

ISSN 2223-3792

Машинное обучение и анализ данных

2015 год

Том 1, номер 12



Машинное обучение и анализ данных

Журнал «Машинное обучение и анализ данных» публикует новые теоретические и обзорные статьи с результатами научных исследований в области искусственного интеллекта, теоретических основ информатики и приложений. Цель журнала — развитие теории машинного обучения, интеллектуального анализа данных и методов проведения вычислительных экспериментов. Принимаются статьи на английском и русском языках.

Журнал включен в российский индекс научного цитирования РИНЦ. Информация о цитировании статей находится на сайте Российского индекса научного цитирования, ISSN 2223-3792, номер свидетельства о регистрации ЭЛ № ФС 77-55486. Журнал зарегистрирован в системе Crossref, doi <http://dx.doi.org/10.21469/22233792>.

- Новостной сайт <http://jmla.org/>
- Электронная система подачи статей <http://jmla.org/papers/>
- Правила подготовки статей <http://jmla.org/papers/doc/authors-guide.pdf>

Тематика журнала:

- классификация, кластеризация, регрессионный анализ;
- алгебраический подход к проблеме синтеза корректных алгоритмов;
- многомерный статистический анализ;
- выбор моделей и сложность;
- глубокое обучение;
- статистическая теория обучения;
- методы прогнозирования временных рядов;
- методы обработки и распознавания сигналов;
- методы оптимизации в задачах машинного обучения и анализа данных;
- методы анализа больших данных;
- методы визуализации данных;
- обработка и распознавание речи и изображений;
- анализ и понимание текста;
- информационный поиск;
- прикладные задачи анализа данных.

Редакционный совет

Ю. Г. Евтушенко, акад.
Ю. И. Журавлёв, акад.
Д. Н. Зорин, проф.
К. В. Рудаков, чл.-корр.

Редколлегия

К. В. Воронцов, д.ф.-м.н.
А. Г. Дьяконов, д.ф.-м.н.
И. А. Матвеев, д.т.н.
Л. М. Местецкий, д.т.н.
В. В. Моттль, д.т.н.
М. Ю. Хачай, д.ф.-м.н.

Координаторы

Ш. Х. Ишкина
М. П. Кузнецов
А. П. Мотренко

Редактор: В. В. Стрижов, д.ф.-м.н. (strijov@ccas.ru)

Вычислительный центр им. А. А. Дородницына ФИЦ ИУ РАН
Московский физико-технический институт
Факультет управления и прикладной математики
Кафедра «Интеллектуальные системы»

Москва, 2015

Journal of Machine Learning and Data Analysis

The journal Machine Learning and Data Analysis publishes original research papers and reviews of the developments in the field of artificial intelligence, theoretical computer science and its applications. The journal aims to promote the theory of machine learning and data mining and methods of conducting computational experiments. Papers are accepted in English and Russian.

The journal is included in the Russian science citation index RSCI. Information about citation to articles can be found at the Russian science citation index website. ISSN 2223-3792. Mass media registration certificate ЭЛ № ФС 77-55486. The Crossref journal doi is <http://dx.doi.org/10.21469/22233792>.

- Journal news and archive <http://jmla.org/>
- Open journal system for papers submission <http://jmla.org/papers/>
- Style guide for authors <http://jmla.org/papers/doc/authors-guide.pdf>

The scope of the journal:

- classification, clustering, regression analysis;
- multidimensional statistical analysis;
- Bayesian methods for regression and classification;
- model selection and complexity;
- deep learning;
- Statistical Learning Theory;
- time series forecasting techniques;
- methods of signal processing and speech recognition;
- optimization methods for solving machine learning and data mining problems;
- methods of big data analysis;
- data visualization techniques;
- methods of image processing and recognition;
- text analysis, text mining and information retrieval;
- applied data analysis problems.

Editorial Council

Yu. G. Evtushenko, acad.
K. V. Rudakov, corr. member
Yu. I. Zhuravlev, acad.
D. N. Zorin, prof.

Editorial Board

A. G. Dyakonov, D.Sc.
M. Yu. Khachay, D.Sc.
I. A. Matveev, D.Sc.
L. M. Mestetskiy, D.Sc.
V. V. Mottl, D.Sc.
K. V. Vorontsov, D.Sc.

Editorial Support

Sh. Kh. Ishkina
M. P. Kuznetsov
A. P. Motrenko

Editor-in-Chief: V. V. Strijov, D.Sc. (strijov@ccas.ru)

Dorodnicyn Computing Centre FRC CSC RAS
Moscow Institute of Physics and Technology
Department of Control and Applied Mathematics
Division “Intelligent Systems”

Moscow, 2015

Содержание

| | |
|---|------|
| <i>И. В. Флоринский, А. Н. Панкратов</i> Цифровое моделирование рельефа с использованием ортогональных полиномов | 1647 |
| <i>П. А. Чочиа</i> Двумерные вариации как средство оценивания сложности изображений | 1660 |
| <i>И. А. Грачева, А. В. Копылов</i> Быстрые алгоритмы обработки изображений на основе гамма-нормальной модели скрытого поля | 1677 |
| <i>Ф. А. Корнилов</i> Исследование влияния рассинхронизации входных изображений на качество работы алгоритмов поиска структурных различий | 1686 |
| <i>Е. В. Медведева, К. А. Карлушин, Е. Е. Курбатова</i> Метод выделения движущихся объектов в видеопотоке на основе оценки их границ | 1696 |
| <i>З. М. Шибзухов, Д. Ю. Чередников</i> О моделях нейронов агрегирующего типа | 1706 |
| <i>А. Н. Трекин, И. А. Матвеев, А. Б. Мурынин, В. Г. Бочкарева</i> Метод повышения разрешения космических изображений с использованием априорной информации в векторной форме для сохранения границ | 1717 |
| <i>Д. Н. Чепрасов, А. А. Маленичев, В. В. Сулимова, О. В. Красоткина, В. В. Моттль, А. А. Марков</i> Восстановление пропущенных данных на ультразвуковых дефектограммах рельсового пути на основе локально-глобального выравнивания | 1731 |
| <i>Е. А. Новиков, И. А. Ваколюк, Р. Д. Ахалкин, И. А. Варчак, И. Г. Шалагинова, Д. А. Швайко, Е. А. Буденкова</i> Автоматизация метода компьютерной окулографии для исследований центральной нервной системы на основе пассивного анализа видеоизображения | 1752 |
| <i>Е. П. Петров, Н. Л. Харина, П. Н. Сухих</i> Метод быстрого сжатия изображений без потерь | 1762 |

Contents

| | |
|--|------|
| <i>I. V. Florinsky and A. N. Pankratov</i> | |
| Digital terrain modeling with orthogonal polynomials | 1647 |
| <i>P. A. Chochia</i> | |
| Two-dimensional variation as an image complexity assessment | 1660 |
| <i>I. A. Gracheva and A. V. Kopylov</i> | |
| Fast image processing algorithms based on the gamma-normal model of hidden field . | 1677 |
| <i>F. A. Kornilov</i> | |
| Research of the impact of misregistration of input images on the accuracy of change detection | 1686 |
| <i>E. V. Medvedeva, K. A. Karlushin, and E. E. Kurbatova</i> | |
| Method of detection of moving objects in video stream on the basis of object boundaries estimation | 1696 |
| <i>Z. M. Shibzukhov and D. Y. Cherednikov</i> | |
| About models of neurons of aggregation type | 1706 |
| <i>A. N. Trekin, I. A. Matveev, A. B. Murynin, and V. G. Bochkareva</i> | |
| A method for upsampling of remote sensing images using vector data for preserving edges | 1717 |
| <i>D. N. Cheprasov, A. A. Malenichev, V. V. Sulimova, O. V. Krasotkina, V. V. Mottl and A. A. Markov</i> | |
| Recovering missing data on ultrasonic rail defectograms via semiglobal warping | 1731 |
| <i>E. A. Novikov, I. A. Vakoliuk, R. D. Akhapkin, I. A. Varchak, I. G. Shalanginova, D. A. Shvaiko, and E. A. Budenkova</i> | |
| Automation method of computer oculoigraphy for research of the central nervous sys- tem based on passive video analysis | 1752 |
| <i>E. P. Petrov, N. L. Kharina, and P. N. Sukhikh</i> | |
| Fast lossless image compression method | 1762 |

Digital terrain modeling with orthogonal polynomials*

I. V. Florinsky and A. N. Pankratov

iflor@mail.ru, pan@impb.ru

Institute of Mathematical Problems of Biology, Russian Academy of Sciences,
1 Vitkevicha st., Pushchino, Moscow Region, Russia

Background: Mathematical problems of digital terrain analysis include interpolation of digital elevation models (DEMs), DEM generalization and denoising, as well as computation of morphometric variables by calculation of partial derivatives of elevation. Traditionally, these procedures are based on numerical treatments of DEMs, two-dimensional (2D) discrete functions of elevation.

Methods: We developed a spectral analytical method and algorithm based on high-order orthogonal expansions using the Chebyshev polynomials of the first kind with the subsequent Fejér summation. The method and algorithm are intended for analytical treatment of regularly spaced DEMs, such as DEM global approximation, generalization, and denoising as well as computation of morphometric variables by analytical calculation of partial derivatives.

Results: To test the method and algorithm, a DEM of the Northern Andes containing 230,880 points (the elevation matrix 480×481) has been used. The DEMs were reconstructed with 480, 240, 120, 60, and 30 expansion coefficients. The first and the second partial derivatives of elevation were analytically calculated from the reconstructed DEMs. The models of 14 local morphometric variables were then computed with the derivatives. A set of maps of elevation and horizontal curvature (k_h) related to different number of expansion coefficients well illustrates data generalization effects, denoising, and removal of artifacts contained in the original DEM.

Concluding Remarks: The test results demonstrated a good performance of the developed method and algorithm. They can be utilized as a universal tool for analytical treatment in digital terrain modeling.

Keywords: *Chebyshev polynomials; Fejér summation; generalization; denoising; partial derivatives*

DOI: 10.21469/22233792.1.12.01

Цифровое моделирование рельефа с использованием ортогональных полиномов*

И. В. Флоринский, А. Н. Панкратов

Институт математических проблем биологии РАН

Математические аспекты цифрового моделирования рельефа включают интерполяцию цифровых моделей высот (ЦМВ), генерализацию и подавление шума в ЦМВ, а также расчет морфометрических характеристик на основе вычисления частных производных высоты. Эти процедуры традиционно основаны на численной обработке ЦМВ – дискретных функций двух переменных. Авторы разработали спектрально-аналитический метод и алгоритм на основе ортогональных разложений высокого порядка с использованием многочленов Чебышёва I рода с последующим суммированием Фейера. Метод и алгоритм предназначены для аналитической обработки ЦМВ, включая глобальную аппроксимацию

*The study was supported by RFBR grant 15-07-02484.

регулярных ЦМВ, генерализацию и подавление шума в ЦМВ, а также расчет морфометрических характеристик на основе аналитического вычисления частных производных. Для тестирования метода и алгоритма использовалась ЦМВ Северных Анд, включающая 230 880 точек (матрица высот 480×481). Цифровые модели высот были восстановлены с использованием 480, 240, 120, 60 и 30 коэффициентов разложения. Первые и вторые частные производные высоты были аналитически рассчитаны по восстановленным ЦМВ. Производные были использованы для расчета моделей горизонтальной кривизны (k_h). Серия карт высоты и k_h , полученных для различного числа коэффициентов разложения, хорошо иллюстрирует генерализацию данных, подавление шума и удаление артефактов, присутствующих в исходной ЦМВ. Результаты тестирования показали высокую работоспособность предлагаемого метода и алгоритма. Они могут применяться как универсальное средство для аналитической обработки данных в цифровом моделировании рельефа.

Ключевые слова: *полиномы Чебышёва; суммирование Фейера; генерализация; подавление шума; частные производные*

DOI: 10.21469/22233792.1.12.01

1 Introduction

Topography is one of the main factors controlling processes taking place in the near-surface layer of the planet. In particular, topography is one of the soil forming factors since it influences: (a) climatic and meteorological characteristics which control hydrological and thermal regimes of soils; (b) prerequisites for gravity-driven overland and intrasoil lateral transport of water and other substances; and (c) spatial distribution of vegetation cover. At the same time, being a result of the interaction of endogenous and exogenous processes of different scales, topography can reflect the geological structure of a terrain. In this connection, digital terrain analysis and digital terrain models (DTMs) are widely used to solve various multiscale problems of geomorphology, hydrology, remote sensing, soil science, geology, geophysics, geobotany, glaciology, oceanology, climatology, planetology, and other disciplines [1–4].

Mathematical issues of quantitative modeling and analysis of the topographic surface can be summarized in three main problems:

- (1) interpolation of irregularly and regularly spaced digital elevation models (DEMs), two-dimensional (2D) discrete functions of an elevation defining the topographic surface as a set of values measured at the grid nodes. This task is commonly carried out by various local interpolation techniques (e.g., piecewise splines) [5–7];
- (2) DEM filtering to denoise, generalize, and decompose DEMs into components of different spatial scales [4]. These tasks are usually attacked by 2D discrete Fourier transform [8–10], 2D discrete wavelet transform [11–13], and 2D singular spectrum analysis [14];
- (3) derivation of local morphometric variables from regularly spaced DEMs to analyze the topographic surface and relationships between topography and other natural and artificial components of geosystems [4]. The list of local morphometric variables includes: slope gradient, slope aspect, minimal curvature, maximal curvature, mean curvature, the Gaussian curvature, unsphericity curvature, horizontal curvature, vertical curvature, difference curvature, accumulation curvature, ring curvature, vertical excess curvature, and horizontal excess curvature. If the topographic surface is defined by a continuous, single-valued function

$$z = f(x, y), \quad (1)$$

where z is the elevation and x and y are the Cartesian coordinates, local morphometric variables are the functions of partial derivatives of elevation:

$$r = \frac{\partial^2 z}{\partial x^2}; \quad t = \frac{\partial^2 z}{\partial y^2}; \quad s = \frac{\partial^2 z}{\partial x \partial y}; \quad p = \frac{\partial z}{\partial x}; \quad q = \frac{\partial z}{\partial y}. \quad (2)$$

For example, horizontal curvature (k_h), one of the most important morphometric attributes, is calculated by the following equation [15]:

$$k_h = -\frac{q^2 r - 2pqs + p^2 t}{(p^2 + q^2) \sqrt{1 + p^2 + q^2}}. \quad (3)$$

Equations of other local morphometric variables can be found elsewhere [4]. To compute r , t , s , p , and q (2) from DEMs based on plane square grids or spheroidal equal angular grids, one can apply methods based on approximation of partial derivatives by finite differences using the 3×3 or 5×5 moving windows [15–19].

However, it is obvious that the three mathematical problems of digital terrain analysis may be resolved within a framework of an analytical treatment of DEMs using global approximation with (orthogonal) polynomials. In the 1960–1980s, there were attempts to apply high-order polynomials for DEM global approximation [20–23]. These attempts have faced several limitations. First, such approaches have required considerable computer resources. Second, practical tasks have demanded to work with increasingly large DEMs containing tens of thousands to a few million points; existed methods and computers could not handle such data amounts. Third, the topographic surface has appeared too complex for wide application of global polynomial approximations. As a result, only low-order orthogonal polynomials have been utilized in trend-surface analysis, a technique to reveal trend and residual components of the topographic surface [24, 25]. However, the current progress in the theory and practice of polynomial approximation [26–28] suggests that it is now possible to solve such problems.

In this paper we describe a spectral analytical method and algorithm based on high-order orthogonal expansions using the Chebyshev polynomials of the first kind with the subsequent Fejér summation. The method and algorithm are intended for the analytical treatment of regularly spaced DEM, such as DEM global approximation, denoising, and generalization as well as computation of morphometric variables by analytical calculation of partial derivatives.

2 Method

Let consider a function of two variables (1) defined in a rectangular domain. To approximate analytically this function, the 2D expansion by the orthogonal Chebyshev polynomials of the first kind

$$z = \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} d_{ij} T_i(x) T_j(y), \quad (4)$$

where $T_i(x)$ and $T_j(y)$ are the Chebyshev polynomials orthogonal in the interval $[-1, 1]$ and d_{ij} are the expansion coefficients, has been used. It is assumed that the domain of the initial function is translated into the domain of the orthogonal polynomials by a linear transformation.

The method and algorithms for the function expansion by the Chebyshev polynomials are described in [29, 30]. To calculate expansion coefficients, we use an operator, or matrix method introduced and studied in [30]. In this study, a 2D approximation of the function (1) is performed as a superposition of one-dimensional approximations by each variable.

We use the following formula for the Chebyshev polynomials:

$$B = \left\{ T_0 = \frac{1}{\sqrt{2}}, T_i = \cos(i \arccos x) \right\},$$

where i is the number of basis function, $i = 1, \dots, l - 1$. The system of functions $T_i(x)$, $i = 0, \dots, l - 1$, satisfies the orthogonality condition in the scalar product defined as follows:

$$(T_i, T_j) = \int_{-1}^1 T_i(x) T_j(x) \frac{1}{\sqrt{1-x^2}} dx = \begin{cases} \pi & \text{if } i = j; \\ 0 & \text{if } i \neq j. \end{cases}$$

In the discrete form, the scalar product on a nonuniform grid of k nodes

$$t_i = \cos\left(\frac{\pi(i-1/2)}{k}\right),$$

which are the zeros of the orthogonal polynomial $T_k(x)$, $i = 1, \dots, k$, has the form:

$$(T_n, T_m) = \frac{2}{k} \sum_{i=1}^k T_n(t_i) T_m(t_i) = \begin{cases} \pi & \text{if } n = m; \\ 0 & \text{if } n \neq m. \end{cases}$$

The expansion coefficients are calculated by the expressions:

$$c_i = \frac{(f, T_i)}{(T_i, T_i)}.$$

Approximations based on orthogonal polynomials always lead to oscillatory artifacts due to the Gibbs phenomenon [31]. To solve this problem, the original approximation was replaced with a smooth representation obtained as an arithmetic mean of all partial sums of an orthogonal series (the Fejér summation). According to the Fejér theorem [32], the initial mean-square approximation becomes the uniform one in this representation. The Fejér summation (or averaging) is the powerful method to suppress or eliminate oscillatory artifacts of the Gibbs phenomenon in data sets approximated with polynomials [31, 33].

Transformation of expansion coefficients corresponding to the arithmetic mean of the partial sums of an orthogonal series has the form:

$$\tilde{c}_i = \frac{l-i}{l} c_i$$

where \tilde{c}_i are the new weighting coefficients of an orthogonal series, $i = 0, \dots, l - 1$.

After the approximation of the function z (1) and the suppression of oscillatory artifacts, it is possible to calculate the partial derivatives r , t , s , p , and q (2) in an analytical representation similar to the function z , that is, in the form of 2D orthogonal series. For the case of the Chebyshev polynomials, calculation of the expansion coefficients of a derivative from the expansion coefficients of the initial function is described in [29]. In general form for an arbitrary basis, a coefficient conversion scheme is presented in [28].

In the present case, these formulas take the following form:

$$p_{l-1} = 0; \quad p_{l-2} = 2(l-2)c_{l-1}; \quad p_j = p_{j+2} + 2j c_{j+1}; \quad p_0 = \frac{p_0}{\sqrt{2}}$$

where p_j are the expansion coefficients of the derivative of an orthogonal series, $j = l - 3, \dots, 0$.

It should also be noted that expansion coefficients of the derivative should be scaled in the case of a linear transformation of function domain:

$$p_j = \frac{2}{T} p_j$$

where T is the length of a function interval, $j = 0, \dots, l - 1$.

Finally, to derive local morphometric variables, all the calculated values of the derivatives are substituted into related equations, for instance, the k_h equation (3).

3 Algorithm

Let the initial array is specified as a matrix A with the dimension $m \times n$ representing the function values at the nodes of a square grid; x_1, \dots, x_m are the values of the grid nodes in the interval $[-1, 1]$ along abscissa, y_1, \dots, y_n are the values of the grid nodes in the interval $[-1, 1]$ along ordinate; and t_1, \dots, t_k are the Gaussian quadrature nodes.

The k values determine the maximum degree of expansion l in Eq. (4). Since $m \approx n$, so, the Gaussian grid is identical for abscissas and ordinates. For a more accurate calculation of the expansion coefficients, it is recommended to choose the k value greater than m and n . In this study, $k = 8 \max(m, n)$.

Let introduce the following notations: L_{xt} is the matrix of linear interpolation from the grid x to the grid t ; T_t is the matrix of the Chebyshev polynomial values $T_i(t_j)$ in the grid t , where $i = 0, \dots, l - 1$, $j = 1, \dots, k$; F is the diagonal matrix with diagonal elements $(l - i)/l$, $i = 0, \dots, l - 1$, to derive the arithmetic mean of the sums of orthogonal series.

The calculation of the expansion coefficients is carried out in two stages. First, all columns of the matrix A are transformed into the expansion coefficients that corresponds to the expansion in the variable y :

$$C = \frac{2}{k} F T_t L_{yt} A.$$

Then, the matrix of the expansion coefficients is transposed and the expansion is repeated that corresponds to the expansion in the variable x :

$$D = \frac{2}{k} F T_t L_{xt} C^T$$

where C and D are the matrices of the expansion coefficients, the intermediate and the resulting ones.

Reconstruction of the approximated function is carried out by a simple summation of orthogonal series:

$$Z = T_x^T D^T T_y.$$

Let E be the differentiation operator in the space of expansion coefficients. Then the coefficient matrices corresponding to the partial derivatives p , r , q , t , and s (2) are as follows:

$$P = ED; \quad R = E^2 D; \quad Q = DE^T; \quad T = D(E^T)^2; \quad S = PE^T.$$

These formulas exist only if the matrices of expansion coefficients are square.

4 Data

To test the method and algorithm proposed, a portion of the Northern Andes measuring $4^\circ \times 4^\circ$, located between 2°S and 2°N , and $78^\circ 30' \text{W}$ and $74^\circ 30' \text{W}$, was selected. The area

covers regions of Ecuador, Colombia, and Peru including the parts of the Coastal plain, the Andean Range, and the Upper Amazon basin (Fig. 1a). A DEM of the study area was extracted from the global DEM GTOPO30 [34]. The DEM contains 230,880 points (the matrix 480×481); the grid spacing is $30''$ (Fig. 1a).

This area and GTOPO30 were selected because this DEM incorporates a high-frequency noise caused by interpolation errors and inaccurate merging of topographic charts having different accuracy. Spatial distribution of the noise in GTOPO30 is uneven and depends on the accuracy of cartographic sources. In particular, the potent noise is typical for forested regions of South America because reasonably detailed and accurate topographic data were unavailable for such areas. Thus, interpolation of sparse contours has been used to compile these portions of GTOPO30. Although DEM noise has no obstacle to produce realistic maps of elevation, it leads to derivation of noisy and unreadable maps of local morphometric variables (computation of the first and the second partial derivatives of elevation dramatically increases the noise [35]). The study area, consisting of two main zones – high mountains and forested foothills – which have different signal-to-noise ratio, is ideally suited to validate the method and algorithm proposed as a tool for DEM analytical treatment. Earlier, this DEM was used to evaluate 2D singular spectrum analysis as a tool to filter DTMs [14].

5 Data Processing

Various numbers of expansion coefficients were evaluated to reconstruct DEMs. Finally, the most expressive variants were selected to illustrate capabilities of the method and algorithm to generalize and denoise DEMs as well as to calculate partial derivatives for further computation of local morphometric attributes. In particular, DEMs were reconstructed with 480, 240, 120, 60, and 30 expansion coefficients (Figs. 1a–1e). The first and the second partial derivatives were analytically calculated from the reconstructed DEMs. Digital models of 14 local morphometric attributes were then computed using the derivatives. The models of slope gradient, slope aspect, minimal curvature, maximal curvature, mean curvature, the Gaussian curvature, unsphericity curvature, horizontal curvature, vertical curvature, difference curvature, accumulation curvature, ring curvature, vertical excess curvature, and horizontal excess curvature were derived. To illustrate efficiency of the method and algorithm, k_h maps are presented (Figs. 1f–1j).

To estimate the approximation accuracy, deviations of the reconstructed DEMs were also calculated from the initial DEM (Fig. 2).

Wide dynamic ranges usually characterize topographic variables. To avoid loss of information on spatial distribution of values of morphometric attributes in mapping, it makes sense to apply a logarithmic transform using the following expression [15]:

$$\Theta' = \text{sign}(\Theta) \ln(1 + 10^n |\Theta|) \quad (5)$$

where Θ is the value of a morphometric variable; and $n = 0$ for elevation and nonlocal variables and $n = 2, \dots, 18$ for local variables. For the DEM analytical treatment, selection of the n value depends on the size of a study area ($n = 8$ was used for k_h mapping). Such a form of logarithmic transformation considers that dynamic ranges of some topographic attributes include both positive and negative values. All DEMs and k_h digital models derived were logarithmically transformed to prepare readable maps of elevation and k_h (Fig. 1).

The DTMs produced had a grid size of $30''$. The plate carrée projection was used for mapping. Data processing and mapping were done by software Matlab R2008b.

