 года и ее последующим изменениям или сопоставимым нормам этики. От каждого из включенных в исследование участников было получено информированное добровольное согласие”. Если в статьях не содержится описания исследований с участием людей или использованием животных и выполненных кем-либо из авторов, в разделе “Соблюдение этических норм” указывается: “Настоящая статья не содержит описания выполненных авторами исследований с участием людей или использованием животных в качестве объектов”.

– **Участие авторов.** В разделе указывается вклад в работу каждого из авторов.

#### – Список литературы.

#### – Резюме на английском языке.

#### – REFERENCES.

Резюме на английском языке включает:

– Заглавие.

– Инициалы и фамилии авторов. Если в статье более одного автора, перед фамилией последнего пишется “and” (I.I. Ivanov, P.P. Petrov, and S.S. Sidorov).

– Название и адрес учреждения. Адрес учреждения обязательно содержит название страны. Если авторы из разных учреждений, то в английском резюме аффилиации отмечаются латинскими буквами (I.I. Ivanov<sup>a</sup>).

– Контактный адрес e-mail одного из авторов. В английском резюме контактный автор отмечается надстрочным знаком # (I.I. Ivanov#).

– Аннотацию (abstract). Содержание аннотации на английском языке должно быть идентичным русскоязычной аннотации.

– Ключевые слова (key words).

#### Пример оформления шапки английского резюме:

##### Paper title

I. I. Ivanov<sup>a, #</sup>, P. P. Petrov<sup>a</sup>, and S. S. Sidorov<sup>b</sup>

<sup>a</sup> First Institution, 127000 Moscow, One lane, 19, Russia

<sup>b</sup> Second Institution, 127000 Moscow, Another lane, 19, Russia

#E-mail: author1@mail.ru

Допускается отклонение от стандартной схемы рукописи, например, объединение разделов “Результаты исследования” и “Обсуждение”. Для обзорных статей более адекватна рубрикация (если необходима), соответствующая тематическим разделам.

**Литературные ссылки в тексте** даются в круглых скобках по фамилии (фамилиям) авторов и, через запятую, году публикации. Если авторов

двоем, их фамилии упоминаются через запятую, без союза. Например: “Предшествующими исследованиями показано (Иванов, Петров, 2000), что...”. В случае трех и более авторов указывается только первая фамилия с добавлением “и др.” (для русскоязычных публикаций) или “et al.” (для англоязычных публикаций). При цитировании нескольких источников, они указываются через точку с запятой: (Рожкова, 2015; Николаев, 2001). Для ссылки на разные работы коллективов с одинаковым первым автором и одного года опубликования добавляются буквы а, б, в. (Рожкова, 2015а; 2015б). В списке литературы буквы а, б, в указываются сразу после года, без пробела. В списке “**REFERENCES**” указывать буквы не нужно.

**Список литературы и список “REFERENCES”** включают все те и только те публикации, на которых имеются ссылки в тексте. Ссылки на неопубликованные работы (с указанием “в печати”, “личное сообщение” и т.п.) не допускаются.

**Обратите внимание, что с 2018 г. список литературы в статьях приводится дважды: один раз – в формате, удобном для читателей, второй – в формате, удобном для поисковых систем.** Под заголовком “Список литературы” русскоязычные публикации указываются в кириллице, англоязычные (и любые другие в латинской транскрипции) – в латинице. Под заголовком “**REFERENCES**” русскоязычные источники указываются в латинице (правила оформления далее), англоязычные (и любые другие в латинской транскрипции) источники полностью дублируются. Порядок публикаций в этих списках одинаков: русскоязычные источники идут перед источниками в латинице и выстраиваются в порядке русского алфавита (даже при транслитерации).

Транслитерация русскоязычных названий для формирования списка “**REFERENCES**” должна выполняться в соответствии с форматом BSI (рекомендуется использовать бесплатный сервис <http://ru.translit.net/?account=bsi/>). Если в русскоязычной публикации имелось англоязычное резюме, или англоязычная версия названия и фамилий авторов, фамилии авторов следует давать в том же написании, как в оригинальной публикации. Если в русскоязычной публикации отсутствовало англоязычное резюме или англоязычная версия названия и фамилий авторов, то для написания фамилий авторов рекомендуется использовать транслитерацию, которую использует сам автор (например, в его англоязычных публикациях). Если авторское написание неизвестно, тогда используется формат транслитерации BSI.

Пожалуйста, аккуратно указывайте названия журналов. Названия журналов должны приводиться либо полностью, либо с сокращениями,

согласно “list of serial titles word abbreviations”. Предпочтительно использовать сокращения.

При наличии у публикации DOI, его указание обязательно.

Если статья написана на русском языке, и в списке цитируемой литературы отсутствуют русскоязычные источники, список дается однократно, при этом предваряется заголовками “Список литературы” и затем, с новой строки, “**REFERENCES**”. Слово “**REFERENCES**” является ключевым для поисковых систем международных баз данных. Заголовок “Список литературы” необходим для всех русскоязычных статей.

#### **Оформление раздела “Список литературы”**

Список начинается на отдельной странице, составляется в алфавитном порядке фамилий авторов (сначала публикации на русском языке, затем – на иностранных языках), без нумерации.