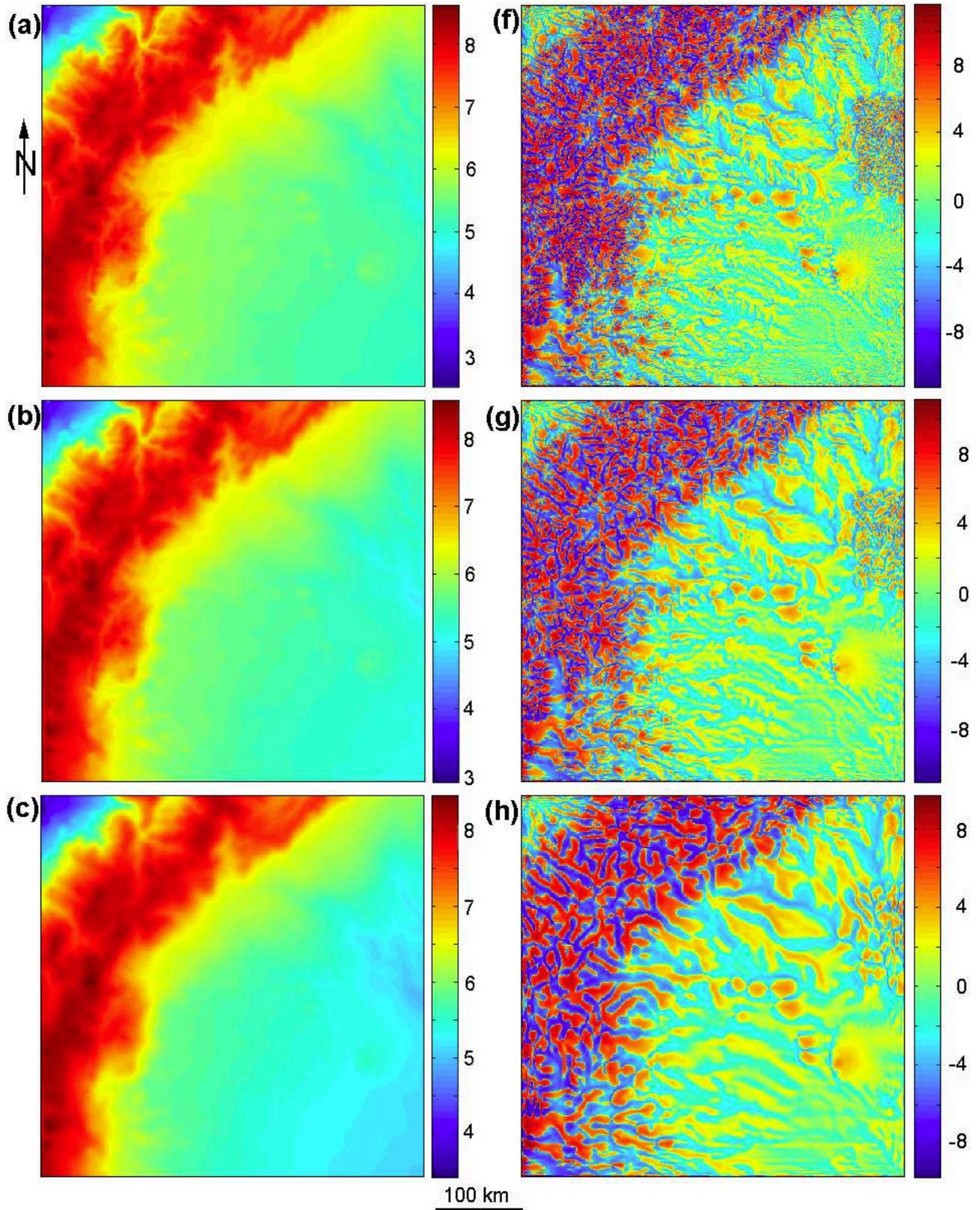


Figure 1 The Northern Andes: elevation maps reconstructed with 480 (a), 240 (b), 120 (c), 60 (d), and 30 (e) expansion coefficients; and k_h maps derived from reconstructed DEMs with 480 (f), 240 (g), 120 (h), 60 (i), and 30 (j) expansion coefficients. The legends are in logarithmic scale (*to be continued*)

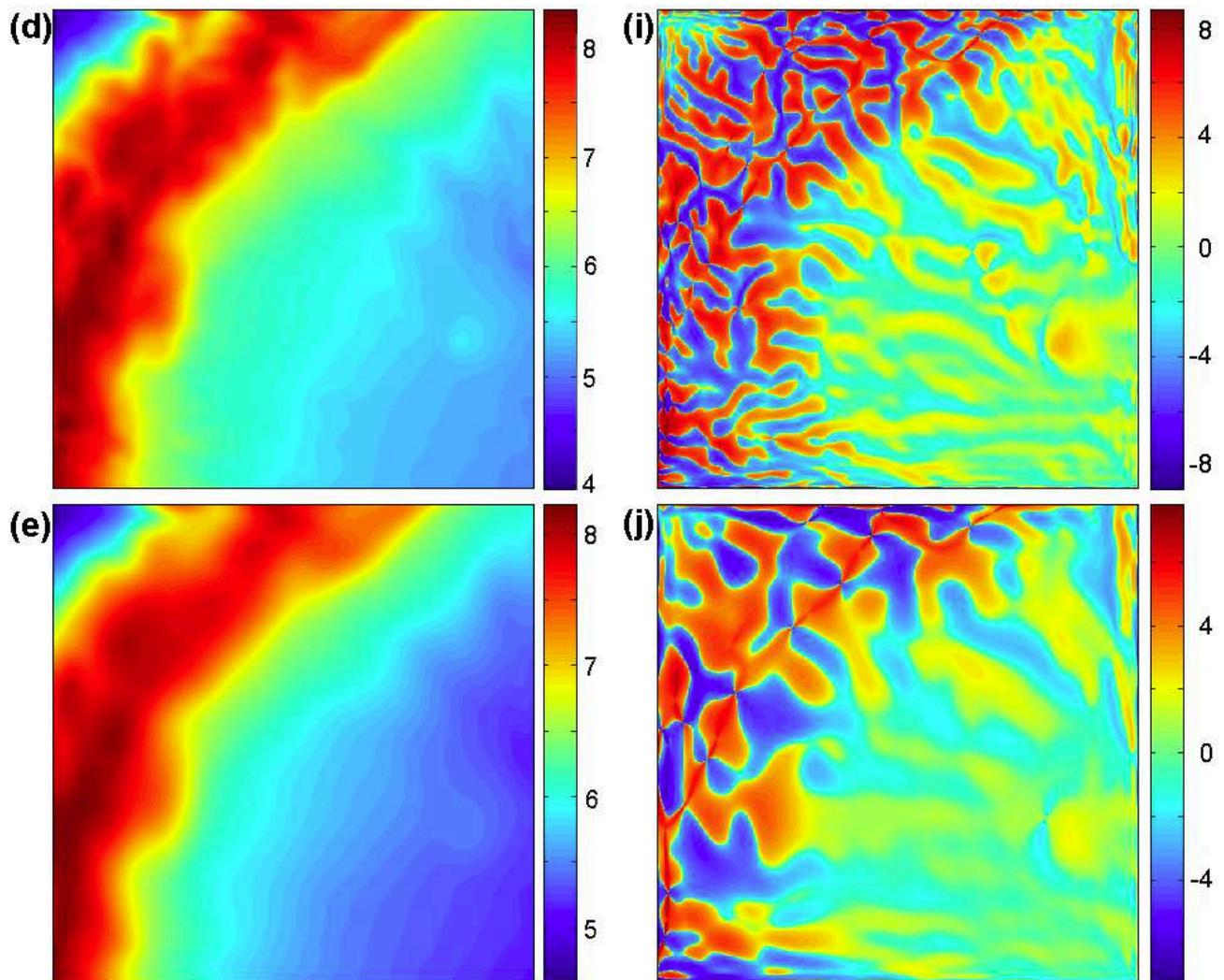


Figure 1 (continued)

6 Results and Discussion

6.1 Visual analysis of maps obtained

A set of elevation maps reconstructed with 480, 240, 120, 60, and 30 expansion coefficients (Figs. 1a–1e) demonstrates a process of DEM generalization, from its minimal level (Fig. 1a) to the maximal one (Fig. 1e). Similarly, a set of k_h maps calculated from DEMs reconstructed with 480, 240, 120, 60, and 30 expansion coefficients (Figs. 1f–1j) shows k_h generalization, from its minimal level (Fig. 1f) to the maximal one (Fig. 1j).

The visual comparison of elevation maps reconstructed with 480, 240, and 120 expansion coefficients (Figs. 1a–1c) allows one to see nothing but marginal changes in image patterns. A cursory examination may lead to an underestimation of results of the DEM generalization. The k_h maps give better insight into the results. A comparison between k_h maps derived from different DEMs (Figs. 1f–1j) shows a pronounced effect of the map generalization. The less number of the expansion coefficients used to reconstruct a DEM, the more smooth and simplified image patterns obtained. One can see the so-called flow structures formed by convergence and divergence areas (negative and positive k_h values, respectively).

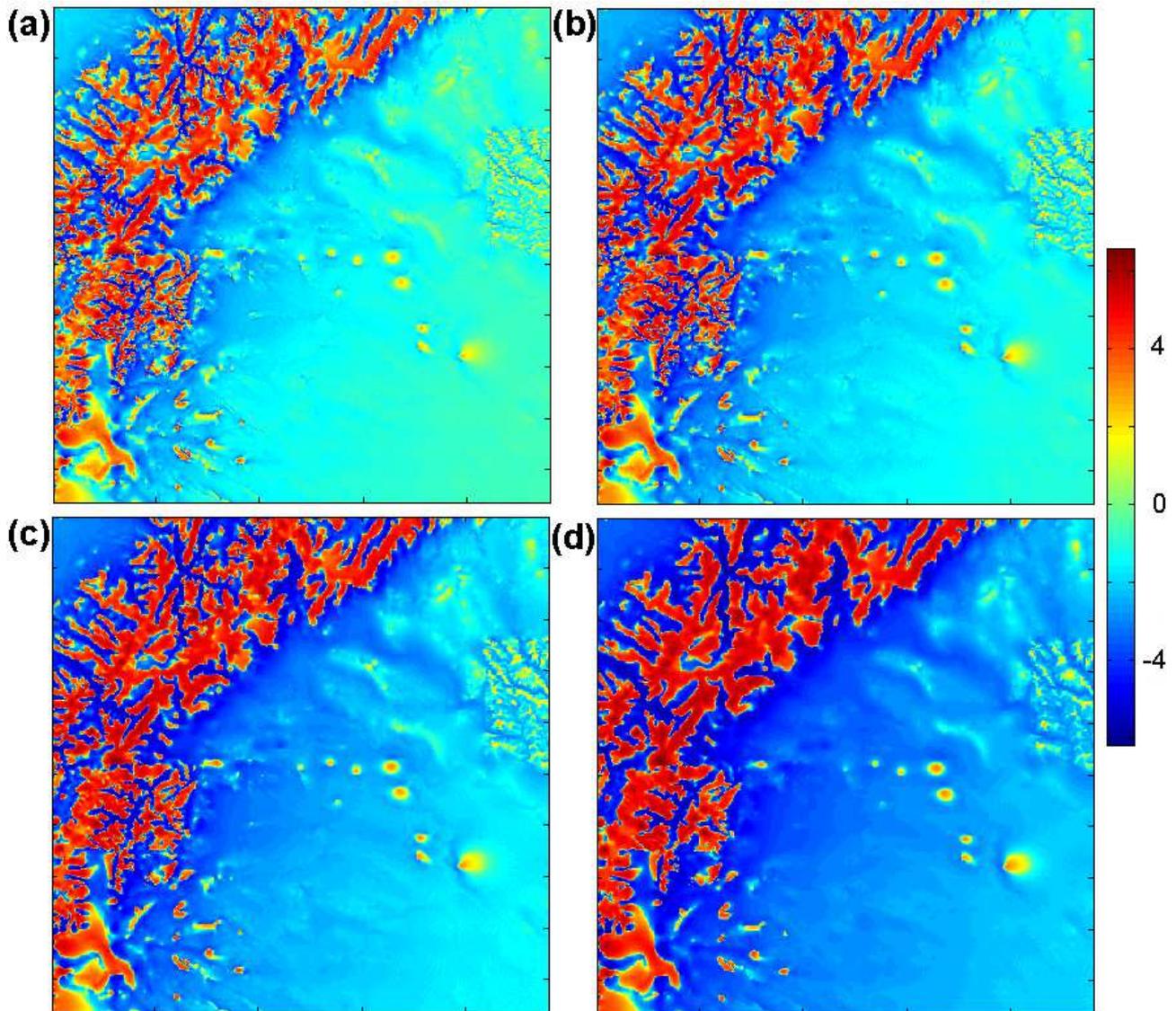


Figure 2 Deviations of the reconstructed DEMs from the initial DEM. Reconstructions were done with 480 (a), 240 (b), 120 (c), and 60 (d) expansion coefficients. The legend is in logarithmic scale

The DEMs reconstructed with 60 and 30 expansion coefficients (Figs. 1d and 1e) are marked by the highest level of generalization. These elevation maps represent a generalized morphostructure of the continental scale, the Andean Range with foothills.

The decrease of the number of the expansion coefficients in DEM reconstruction acts as a high-frequency filtering. Indeed, manifestation of high-frequency noise can be found on the k_h map derived from the DEM reconstructed with 480 expansion coefficients (see the bottom right corner in Fig. 1f). This noise is typical for the Andean foothills covered by dense rain forests. However, there are no traces of this noise on k_h maps derived from DEMs reconstructed with less number of expansion coefficients (Figs. 1g–1j). One can also observe the removal of an artifact of other sort: a rectangular feature along the northeastern border of several k_h maps (Figs. 1f–1h). This is a trace of the inaccurate merging of adjacent topographic charts marked by different accuracy during the compilation of GTOPO30. All traces of the artifact disappear on k_h maps calculated from DEMs reconstructed with 60 and 30 expansion coefficients (Figs. 1i and 1j).

It is well known that orthogonal polynomial approximation produces boundary effects which cannot be completely eliminated. In the present case, they appear as linear artifacts at the boundaries of k_h maps (Figs. 1f–1h). A problem with map boundaries also arises using the finite difference algorithms: it is impossible to estimate partial derivatives for boundary columns and rows of a DEM, because they are estimated for the center point of the moving window.

Note that the above mentioned noise and artifacts are weakly expressed: they are not visible on the elevation maps (Figs. 1a–1e) and become clearly visible only on k_h maps (Figs. 1f–1j) after the calculation of partial derivatives of elevation which reinforce their expression.

6.2 Accuracy of DEM approximation and morphometric calculation

In digital terrain modeling, it is undesirable to use statistical metrics describing the accuracy of a DTM by a single value. Indeed, the values of elevation and other morphometric variables can be unevenly distributed in space that leads to the spatial variability of the DTM accuracy. To estimate the DTM accuracy, it is appropriate to produce the maps of statistical metrics characterizing DTM accuracy or errors at each point of a DTM (e.g., Fig. 2).

The analysis of Fig. 2 shows that maximum deviations of reconstructed DEMs predictably occur in areas of the contrast mountainous topography with dramatic elevation changes. The use of the global approximation with the Fejér summation leads to a strong smoothing of the elevation function within such areas. However, it is necessary to remember that the goal is twofold: on the one hand, it is required to hold a sufficient approximation accuracy of elevation values using large sets of expansion coefficients in DEM reconstruction and, on the other hand, it is required to suppress high-frequency noise during DEM reconstruction. Since these goals to some extent contradict one another, the key criterion is a sufficient plausibility of reconstructed elevation maps with a simultaneous satisfactory noise reduction.

It is clear that the smaller the number of expansion coefficients used to reconstruct DEMs, the greater the deviation of a DEM reconstructed from the initial DEM (Figs. 2b–2d). We generalize the topographic surface reducing the number of expansion coefficients (Figs. 1c–1e), so the approximation accuracy problem becomes irrelevant.

The local morphometric variables are the functions of measured variables (elevations). Thus, to estimate the accuracy of calculation of local morphometric attributes, one can use the criterion of root mean square error (RMSE) of a function of measured variables. The RMSE equation for k_h is as follows [4, 36]:

$$m_{k_h} = \frac{1}{p^2 + q^2} \left(\frac{1}{1 + p^2 + q^2} \left\{ m_p^2 \left[p (q^2 r - 2pqs + p^2 t) \left(\frac{2}{p^2 + q^2} + \frac{1}{1 + p^2 + q^2} \right) + 2 (qs - pt) \right]^2 \right. \right. \right. \\ \left. \left. + m_q^2 \left[q (q^2 r - 2pqs + p^2 t) \left(\frac{2}{p^2 + q^2} + \frac{1}{1 + p^2 + q^2} \right) + 2 (ps - qr) \right]^2 \right. \right. \\ \left. \left. + m_r^2 q^4 + 4m_s^2 p^2 q^2 + m_t^2 p^4 \right\} \right)^{1/2} \quad (6)$$

where $m_p, m_q, m_r, m_s,$ and m_t are the RMSEs of $p, q, r, s,$ and t (2), respectively. The RMSE equations of other local morphometric variables can be found elsewhere [4].

Equation (6) can be applied to produce a digital model and map of m_{k_h} . Earlier, analysis of m_{k_h} maps allowed to determine some regularity for spatial distribution of m_{k_h} values. Generally, the flatter topography, the higher m_{k_h} value [18, 36]. The m_{k_h} values can be in excess of maximum absolute values of k_h within flat areas. However, this does not mean that there are errors in DTMs within such areas: m_{k_h} is a statistical property of a function k_h ; that is, m_{k_h} values represent a possibility for errors.

Table 1 The average processing time to reconstruct a DEM and to derive a morphometric variable (seconds) depending on the number of expansion coefficients used

| Number of coefficients | DEM reconstruction | DEM reconstruction + k_h derivation |
|------------------------|--------------------|---------------------------------------|
| 30 | 0.05 | 0.06 |
| 60 | 0.06 | 0.07 |
| 120 | 0.07 | 0.08 |
| 240 | 0.08 | 0.09 |
| 480 | 1.40 | 1.50 |
| 960 | 3.10 | 3.20 |
| 1200 | 4.40 | 4.50 |
| 2400 | 13.60 | 13.70 |
| 3000 | 20.50 | 20.60 |

6.3 Efficiency of the method and algorithm

To estimate the efficiency of the method and algorithm developed, an office computer with modest facilities (Intel Celeron, CPU G460, 1.80 GHz, 1.95 GB RAM) was used. Test results are presented in Table 1. It is clearly seen that most of the time is spent on the DEM reconstruction; it nonlinearly depends on the number of expansion coefficients used. In this regard, the time for calculation of partial derivatives and local morphometric variables is absolutely negligible.

Earlier, the initial DEM was used to evaluate 2D singular spectrum analysis as a tool to generalize, smooth, and denoise DTMs; the software 2D-SSA was utilized [14]. The processing time and DEM treatment results were, in general, comparable to those presented in this paper. However, 2D singular spectrum analysis does not allow one to calculate analytically partial derivatives and morphometric variables.

The processing time to derive a local morphometric variable by a finite difference algorithm from the initial DEM was estimated as about 0.5 s (the software LandLord [4] was applied). So, the efficiency of the developed method and algorithm is comparable with that of existing finite difference ones.

The advantage of the developed method and algorithm is the possibility to apply several procedures of DTM treatment — DEM approximation, generalization, smoothing, as well as calculation of partial derivatives and morphometric variables — within the framework of a single approach. These procedures are usually implemented in individual algorithms, utilities, or even programs. Such integration allows one to save time for DTM processing.

7 Concluding Remarks

The study demonstrated a good performance of the developed method and algorithm. They can be utilized as a universal tool for analytical treatment of regularly spaced DEMs including DEM global approximation, denoising, and generalization as well as derivation of local morphometric variables from DEMs using analytical calculation of partial derivatives. Further development of the method will include: (a) incorporation of other high-order orthogonal polynomials (e. g., the Fourier and Legendre ones); (b) their comparative analysis in terms of efficiency and accuracy of DEM treatment; (c) evaluation of calculation accuracy using appropriate statistical metrics (e. g., RMSE of a function of the measured values); and (d) comparison of the developed method and algorithm with traditional techniques applied in digital terrain modeling.

References

- [1] Wilson, J.P., and J. C. Gallant, eds. 2000. *Terrain analysis: Principles and applications*. New York, NY: Wiley. 479 p.
- [2] Li, Z., Q. Zhu, and C. Gold. 2005. *Digital terrain modeling: Principles and methodology*. New York, NY: CRC Press. 323 p.
- [3] Hengl, T., and H. I. Reuter, eds. 2009. *Geomorphometry: Concepts, software, applications*. Amsterdam: Elsevier. 796 p.
- [4] Florinsky, I. V. 2012. *Digital terrain analysis in soil science and geology*. Amsterdam: Elsevier / Academic Press. 379 p.
- [5] Schut, G.H. 1976. Review of interpolation methods for digital terrain models. *Can. Surv.* 30(5):389–412.
- [6] Watson, D. 1992. *Contouring: A guide to the analysis and display of spatial data*. Oxford: Pergamon. 340 p.
- [7] Mitas, L., and H. Mitasova. 1999. Spatial interpolation. *Geographical information systems: Principles, techniques, management and applications*. 2nd ed. Hoboken: Wiley. 481–492.
- [8] Papo, H. B., and E. Gelbman. 1984. Digital terrain models for slopes and curvatures. *Photogramm Eng. Remote Sens.* 50(6):695–701.
- [9] Harrison, J.M., and C.-P. Lo. 1996. PC-based two-dimensional discrete Fourier transform programs for terrain analysis. *Comp. Geosci.* 22(4):419–424.
- [10] Arrell, K., S. Wise, J. Wood, and D. Donoghue. 2008. Spectral filtering as a method of visualising and removing striped artefacts in digital elevation data. *Earth Surf. Proc. Landforms* 33(6):943–961.
- [11] Bergbauer, S., T. Mukerji, and P. Hennings. 2003. Improving curvature analyses of deformed horizons using scale-dependent filtering techniques. *Amer. Assoc. Petrol Geol. Bull.* 87(8):1255–1272.
- [12] Bjørke, J. T., and S. Nilsen. 2003. Wavelets applied to simplification of digital terrain models. *Int. J. Geogr. Inf. Sci.* 17(7):601–621.
- [13] Wu, F. 2003. Scale-dependent representations of relief based on wavelet analysis. *Geo-Spat. Inf. Sci.* 6(1):66–69.
- [14] Golyandina, N. E., K. D. Usevich, and I. V. Florinsky. 2007. Filtering of digital terrain models by two-dimensional singular spectrum analysis. *Int. J. Ecol. Devel.* 8(F07):81–94.
- [15] Shary, P. A., L. S. Sharaya, and A. V. Mitusov. 2002. Fundamental quantitative methods of land surface analysis. *Geoderma* 107(1-2):1–32. doi:10.1016/S0016-7061(01)00136-7.
- [16] Evans, I. S. 1979. *Statistical characterization of altitude matrices by computer. An integrated system of terrain analysis and slope mapping*. Final Report on Grant DA-ERO-591-73-G0040. Durham: Department of Geography, University of Durham. 192 p.
- [17] Zevenbergen, L. W., and C. R. Thorne. 1987. Quantitative analysis of land surface topography. *Earth Surf. Proc. Landforms* 12(1):47–56. doi:10.1002/esp.3290120107.
- [18] Florinsky, I. V. 1998. Derivation of topographic variables from a digital elevation model given by a spheroidal trapezoidal grid. *Int. J. Geogr. Inf. Sci.* 12(8):829–852. doi:10.1080/136588198241527.
- [19] Florinsky, I. V. 2009. Computation of the third-order partial derivatives from a digital elevation model. *Int. J. Geogr. Inf. Sci.* 23(2):213–231. doi:10.1080/13658810802527499.
- [20] Heifetz, B. S. 1964. Approximation of the topographic surface by the Chebyshev orthogonal polynomials. *Izv. Vuzov. Geod. Aerofoto* 2:78–86. (In Russian.)
- [21] Van Rossel, J. 1972. Digital hypsographic map compilation. *Photogramm Eng.* 38(11):1106–1116.

- [22] Yagodina, L. L. 1972. Modelling of the topographic field by the Chebyshev orthogonal polynomials. *Vest. Leningrad Univ. Ser. Geogr.* 24(4):136–142 (In Russian.)
- [23] Segu, W. P. 1985. Terrain approximation by fixed grid polynomial. *Photogramm Rec.* 11(65):581–591. doi:10.1111/j.1477-9730.1985.tb00525.x.
- [24] Chorley, R. J., and P. Huggett. 1965. Trend-surface mapping in geographical research. *Trans. Inst. Brit. Geogr.* 37:47–67. doi:10.2307/621689.
- [25] Tobler, W. R. 1969. Geographical filters and their inverses. *Geogr. Anal.* 1(3):234–253. doi:10.1111/j.1538-4632.1969.tb00621.x.
- [26] Dedus, F. F., S. A. Makhortykh, M. N. Ustinin, and A. F. Dedus. 1999. *Generalized spectral analytical method for data set processing. Problems of image analysis and pattern recognition.* Moscow: Mashinostroenie. 356 p. (In Russian, with English abstract.)
- [27] Dedus, F. F., L. I. Kulikova, A. N. Pankratov, and R. K. Tetuev. 2004. *Classical orthogonal bases in problems of analytical description and processing of information signals.* Moscow: Moscow State Univ. 168 p. (In Russian.)
- [28] Tetuev, R. K., and F. F. Dedus. 2007. *Classical orthogonal polynomials. Application to data processing problems.* Pushchino: Institute of Mathematical Problems of Biology, Russian Academy of Sciences. 60 p. (In Russian.)
- [29] Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical recipes in C. The art of scientific computing.* 2nd ed. Cambridge: Cambridge Univ. Press. 994 p.
- [30] Pankratov, A. N. 2004. On the implementation of algebraic operations on orthogonal function series. *Comp. Math. Math. Phys.* 44(12):2017–2023.
- [31] Jerri, A. J. 1998. *The Gibbs phenomenon in Fourier analysis, splines and wavelet approximations.* Boston: Kluwer. 336 p. doi:10.1007/978-1-4757-2847-7.
- [32] Courant, R., and D. Hilbert. 1989. *Methods of mathematical physics.* Vol. 1. New York, NY: Wiley. 560 p. doi:10.1002/9783527617210.
- [33] Pankratov, A. N., and L. I. Kulikova. 2006. On the derivatives calculation under uniform and mean square approximation of signals. *Computer Applications in Scientific Researches Proceedings, IVTN-2006 Session.* Moscow. 56.
- [34] U. S. Geological Survey. 1996. GTOPO30, a 30-arc seconds global digital elevation model. Available at: <https://1ta.cr.usgs.gov/GTOPO30> (accessed August 12, 2015).
- [35] Florinsky, I. V. 2002. Errors of signal processing in digital terrain modelling. *Int. J. Geogr. Inf. Sci.* 16(5):475–475. doi:10.1080/13658810210129139.
- [36] Florinsky, I. V. 1998. Accuracy of local topographic variables derived from digital elevation models *Int. J. Geogr. Inf. Sci.* 12(1):47–61. doi:10.1080/136588198242003.

Received June 15, 2015

Двумерные вариации как средство оценивания сложности изображений*

П. А. Чочиа

chochia@iitp.ru

Институт проблем передачи информации РАН, Москва 127051, Большой Каретный пер., д. 19

Исследованы вопросы оценки *сложности* изображения, которая интерпретируется как характеристика, отражающая число, размеры и заметность его деталей. Исследованы возможности использования для этого двумерных вариаций. Рассмотрены модификации известных двумерных вариаций в применении к дискретным изображениям. Предложена оценка, названная *показатель размеров объектов*. Теоретические выводы подтверждены экспериментальными исследованиями. Проведен анализ изменения значений двумерных вариаций изображения при увеличении шума, а также при сглаживании и декомпозиции. Показано, что предложенная комбинация двумерных вариаций отражает морфологическую структуру изображения и может служить оценкой его сложности.

Ключевые слова: *двумерные вариации; сложность изображения; обработка видеoinформации*

DOI: 10.21469/22233792.1.12.02

Two-dimensional variation as an image complexity assessment*

P. A. Chochia

Institute for Information Transmission Problems RAS, 19 Bolshoy Karetny, Moscow, Russia
The questions of *image complexity* assessment and using of two-dimensional variations are studied. The image complexity is interpreted as some attribute which is specified by the quantity, the sizes, and the visibility of image details. Different known two-dimensional variation methods are considered in their application to digital images. The modified assessment is proposed that is named as the *component size index*. The change of variations under different image transformations are analyzed. Theoretical conclusions are confirmed by the experimental explorations. The proposed combination of two-dimensional variations was demonstrated to reflect morphological structure of an image and to assess its complexity.

Keywords: *two-dimensional variation; image complexity; image processing*

DOI: 10.21469/22233792.1.12.02

1 Введение

Применение методов обработки видеoinформации предполагает оценку получаемого результата. В задачах, которые удается формулировать как задачи восстановления или фильтрации сигнала, для этой цели часто используются различные варианты отклонений (MSE — mean square error, PSNR — peak signal-to-noise ratio, SSIM — structural similarity и др.), которые удобны в теоретических разработках, а также в случаях, когда исходный сигнал известен. В реальных же ситуациях неискаженный сигнал недоступен. При

*Работа выполнена при финансовой поддержке Российского научного фонда, проект № 14-50-00150.

таких условий хотелось бы найти численную оценку результата, которая была бы адекватна изображению как двумерной функции и при этом не использовала бы оригинальный сигнал.

Для решения подобной задачи, как правило, предлагаются методы, использующие один из следующих двух подходов. Первый заключается в попытке решения задачи оценивания качества изображения [1, 2]. Второй чаще всего связан с независимым оцениванием сложности сигнала, но обычно его используют при существенных ограничениях, например в задаче сжатия [3] или в задаче автоматического опознавания целей [4].

Применительно к изображению, которое является двумерным сигналом, также хотелось бы выбрать оценку, характеризующую его пространственную изменчивость. Плодотворным здесь представляется подход, основанный на рассмотрении изображения как функции с ограниченной вариацией [5] и оценивании ее изменчивости при помощи механизма двумерных и многомерных вариаций [6–9]. Вопрос возможности применения двумерных вариаций для оценки изображения рассматривался в статьях [10–12]; настоящая работа является продолжением указанных исследований.

В работе рассматриваются варианты реализации различных двумерных вариаций в применении к дискретным сигналам и исследованы возможности их использования для оценки сложности изображений. Сложность изображения интерпретируется как характеристика, отражающая число, размеры и заметность (контраст) его деталей. Также полагается, что при увеличении числа и контраста деталей сложность изображения должна возрастать, а при увеличении размеров деталей — уменьшаться.

Понятие сложности трактуется по-разному для сигналов, функций и потока сообщений. Одномерный сигнал, являющийся действительной функцией времени, характеризуется энергией, длительностью и шириной спектра [13]. Сложность аналитической функции многих переменных принято характеризовать числом ее переменных n , а для s раз дифференцируемых функций — отношением n/s [14]. Теоретико-информационный подход к оценке сложности потока сообщений, представимого в дискретном виде, основан на вычислении энтропии $H = -\sum p(z) \log(p(z))$, где $p(z)$ — вероятность элемента изображения со значением z . Иногда для оценивания сложности дискретного сигнала используется коэффициент сокращения объема данных (количества бит), который требуется для восстановления. Такой подход, идеологически близкий понятию колмогоровской сложности [15], связан с необходимостью выбора того или иного алгоритма сжатия и по сути является вариантом энтропийного оценивания.

Каждый из указанных способов оценки сложности по тем или иным причинам мало или даже совсем не применим к изображению как двумерной функции яркости пространственных координат, отражающей некоторую физическую характеристику наблюдаемой сцены. В частности, существенный общий недостаток перечисленных вариантов оценок заключается в том, что они не являются метрическими характеристиками и не зависят от амплитуды сигнала.

2 Двумерные вариации и их дискретные аналоги

Распространенной метрической характеристикой изменчивости и сложности одномерной функции на отрезке $[a, b]$ является вариация. Считая функцию $f(x)$ отражением некоторой реальной физической величины (например, энергии или плотности), можно полагать, что $f(x)$ на наблюдаемом отрезке ограничена и имеет конечное число точек разрыва первого рода. Тогда $f(x)$ является функцией с ограниченной полной вариацией, которая по определению есть величина

$$V_a^b(z(x)) \triangleq \sup_P \sum_{k=0}^{K-1} \|z(x_{k+1}) - z(x_k)\|, \quad (1)$$

т. е. точная верхняя грань по всем возможным разбиениям P отрезка $[a, b]$. В контексте кусочно-постоянной модели изображения [16] существенно отметить, что значение вариации не изменяется при «вставке» в функцию $z(x)$ участков вида $z(x) = \text{const}$, $\{x \in [c, d] \subset [a, b]\}$, если при этом не добавляются разрывы первого рода.

Полная вариация $V_a^b(z(x))$ может служить удовлетворительной оценкой сложности одномерной функции. Предложено много ее обобщений на случай функции многих переменных (вариации Витали, Арцела, Фреше, Тонелли и др.). Все они сводятся к тому, что определяется некоторый функционал, ограниченность которого гарантирует наличие у функции ряда свойств, аналогичных свойствам одномерной функции с конечным изменением. Однако перечень таких свойств оказывается ощутимо беднее набора свойств вариации для одномерных функций, хотя при этом и удается формулировать отдельные выводы и теоремы [8]. К тому же возникает определенная неоднозначность многомерных вариаций, в частности связанная с тем, что они существенно зависят от выбора системы координат.

Многомерные вариации (по аналогии с одномерными) формулируются для непрерывных функций как точные верхние грани выбранных функционалов по множеству допустимых разбиений носителя текущими гиперплоскостями на элементарные параллелепипеды. В случае дискретной функции неприятность состоит в том, что минимальное разбиение ограничено снизу пространственной точностью ее представления. Поэтому формулы вариации дискретных функций являются лишь соответствующими аналогами и приближениями формул вариации непрерывных функций, что в некоторых случаях может вызвать определенные неточности. Также необходимо отметить, что далеко не все операции могут быть применимы к дискретной функции без потери информации: например, операции поворота на произвольный угол или операции сжатия как в области носителя, так и в пространстве значений.

Рассмотрим, как будут выглядеть наиболее известные определения вариаций [8, 9, 17] в применении к двумерной дискретной функции $\mathbf{F} = \{f(i, j)\}$ ($i = 1, \dots, I; j = 1, \dots, J$) на прямоугольном носителе D с соответствующим разбиением.

Арцела вариация [8, 17] по сути является суммой модулей градиентов по координатным осям и представляется формулой:

$$A(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} (|f(i+1, j) - f(i, j)| + |f(i, j+1) - f(i, j)|). \quad (2)$$

Данный способ оценки изменчивости двумерной функции используется в большинстве теоретических и практических разработок.

Витали вариация [8, 17] формально записывается как

$$V(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} |f(i+1, j+1) + f(i, j) - f(i, j+1) - f(i+1, j)|. \quad (3)$$

В качестве существенного недостатка данного *дискретного* представления вариации следует отметить, что на участках функции $f(i, j)$, где наблюдаются перепады ее значений в направлении только одной из координатных осей (i или j), отклик оператора, соответствующего формуле (3), будет равен нулю.

Пьерпонта вариация [17, 18] предполагает равномерное разбиение каждой из сторон прямоугольника D на равные между собой отрезки, что идеологически соответствует равномерной дискретизации носителя. В каждом образуемом элементарном параллелепипеде $\delta(i, j)$ измеряется величина колебания $\omega(\mathbf{F}, \delta(i, j))$, а вариация вычисляется как

$$P(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} \omega(\mathbf{F}, \delta(i, j)). \tag{4}$$

Согласно [18], колебание ω есть $\omega(\mathbf{F}, \delta(i, j)) = \max\{f(x, y) \in \delta(i, j)\} - \min\{f(x, y) \in \delta(i, j)\}$, что в дискретном случае можно определить как

$$\omega(\mathbf{F}, \delta(i, j)) = \max\{|f(i+1, j) - f(i, j)|, |f(i, j+1) - f(i, j)|\}.$$

Тонелли вариация [8, 17] определяется как сумма одномерных вариаций (1) по всем значениям i и j каждой из координат при постоянном значении второй координаты. Пусть $\psi_1(i)$ — одномерная вариация функции $f(i, j)$ по j при $i = \text{const}$, а $\psi_2(j)$ — одномерная вариация функции $f(i, j)$ по i при $j = \text{const}$. Тогда

$$T(\mathbf{F}, D) = \sum_{i=1}^I \psi_1(i) + \sum_{j=1}^J \psi_2(j). \tag{5}$$

Нетрудно убедиться, что значения, вычисляемые по формулам (5) и (2) будут равны, т. е. вариации Арцела и Тонелли для дискретных функций совпадают.

Фреше вариация [17] в формулировке для двумерной дискретной функции совпадает с Харди вариацией [8, 17] и вычисляется следующим образом:

$$\begin{aligned} \Delta_1(f; (i, j)) &= f(i+1, j) - f(i, j); \\ \Delta_2(f; (i, j)) &= \Delta_2(\Delta_1(f; (i, j))) = (f(i+1, j+1) - f(i, j+1)) - (f(i+1, j) - f(i, j)); \end{aligned}$$

$$F(\mathbf{F}, D) = H(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} |\Delta_2(\Delta_1(f; (i, j)))|. \tag{6}$$

Как легко заметить, формулы (6) и (3) совпадают, т. е. вариации Витали, Фреше и Харди для дискретных функций будут между собой равны, в то время как для непрерывных функций они различаются.

Рассмотренные вариации так или иначе имеют в своей основе значение модуля градиента функции в точке, поэтому их величины и поведение близки между собой.

Кронрода вариация [7] в ряду вариаций многомерных функций стоит особняком по той причине, что не использует градиентные характеристики функции, а также дает в результате значения не одного, а нескольких функционалов. Ее рассмотрению посвящен следующий раздел.

3 Двумерная вариация Кронрода

Обобщения отдельных выводов и теорем, которые удавалось формулировать, используя определения вышеперечисленных вариаций (для непрерывных функций), привели к выводу, что функция многих переменных должна характеризоваться не одним, а несколькими функционалами, которые в определенном смысле независимы. Данный

тезис был обоснован А. С. Кронродом при изучении функций двух переменных [7]. Основываясь на понятии множеств уровня, он предложил для функции двух переменных использовать два функционала, определяемые следующим образом:

$$w_1(f) = \int_{-\infty}^{\infty} v_0(e_t) dt; \quad w_2(f) = \int_{-\infty}^{\infty} v_1(e_t) dt. \quad (7)$$

Здесь множество e_t — это t -уровень функции $f(x, y)$, т. е. множество тех точек (x, y) , в которых $f(x, y) = t$; $v_0(e_t)$ — число компонент множества e_t ; $v_1(e_t)$ — длина e_t (по Хаусдорфу). Тем самым значение w_1 характеризует число локальных экстремумов функции.

Витушкиным было найдено удачное определение вариации множества, причем данный подход, также основанный на рассмотрении множеств уровня, был расширен для функций многих переменных [8]. Для плоского множества e_t , которое является t -уровнем функции $f(x, y)$, вариация задается не одним, а тремя значениями: v_0 , v_1 и v_2 , которые определяются следующим образом.

Пусть E — замкнутое ограниченное множество на плоскости; тогда вариация $v_0(E)$ есть число компонент E , вариация $v_2(E)$ есть суммарная площадь компонент E , а значение вариации $v_1(E)$ определяется следующим выражением [8, 9]:

$$v_1(E) = c \int_0^{2\pi} v_0(E, L_\alpha) d\alpha,$$

где

$$v_0(E, L_\alpha) = \int_{L_\alpha} v_0\left(E \cap (L_\alpha)^\perp_z\right) dz.$$

Здесь L_α — прямая $x \cos \alpha + y \sin \alpha = 0$, а $(L_\alpha)^\perp_z$ — прямая, проходящая через точку $z \in L_\alpha$ перпендикулярно к L_α . Константа c выбирается таким образом, чтобы для отрезка I единичной длины на оси Ox выполнялось соотношение: $v_1(I) = 1$. Для случая конечной $v_0(e_t)$, вариацию $v_1(e_t)$ можно интерпретировать как длину границ компонент e_t . Вариации v_0 , v_1 и v_2 названы соответственно *нулевой*, *линейной* и *плоской* согласно тому, как изменяются их значения при гомеоморфизме: v_0 не меняется, v_1 меняется линейно с изменением коэффициента растяжения K , а v_2 меняется пропорционально K^2 , т. е. как площадь плоской фигуры.

Важно отметить, что вариация v_0 множества уровня e_t , входящая в определение вариации $w_1(f)$, является не метрической, а в определенном смысле топологической характеристикой функции, основанной на понятии связности и не изменяющейся при гомеоморфизме. Касательно значения вариации $w_2(f)$, в [5] было показано, что для непрерывно дифференцируемой на D функции $f(x, y)$ справедливо равенство:

$$w_2(f) = \iint_D |\text{grad}(f(x, y))| dx dy,$$

т. е. $w_2(f)$ является интегралом модуля градиента функции по области носителя. В нашем случае $f(x, y)$ есть значение сигнала (яркости) изображения, которое можно считать неотрицательным и ограниченным, а значит, пределы интегрирования в (7) можно ограничить диапазоном $[0, T]$, где T — максимальное допустимое значение.

В отличие от v_0 и v_1 характеристика $v_2(e_t)$ не представляет интереса для оценки сложности функции. Рассматривая $v_2(e_t)$ как сумму площадей всех компонент сечения уровня t функции яркости $f(x, y)$, разделенных границами, очевидно, что $v_2(e_t) = S(D)$, т. е. равно площади всего множества D , и тем самым дополнительной информации не несет.

Можно предположить, что полезным окажется использование подмножества $\tilde{v}_2(e_t)$ ($\tilde{v}_2(e_t) \subset v_2(e_t)$), состоящего, скажем, только из тех точек (x, y) , для которых выполняется $f(x, y) \geq t$. По аналогии с (7) определим третий функционал как интеграл плоской вариации

$$w_3(f) = \int_0^T \tilde{v}_2(e_t) dt,$$

значение которого будет равно объему трехмерной фигуры, ограниченной функцией $f(x, y)$ и плоскостью $z(x, y) = 0$. Поскольку $f(x, y)$ задана на ограниченном двумерном множестве $D \ni (x, y)$, справедливо выражение $\tilde{v}_2(e_t)/S(D) = 1 - F_f(t)$, где $S(D)$ — площадь D , а $F_f(t)$ — функция распределения значений $f(x, y)$. Отсюда

$$\tilde{v}_2(e_t) = S(D) \left(1 - \int_0^t h(z) dz \right) = S(D) \int_t^T h(z) dz,$$

где $h(z) = p\{f(x, y) = z\}$ есть плотность вероятности значений функции $f(x, y)$. Легко видеть, что при этом соотношение $w_3(f)/S(D)$ будет равно среднему значению $f(x, y)$ на D , или, в интерпретации изображения, его средней яркости. Это означает, что функционал $w_3(f)$ также не содержит информации о структуре $f(x, y)$. Тем самым, для оценки сложности функции $f(x, y)$ остаются только две вышеприведенные вариации $w_1(f)$ и $w_2(f)$.

В случае дискретной функции $f(i, j)$ двумерные вариации (7) выражаются следующими формулами:

$$w_1(f) = \sum_{x=0}^{T-1} \frac{v_0(e_t)}{T}; \quad w_2(f) = \sum_{x=0}^{T-1} \frac{v_1(e_t)}{T}. \tag{8}$$

где T — общее число возможных значений функции $f(i, j)$ (для изображения — число градаций яркости). При вычислении вариаций в качестве множества e_t выступает бинарная матрица $\mathbf{B} = \{b(i, j)\}$, строящаяся следующим образом:

$$b(i, j) = \begin{cases} 0, & \text{если } f(i, j) < t; \\ 1, & \text{если } f(i, j) \geq t. \end{cases}$$

Дискретизация изображения осуществляется, как правило, на квадратной решетке. При этом значение $v_0(e_t)$ определяется как сумма числа связных компонент, составленных из соседствующих (в смысле 4-соседства [19]) элементов изображения со значением $b(i, j) = 0$ и числа аналогичных связных компонент из элементов с $b(i, j) = 1$. Значение $v_1(e_t)$ есть суммарная длина границ компонент $v_0(e_t)$. В случае дискретной функции $f(i, j)$ значение $v_1(e_t)$ зависит от способа вычисления длины границы. На квадратной решетке длину границ чаще всего измеряют в метрике L_1 , определяя $v_1(e_t)$ как суммарное число сегментов решетки, разделяющих элементы с несовпадающими значениями.

Формальное вычисление значений вариаций согласно (8) на ограниченном носителе D предполагает, что сама область D также учитывается в качестве отдельной компоненты.

Это приводит к следующему противоречию. Вариации w_1 и w_2 для функции постоянного значения $f(i, j) = \text{const}$, для которой они, очевидно, должны быть равны нулю, оказываются ненулевыми: $w_1(f) = 1$, а $w_2(f) = P(D)$ — периметру области D . Для устранения данного противоречия следует модифицировать формулы (8) следующим образом:

$$w_1(f) = \left(\sum_{x=0}^{T-1} \frac{v_0(e_t)}{T} \right) - 1; \quad w_2(f) = \left(\sum_{x=0}^{T-1} \frac{v_1(e_t)}{T} \right) - P(D). \quad (9)$$

В дальнейшем значения w_1 и w_2 для дискретной функции $f(i, j)$ на ограниченном носителе будут пониматься именно в смысле (9). Будучи нормированным на величину диапазона T , значение w_1 является характеристикой числа и амплитуды объектов изображения, образующих в сечениях отдельные компоненты. Поэтому значение w_1 удобно именовать *показателем числа объектов* изображения. Можно также показать, что значение второй вариации $w_2(f)$ в (8) будет совпадать со значением дискретной вариации Тонелли (5).

4 Показатель размеров объектов изображения

Получаемые значения вариаций Кронрода позволяют оценить некоторые параметры изображения, которые отражают как амплитудные, так и морфологические его характеристики. Значение первой вариации w_1 отражает число и контраст деталей на изображении, а значение второй, w_2 , — сумму периметров деталей. Наряду с этим, важным выглядит соотношение вариаций $q_w(f) = w_2(f)/w_1(f)$, которое отражает средний периметр деталей на изображении (здесь предполагаем, что $w_1 > 0$, иначе $q_w(f) = 0$).

В дискретном изображении наименьшим возможным объектом является один элемент (пиксель). Это значит, что минимально возможная компонента из числа входящих в множество уровня e_t имеет линейный размер, равный одному шагу дискретизации. Периметр такой компоненты из одного элемента будет равен 4 — значению, которое есть аналог числа π в метрике L_1 и является минимально возможным для соотношения периметр/площадь объекта. Предполагая, что в дискретном случае длина границ компонент, составляющих $v_1(e_t)$, измеряется в метрике L_1 , очевидно, что и для $q_w(f)$ также существует минимально возможное значение, которое тоже равно 4, причем минимум достигается в случае, когда все компоненты множеств уровня e_t для всех t являются одноэлементными. В связи со сказанным удобно ввести характеристику

$$d(f) = \frac{q(f)}{4}, \quad (10)$$

которую по аналогии с показателем числа объектов w_1 в (9) естественно назвать *показателем размеров объектов* изображения; для нее будет выполняться соотношение $d(f) \geq 1$.

5 Двумерная вариация как оценка сложности изображения

Для изображений с преобладанием мелких деталей, которые на множествах уровня отображаются компонентами небольших размеров, значения d будут малыми, а для изображений с крупными деталями — большими. Отметим, что к деталям изображения относятся также и шумы, являющиеся локальными выбросами сигнала и проявляющиеся на множествах уровня как самостоятельные компоненты с минимально возможным периметром. В связи с этим значение d должно убывать при увеличении числа и амплитуды шумовых выбросов. Ниже этот факт подтверждается экспериментально.

Особый интерес представляет поведение значений $w_1(f)$ и $d(f)$ при возможных преобразованиях функции $f(x, y)$. Для некоторых важных частных случаев можно сформулировать следующие утверждения.

1. Значение w_1 не зависит от выбора системы координат, а d зависит от выбора или поворота системы координат относительно носителя D лишь в степени точности дискретизации $f(x, y)$.
2. При линейных амплитудных преобразованиях вида $Cf(x, y)$, где C — константа, значение w_1 увеличивается пропорционально C , а значение d не изменяется.
3. При линейном пространственном растяжении носителя D в K раз значение w_1 не изменяется, а значение d увеличивается также в K раз.
4. Расширим D добавлением области U , на которой $f(x, y)$ имеет постоянное значение, т. е. $f(x, y) = \text{const}$ $\{(x, y) \in U \subset \tilde{D} = D \cup U\}$, с условием, что при расширении не возникает новых участков разрыва первого рода. Такое расширение не приводит к изменению значений w_1 и d .
5. При увеличении количества деталей на изображении (но при сохранении необходимых статистических соотношений и распределений) значение w_1 пропорционально увеличивается, а d не изменяется.
6. Пусть область D прямоугольна, и по выбранной оси, скажем, оси Y , ограничена отрезком $[a, b]$. Расширим ее областью U , заданной на полуотрезке $(b, c]$, на котором $f(x, y)$ имеет зеркальное продолжение: $f(x, b+y) = f(x, b-y)$. При таком расширении значение w_1 возрастает пропорционально изменению площади, тогда как значение d не меняется.

Таким образом, при указанных изменениях двумерной функции $f(x, y)$ как минимум одно из значений w_1 и d оказывается инвариантом, второе же значение является предсказуемой функцией преобразования. Применительно к изображениям утверждение 2 соответствует линейному изменению контраста; утверждение 3 — линейной геометрической трансформации; утверждение 4 — добавлению/удалению участков изображения с ровным фоном без деталей; утверждение 5 близко предыдущему и соответствует концентрации/разреженности деталей на изображении; утверждение 6 — вариант зеркальной пролонгации изображения за рамки области D , часто используемый алгоритмами локального анализа. Кроме того, повторение свойств части изображения на всем изображении по существу означает однородность (стационарность), т. е. достаточность определения значений w_1 и d на доверительном участке и интерпретации их для изображения в целом.

Отметим важную особенность введенного показателя размеров объектов $d(f)$, имеющую следствие в утверждениях 2 и 5. Характеристика $d(f)$ обладает тем свойством, что отражает лишь средние размеры объектов на изображении и не зависит от их количества и контраста самих объектов — эту часть информации несет вариация $w_1(f)$.

Из сказанного ясно, что первая вариация $w_1(f)$ вместе с характеристикой $d(f)$ отражают пространственную и яркостную изменчивость изображения и в совокупности могут служить показателями сложности изображения в соответствии с требованиями, сформулированными во введении. Различное поведение указанных характеристик подтверждает ту изначальную гипотезу Кронрода, что для описания изменчивости двумерных функций какой-то одной вариации недостаточно и следует использовать два независимых функционала.