Обратите внимание, что теперь фамилии и инициалы пишутся прямым шрифтом, название книги/журнала – курсивом, знак “//” в списке литературы не используется.

Список составляется по следующему образцу.

#### *Книги*

Фамилия И.О. Заголовок книги. Город (Москва и Санкт-Петербург - сокращенно: М., СПб.). Изд-во, 1995. 351 с.

*Статьи и главы в книгах, непериодических изданиях, сборниках трудов*

Фамилия И.О. Заголовок статьи. *Название книги*. Под ред. Фамилия И.О. Город. Изд-во, 1995. Ч. I. С. 22-35.

#### *Статьи в журналах*

Фамилия И.О. Заголовок статьи. *Название журнала*. 1995. Т. 1. № 1. С. 22-27. DOI: xxxxxxx.

#### *Диссертации и авторефераты диссертаций*

Мы рекомендуем вместо ссылок на диссертации и авторефераты давать ссылки на статьи или монографии по теме диссертации: такие публикации легче найти, если читатель хочет ознакомиться с материалом, и такие цитирования лучше учитываются поисковыми системами.

Фамилия И.О. *Название диссертации*. Дисс. канд. (докт.) биол. наук. Город. 2000. 351 с.

#### *Интернет-ресурсы*

Фамилия И.О. *Название статьи*. URL: [http://example\\_url.html](http://example_url.html) (дата обращения: 23.06.2013).

#### *ГОСТы*

ГОСТ 8.586.5-2005. *Название ГОСТа*. М. Центр стандартизации, 2007. 10 с.

#### *Патенты*

Фамилия И.О. *Название патента*. Патент РФ. № 2486597. 2013.

### Оформление раздела “REFERENCES”

Если в цитируемой русскоязычной публикации имелись англоязычные версии заголовка и фамилий авторов, то в разделе “References” они должны быть даны в точном соответствии с тем, как представлены в публикации. Если в русскоязычной публикации эти данные отсутствовали, название, фамилии авторов и библиографические данные должны быть транслитерированы в соответствии с форматом BSI. После транслитерации в квадратных скобках должен быть указан перевод (без выделения курсивом). В самом конце библиографического описания в круглых скобках помещают указание на исходный язык публикации, например, “(in Russian)”. Издательство приводится транслитерацией, если не имеет собственного названия на английском языке.

#### *Книги*

Polyak S.L. *The Retina*. Chicago, The University of Chicago Press. 1941. 607 p.

Byzov A.L. *Elektrofiziologicheskie issledovaniya set-chatki* [Electrophysiological studies of the retina]. Moscow. Nauka Publ. 1966. 196 p. (in Russian).

*Статьи и главы в книгах, непериодических изданиях, сборниках трудов*

Rozhkova G.I. Binokulyarnoe zrenie [Binocular vision]. *Rukovodstvo po fizioligii. Fiziologiya zreniya* [Handbook on physiology. Vision physiology] Moscow. Nauka, 1992. P. 586–664 (in Russian).

#### *Статьи в журналах*

Gladkov A.P., Kuznetsova E.G., Gladilin S.A., Gracheva M.A. Adaptivnaya stabilizatsiya yarkosti izobrazheniya v tekhnicheskoi sisteme raspoznavaniya krupnykh dvizhushchikhsya ob'ektorov [Adaptive image brightness stabilization for the industrial system of large moving object recognition]. *Sensornye sistemy*

[Sensory systems]. 2017. V. 31 (3). P. 247–260. DOI: XXXXXX (in Russian).

#### *Диссертации и авторефераты докторской*

Grigor'ev Iu.A. *Razrabotka nauchnykh osnov proektirovaniia arkhitektury raspredelenykh sistem obrabotki dannykh*. Diss. dokt. tekhn. nauk [Development of scientific bases of architectural design of distributed data processing systems. Dr. eng. sci. diss.]. Moscow. 1996. 243 p. (in Russian).

#### *Интернет-ресурсы*

Kondrat'ev V.B. *Global'naya farmatsevticheskaya promyshlennost'* [The global pharmaceutical industry]. URL: [http://perspektivy.info/rus/ekob/globalnaja-farmacevticheskaja\\_promyshlennost\\_2011-07-18.html](http://perspektivy.info/rus/ekob/globalnaja-farmacevticheskaja_promyshlennost_2011-07-18.html) (accessed 23.06.2013) (in Russian).

#### *ГОСТы*

GOST 8.586.5–2005. *Izmerenie raskhoda i kolichestva zhidkosteii i gazov s pomoshch'iu standartnykh suzhaiushchikh ustroistv* [State Standard 8.586.5 – 2005. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow, Standartinform Publ., 2007. 10 p.

#### *Патенты*

Nikolaev D.P., Postnikov V.V., Khanipov T.M., Usilin S.A., Grigoryev A.S. *Sposob avtomaticheskoi klassifikatsii transportnykh sredstv* [Method of automatic classification of vehicles]. Patent RF. No. 2280590. 2006.

Рукописи следует направлять по электронной почте на адрес editor@sensorysystems.ru и дублировать на адрес sensys-li@mail.ru.

С правилами для авторов можно также ознакомиться по адресу: [http://sensorysystems.ru/ru/article\\_submission.html](http://sensorysystems.ru/ru/article_submission.html)