6 О сравнении и интерпретации оценок

Поскольку характеристик сложности оказывается две, причем они, вообще говоря, независимы, то возникает естественный вопрос, как с их помощью сравнивать различные

изображения между собой. Достаточно очевидно, что если характеристика $d(f)$ постоянна, то с увеличением $w_1(f)$ сложность изображения будет возрастать; также сложность будет возрастать и с уменьшением $d(f)$ при постоянном $w_1(f)$. Тогда, если для пары изображений f_1 и f_2 выполняются соотношения $w_1(f_1) < w_1(f_2)$ и $d(f_1) > d(f_2)$, то естественно полагать, что изображение f_2 является более сложным, чем f_1 . Остается вопрос, как сравнивать другие случаи.

Рассмотрим изображение с малым значением w_1 и большим значением d ; такие соотношения означают, что на изображении имеется малое число деталей больших размеров (с большими периметрами). Если значения w_1 и d малы, это свидетельствует о том, что деталей на изображении содержится мало, причем они небольшого размера. Случай одновременно больших значений w_1 и d , означающий наличие большого числа деталей больших размеров, как очевидно, ограничен размерами самой области D . Во всех подобных случаях обычно принято говорить, что оценка сложности зависит от конкретной задачи.

В некоторых случаях оптимизация в многомерном пространстве затруднительна и приходится из нескольких функционалов тем или иным образом формировать единое обобщенное значение. В качестве одного из простейших вариантов обобщения можно предложить следующую линейную комбинацию:

$$w(z) = w_1(z) + Cw_2(z) = w_1(z)(1 + Cd(z)), \quad (11)$$

которая может, например, применяться в задачах восстановления в качестве стабилизирующего функционала $\Omega(z)$.

7 Использование оценки сложности в задачах восстановления

Численные оценки сигнала актуальны при решении различных задач, в частности классической оптимизационной задачи восстановления искаженного сигнала методом регуляризации [20], которая формулируется следующим образом. Пусть исходный сигнал z искажен оператором A (линейным или нелинейным) и шумом ξ , в результате чего наблюдается сигнал $u = A(z) + \xi$. Полагая оператор A с определенной точностью известным, строится восстанавливающий оператор \hat{A} (алгоритм) $\hat{A}(u) \Rightarrow \{z_\alpha\}$, отображающий u на множество допустимых результатов $z_\alpha \in Z$. Вычисляется функционал

$$\varphi_\alpha(u, \lambda) = \|Az_\alpha - u\| + \lambda\Omega(z_\alpha), \quad (12)$$

где $\Omega(z_\alpha)$ — неотрицательный стабилизирующий функционал, а λ — регуляризирующий параметр. Оптимальный на Z результат z^* находится путем минимизации функционала $\varphi_\alpha(u, \lambda)$ при выбранном значении λ :

$$\varphi_{z^*}(u, \lambda) = \inf_{z_\alpha \in Z} \varphi_\alpha(u, \lambda). \quad (13)$$

Восстанавливающий оператор \hat{A} строится исходя из предположений о свойствах искажающего оператора A и зависит от конкретных условий. В каких-то случаях, как в задаче линейного восстановления [20], он может быть обратным оператором $\hat{A} = A^{-1}$, в других, как, например, в задаче фильтрации шума на основе полной вариации [21], — единичным оператором $\hat{A} = E$ или каким-то другим.

Важным является то, что в отличие от восстанавливающего оператора \hat{A} стабилизирующий функционал $\Omega(z)$, вообще говоря, не зависит от искажающего преобразования и выбирается как некоторая характеристика сигнала z , которая может задаваться исходя

из модели сигнала. Для одномерных функций обычно предлагается использовать оценки типа нормы или вариации [20]. В контексте нашей задачи функционал $\Omega(z)$ можно интерпретировать как оценку сложности сигнала. В качестве такового может использоваться значение $w(z)$ из (11).

8 Значения оценок сложности для различных изображений

Представляет интерес, насколько измеряемые значения сложности $w_1(f)$ и $d(f)$ могут меняться от изображения к изображению, как они зависят от уровня шумов и как изменяются при сглаживании. При проведении подобных экспериментов требуется сравнение сигнала, получаемого при преобразовании, с неким эталонным и неискаженным. Сформировать такое эталонное изображение можно, основываясь на двухмасштабной многокомпонентной модели изображения [16]. Согласно данной модели, изображение $f(i, j)$ представляется как аддитивная смесь двух компонент: кусочно-гладкой $S(i, j)$ и текстурно-шумовой $\tau(i, j)$. Последняя в свою очередь несет информацию о текстуре, мелких деталях $t(i, j)$ и шуме $\xi(i, j)$:

$$f(i, j) = S(i, j) + \tau(i, j) = S(i, j) + t(i, j) + \xi(i, j). \tag{14}$$

Для локальных алгоритмов обработки результат преобразования в каждой точке (i, j) зависит лишь от элементов, попадающих в ограниченную и сравнительно небольшую область анализа $R(i, j)$. В [16] показано, что в таком случае для большинства реальных изображений можно считать, что $S(u, v) = S(i, j)$ при условии $(u, v) \subset R(i, j)$. Таким образом в пределах области $R(i, j)$ значение $S(u, v)$ полагается постоянным. Это позволяет

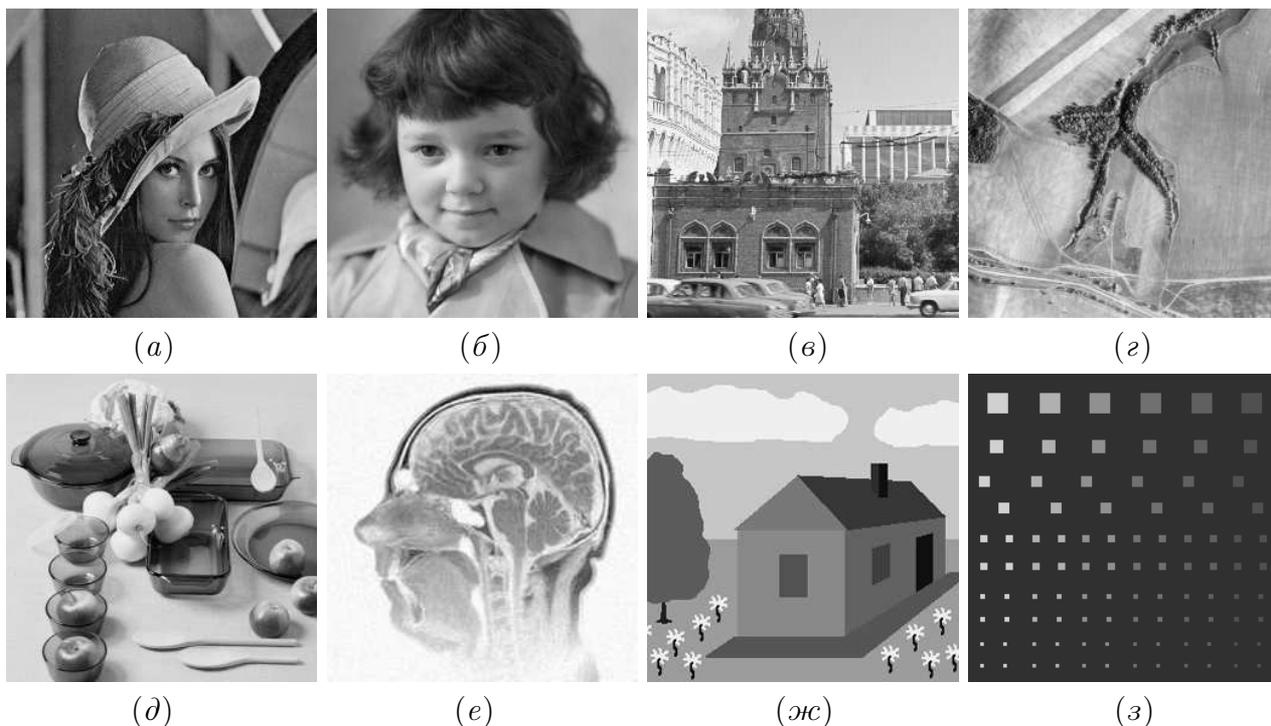


Рис. 1 Тестовые изображения: (а) и (б) портреты с мелкими и крупными деталями; (в), (г) и (д) городской ландшафт, аэрофотоснимок и натюрморт соответственно; (е) компьютерная томограмма; (ж) синтезированный рисунок с протяженными кусочно-постоянными объектами; (з) тестовый сигнал с деталями различных размеров и яркости на ровном фоне

Таблица 1 Значения двумерных вариаций, показателя размеров объектов (d) и энтропии (E) для изображений на рис. 1

| Изображение | A/T | P | $V/F/H$ | $K-w_1$ | $K-w_2$ | d | E |
|----------------|---------|---------|---------|---------|---------|-------|------|
| <i>a</i> | 5242,72 | 3715,70 | 2859,43 | 469,85 | 5250,25 | 2,80 | 7,59 |
| <i>б</i> | 2583,56 | 1825,14 | 1038,43 | 129,23 | 2586,58 | 5,01 | 7,60 |
| <i>в</i> | 7928,75 | 6072,46 | 3497,91 | 618,79 | 7935,70 | 3,21 | 7,59 |
| <i>г</i> | 4178,20 | 2954,15 | 1981,67 | 332,67 | 4184,39 | 3,16 | 7,19 |
| <i>д</i> | 3256,29 | 2389,90 | 1562,20 | 256,84 | 3260,76 | 3,18 | 7,34 |
| <i>e</i> | 5107,64 | 3642,48 | 2499,84 | 419,62 | 5108,65 | 3,05 | 6,76 |
| <i>ж</i> | 1160,50 | 1021,17 | 528,25 | 11,34 | 1160,48 | 25,58 | 3,11 |
| <i>з</i> | 811,67 | 778,54 | 132,52 | 33,13 | 811,67 | 6,12 | 0,58 |
| <i>з + шум</i> | 1944,96 | 1539,15 | 945,04 | 233,96 | 1944,93 | 2,08 | 3,63 |

формировать эталонное изображение как кусочно-постоянный двумерный сигнал, который впоследствии будет подвергаться искажениям и процедурам сглаживания.

Для экспериментов был отобран набор из реальных и синтезированных изображений различной детальности и содержания. Часть изображений представлена на рис. 1.

В табл. 1 и 2 представлены значения двумерных вариаций, измеренных для тестовых изображений на рис. 1. Двумерные вариации Арцела (A), Тонелли (T), Пьерпонта (P), Витали (V), Фреше (F), Харди (H), Кронрода ($K-w_1$ и $K-w_2$) вычислялись согласно формулам (2)–(6) и (9). Приведены также соответствующие показатели размеров объектов (d) согласно (10) и значения энтропии (E).

Как видно, значения одних и тех же вариаций для реальных изображений могут различаться в несколько раз, тогда как значения энтропии отличаются очень мало. Это подтверждает сделанный ранее вывод, что энтропия не является подходящей оценкой сложности изображения. Сравнение значений вариаций в табл. 1 и 2 демонстрирует достаточно точное совпадение вариаций Арцела/Тонелли (A/T) и второй вариации Кронрода ($K-w_2$) для всех изображений.

Для синтезированных изображений (*ж*) и (*з*) значения вариаций и энтропии оказываются существенно отличающимися от соответствующих значений для реальных изображений. Объясняется это тем, что данные изображения являются кусочно-постоянными и не содержат шума и мелких деталей. Добавление же даже небольшого шума, который обозначен составляющей $\xi(i, j)$ в (14), резко меняет значения оценок и делает их мало отличимыми от тех, которые были получены для реальных изображений. Это демонстрируется в последней строке табл. 1, соответствующей изображению (*з*), искаженному аддитивным гауссовым шумом с $\sigma = 2$ градации яркости. Более подробно результаты добавления шума и фильтрации рассматриваются ниже.

9 Влияние шума на оценку сложности изображения

Для определения влияния шума на оценку сложности изображения было выбрано синтезированное изображение с кусочно-постоянными объектами разных размеров и контраста (см. рис. 1, *ж*). На него наносился нормально распределенный аддитивный шум $N(0, \sigma^2)$ с величиной среднеквадратического отклонения (СКО) $\sigma = 1, 2, 3$ и 5 градаций яркости. Полученные значения различных двумерных вариаций, показателя размеров объектов (d) и энтропии (E) представлены графически на рис. 2, а и 2, б, а численно — в табл. 2.

Таблица 2 Значения двумерных вариаций, показателя размеров объектов (d), энтропии (E) и их зависимости от уровня шума для изображения на рис. 1, \mathcal{E}

| Шум, градации яркости | A/T | P | $V/F/H$ | $K-w_1$ | $K-w_2$ | d | E |
|-----------------------|---------|---------|---------|---------|---------|-------|------|
| $\sigma = 0$ | 1160,50 | 1021,17 | 528,25 | 11,34 | 1160,48 | 25,58 | 3,11 |
| $\sigma = 1$ | 1641,15 | 1327,58 | 909,12 | 110,27 | 1644,71 | 3,73 | 5,06 |
| $\sigma = 2$ | 2198,37 | 1696,72 | 1303,17 | 210,29 | 2204,56 | 2,62 | 5,94 |
| $\sigma = 3$ | 2758,17 | 2072,65 | 1697,21 | 309,89 | 2766,76 | 2,23 | 6,43 |
| $\sigma = 5$ | 3867,95 | 2816,74 | 2487,14 | 511,75 | 3892,58 | 1,90 | 7,01 |

Из приведенных данных и графиков можно сделать следующие выводы: (а) с увеличением уровня шума (для выбранного диапазона шумов) значения всех вариаций растут практически линейно; (б) показатель размеров объектов (d) сначала резко падает, затем убывает медленно, стремясь к предельному значению 1; (в) энтропия (E) асимптотически возрастает до предельного значения 8, равного числу битов в двоичной записи значений яркости; (г) в подтверждение сказанному выше значения вариаций Арцела/Тонелли (A/T) и второй вариации Кронрода ($K-w_2$) с большой степенью точности совпадают.

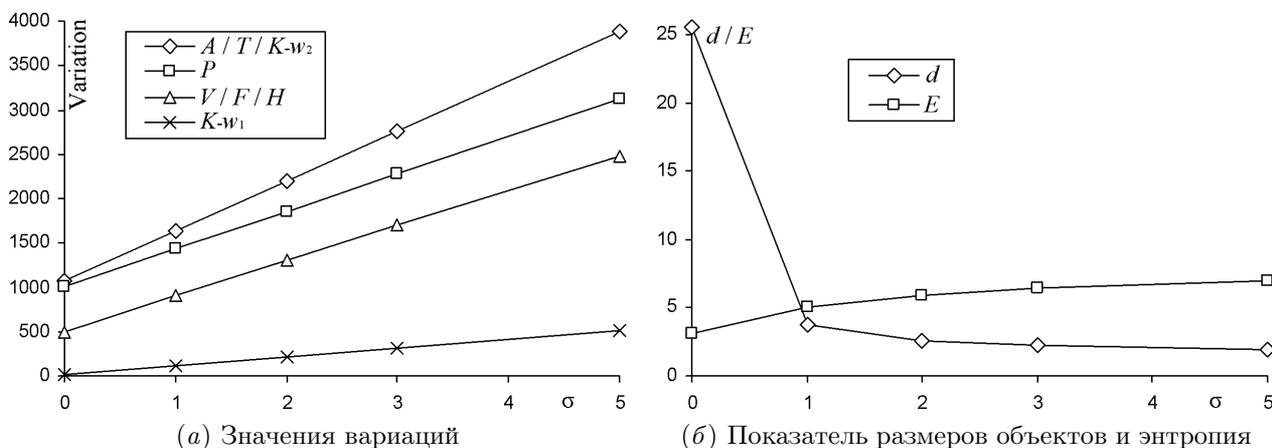


Рис. 2 Зависимость вариаций ($Variation$), показателя размеров объектов (d) и энтропии (E) от величины нанесенного гауссова шума (σ , градации яркости) для изображения на рис. 1, \mathcal{E}

10 Влияние сглаживания на оценку сложности изображения

Для исследования изменений оценки сложности изображения при сглаживании было сформировано кусочно-постоянное изображение с множеством объектов разных размеров и контраста относительно фона (см. рис. 1, \mathcal{E}). Изображение искажалось аддитивным гауссовым шумом с $\sigma = 5$ градаций яркости, после чего для удаления шума и восстановления изображения применялись сглаживающие алгоритмы локального среднего и локальной медианы [19], а также алгоритм декомпозиции изображения [22].

Алгоритмы сглаживания формулируются следующим образом. Определяется фрагмент анализа $R(i, j)$ размерами $L \times L$ с центром в точке (i, j) . Локальное среднее по точкам (m, n) фрагмента $R(i, j)$ задается формулой:

Таблица 3 Значения СКО, двумерных вариаций ($K-w_1$ и $K-w_2$), показателя размеров объектов (d) и энтропии

| № | Вид изображения | СКО | $K-w_1$ | $K-w_2$ | d | Энтропия |
|---|--|-------|---------|---------|-------|----------|
| 1 | Исходное | 0,00 | 33,13 | 811,87 | 6,13 | 0,58 |
| 2 | С шумом, $\sigma = 5$ градаций яркости | 5,00 | 533,71 | 3644,04 | 1,71 | 4,93 |
| 3 | Локальное среднее | 19,30 | 8,83 | 463,95 | 13,13 | 4,08 |
| 4 | Локальная медиана | 20,26 | 6,80 | 253,46 | 10,51 | 2,09 |
| 5 | Декомпозиция (1 итерация) | 0,64 | 36,55 | 872,95 | 5,97 | 1,68 |
| 6 | Декомпозиция (3 итерации) | 0,63 | 32,89 | 822,22 | 6,25 | 1,62 |

$$A_f(i, j) = \frac{1}{L^2} \sum_{(m,n) \in R(i,j)} f(m, n). \quad (15)$$

Локальная медиана по множеству точек того же фрагмента записывается как

$$M_f(i, j) = \text{med}\{f(m, n) \mid (m, n) \in R(i, j)\}. \quad (16)$$

Согласно многокомпонентному представлению (14), изображение $f(i, j)$ представляет собой сумму кусочно-гладкой компоненты $S(i, j)$ и текстурно-шумовой компоненты $\tau(i, j)$, несущей информацию о текстуре, мелких деталях и шуме. Декомпозиция [22] имеет целью разделение изображения $f(i, j)$ на указанные компоненты. Из них в данном случае нас интересует сглаженная компонента $S(i, j)$. Алгоритм заключается в том, что для каждой точки изображения (i, j) производится последовательный локальный анализ сначала по внутренней окрестности $r(i, j)$ малого размера $l \times l$, а затем по окружающему ее фрагменту $R(i, j)$ большого размера $L \times L$ ($r \subset R, l < L$). Для анализа множества значений элементов, попадающих в соответствующие фрагменты $r(i, j)$ и $R(i, j)$, используются методы, основанные на порядковых статистиках. Применяемые методы близки известному сигма-фильтру [23], который в свою очередь является частным случаем парзеновского окна [24]. Подробно алгоритм декомпозиции описан в [22, 25].

Результаты преобразований в виде значений двумерных вариаций w_1 и w_2 , а также показателя размеров объектов d приведены в табл. 3. Строка 1 соответствует эталонному изображению; 2 — искаженному шумом. Остальные строки отражают результаты фильтрации зашумленного изображения различными алгоритмами сглаживания при одинаковом размере фрагмента анализа (15×15 элементов): 3 — локальное среднее (15); 4 — локальная медиана (16); 4 — сглаженная компонента $S(i, j)$ в (14) после одной итерации декомпозиции; 5 — сглаженная компонента $S(i, j)$ после 3 итераций декомпозиции. Остальные значения вариаций и показателя размеров объектов, получаемые при использовании фрагментов сглаживания других размеров, представлены графически на рис. 3 и 4.

На рис. 3 и 4 показаны зависимости первой и второй вариаций Кронрода (w_1 и w_2), а также показателя размеров объектов (d) и СКО результата фильтрации от неискаженного оригинала в зависимости от размеров фрагмента сглаживания (величина Δ изменяется от 3×3 до 15×15 элементов) для результатов фильтрации алгоритмами локального среднего (Mean), локальной медианы (Median) и декомпозиции (Decomp).

Анализируя полученные результаты, отметим следующее.

1. С увеличением размеров фрагмента сглаживания значения первой (w_1) и второй (w_2) вариаций Кронрода, а также показателя размеров объектов (d) для алгоритма деком-

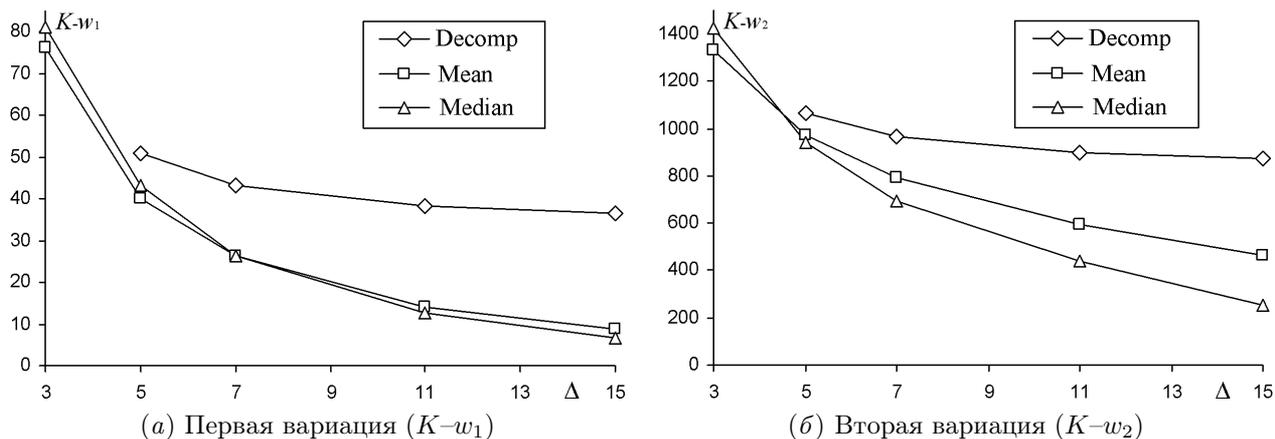


Рис. 3 Зависимость первой ($K-w_1$) и второй ($K-w_2$) вариаций Кронрода от размеров фрагмента сглаживания (Δ)

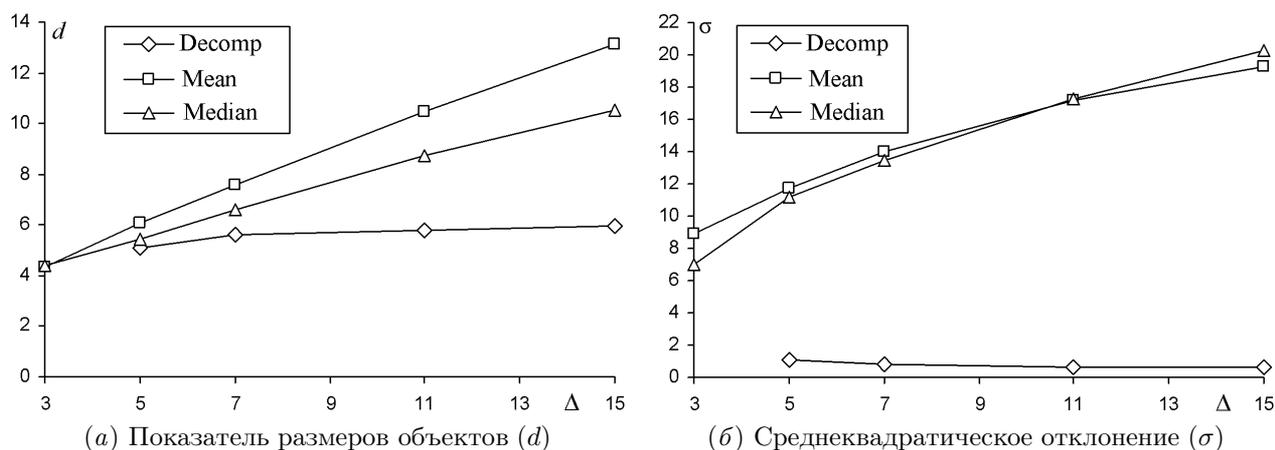


Рис. 4 Изменение показателя размеров объектов (d) и СКО (σ) результата фильтрации от искаженного оригинала в зависимости от размеров фрагмента сглаживания (Δ)

позиции стремятся к постоянным значениям, в то время как для алгоритмов локального среднего и локальной медианы они либо стремятся к нулю, либо возрастают почти линейно (см. рис. 3 и 4, а).

2. Точность восстановления изображения, искаженного шумом, при использовании алгоритма декомпозиции возрастает с увеличением размера фрагмента сглаживания и числа итераций, а алгоритмы локального среднего или локальной медианы при этом дают все более отличающиеся от оригинала результаты (см. табл. 3 и рис. 4, б).
3. Значения вариаций (w_1 и w_2) и показателя размеров объектов (d) в результате декомпозиции возвращаются к значениям, близким тем, которые соответствуют исходному неискаженному изображению (см. табл. 3). То, что данные значения выходят на плато одновременно с асимптотическим приближением к нулю величины отклонения от оригинала (см. рис. 4, б), говорит о стабилизации результата преобразования вблизи исходного неискаженного изображения.
4. Стремление для обычных вариантов сглаживания значений w_1 и w_2 к нулю, а значений d к возрастанию означает уменьшение сложности изображения, что говорит о раз-

рушении сигнала изображения. Выход тех же значений на плато в случае алгоритма декомпозиции означает, что основная структура деталей изображения сохраняется.

5. Малое значение энтропии (2,09) в строке 4 табл. 3 объясняется особенностью медианной фильтрации сигнала, имеющего ограниченное число значений.

Полученные результаты позволяют сделать вывод, что при условии справедливости модели [16] выбранная совокупность значений w_1 и d достаточно хорошо отражает морфологическую структуру изображения, в частности такие важные параметры, как число, размеры и заметность деталей изображения, и поэтому может являться оценкой сложности изображения. Определяемая таким способом характеристика может быть использована в качестве оценки сложности изображения, например как стабилизирующий функционал при решении задач восстановления сигнала.

11 Заключение

Исследована возможность оценки сложности изображения при помощи двумерных вариаций. Показана редукция многомерных вариаций непрерывных функций к двумерным дискретным функциям. В данном контексте проведено сравнение различных двумерных вариаций; подтверждено преимущество использования двумерной вариации Кронрода. На основе двумерных вариаций введена характеристика, соответствующая среднему значению периметров деталей на изображении. Предложена модификация данной характеристики, названная *показатель размеров объектов*.

Проанализированы изменения значений вариаций при различных трансформациях изображения. Теоретические выводы подтверждены экспериментальными исследованиями. Продемонстрировано влияние преобразований на сложность изображения, которая интерпретируется как характеристика, отражающая число, размеры и заметность деталей на изображении. Показано, что обычные варианты сглаживания приводят к разрушению сигнала изображения, а алгоритм декомпозиции сохраняет основную структуру деталей изображения.

Предложен вариант редукции двух значений вариаций к единственному обобщенному функционалу, который может использоваться в классических оптимизационных схемах.

Проведенные исследования подтверждают, что комбинация показателя числа объектов w_1 и показателя размеров объектов d отражает морфологическую структуру изображения и может служить оценкой его сложности.

Литература

- [1] Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. Image quality assessment: From error visibility to structural similarity // IEEE Trans. Image Proc., 2004. Vol. 13. No. 4. P. 600–612.
- [2] Wang Z., Bovik A. C. Modern image quality assessment. — New York, NY, USA: Morgan and Claypool, 2006. 156 p.
- [3] Yu H., Winkler S. Image complexity and spatial information // 5th Workshop (International) on Quality of Multimedia Experience (QoMEX) Proceedings. — Klagenfurt, Austria, 2013. P. 12–17.
- [4] Peters R. A., Strickland R. N. Image complexity metrics for automatic target recognizers // Automatic Target Recognizer System and Technology Conference. — Silver Spring, MD, 1990. P. 30–31.
- [5] Миллюкова О. П. Изображение как функция с ограниченной полной вариацией // Иконика. Цифровая обработка видеoinформации, 2004. — М.: Наука, 1989. С. 19–25.
- [6] Adams C. R., Clarkson J. A. On definitions of bounded variation for functions of two variables // Trans. Am. Math. Soc., 1933. Vol. 36. P. 824.

- [7] Кронрод А. С. О функциях двух переменных // УМН, 1950. Т. 5. № 1. С. 24–134.
- [8] Витушкин А. Г. О многомерных вариациях. — М.: Гостехиздат, 1955. 220 с.
- [9] Иванов Л. Д. Вариации множеств и функций. — М.: Наука, 1975. 352 с.
- [10] Chochia P. A., Milukova O. P. Two-dimensional variation and image decomposition // 6th Workshop (International) on Image Processing and Computer Optics (DIP-97) Proceedings SPIE, 1998. Vol. 3346. P. 329–339.
- [11] Милукова О. П., Чочиа П. А. Об оценке сложности изображений с помощью двумерных вариаций // Информационные процессы, 2012. Т. 12. № 4. С. 362–371.
- [12] Чочиа П. А., Милукова О. П. Сравнение двумерных вариаций в контексте оценки сложности дискретных изображений // Информационные процессы, 2015. Т. 15. № 2. С. 169–182.
- [13] Варакин Л. Е. Теория сложных сигналов. — М.: Советское радио, 1970. 376 с.
- [14] Витушкин А. Г. Оценка сложности задачи табулирования. — М.: Физ.-мат. лит., 1959. 228 с.
- [15] Колмогоров А. Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации, 1965. Т. 1. Вып. 1. С. 3–11.
- [16] Чочиа П. А. Двухмасштабная модель изображения // Кодирование и обработка изображений. — М.: Наука, 1988, С. 69–87.
- [17] Математическая энциклопедия. Т. 1–5. — М.: Советская энциклопедия, 1977.
- [18] Pierpont J. Lectures on the theory of functions of real variables. — New York, NY, USA: Dover Publications, 1959. Vol. 1. 554 p.
- [19] Гонсалес Р., Вудс Р. Цифровая обработка изображений. — М.: Техносфера, 2012.
- [20] Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. — М.: Наука, 1979. 285 с.
- [21] Rudin L. I., Osher S., Fatemi E. Nonlinear total variation based noise removal algorithms // Physica D Nonlinear Phenomena, 1992. Vol. 60. P. 259–268.
- [22] Чочиа П. А. Сглаживание изображения при сохранении контуров // Кодирование и обработка изображений. — М.: Наука, 1988. С. 87–98.
- [23] Lee J.-S. Digital image smoothing and the sigma filter // Computer Vision Graphics Image Proc., 1983, Vol. 24. No. 2. P. 255–269.
- [24] Parzen E. On the estimation of a probability density function and mode // Ann. Math. Stat., 1962. Vol. 33. P. 1065–1076.
- [25] Чочиа П. А. Переход от 2D- к 3D-изображениям: модификация двухмасштабной модели и алгоритмов обработки // Информационные процессы, 2014. Т. 14. № 3. С. 242–255.

Поступила в редакцию 25.05.2015

References

- [1] Wang, Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Proc.* 13(4):600–612.
- [2] Wang, Z., and A. C. Bovik. 2006. *Modern image quality assessment*. New York, NY: Morgan and Claypool. 156 p.
- [3] Yu, H., and S. Winkler. 2013. Image complexity and spatial information. *5th Workshop (International) on Quality of Multimedia Experience (QoMEX)*. Klagenfurt, Austria. 12–17.
- [4] Peters, R. A., and R. N. Strickland. 1990. Image complexity metrics for automatic target recognizers. *Automatic Target Recognizer System and Technology Conference*. Silver Spring, MD. 30–31.
- [5] Milukova O. P. 2004. Izobrazenie kak funktsiya s ogranichennoy polnoy vaeiatsiey [Image as a function with limited total variation]. *Ikonika. Tsifrovaya obrabotka videoinformatsii* [Iconika. Digital processing of video information]. Moscow: Nauka. 19–25.

- [6] Adams, C. R., and J. A. Clarkson. 1933. On definitions of bounded variation for functions of two variables. *Trans. Am. Math. Soc.* 36:824.
- [7] Kronrod, A. S. 1950. O funktsiyakh dvukh peremennykh [About two-variable functions]. *Uspekhi matematicheskikh nauk* 5(1):24–134.
- [8] Vitushkin, A. G. 1955. *O mnogomernykh variatsiyakh* [About multidimensional variations]. Moscow: Gostekhizdat. 220 p. (In Russian.)
- [9] Ivanov, L. D. 1975. *Variatsii mnozestv i funktsiy* [Variations of sets and functions]. Moscow: Nauka. 352 p. (In Russian.)
- [10] Chochia, P. A., and O. P. Milukova. 1998. Two-dimensional variation and image decomposition. *6th Workshop (International) on Image Processing and Computer Optics (DIP-97) Proceedings SPIE.* 3346:329–339.
- [11] Milukova, O. P., and P. A. Chochia. 2013. On estimation of the image complexity by two-dimensional variations. *J. Communications Technology Electronics.* 58(6):628–635.
- [12] Chochia, P. A., and O. P. Milukova. 2015. Svravnenie dvumetnykh variatsiy v kontekste otsenki sloznosti diskretnykh izobrazheniy [Comparison of two-variable variations in a context of image complexity assessment]. *Informatsionnye protsessy* 15(2):169–182.
- [13] Varakin, L. E. 1970. *Teoriya sloznykh signalov* [The theory of complex signals]. Moscow: Sovetskoe Radio. 376 p. (In Russian.)
- [14] Vitushkin, A. G. 1959. *Otsenka sloznosti zadachi tabulirovaniya* [The assesment of tabulation task complexity]. Moscow: Fiz.-mat. lit. 228 p. (In Russian.)
- [15] Kolmogorov, A. N. 1965. *Tri podkhoda k opredeleniyu ponyatiya “kolichestvo informatsii”* [Three approaches to difinition of the “information quantity” conception]. *Problems Information Transmission* 1(1):3–11.
- [16] Chochia, P. A. 1988. *Dvukhmasshtabnaya model’ izibrazeniya* [Two-scale image model]. *Kodirovanie i obrabotka izobrazheniy* [Image coding and processing]. Moscow: Nauka. 69–87.
- [17] *Matematicheskaya entsiklopediya* [Mathematical encyclopedia]. 1977. Moscow: Sovetskaya entsikloprdiya. Vols. 1–5.
- [18] Pierpont, J. 1959. *Lectures on the theory of functions of real variables*. New York, NY: Dover Publications. Vol. 1. 554 p. (In Russian.)
- [19] Gonzalez, R. C., and R. E. Woods. 2008. *Digital image processing*. Reading, Mass.: Addison-Wesley. 954 p.
- [20] Tikhonov, A. N., and V. Ya. Arsenin. 1979. *Metody resheniya nekorrektnykh zadach* [The methods for ill-posed problems decision]. Moscow: Nauka. 285 p. (In Russian.)
- [21] Rudin, L. I., S. Osher, and E. Fatemi. 1992. Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena* 60:259–268.
- [22] Chochia, P. A. 1988. Sglazivanie izobrazheniy pri sokhraneni konturov [Image smoothing under preservation of contours]. *Kodirovka i obrabotka izobrazheniy* [Image coding and processing]. Moscow: Nauka. 87–98.
- [23] Lee, J.-S. 1983. Digital image smoothing and the sigma filter. *Computer Vision Graphics Image Processing* 24(2):255–269.
- [24] Parzen, E. 1962. On the estimation of a probability density function and mode. *Ann. Math. Stat.* 33:1065–1076.
- [25] Chochia, P. A. 2015. Transition from 2D- to 3D-images: Modification of two-scale image model and image processing algorithms. *J. Communications Technology Electronics* 60(6):678–687.

Received May 25, 2015

Быстрые алгоритмы обработки изображений на основе гамма-нормальной модели скрытого поля*

И. А. Грачева, А. В. Копылов

gia1509@mail.ru, And.Kopylov@gmail.com

Тульский государственный университет, Тула, Россия

В рамках байесовского подхода к проблеме обработки изображений задача восстановления может быть выражена как задача поиска скрытой марковской компоненты двух-компонентного случайного поля, роль наблюдаемой компоненты которого играет анализируемое изображение. Однако для некоторых типов задач обработки изображений, таких как удаление тумана на изображении, сжатие контраста HDR (high dynamic range) изображений, восстановление структуры изображения, такая постановка задачи не подходит. В данной работе предлагается расширение постановки задач при байесовском подходе к проблеме обработки изображений.

Ключевые слова: *удаление тумана на изображении; сжатие контраста HDR изображений; гамма-нормальная модель скрытого поля*

DOI: 10.21469/22233792.1.12.03

Fast image processing algorithms based on the gamma-normal model of hidden field*

I. A. Gracheva and A. V. Kopylov

Tula State University, 92 Lenina pr., Tula, Russia

Within the Bayesian approach, the problem of reconstruction can be expressed as the problem of finding the hidden Markov component of random field of two-component, where observed component is the analyzed image. However, for certain types of image processing problems, such as image haze removal, HDR (high dynamic range) image compression, structure-transferring of image, this formulation of the problem is not the solution. In this paper, an extension of the problem formulations at the Bayesian approach to the problem of image processing.

Keywords: *image haze removal; HDR image compression; gamma-normal model of hidden field*

DOI: 10.21469/22233792.1.12.03

1 Введение

Предварительная обработка является важным этапом в работе практически любой системы анализа изображений или компьютерного зрения. Данный этап играет ключевую роль в повышении качества дальнейшего анализа и интерпретации данных. Задачей обработки изображения может быть как улучшение (восстановление, реставрация) изображения по какому-то определенному критерию, так и специальное преобразование, кардинально меняющее изображение. Примерами обработки изображений могут служить

* Работа выполнена при финансовой поддержке РФФИ, проект № 13-07-00529.

повышение контраста, резкости, коррекция цветов, сглаживание, сжатие контраста, выделение границ объектов, удаление тумана или шума.

Принятый в данной работе байесовский подход часто используется как универсальный метод для целого класса задач обработки изображений. Результат обработки в этом случае может быть представлен как преобразование исходного изображения, определенного на дискретном множестве, во вторичный массив данных, который определен на том же множестве аргументов и принимает значения из некоторого подходящего множества, в зависимости от решаемой задачи. Анализируемое изображение и результат обработки в рамках данного подхода рассматриваются как наблюдаемая и скрытая компоненты двухкомпонентного случайного поля. Гамма-нормальная модель скрытого поля, предложенная в работе [1], хорошо показала себя в задаче удаления шума на изображении. Однако для таких задач, как удаление тумана на изображении, сжатие контраста HDR изображений, восстановление структуры изображения, данная постановка не подходит.

В настоящее время удаление тумана с изображений является важной задачей, так как объекты на изображениях, сделанных в условиях дымки или тумана, плохо различимы. Для дальнейшей работы с такими изображениями требуется увеличить их контрастность и улучшить цветопередачу. Часто требуется улучшить качество не одного изображения, а целого набора изображений, причем в режиме реального времени. Современные методы удаления тумана [2–4] строятся на различных предположениях о глубине текстуры или цвета на изображении. Однако все эти методы слишком медленные и не могут обрабатывать большие потоки данных в режиме реального времени.

Видение реального мира в нашей повседневной жизни часто имеет очень широкий диапазон значений яркости. Зрительная система человека способна воспринимать сцены с динамическим диапазоном свыше пяти порядков и может постепенно адаптироваться к сценам с динамическим диапазоном свыше девяти порядков. В цифровой обработке изображений такие сцены получаются при использовании технологии HDR [5]. HDR изображения имеют много преимуществ по сравнению со стандартными изображениями и являются чрезвычайно полезными в некоторых приложениях. Тем не менее отображение HDR изображения представляет собой сложную задачу, потому что динамический диапазон различных устройств отображения (мониторы, принтеры и др.) намного меньше, чем динамический диапазон, обычно встречающийся в реальных сценах. Таким образом, возникает задача сжатия HDR изображений при максимальном сохранении их визуального содержимого.

В рамках данной работы оценивание скрытого случайного поля происходит по некоторому изображению, специфичному для каждой конкретной задачи. Например, в задаче удаления тумана в роли такого изображения выступает карта рассеивания, в задаче сжатия контраста — слабоконтрастное изображение, в задаче восстановления структуры — грубая оценка границы. Результирующее изображение получается путем уточнения этого исходного изображения с помощью оценок параметров модели скрытого поля, найденных с учетом управляющего изображения, в роли которого может выступать исходное анализируемое изображение.

2 Нестационарная гамма-нормальная модель скрытого случайного поля

Как уже отмечалось ранее, при расширенной постановке задачи восстановления изображений при байесовском подходе оценивание скрытой марковской компоненты $X = (x_t, t = 1, \dots, N)$, $T = \{t = (t_1, t_2) : t_1 = 1, \dots, N_1, t_2 = 1, \dots, N_2\}$ происходит не по

одному наблюдаемому изображению $Y = (y_t, t \in T)$, как это было в задаче удаления шума на изображении [1], а еще и по некоторому изображению $Y^g = (y_t^g, t \in T)$, специфичному для каждой конкретной задачи. Вероятностные свойства двухкомпонентного случайного поля (X, Y) полностью определяются совместной условной плотностью вероятности $\Phi(Y|X, \delta)$ исходной функции $Y = (y_t, t \in T)$ по отношению к скрытой последовательности $X = (x_t, t \in T)$, и совместным априорным распределением $\Psi(X|\Lambda, \delta)$ скрытой последовательности $X = (x_t, t \in T)$.

Совместная условная плотность вероятности $\Phi(Y|X, \delta)$ имеет вид Гауссовского распределения:

$$\Phi(Y|X, \delta) = \frac{1}{\delta^{N/2}(2\pi)^{N/2}} \exp\left(-\frac{1}{2\delta} \sum_{t \in T} (y_t - x_t)^2\right), \tag{1}$$

где $E(e_t^2) = \delta$ — дисперсия шума в наблюдении, которая считается неизвестной.

В данной работе не используются преобразования, стабилизирующие дисперсию так, чтобы шум становился приблизительно Гауссовским с постоянной дисперсией. Наоборот, предполагается, что общая дисперсия независимых компонент шума модели изменяется $E(\xi_t^2) = r_t$. Неизвестные дисперсии $(r_t, t \in T)$ считаются пропорциональными дисперсии шума в наблюдении $r_t = \lambda_t \delta$ с коэффициентами пропорциональности λ_t , действующими как факторы неизвестного мгновенного изменения скрытой компоненты $X = (x_t, t \in T)$, которые также неизвестны.

В соответствии с этим предположением совместное априорное распределение оценок скрытой последовательности $X = (x_t, t \in T)$ условно нормально по отношению к последовательности факторов $\Lambda = (\lambda_t, t \in T)$. Таким образом, приходим к совместному априорному распределению вида:

$$\Psi(X|\Lambda, \delta) \propto \frac{1}{\left(\prod_{t \in T} \delta \lambda_t\right)^{1/2} (2\pi)^{n(N-1)/2}} \exp\left(-\frac{1}{2} \sum_{t', t'' \in V} \frac{1}{\delta \lambda_t} (x_{t'} - x_{t''})^2\right),$$

где V — граф соседства элементов изображения, имеющий вид решетки.

Наконец, предположим, что обратные факторы $1/\lambda_t$ должны быть априори независимы и одинаково гамма-распределены

$$\gamma\left(\frac{1}{\lambda_t} \mid \delta, \lambda, \mu\right) \propto \left(\frac{1}{\lambda_t}\right)^{(2\mu+1)/(2\delta\mu)} \exp\left(-\frac{\lambda}{2\delta\mu} \frac{1}{\lambda_t}\right),$$

на положительной полуоси $\lambda_t \geq 0$. Математическое ожидание и дисперсия гамма-распределения вычисляются как

$$E\left(\frac{1}{\lambda_t}\right) = \frac{(1 + \delta)\mu + 1}{\lambda}; \quad \text{Var}\left(\frac{1}{\lambda_t}\right) = 2\delta\mu \frac{(1 + \delta)\mu + 1}{\lambda^2}.$$

Независимые априорные распределения каждого мгновенно обработанного фактора $1/\lambda_t$ почти полностью сконцентрированы вокруг математического ожидания $1/\lambda$, если $\mu \rightarrow 0$. При $\mu \rightarrow \infty$ коэффициент $1/\lambda$ имеет тенденцию к почти равномерному распределению. Таким образом, приходим к априорной плотности распределения вида:

$$G(\Lambda|\delta, \lambda, \mu) = \exp\left[-\frac{1}{2\delta\mu} \sum_{t \in T} \left(\lambda \frac{1}{\lambda_t} + \frac{1}{\lambda} \ln \lambda_t\right)\right].$$

Совместное априорное нормальное гамма-распределение обеих последовательностей $X = (x_t, t \in T)$ и $\Lambda = (\lambda_t, t \in T)$ имеет вид:

$$H(X, \Lambda | \delta, \lambda, \mu) = \Psi(X | \Lambda, \delta) G(\Lambda | \delta, \lambda, \mu).$$

В сочетании с совместной условной плотностью вероятности (1) это дает основу для байесовского оценивания скрытой последовательности $X = (x_t, t \in T)$.

Байесовские оценки независимы от дисперсии шума в наблюдениях δ :

$$\left\{ \begin{array}{l} (\hat{X}, \hat{\Lambda} | \lambda, \mu) = \arg \min_{X, \Lambda} J(X, \Lambda | Y, \lambda, \mu); \\ J(X, \Lambda | Y, \lambda, \mu) = \sum_{t \in T} (y_t - x_t)^2 + \sum_{t', t'' \in V} \left\{ \frac{1}{\lambda_t} \left[(x_{t'} - x_{t''})^2 + \frac{\lambda}{\mu} \right] + \left(1 + \frac{1}{\mu} \right) \ln \lambda_t \right\}. \end{array} \right. \quad (2)$$

Растущее значение параметра μ наделяет этот критерий выраженной склонностью к сохранению большинства оценок факторов $\hat{\lambda}_t$, находящихся близко к значению λ , и допускает большие разрывы, открывая тем самым скрытые события в первую очередь гладкости исходной последовательности.

Условно оптимальные факторы $\hat{\Lambda}(X, \lambda, \mu) = [\hat{\lambda}_t(X, \lambda, \mu), t \in T]$ определяются независимо друг от друга:

$$\hat{\Lambda}(X, \lambda, \mu) = \arg \min_{\Lambda} J(\Lambda | X, \lambda, \mu) : \frac{\partial}{\partial \lambda_t} \left\{ \frac{1}{\lambda_t} \left[(x_{t'} - x_{t''})^2 + \frac{\lambda}{\mu} \right] + \left(1 + \frac{1}{\mu} \right) \ln \lambda_t \right\} = 0.$$

Нулевые условия для производных, за исключением тривиальных решений $\lambda_t \rightarrow \infty$, приводят к равенству

$$\hat{\lambda}_t(X, \lambda, \mu) = \lambda \frac{(1/\lambda)(x_{t'} - x_{t''})^2 + 1/\mu}{1 + 1/\mu}. \quad (3)$$

Подстановка (3) в (2) дает эквивалентную форму, которая позволяет избежать немедленного нахождения самих факторов:

$$\left\{ \begin{array}{l} (\hat{X}, \hat{\Lambda} | \lambda, \mu) = \arg \min_{X, \Lambda} J(X, \Lambda | Y, \lambda, \mu); \\ J(X, \Lambda | Y, \lambda, \mu) = \sum_{t \in T} (y_t - x_t)^2 + \sum_{t', t'' \in V} \left\{ \left(1 + \frac{1}{\mu} \right) \ln \frac{(1/\lambda)(x_{t'} - x_{t''})^2 + 1/\mu}{1 + 1/\mu} \right\}. \end{array} \right. \quad (4)$$

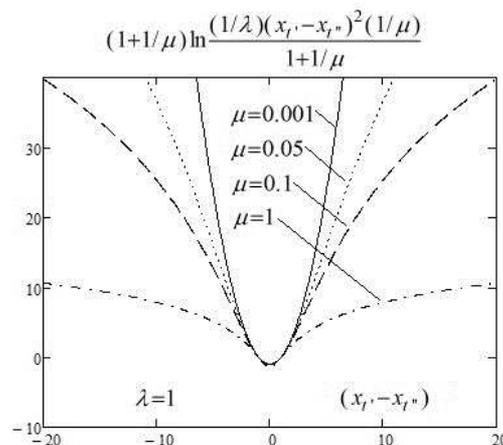


Рис. 1 «Эффект насыщения» из-за штрафа при достаточно большом значении параметра μ и фиксированном значении $\lambda = 1$

Это квадратичная в непосредственной близости от нулевой точки функция $(x_{t'} - x_{t''})^2 = 0$ и остается таковой практически на всей числовой оси, если значение μ достаточно мало, как показано на рис. 1. Но при росте μ первоначальный квадратичный штраф подвергается все более заметному влиянию насыщения на некотором расстоянии от нуля. Это означает, что критерий сильно наказывает оценки исходной функции, но становится более снисходительным к резким разрывам.

3 Алгоритм удаления тумана на изображении

Удаление тумана на изображении является проблемой, которая требует метода вывода или предварительных знаний о сцене. Количество тумана, наблюдаемого на изображении, зависит от расстояния объекта до камеры, длины волны света и размера рассеивающих частиц в атмосфере. Учитывая местоположение туманного пикселя $t \in T$ на изображении $Y = (y_t, t \in T)$, $Y \in R^3$, обработанный пиксель изображения без тумана $X = (x_t, t \in T)$, $X \in R^3$ связан с туманной версией изображения через модель атмосферного рассеивания [6]

$$y_t = \text{map}_t x_t + (1 - \text{map}_t) a, \quad (5)$$

где $a \in R^3$ — атмосферный свет; $\text{Map} = (\text{map}_t, t \in T)$, $\text{Map} \in R^3$, — карта светопередачи (средство передачи части света, который не рассеялся и достиг камеры).

Начнем с предположения о том, как оценить значение атмосферного света $a \in R^3$. В большинстве методов удаления тумана с изображений в качестве атмосферного света используется пиксель с самой высокой интенсивностью. Но на реальных изображениях самый яркий пиксель может находиться на белом автомобиле или белом здании.

В данной работе будем использовать темный канал для улучшения оценки атмосферного света [5]. Темный канал основан на наблюдении о затемнении исходного изображения. Изображение темного канала как раз и будет играть роль дополнительного изображения $Y^g = (y_t^g, t \in T)$. Для получения темного канала разбиваем исходное изображение на блоки любого размера (чем больше размер блока, тем более затемненным будет изображение; оптимальным является размер 5×5). В большинстве таких блоков некоторые пиксели, так называемые темные пиксели, имеют очень низкую интенсивность, по крайней мере, одного из трех цветовых компонент (RGB). Низкие интенсивности в темном канале возникают в основном из-за трех факторов: теней; красочных предметов или поверхностей, которым не хватает цвета в любом из трех цветовых компонент (RGB); темные объекты или поверхности, например темный ствол дерева. Таким образом, для исходного изображения $Y = (y_t, t \in T)$

$$y_t^g = \min_{c \in (r, g, b)} (\min_{j \in \Omega(t)} (y_j^c)), \quad (6)$$

где $Y^g = (y_t^g, t \in T)$ — темный канал изображения $Y = (y_t, t \in T)$; y^c — значение цветового канала изображения $Y = (y_t, t \in T)$; $\Omega(t)$ — локальный блок с центром в t .

Это наблюдение говорит, что в некотором блоке интенсивность низка и, как правило, равна нулю, если y_t представляет собой затемненный пиксель на исходном изображении. Таким образом, эти темные пиксели могут обеспечить точную оценку туманных пикселей на изображении.

Для получения значений атмосферного света по каждому из трех цветовых компонент (RGB) выбираем первые 0,1% самых ярких пикселей в темном канале. Среди этих пикселей значения пикселей с самой высокой интенсивностью по каждому из трех цветовых компонент на исходном изображении выбираются в качестве атмосферного света. Следует отметить, что эти пиксели могут быть не самыми яркими на исходном изображении. Этот

простой метод, основанный на темном канале, более надежен, чем метод «яркого пикселя». Следующим шагом является получение карты светопередачи $\text{Map} = (\text{map}_t, t \in T)$,

$$\text{map}_t = 1 - \omega \hat{\lambda}_t, \quad (7)$$

где ω — коэффициент уровня белого (эмпирическим путем было выявлено, что оптимальное значение коэффициента $\omega = 0,9$), а $\hat{\lambda}_t$ найдены по (3) в результате одной итерации процедуры Гаусса–Зайделя.

И, наконец, найдя все необходимые параметры из (6) и (7), получаем изображение без тумана из (5)

$$x_t = \frac{y_t - a}{\text{map}_t} + a.$$

Данный алгоритм имеет линейную вычислительную сложность относительно размера анализируемого изображения.

4 Алгоритм сжатия контраста HDR изображений

Предлагаемый в данной работе алгоритм — это попытка математически имитировать известную художественную технику для создания сцены с высокой контрастностью. Когда художники создают живопись, они пытаются охватить внешний вид «грубой» последовательностью границ и теней. Многие начинают с эскиза крупных, важных особенностей сюжета, а затем постепенно добавляют более тонкие детали. Первоначальные эскизы содержат четко определяемые границы вокруг больших, плавно затененных областей с высокой контрастностью и наиболее важными особенностями сцены. Далее художник добавляет больше оттенков и границ, чтобы создать более мелкие детали, заполняющие визуально пустые области. Этот метод особенно хорошо работает для сюжетов с высокой детализацией, так как это позволяет настроить отдельные настройки контрастности на каждом этапе с помощью увеличения детализации. Художник может также подчеркнуть или заглушить некоторые компоненты сцены, чтобы контролировать их видимость и обратить внимание зрителя на них.

Также предлагаемый подход опирается на широко распространенное допущение, что зрительная система человека не очень чувствительна к абсолютным яркостям, достигающим сетчатки, а с учетом местных изменений соотношений интенсивностей это уменьшает эффект глобальных различий, которые могут быть связаны с различиями в освещении [7]. Алгоритм основан на использовании нестационарной гамма-нормальной модели в рамках байесовского подхода к проблеме обработки изображений, описанной в разд. 1. Точно так же, как и для решения задачи удаления шума на изображении, можно оценить значения скрытого поля в соответствии с (3) и (4), используя процедуру Гаусса–Зайделя и процедуру фильтрации-интерполяции Камана, получая тем самым промежуточное слабоконтрастное изображение $Y^g = (y_t^g, t \in T)$. А затем для сжатия контраста HDR изображений применяем простую процедуру пересчета элементов изображения:

$$x_t = y_t - \alpha (y_t^g + \text{mean}(y^g)),$$

где $X = (x_t, t \in T)$ — результат сжатия контраста исходного HDR изображения; $Y = (y_t, t \in T)$ — анализируемое HDR изображение; $Y^g = (y_t^g, t \in T)$ — промежуточное слабоконтрастное изображение; α — масштабирующий коэффициент; $\text{mean}(y^g)$ — среднее значение интенсивности слабоконтрастного изображения.

В результате любое резкое изменение яркости высокого динамического диапазона изображения порождает увеличение значения яркости дисперсии шума. С другой стороны, мелкие детали, такие как текстуры, соответствуют дисперсии шума гораздо меньшей величины. Большие значения дисперсии шума критерий наказывает в большей степени, чем мелкие, таким образом, сжимая кардинальные изменения яркости, сохраняя мельчайшие детали.

5 Экспериментальные результаты

Результаты по удалению тумана с изображений, представленные на рис. 2, показывают, что предложенный метод достигает хорошего восстановления контраста и цветопередачи, в результате чего происходит значительное улучшение видимости объектов на изображении. А высокая скорость работы алгоритма позволит обрабатывать поток изображений в режиме реального времени.

Разработанная процедура имитирует схему технологического процесса художника в обратном направлении — избирательно удаляет детали из сцены, оставляя только плавно затененные области, разделенные резкими границами. Можно легко восстановить удаленные детали путем вычитания сглаженного изображения из исходного. А затем, следуя схеме прорисовки деталей, которую используют художники, повысить контраст, как показано на рис. 3.

В зависимости от степени сглаживания изображения и значения масштабирующего коэффициента регулируется степень детализации изображений.

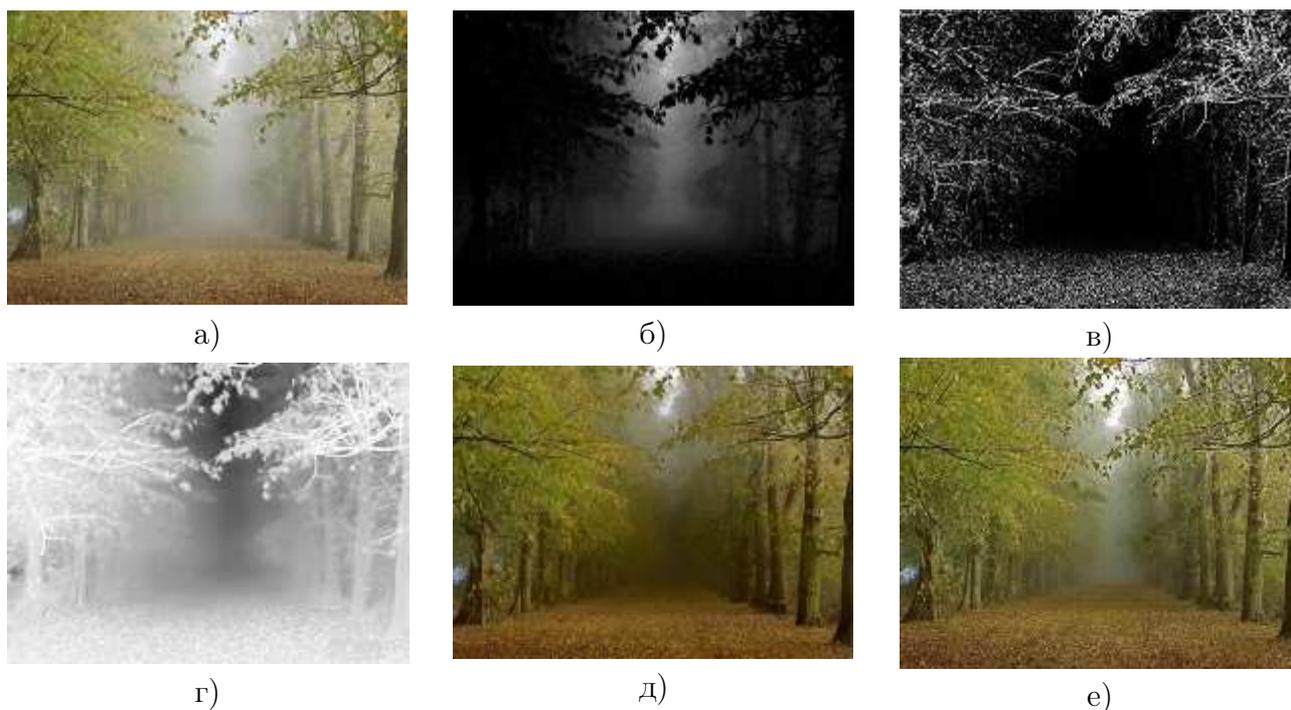


Рис. 2 Результаты удаления тумана на изображении Forest 800×600 : (а) исходное изображение; (б) выделение темного канала на исходном изображении; (в) границы объектов на исходном изображении; (г) карта светопередачи; (д) результат удаления тумана предложенным способом; (е) результат удаления тумана методом [4]

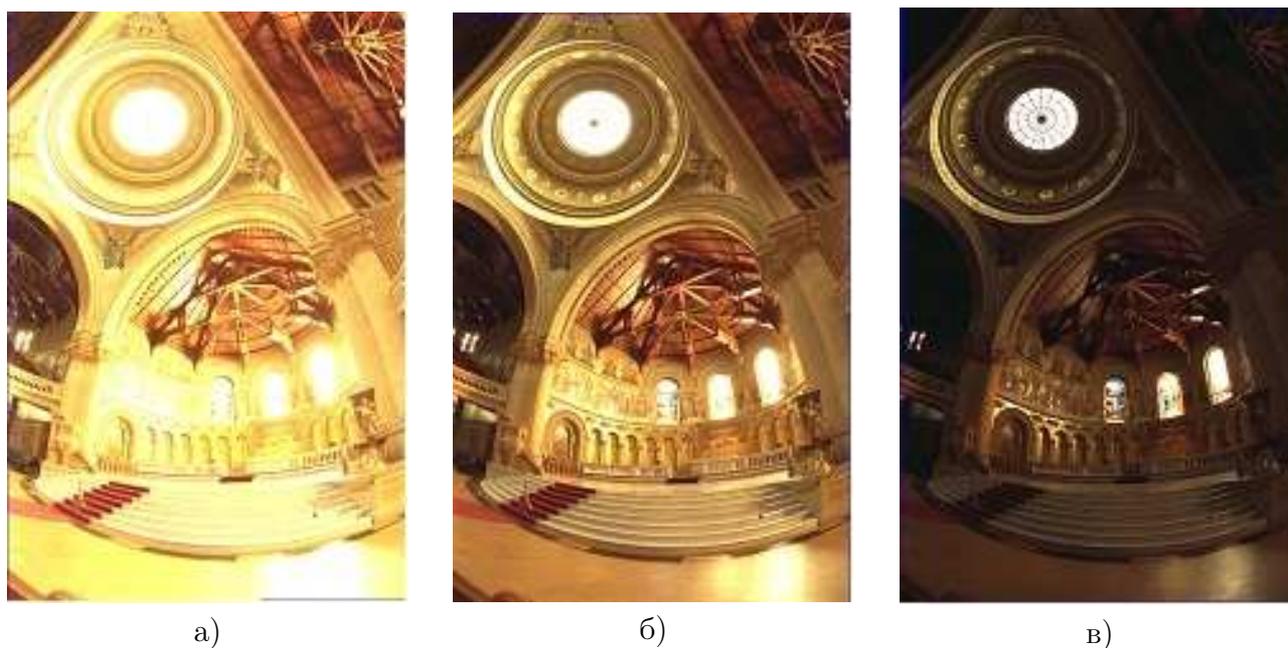


Рис. 3 Результаты сжатия контраста HDR изображения: (а) исходное изображение; (б) сжатие контраста исходного изображения при $\alpha = 0,8$; (в) сжатие контраста исходного изображения при $\alpha = 0,5$

6 Заключение

Применение нестационарной гамма-нормальной модели позволило разработать эффективные алгоритмы решения поставленных задач на основе процедуры Гаусса–Зайделя. Граф смежности элементов изображения представляется в виде решетки и затем аппроксимируется последовательностью деревьев, каждое из которых связано с отдельным столбцом изображения. Оценивание скрытых переменных вторичного массива данных в каждом дереве осуществляется за счет процедуры фильтрации-интерполяции Калмана, что позволяет достичь линейной зависимости времени работы алгоритма от размера исходного изображения.

Литература

- [1] Грачева И. А., Копылов А. В., Красоткина О. В. Адаптивная процедура обобщенного сглаживания изображений на основе статистического подхода // Известия ТулГУ. Технические науки, 2014. Вып. 11. № 1. С. 390–401.
- [2] Tan K., Oakley P. J. Physics-based approach to color image enhancement in poor visibility conditions // Optical Society of America, 2001. Vol. 18. P. 2460–2467.
- [3] Kratz L., Nishino K. Factorizing scene albedo and depth from a single foggy image // IEEE Conference (International) on Computer Vision (ICCV), 2009. P. 1701–1708.
- [4] Kaiming H., Jian S., Xiaoou T. Single image haze removal using dark channel prior // IEEE Conference (International) on Computer Vision and Pattern Recognition (CVPR), 2009. P. 1956–1963. [Прим.ред.: Рис.2 е) опубликован согласно лицензионному соглашению <https://github.com/sjtrny/Dark-Channel-Haze-Removal/blob/master/LICENSE>]
- [5] Aggarwal M., Huja A. High dynamic range panoramic imaging // IEEE Conference (International) on Computer Vision (ICCV), 2009. Vol. 1. P. 2–9.
- [6] Kristofor B., Gibson B., Truong Q. N. Fast single image fog removal using the adaptive Wiener filter // Conference (International) on Information Processing (ICIP), 2013.

- [7] DiCarlo J., Wandell B. Rendering high dynamic range images // SPIE Electronic Imaging '2000 Conference Proceedings, 2000. Vol. 3965. P. 392–401.

Поступила в редакцию 18.06.2015

References

- [1] Gracheva, I. A., A. V. Kopylov, and O. V. Krasotkina. 2014. Adaptive procedure for generalized smoothing of images on the basis of statistical approach. *News of the Tula State University. Technical Sciences* 11(1):390–401.
- [2] Tan, K., and P. J. Oakley. 2001. Physics-based approach to color image enhancement in poor visibility conditions. *Optical Society of America* 18:2460–2467.
- [3] Kratz, L., and K. Nishino. 2009. Factorizing scene albedo and depth from a single foggy image. *IEEE Conference (International) on Computer Vision (ICCV)*. 1701–1708.
- [4] Kaiming, H., S. Jian, and T. Xiaoou. 2009. Single image haze removal using dark channel prior. *IEEE Conference (International) on Computer Vision and Pattern Recognition (CVPR)*. 1956–1963. [Fig.2 e) is published according to the copyright notice <https://github.com/sjtrny/Dark-Channel-Haze-Removal/blob/master/LICENSE> - Ed.]
- [5] Aggarwal, M., and A. Huja. 2009. High dynamic range panoramic imaging. *IEEE Conference (International) on Computer Vision (ICCV)* 1:2–9.
- [6] Kristofor, B., V. Gibson, and Q. N. Truong. 2013. Fast single image fog removal using the adaptive Wiener filter. *Conference (International) on Information Processing (ICIP)*.
- [7] DiCarlo, J., and B. Wandell. 2000. Rendering high dynamic range images. *SPIE Electronic Imaging '2000 Conference (Proceedings)*. 3965:392–401.

Received June 18, 2015

Исследование влияния рассинхронизации входных изображений на качество работы алгоритмов поиска структурных различий*

Ф. А. Корнилов

fakornilov@mail.ru

Институт математики и механики им. Н. Н. Красовского Уральского отделения Российской академии наук (ИММ УрО РАН), Екатеринбург, Россия

Рассматривается влияние рассинхронизации входных изображений на качество работы алгоритмов поиска структурных различий, под которыми понимаются появившиеся или исчезнувшие объекты местности, для одновременных космических снимков земной поверхности. Работа алгоритмов оценивается на паре изображений, представляющих собой смещенные относительно друг друга копии одного и того же снимка, что позволяет выяснить степень устойчивости методов к геометрическому несоответствию входных данных с исключением влияния структурных различий или случайного шума. Приводится описание нового метода сравнения структур изображений, использование которого совместно с алгоритмами поиска структурных различий позволяет уменьшить количество ложных тревог при рассинхронизации входных изображений. Эксперименты на данных, подготовленных путем взаимного сдвига пары изначально геометрически выровненных космоснимков, показывают, что предложенная модификация алгоритмов подходит для прикладного применения.

Ключевые слова: *обработка изображений; структурные различия; рассинхронизация изображений*

DOI: 10.21469/22233792.1.12.04

Research of the impact of misregistration of input images on the accuracy of change detection*

F. A. Kornilov

Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences,
16 S. Kovalevskaja st., Ekaterinburg, Russia

Background: The paper is devoted to studying the impact of misregistration of the input images on the performance of the structural changes detection algorithms for multitemporal Earth's surface satellite images. Here, the structural changes mean appeared or disappeared ground's objects.

Methods: Algorithms performance is estimated using a pair of images, where the second image is a geometrically shifted copy of the first image; such kind of testing lets to obtain degree of robustness of the methods to geometrical misregistration of the input data without influence of the structural changes or random noise.

*Работа выполнена при частичной финансовой поддержке госбюджетной темы «Экстремальные задачи и алгоритмы теории приближения функций с приложением к проблемам навигации по геофизическим полям и управления антенными системами» (ГР №01201367475).

Results: The new method for comparing images structures is introduced; applied together with other structural changes algorithm, it leads to decrease the number of false alarms in the presence of misregistration of input images.

Concluding Remarks: The experiments on the image pairs, obtained by shifting originally aligned satellite images, show that considered algorithms modification is valuable for real applications.

Keywords: *image processing; structural differences; misregistration of images*

DOI: 10.21469/22233792.1.12.04

1 Введение

Задача поиска различий на двух изображениях возникает в различных областях компьютерного зрения, таких как сжатие видеоданных, системы видеонаблюдения и др. В данной работе рассматривается поиск различий на космических снимках земной поверхности. Для их автоматического анализа особый интерес представляет обнаружение не всех изменений, а так называемых *структурных различий*, которые заключаются в существенном изменении наблюдаемой сцены. Таковыми будем считать появление, исчезновение или изменение формы объектов: домов, дорог, участков леса и т.д. При этом изменения освещенности и цвета объектов структурными различиями не считаются.

С точки зрения требований к входным данным в задаче поиска структурных различий рассматриваются два полутоновых разновременных космических снимка одного и того же участка земной поверхности. Предполагается, что они имеют одинаковый размер в пикселях и геометрически выровнены, т.е. каждый элемент (x, y) на обоих снимках имеет одни и те же координаты на местности. Разумеется, точность такой привязки к местности будет напрямую влиять на качество работы алгоритмов, что привело к появлению статей, исследующих влияние недостаточной геометрической выровненности (рассинхронизации) изображений на результат поиска различий.

В работе [1] приведено описание эксперимента по оценке влияния рассинхронизации для изображения и его копии, благодаря чему исключено влияние структурных различий, и та же оценка для рассинхронизированных разновременных изображений. Рассинхронизация тождественна сдвигу изображения по обеим осям. Для обоих экспериментов построены графики зависимости величины, представляющей собой относительную сумму квадратов попиксельной разности исходного и сдвинутого изображений (некоторый аналог величины структурного различия, называемый полудисперсией), от величины сдвига. Графики в обоих случаях получились похожими на график степенной функции с показателем 0,5. Результаты таковы: ошибка поиска различий, связанная с рассинхронизацией изображений, составляющая 10% от величины собственно структурных различий, может быть вызвана сдвигом в 0,2 пикселя. При сдвиге в 1 пиксель эта ошибка достигает 50%. Этот подход был развит в работе [2], где авторы выбирали разные интервалы длин волн для эксперимента. Результаты аналогичны статье предшественников [1]: ошибка в регистрации 0,1934 пикселя позволяет найти 90% истинных структурных различий среди всех обнаруженных, а ошибка в 1 пиксель приводит к резкому падению точности.

Данная работа состоит из двух частей. В первой будет приведено исследование влияния рассинхронизации входных изображений на качество работы алгоритмов, описанных в [3], на основе подхода, аналогичного используемому в [1], но для более существенных сдвигов в несколько пикселей: подобное исследование интересно с точки зрения несовершенства (или отсутствия) геопривязки входных снимков или, что более вероятно, разного угла

съемки. Во второй части будет предложен специальный алгоритм, позволяющий частично устранить влияние подобной рассинхронизации, и результаты его применения к реальным снимкам.

2 Алгоритмы поиска структурных различий

В данном разделе приводится краткое описание разработанных алгоритмов поиска структурных различий изображений и их устойчивость к рассинхронизации входных снимков.

2.1 Общая схема алгоритмов

Предлагаемые в данной работе методы опираются на использование функций преобразования яркости — попиксельно действующих операторов относительной радиометрической коррекции исходных изображений, которые строятся для каждого положения сканирующего окна на основе наборов яркостей изображений в нем и применяются для «выравнивания» яркости фрагментов с сохранением их структуры. Такой подход позволяет игнорировать не считающиеся структурными различиями объекты, изменившие свой цвет, но сохранившие форму. Алгоритмы состоят из следующих шагов.

1. Исходные изображения сканируются с шагом 1 локальным окном заданного размера $d \times d$; ограниченные данным окном фрагменты будем обозначать f и g . Центральную точку этого окна будем обозначать x_c .
2. Для каждого положения окна строятся две функции преобразования яркости F_{fg} и F_{gf} . Использование схемы с двумя функциями позволяет добиться симметризации результата, т. е. становится неважным, сравнивается первое изображение со вторым или наоборот.
3. С помощью функций F_{fg} и F_{gf} строятся преобразованные фрагменты изображений $f' = F_{fg}(f)$ и $g' = F_{gf}(g)$, при этом яркость f' «выровнена» по яркости фрагмента g с сохранением структуры f , и аналогично — для фрагмента g' .
4. Строится результирующее разностное изображение:

$$R(x_c) = \max(|f'(x_c) - g(x_c)|, |g'(x_c) - f(x_c)|).$$

Для этого изображения яркость точки характеризует величину структурного несоответствия исходных изображений, т. е. чем ярче точка, тем вероятнее, что в ней присутствует структурное различие.

5. Производится пороговая обработка изображения R . Таким образом, обработка каждого сканирующего окна дает на результирующем изображении одну точку. Далее из этих точек формируются связные области структурных различий — конечный результат. Для борьбы со случайными выбросами можно фильтровать найденные области, отбрасывая те, площадь которых незначительна.

В качестве функций преобразования яркости рассматриваются:

- (1) морфологический проектор [4]:

$$f'(x) = \sum_i \frac{\sum_{x' \in X} g(x') \chi_i^f(x')}{\sum_{x' \in X} \chi_i^f(x')} \chi_i^f(x),$$

где

$$\chi_i^f(x) = \begin{cases} 1, & \text{если } f(x) = i; \\ 0, & \text{иначе,} \end{cases}$$

суммирование ведется по всем уровням яркости i фрагмента f первого изображения;

- (2) регуляризованный вариант морфологического проектора, основанный на использовании билатерального фильтра [5]:

$$f'(x) = \sum_i \left(\frac{\sum_{x' \in X} g(x') \tilde{\chi}_i^f(x')}{\sum_{x' \in X} \tilde{\chi}_i^f(x')} \frac{\tilde{\chi}_i^f(x)}{\sum_j \tilde{\chi}_i^f(x)} \right),$$

где $\tilde{\chi}_i^f$ вычисляется по формуле

$$\tilde{\chi}_i^f(x) = \exp \left(-\frac{(f(x) - i)^2}{\sigma_c^2} \right) \exp \left(-\frac{(x - x_c)^2}{\sigma_d^2} \right);$$

- (3) линейная функция преобразования яркости:

$$f'(x) = kf(x) + b,$$

где коэффициенты находятся с помощью метода наименьших квадратов (МНК) по двум фрагментам изображений, ограниченных сканирующим окном, из соотношения

$$\sum_{x \in S_d} (g(x) - (kf(x) + b))^2 \rightarrow \min_{k,b}.$$

Здесь S_d — множество точек, ограниченных сканирующим окном;

- (4) квадратичная функция преобразования яркости:

$$f'(x) = af^2(x) + bf(x) + c,$$

где коэффициенты находятся с помощью МНК по двум фрагментам изображений, ограниченных сканирующим окном, из соотношения

$$\sum_{x \in S_d} (g(x) - (af^2(x) + bf(x) + c))^2 \rightarrow \min_{a,b,c}.$$

2.2 Влияние рассинхронизации изображений на результат работы алгоритмов

Для оценки влияния рассинхронизации входных изображений будет использован подход, аналогичный предложенному в [1]. Для этого будет взят космический снимок земной поверхности (городская застройка), а в качестве второго изображения будет рассмотрена копия первого, сдвинутого на определенную величину по горизонтали и по вертикали. Полученная пара изображений, разумеется, не содержит никаких структурных различий, но из-за отсутствия геометрической выровненности таковыми станут сместившиеся границы объектов и т. д. Следовательно, относительная площадь обнаруженных различий будет эквивалентна вероятности ложной тревоги. На рис. 1 представлены полученные графики зависимости подобной ошибки от величины сдвига изображений относительно друг друга.

Отложенные на горизонтальной оси значения sd не целые, так как сдвиг осуществляется на целое значение по горизонтали и вертикали, а sd — это расстояние между пикселями после сдвига. Можно видеть, что наиболее устойчивым к возмущению входных данных оказался алгоритм, основанный на регуляризованном варианте морфологического проектора. Далее идет алгоритм, основанный на морфологическом проекторе, а линейные и квадратичные функции продемонстрировали практически идентичные результаты.

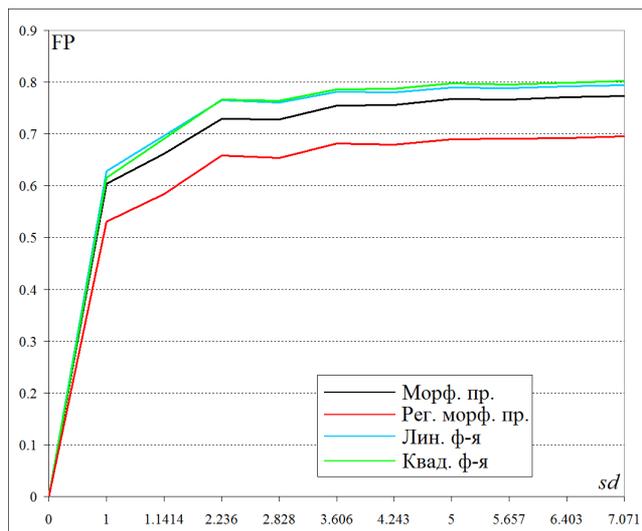


Рис. 1 Графики зависимости уровня ложной тревоги (false positive, FP) от величины рассинхронизации изображений sd , изначально не содержащих структурных различий, для различных алгоритмов

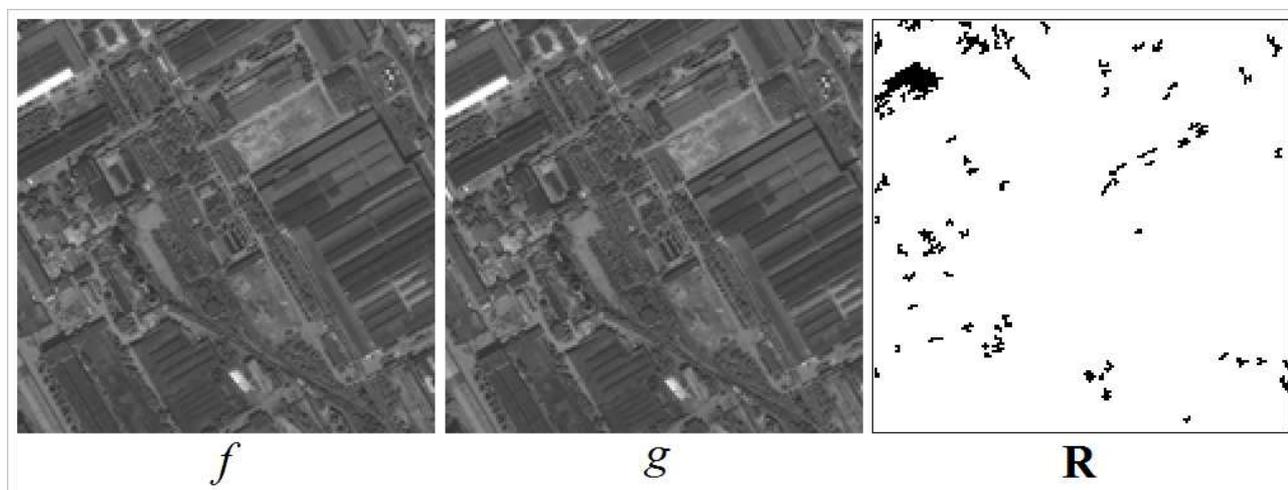


Рис. 2 Исходные изображения f и g , не содержащие структурных различий, но сдвинутые относительно друг друга на 3 пикселя по горизонтали и по вертикали, и результат поиска различий для них с помощью алгоритма, основанного на регуляризованном варианте морфологического проектора. Здесь все найденные различия (отмеченные черным цветом) являются ложной тревогой

На рис. 2 представлены сгенерированные снимки, полученные сдвигом исходного изображения на 3 пикселя по горизонтали и по вертикали, и результат поиска структурных различий для них с помощью алгоритма, основанного на регуляризованном варианте морфологического проектора. Полученное результирующее разностное изображение содержит

значительное количество ложных тревог в тех областях, где присутствуют значительные перепады яркости — таких, как границы объектов.

3 Поиск структурных различий с устранением рассинхронизации входных изображений

В соответствии с предложенным в [6] определением структура изображения — это набор его множеств уровня, а структурное различие — это набор точек, в которых яркости пары выровненных по яркости изображений существенно различаются. Работа алгоритмов их обнаружения основана на анализе яркостей в сканирующем окне и последующей оценке степени различия для его центральной точки x_c . Соответственно, при таком подходе к поиску различий на первом изображении можно рассматривать лишь уровень яркости, которому принадлежит центральная точка. Основываясь на этом рассуждении, можно предложить подход, анализирующий на каждом из изображений яркости пикселей сканирующего окна относительно яркости центральной точки x_c , благодаря чему становится возможным не использовать яркостную коррекцию снимков. Такой алгоритм получил название «степени биективности».

3.1 Анализ «степени биективности» изображений

В основе вычисления «степени биективности» SB лежит следующая идея: при отсутствии структурных различий пиксели на первом изображении близкие по яркости к $f(x_c)$ (попадающие в некоторую окрестность яркости точки x_c), на втором должны оказаться близки по яркости к $g(x_c)$. Необходимость задания окрестности связана с задачей подавления шума на изображениях. Формально это можно определить с помощью понятия штрафа. Для этого задаются два радиуса окрестностей — δ для первого и ε для второго изображения. Пиксели, яркость которых лежит за пределами δ -окрестности яркости центральной точки x_c первого изображения, а на втором изображении попадают в ε -окрестность яркости центральной точки, штрафуются на величину α , и аналогично, пиксели, яркость которых лежит в δ -окрестности яркости центральной точки x_c первого изображения, а на втором изображении не попадает в ε -окрестность яркости центральной точки, штрафуются на величину β (рис. 3). Итоговая оценка в данной точке — «степень биективности» — есть сумма штрафов для всех пикселей локального окна.

Дадим более строгое определение «степени биективности». Пусть уровни яркости фрагментов изображений f и g относительно центральной точки x_c определяются соотношениями:

$$L_f^0(x_c) = \{x \in S_d \mid |f(x) - f(x_c)| < \delta\};$$

$$L_g^0(x_c) = \{x \in S_d \mid |g(x) - g(x_c)| < \varepsilon\},$$

где S_d — множество пикселей сканирующего окна. Зададим функцию штрафа $h: S_d \rightarrow \mathbb{R}$. В частном случае каждая оказавшаяся за пределами заданной окрестности точка штрафуются на 1, но можно использовать и другие весовые функции, например функцию Гаусса, убывающую к границе окна. Тогда «степень биективности» вычисляется по формуле:

$$SB(x_c) = \sum_{x \in L^0} h(x),$$

где L^0 — симметрическая разность двух множеств: $L^0 = L_f^0(x_c) \Delta L_g^0(x_c)$.

Преимущества данного метода проявляются в следующих случаях.

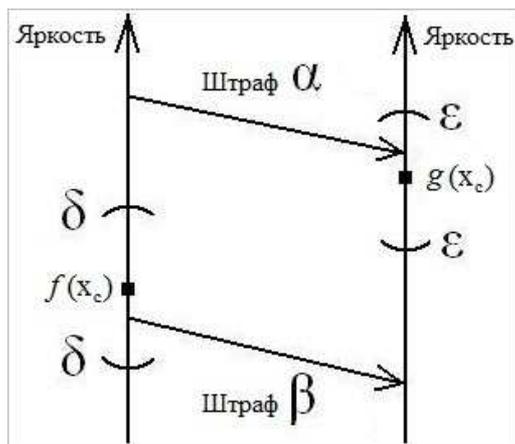


Рис. 3 Принцип вычисления «степени биективности» для двух изображений

- (1) Поиск структурных различий, размеры которых составляют несколько пикселей. При анализе космосъемки такие выбросы почти наверняка являются шумом, и все вышеперечисленные алгоритмы будут стремиться их подавить. Однако возможно появление практической задачи, где ключевую роль будут играть именно такие незначительные по площади различия, и вычисление «степени биективности» позволит их точно локализовать.
- (2) Локальная геометрическая коррекция исходных снимков, которая будет рассмотрена в следующем пункте.

3.2 Алгоритм поиска структурных различий с локальной геометрической коррекцией

Предлагается следующая стратегия использования «степени биективности»: алгоритмы работают со сканирующим окном, размеры которого относительно невелики, и можно считать, что присутствующие искажения значительно проявляются на размерах больше, чем величина стороны локального окна d , и, следовательно, в пределах выбранной окрестности локальные сдвиги пикселей являются параллельными. Тогда, прежде чем строить функции преобразования яркости, можно вычислить «степень биективности» для текущего положения окна и для всех окон со сдвигом по небольшой окрестности.

Таким образом, первый пункт приведенного выше алгоритма поиска структурных различий будет выглядеть следующим образом: исходные изображения сканируются с шагом 1 локальным окном заданного размера $d \times d$. Для каждого положения окна перебираются все возможные сдвиги из отрезка $[-sd, sd]$ по каждой координате (отдельно), т. е. перебираются все возможные положения окна из квадрата от $(-sd, -sd)$ до (sd, sd) , и вычисляется величина «степени биективности» SB . Среди полученных значений выбирается наименьшее, и для соответствующего ему положению окна строятся функции преобразования яркости и разностные изображения. При этом результат обработки записывается по-прежнему в центр окна в исходном положении.

3.3 Результаты работы алгоритмов с локальной геометрической коррекцией

Как и в первой части статьи, возьмем два геометрически выровненных космоснимка городской застройки и добавим в них локальный сдвиг, после чего применим к полученным изображениям описанные выше алгоритмы поиска структурных различий — без использования «степени биективности» и с ней. Полученные результаты сравним с ручной

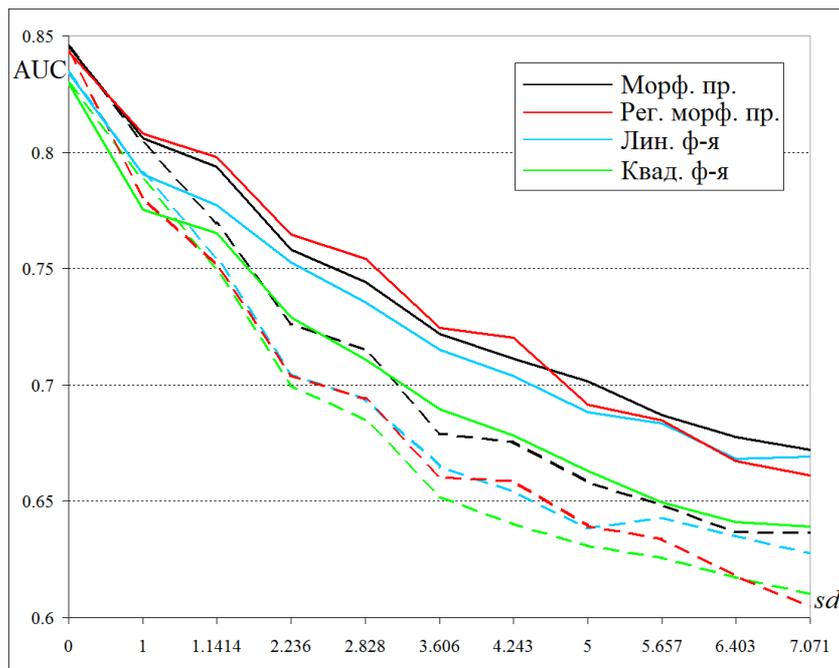


Рис. 4 Графики зависимости AUC(sd) для различных алгоритмов поиска структурных различий изображений без геометрической коррекции (пунктирные линии) и с ее использованием (сплошные линии). Размер сканирующего окна для всех алгоритмов $d = 21$

разметкой; критерием качества работы станет анализ ROC-кривых [7] (ROC — Receiver Operating Characteristic, операционная характеристика приемника). Количественную интерпретацию ROC-кривой дает показатель AUC (*Area Under the Curve* — площадь под ROC-кривой) — площадь, ограниченная ROC-кривой и осью доли ложных положительных классификаций. Чем выше показатель AUC, тем качественнее классификатор, при этом значение 0,5 соответствует случайному решению.

На рис. 4 изображены графики зависимости значения AUC от величины сдвига для алгоритмов без геометрической коррекции (пунктирные линии) и с ее использованием (сплошные линии). Во всех случаях применение «степени биективности» позволило повысить качество работы.

На рис. 5 представлено сравнение результатов работы алгоритма поиска структурных различий изображений на основе регуляризованного варианта морфологического проектора при отсутствии локальной геометрической коррекции и в ее присутствии для изображений, представляющих собой реальные снимки городской застройки, в которые была внесена погрешность их привязки друг к другу.

Параметры выбраны следующим образом: функция штрафа $h(\cdot) \equiv 1$, $\delta = 30$ и $\varepsilon = 30$ (диапазон возможных яркостей — от 0 до 255). Сдвиг изображений был сделан вручную на три пикселя вправо и три пикселя вниз; его можно оценить по смещению ярко-белого дома в левом верхнем углу снимков. Можно видеть, что без применения геометрической коррекции алгоритм нашел очень много лишних объектов — по сути, структурным различием стал любой значительный перепад яркости на снимках. В то же время тот же алгоритм, использующий вычисление «степени биективности» для геометрического сопоставления локальных окон, обнаружил только структурные различия. Таким образом, можно говорить о применимости предложенного метода для практического применения.

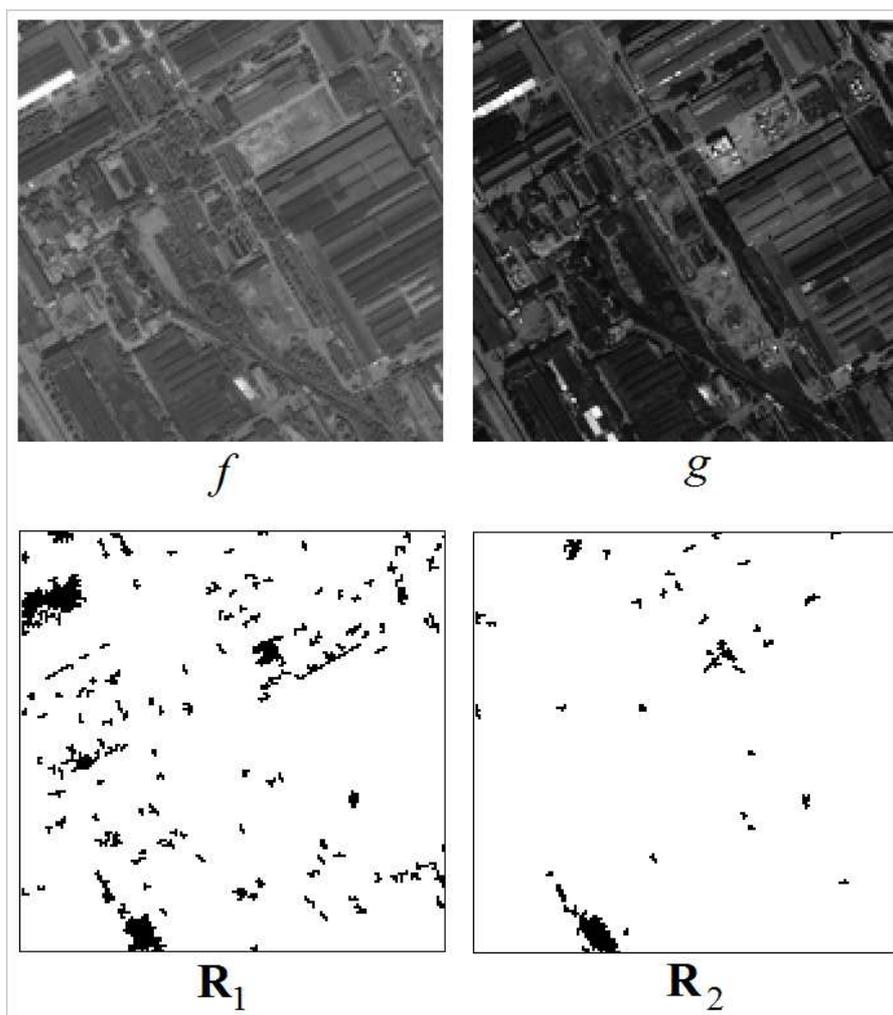


Рис. 5 Сравнение результатов работы алгоритма поиска структурных различий изображений f и g на основе регуляризованного варианта морфологического проектора при отсутствии локальной геометрической коррекции (\mathbf{R}_1) и в ее присутствии (\mathbf{R}_2). Размер сканирующего окна $d = 21$, порог $T = 25$

4 Заключение

Алгоритмы поиска структурных различий изображений, как и любые другие методы анализа изображений, очень чувствительны к качеству входных данных. При этом геометрическая выровненность изображений играет ключевую роль, ведь смещение объекта даже на один пиксель порождает сложное структурное различие. В данной работе приведены результаты оценки устойчивости алгоритмов к подобным возмущениям — разумеется, качество работы падает весьма ощутимо — и описывается метод, называемый анализом «степени биективности», чье использование в алгоритмах поиска различий позволяет значительно уменьшить количество возникающих ошибок.

Литература

- [1] Townshend J. R. G., Justice C. O., Gurney C. The impact of misregistration on change detection // IEEE Trans. Geosci. Remote Sensing, 1992. Vol. 30. Iss. 5. P. 1054–1060.

- [2] *Dai X., Khorram S.* Quantification of the impact of misregistration on the accuracy of remotely sensed change detection // *IEEE Geoscience and Remote Sensing (IGARSS 1997)*, 1997. Vol. 4. P. 1763–1765.
- [3] *Корнилов Ф. А., Перевалов Д. С.* Задача обнаружения структурных различий изображений // *Алгоритмы и програм. средства парал. вычислений.* — Екатеринбург: УрО РАН, 2011. Вып. 11. С. 37–56.
- [4] *Пытьев Ю. П., Чуличков А. И.* Методы морфологического анализа изображений. — М.: Физ-МатЛит, 2010. 336 с.
- [5] *Tomasi C., Manduchi R.* Bilateral filtering for gray and color images // *6th Conference (International) on Computer Vision Proceedings.* Bombay, India: IEEE, 1998. P. 839–846.
- [6] *Корнилов Ф. А.* Исследование алгоритма поиска структурных различий изображений // *Тр. Ин-та математики и механики УрО РАН.* — Екатеринбург: ИММ УрО РАН, 2013. Т. 19. № 1. С. 144–159.
- [7] *Green D. M., Swets J. M.* Signal detection theory and psychophysics. — New York, NY, USA: John Wiley and Sons Ltd., 1966. 455 p.

Поступила в редакцию 15.06.2015

References

- [1] Townshend, J. R. G., C. O. Justice, and C. Gurney. 1992. The impact of misregistration on change detection. *IEEE Trans. Geosci. Remote Sensing* 30(5):1054–1060.
- [2] Dai, X., and S. Khorram. 1997. Quantification of the impact of misregistration on the accuracy of remotely sensed change detection. *IEEE Geoscience and Remote Sensing (IGARSS 1997)*. 4:1763–1765.
- [3] Kornilov, F. A., and D. S. Perevalov. 2011. The problem of detecting structural differences on the images. *Algorithms and software for parallel computing.* Ekaterinburg: IMM UB RAS. 11:37–56.
- [4] Pitev, Y. P., and A. I. Chulichkov. 2010. *The methods of morphological analysis of images.* Moscow: FizMatLit. 336 p.
- [5] Tomasi, C., and R. Manduchi. 1998. Bilateral filtering for gray and color images. *6th International Conference on Computer Vision Proceedings.* Bombay, India: IEEE. 839–846.
- [6] Kornilov, F. A. 2013. Study algorithm detection structural differences on the images. *Trudy Instituta Matematiki i Mekhaniki UrO RAN.* Ekaterinburg: IMM UB RAS. 19(1):144–159.
- [7] Green, D. M., and J. M. Swets. 1966. *Signal detection theory and psychophysics.* New York, NY: John Wiley and Sons Ltd., 1966. 455 p.

Received June 15, 2015

Метод выделения движущихся объектов в видеопотоке на основе оценки их границ

Е. В. Медведева, К. А. Карлушин, Е. Е. Курбатова
emedv@mail.ru

Вятский Государственный университет, г. Киров, Россия

Предложен метод выделения движущихся объектов в видеопотоке. Метод основан на представлении последовательности изображений трехмерной цепью Маркова и оценки контуров движущихся объектов по вычисленной величине количества информации в элементах изображения. Для решения задачи выделения объекта интереса по полученным контурным точкам использован плотностный алгоритм кластеризации DBSCAN. Проведена оценка точности определения координат движущихся объектов. Разработанный метод требует небольших вычислительных ресурсов, что делает возможным его применение при обработке данных в реальном масштабе времени.

Ключевые слова: *выделение движущихся объектов; выделение контуров; трехмерная цепь Маркова; кластеризация; оценка точности определения координат*

DOI: 10.21469/22233792.1.12.05

Method of detection of moving objects in video stream on the basis of object boundaries estimation

E. V. Medvedeva, K. A. Karlushin, and E. E. Kurbatova
Vyatka State University, 36 Moskovskaya st., Kirov, Russia

The purpose of the research is to develop a new method of detection of moving objects in the frame sequence using practically stationary background. The method is based on the estimation of moving object boundaries by calculation of the value of information quantity and requires less computational resources than the existing well-known methods.

The method uses approximation of the digital halftone images (DHI) sequence by a three-dimensional Markov chain with several states and representation of the DHI by g -digit binary images (DBI). To find contours, the value of information quantity in each DBI element is calculated in accordance with various combinations of neighborhood elements. Then, the calculated value of information quantity is compared with a threshold to define whether the pixel belongs to the contour.

To define an object of interest by obtained contour points, DBSCAN density clustering algorithm has been used.

The proposed method of moving objects contours definition requires small computational resources as for each element, only operations of comparison with three neighboring elements are carried out. The developed method is faster in comparison with the known method in 2–5.6 times.

The results of modeling of the developed method are shown. The gain in root mean-squared error of coordinates determination accuracy for the developed method in comparison with the known subtraction method is 1.5–2.5.

The developed method requires small computational resources; thus, it can be applied in real-time data processing. The range of variation of object dimensions in videosequence can be wide and the number of the moving objects can be priori unknown.

Keywords: *moving objects detection; three-dimensional Markov chains; contours detection; clustering; estimation of coordinates determination accuracy*

DOI: 10.21469/22233792.1.12.05

1 Введение

В некоторых системах видеонаблюдения, применяемых, например, для автоматического обнаружения событий и контроля транспортных потоков, необходимо выделить и оценить параметры движущихся объектов на сравнительно неподвижном фоне. Как правило, такие системы должны с большой вероятностью обнаруживать объекты интереса. При этом диапазон варьирования размеров объектов интереса может быть большим, а количество движущихся объектов — априорно неизвестным. Поскольку характер условий наблюдения, определяемый конкретным местом установки видеодатчиков, как правило, мало предсказуем, а вычислительные ресурсы ограничены, возникает потребность в разработке устойчивых и вместе с тем достаточно быстродействующих алгоритмов анализа изображений. Несмотря на значительные успехи в этой области [1–4], универсальных подходов, одинаково эффективных при всем многообразии условий наблюдения, до сих пор не найдено. Поэтому разработка эффективных методов выделения объектов движения в видеопотоке, позволяющих в реальном масштабе времени выделять объекты интереса с высокой точностью, остается актуальной задачей.

Цель данной работы — разработка метода выделения движущихся объектов в последовательности кадров на сравнительно неподвижном фоне на основе оценки их границ по вычисленной величине количества информации, реализация которого достигается меньшими вычислительными ресурсами, чем известными методами и высокой точностью.

2 Математическая модель видеопоследовательности

Поскольку последовательности видеок кадров представляют собой трехмерные сигналы, для выделения движущихся объектов необходимо использовать пространственно-временную математическую модель изображения. Пространственно-временная модель должна быть адекватна реальным последовательностям изображений и обладать высокой вычислительной эффективностью, чтобы алгоритмы обработки изображений, синтезируемые на ее основе, требовали небольших вычислительных ресурсов.

Такой математической моделью является модель на основе трехмерных цепей Маркова [5].

Предполагается, что цифровые полутонные изображения (ЦПИ) в видеопоследовательности представлены g -разрядными двоичными числами. Это позволяет разбить последовательность ЦПИ на последовательности g разрядных двоичных изображений (РДИ), каждое из которых является суперпозицией трех простых цепей Маркова с двумя равновероятными дискретными состояниями $M_1^{(l)}$ и $M_2^{(l)}$ и матрицами вероятностей переходов по горизонтали ${}^1\Pi = \left\| \pi_{ij}^{(l)} \right\|_{2 \times 2}$, вертикали ${}^2\Pi = \left\| \pi_{ij}^{(l)} \right\|_{2 \times 2}$ и времени ${}^4\Pi = \left\| \pi_{ij}^{(l)} \right\|_{2 \times 2}$ ($l = 1, \dots, g, i, j = \overline{1, 2}$).

На приведенном фрагменте последовательности РДИ (рис. 1) элемент изображения $\nu_4^{(l)}$ зависит от состояний трех окрестных элементов $\Lambda_{i,j,k}^{(l)} = \left\{ \nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)} \right\}$, расположенных в k -м и $(k - 1)$ -м кадрах.

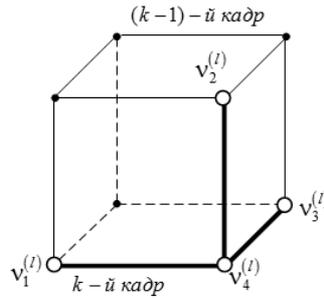


Рис. 1 Фрагмент последовательности РДИ

Количество информации в элементе $\nu_4^{(l)}$ относительно состояний элементов окрестности $\Lambda_{i,j,k}^{(l)} = \{\nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)}\}$ можно определить как разность собственной информации в элементе $\nu_4^{(l)}$ и взаимной информации, полученной от элементов окрестности $\Lambda_{i,j,k}^{(l)}$ [5, 6]:

$$\begin{aligned}
 I\left(\nu_4^{(l)} \mid \nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)}\right) &= I\left(\nu_4^{(l)}\right) - I\left(\nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)}, \nu_4^{(l)}\right) = \\
 &= - \left[\log p\left(\nu_4^{(l)}\right) + \log \frac{\prod p\left(\nu_i^{(l)}, \nu_j^{(l)}\right) p\left(\nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)}, \nu_4^{(l)}\right)}{\prod_{i=1}^4 p\left(\nu_i^{(l)}\right) \prod p\left(\nu_i^{(l)}, \nu_j^{(l)}\right)} \right] = \\
 &= - \log \frac{w\left(\nu_4^{(l)} \mid \nu_1^{(l)}\right) w\left(\nu_4^{(l)} \mid \nu_2^{(l)}\right) w\left(\nu_4^{(l)} \mid \nu_3^{(l)}\right) w\left(\nu_4^{(l)} \mid \nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)}\right)}{w\left(\nu_4^{(l)} \mid \nu_1^{(l)}, \nu_2^{(l)}\right) w\left(\nu_4^{(l)} \mid \nu_1^{(l)}, \nu_3^{(l)}\right) w\left(\nu_4^{(l)} \mid \nu_2^{(l)}, \nu_3^{(l)}\right)}, \quad (1)
 \end{aligned}$$

где произведения, обозначенные символом \prod , вычисляются по всем возможным несовпадающим комбинациям различных подстрочных индексов трехмерного случайного поля $p\left(\nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)}, \nu_4^{(l)}\right)$, $p\left(\nu_i^{(l)}, \nu_j^{(l)}, \nu_k^{(l)}\right)$, $p\left(\nu_i^{(l)}, \nu_j^{(l)}\right)$, $i = j = k = 1, \dots, 4$; $i \neq j \neq k$ — совместные плотности вероятностей значений элементов; $p\left(\nu_i^{(l)}\right)$, $i = 1, \dots, 4$ — априорные плотности вероятностей значений элементов; $w\left(\nu_4^{(l)} \mid \nu_i^{(l)}\right)$, $i = 1, \dots, 3$ — одномерные плотности вероятностей переходов; $w\left(\nu_4^{(l)} \mid \nu_i^{(l)}, \nu_j^{(l)}\right)$, $i = j = 1, \dots, 3$; $i \neq j$ — плотности вероятностей переходов в двумерной цепи Маркова; $w\left(\nu_4^{(l)} \mid \nu_1^{(l)}, \nu_2^{(l)}, \nu_3^{(l)}\right)$ — плотность вероятности переходов в трехмерной цепи Маркова.

Плотность вероятности переходов в трехмерной двоичной цепи Маркова $w\left(\nu_4^{(l)} \mid \Lambda_{i,j,k}^{(l)}\right)$ можно представить в форме:

$$\begin{aligned}
 w\left(\nu_4^{(l)} \mid \Lambda_{i,j,k}^{(l)}\right) &= \sum_{i, \dots, r=1}^2 \pi\left(\nu_4^{(l)} = M_i^{(l)} \mid \nu_1^{(l)} = M_j^{(l)}; \nu_2^{(l)} = M_k^{(l)}; \nu_3^{(l)} = M_r^{(l)}\right) \times \\
 &\quad \times \delta\left(\nu_1^{(l)} - M_j^{(l)}\right) \delta\left(\nu_2^{(l)} - M_k^{(l)}\right) \delta\left(\nu_3^{(l)} - M_r^{(l)}\right), \quad (2)
 \end{aligned}$$

где $\delta(\cdot)$ — дельта-функция.

В трехмерной математической модели вероятности смены состояний элемента $\nu_4^{(l)}$ зависят от различных сочетаний состояний элементов окрестности $\Lambda_{i,j,k}^{(l)}$ и их корреляции

с элементом $\nu_4^{(l)}$. Все возможные значения этих вероятностей для различных комбинаций состояний элементов окрестности $\Lambda_{i,j,k}^{(l)}$ могут быть вычислены в соответствии с аргументом выражения (1) и сведены в матрицу Π вида:

$$\Pi = \begin{pmatrix} \pi_{iii}^{(l)} & \pi_{iiij}^{(l)} \\ \pi_{iiji}^{(l)} & \pi_{ijjj}^{(l)} \\ \vdots & \vdots \\ \pi_{jjji}^{(l)} & \pi_{jjjj}^{(l)} \end{pmatrix} = \begin{pmatrix} \alpha_1^{(l)} & \alpha_1'^{(l)} \\ \alpha_2^{(l)} & \alpha_2'^{(l)} \\ \vdots & \vdots \\ \alpha_8^{(l)} & \alpha_8'^{(l)} \end{pmatrix}, \quad i, j = 1, 2; \quad i \neq j, \quad (3)$$

элементы которой симметричны и удовлетворяют условию нормировки $\alpha_q + \alpha_q'^{(l)} = 1$.

Например, выражения для вычисления элементов первой строки матрицы Π имеют вид:

$$\alpha_1^{(l)} = \pi_{iii}^{(l)} = \frac{{}^1\pi_{ii}^{(l)} \cdot {}^2\pi_{ii}^{(l)} \cdot {}^4\pi_{ii}^{(l)} \cdot {}^7\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)} \cdot {}^5\pi_{ii}^{(l)} \cdot {}^6\pi_{ii}^{(l)}}; \quad \alpha_1'^{(l)} = \pi_{iiij}^{(l)} = \frac{{}^1\pi_{ij}^{(l)} \cdot {}^2\pi_{ij}^{(l)} \cdot {}^4\pi_{ij}^{(l)} \cdot {}^7\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)} \cdot {}^5\pi_{ii}^{(l)} \cdot {}^6\pi_{ii}^{(l)}}, \quad (4)$$

где ${}^r\pi_{ii}^{(l)}$ ($r = 1, \dots, 7$) — элементы матриц вероятностей переходов в одномерных цепях Маркова с двумя состояниями по горизонтали ${}^1\Pi^{(l)}$, вертикали ${}^2\Pi^{(l)}$, времени ${}^4\Pi^{(l)}$ и четырех сопутствующих матриц вероятностей переходов:

$${}^3\Pi^{(l)} = {}^1\Pi^{(l)} \times {}^2\Pi^{(l)}; \quad {}^5\Pi^{(l)} = {}^1\Pi^{(l)} \times {}^4\Pi^{(l)}; \quad {}^6\Pi^{(l)} = {}^2\Pi^{(l)} \times {}^4\Pi^{(l)}; \quad {}^7\Pi^{(l)} = {}^3\Pi^{(l)} \times {}^4\Pi^{(l)}. \quad (5)$$

Вычисление остальных элементов матрицы $\Pi^{(l)}$ осуществляется в соответствии с состоянием элементов окрестности $\Lambda_{i,j,k}^{(l)}$.

3 Метод выделения движущихся объектов

В данной работе для обнаружения объектов интереса предложено выделять их контуры. Контуры являются наиболее информативными составляющими изображения. По контурам можно определить размер, форму, положение объекта, повысить точность определения координат движущихся объектов и т. д. За счет того что контурные точки составляют незначительную часть всех точек на изображениях, работа с ними позволяет сократить объем запоминающих устройств, повысить быстродействие цифровых систем обработки изображений, улучшить массогабаритные характеристики и снизить энергопотребление.

На рис. 2 представлены зависимости вероятностей переходов между состояниями элементов в последовательности РДИ по горизонтали, вертикали и между кадрами от номера разряда ЦПИ. Графики получены усреднением статистических характеристик по большому количеству реальных изображений, полученных стационарной видеокамерой, аналогичных рис. 4 (см. разд. 4).

Из приведенного графика видно, что наибольшей корреляцией обладают элементы, расположенные в старших разрядах ЦПИ. Поэтому контуры основных детальных областей можно выделить по РДИ старших разрядов ЦПИ.

Для выделения контуров движущихся объектов в каждом элементе РДИ в соответствии с состояниями элементов окрестности вычисляют величину количества информации (1). Для видеопоследовательностей, полученных стационарной видеокамерой при неизменном уровне освещенности, будет характерно незначительное изменение вероятностей переходов между элементами изображения от кадра к кадру в течение длительного промежутка времени. Поэтому предполагалось, что вероятности переходов между

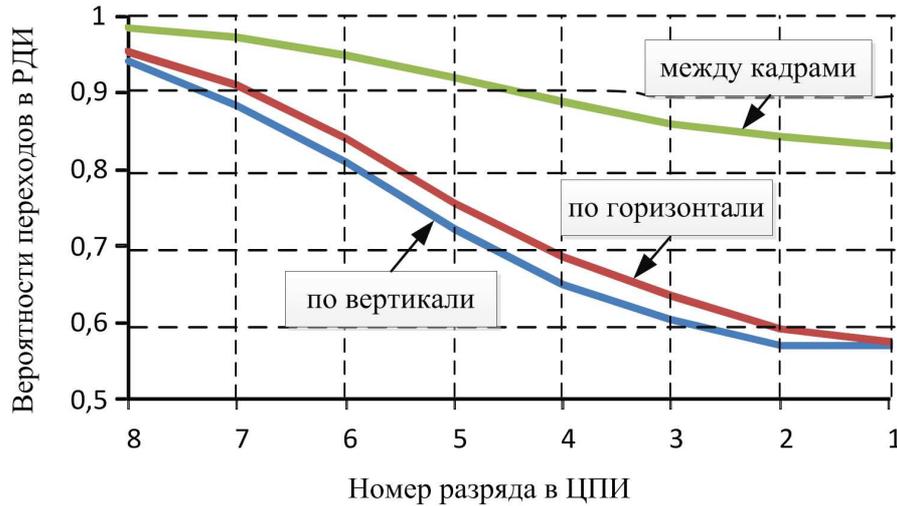


Рис. 2 Зависимости вероятностей переходов между состояниями элементов в последовательности РДИ (8 — старший разряд; 1 — младший разряд)

элементами видеопоследовательности постоянны и априорно известны, причем корреляция элементов между кадрами (${}^4\pi_{ij}$) больше, чем внутри кадра (${}^1\pi_{ii}$, ${}^2\pi_{ii}$) (см. рис. 2).

При отсутствии объекта (контрастной области) состояния окрестных элементов $\Lambda_{i,j,k}^{(l)}$ будут совпадать с $\nu_4^{(l)}$ и количество информации в элементе $\nu_4^{(l)}$ будет минимально.

При наличии неподвижного объекта, на его границе состояние элемента $\nu_4^{(l)}$ будет совпадать с состоянием элемента $\nu_4^{(l)}$ в $(k - 1)$ -м кадре и не совпадать с одним или двумя окрестными элементами в k -м кадре и количество информации в $\nu_4^{(l)}$ возрастет.

При смещении объекта элемент $\nu_4^{(l)}$ будет иметь различные состояния с элементом $\nu_3^{(l)}$ в $(k - 1)$ -м кадре и с одним или двумя окрестными элементами в k -м кадре и количество информации в элементе $\nu_4^{(l)}$ еще более увеличится. При этом чем больше корреляция между соседними кадрами, что характерно для видеопоследовательностей, полученных неподвижным датчиком, тем больше информации будет в элементе $\nu_4^{(l)}$ при смещении объектов.

Сравнивая значения вычисленной величины количества информации в элементе изображения с порогом, определяем принадлежность данной точки контуру. Значение порога h определяют по вычисленной величине количества информации в элементе $\nu_4^{(l)}$ l -го РДИ:

$$h = \frac{I\left(\nu_4^{(l)} = M_i^{(l)} \mid \nu_1^{(l)} = M_i^{(l)}, \nu_2^{(l)} = M_j^{(l)}, \nu_3^{(l)} = M_j^{(l)}\right)}{2} + \frac{I\left(\nu_4^{(l)} = M_i^{(l)} \mid \nu_1^{(l)} = M_j^{(l)}, \nu_2^{(l)} = M_j^{(l)}, \nu_3^{(l)} = M_i^{(l)}\right)}{2}. \quad (6)$$

Предложенный метод выделения контуров движущихся объектов требует небольших вычислительных ресурсов, определяемых операциями сравнения с тремя окрестными элементами. На рис. 3 приведено сравнение во времени обработки изображений разной размерности предложенным методом на основе трехмерных цепей Маркова, позволяющим определить контуры движущихся объектов, а также методами, выделяющими контуры подвижных и неподвижных объектов: на основе двумерных цепей Маркова [7] и известными методами Робертса, Превитта, Собела и Канни [1–4, 8].

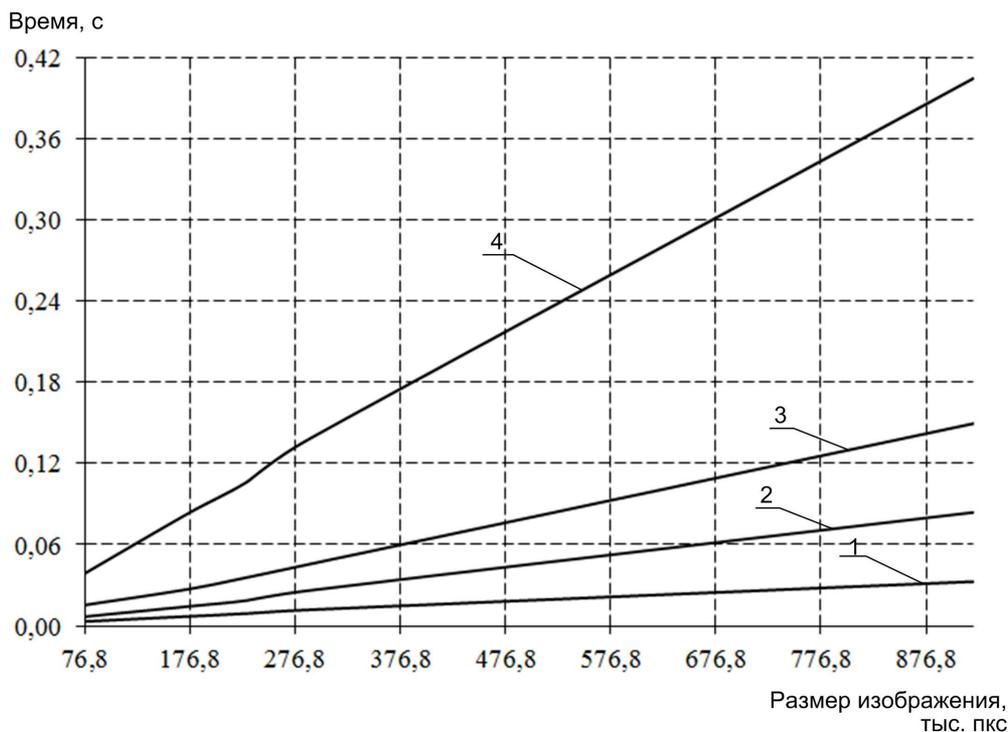


Рис. 3 Время обработки изображений методами: 1 — на основе двумерной цепи Маркова; 2 — на основе трехмерной цепи Маркова; 3 — Робертса, Превитта и Собела; 4 — Канни

Исследование проводилось в операционной системе Windows 7 на процессоре Intel Core i7-3770, 3,4 ГГц. Полученные оценки по времени разработанного метода свидетельствуют о его преимуществах по сравнению с распространенными методами выделения границ (до 2–5,6 раз). Следует отметить, что двумерные алгоритмы выделяют контуры всех контрастных областей на изображении, поэтому для обнаружения подвижных объектов потребуются еще дополнительные преобразования изображения, что приведет к увеличению вычислительных и временных ресурсов.

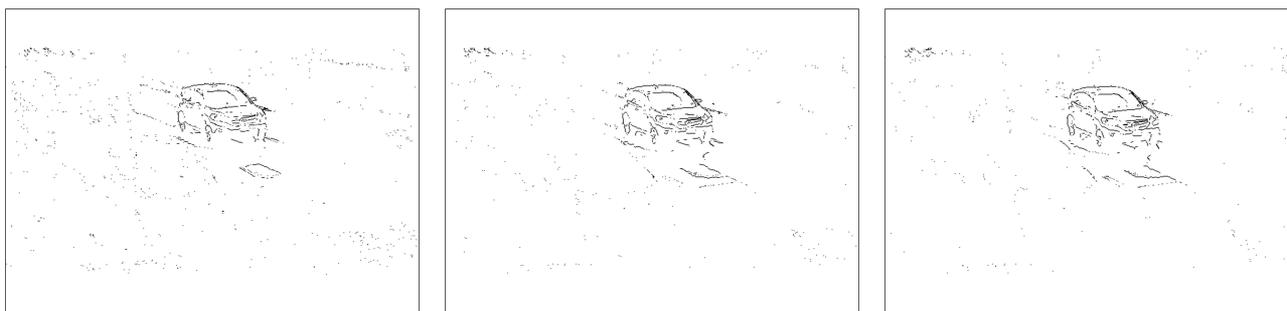
Результатом обработки изображений алгоритмом на основе трехмерных цепей Маркова является совокупность контурных точек, в которых происходит движение одного или нескольких объектов интереса, а также точек, представляющих собой шум, вызванный, например, движением листьев на деревьях, бликами и т. п. При этом контурные точки, принадлежащие объекту интереса, являются связными и имеют большую протяженность в отличие от контурных точек, представляющих собой шум.

Для решения задачи выделения объекта интереса необходимо выделить точки, которые наилучшим образом представляют объект. В данной работе использован плотностный алгоритм кластеризации DBSCAN [9]. В отличие от других известных алгоритмов кластеризации, например k-means, BIRCH, Fuzzy C-means, CURE, алгоритм DBSCAN не требует задания числа кластеров, что позволяет обнаруживать произвольное число объектов. Также алгоритм позволяет выделять кластеры произвольной формы в отличие от большинства алгоритмов, где форма создаваемых кластеров близка к сферической. Для работы алгоритма требуется задание двух параметров: максимальный размер области ϵ и требуемое количество точек, находящихся на расстоянии ϵ от текущей точки.

Координаты движущегося объекта соответствуют центру прямоугольного окна, в которое вписывается объект интереса.



Рис. 4 Совмещенные кадры видеопоследовательности



(а) Контурные изображения



(б) Изображения с выделенными объектами

Рис. 5 Кадры видеопоследовательности с выделенными объектами

4 Результаты исследований

При выполнении исследования эффективности разработанного метода использовались несколько видеопоследовательностей средней длительностью 300 кадров. Для оценки качества выделения движущегося объекта вычислялась вероятность правильного выделения объекта интереса и среднеквадратическая ошибка (СКО) определения координат объекта интереса.

На рис. 4 показан фрагмент видеопоследовательности, в котором три соседних кадра совмещены для визуальной оценки передвижения объекта. На рис. 5 приводятся контурные (рис. 5, а) и выделенные прямоугольным окном объекты (рис. 5, б).

В результате выделения объектов интереса возникают ошибки двух типов: на контурном изображении объект интереса определен (как совокупность точек шума), в то время

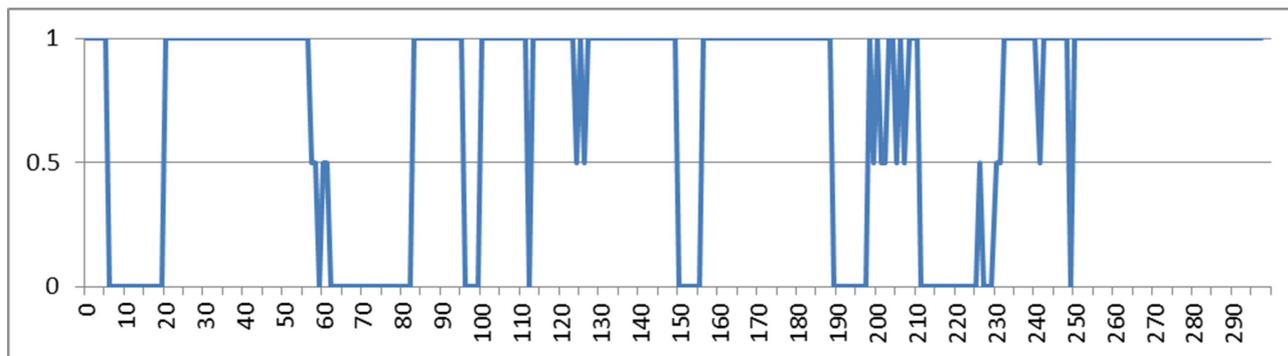


Рис. 6 Зависимость вероятности правильного выделения объекта от номера кадра

как на реальном изображении он отсутствует; на контурном изображении объект не выделен (например, в результате частичного заслонения другим объектом), а на реальном изображении он имеется.

Вероятность правильного выделения объекта интереса определялась как отношение количества верно выделенных объектов (N_B) к общему количеству движущихся объектов (N_O) в видеопоследовательности:

$$P_{\text{пр}} = \frac{N_B}{N_O}. \quad (7)$$

Для приведенной видеопоследовательности общая вероятность правильного определения объекта интереса составила 0,75, что вызвано скрыванием объекта интереса препятствием. Размеры объектов интереса варьировались от 30×30 до 120×120 пикселей. Количество одновременно присутствующих объектов в кадре менялось на протяжении всей видеопоследовательности, но не превышало двух. Зависимость вероятности правильного выделения объекта в тестовой видеопоследовательности от номера кадра представлена на рис. 6.

Кадры, в которых вероятность правильного выделения объекта интереса равна нулю, соответствуют ситуации с заслонением объекта интереса. Для других видеопоследова-

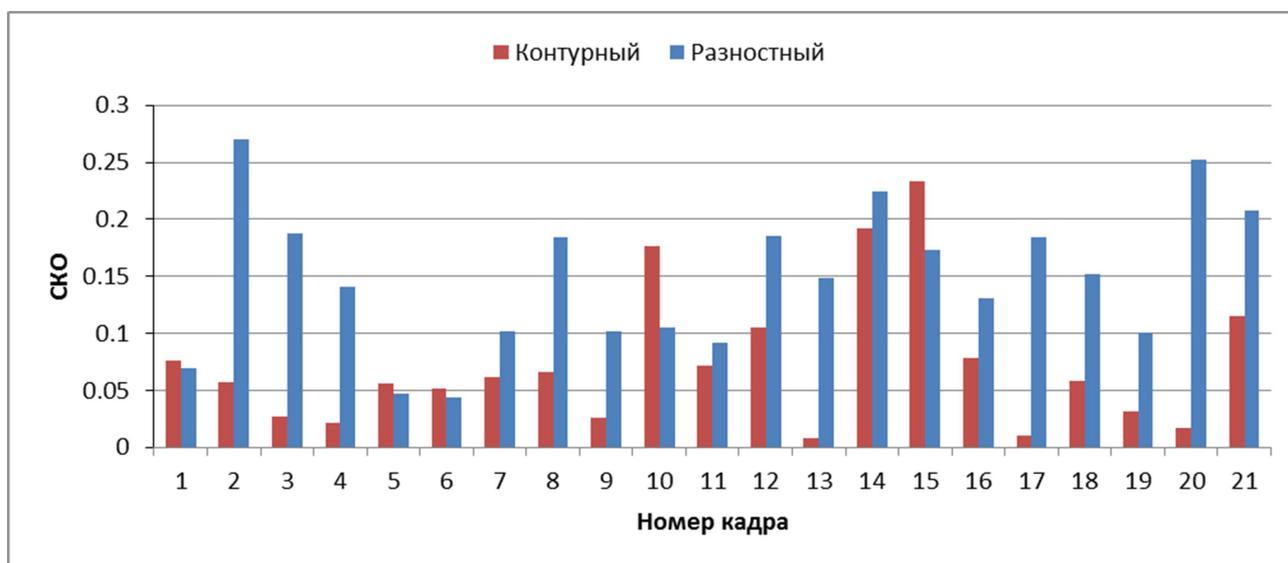


Рис. 7 Сравнение СКО определения координат движущихся объектов разными методами

тельность, в которых не происходит скрытие объектов интереса, величина правильного выделения достигает 0,91–0,95.

Также для оценки качества работы алгоритма вычислялась нормированная СКО оценки координат движущегося объекта:

$$\text{СКО} = \frac{1}{L} \sqrt{\sigma_x^2 + \sigma_y^2}, \quad (8)$$

где L — половина диагонали прямоугольного окна, обрамляющего объект интереса в идеальном случае; σ_x и σ_y — смещение координат центра движущегося объекта по горизонтали и вертикали от их истинного значения, которое определялось по выделенному пользователем объекту (т. е. вручную) как пересечение диагоналей прямоугольного окна.

При выполнении исследования сравнивались результаты СКО оценки координат разработанным и известным разностным методом [2]. Результаты сравнения СКО двумя методами на нескольких кадрах последовательности (см. рис. 4) представлены на рис. 7. При оценке всей видеопоследовательности из 300 кадров СКО разработанного контурного метода меньше разностного в 2 раза. Для большинства видеопоследовательностей выигрыш в СКО для разработанного метода по сравнению с разностным составляет 1,5–2,5.

5 Заключение

Из приведенных результатов следует, что предложенный метод выделения движущихся объектов в видеопотоке, основанный на вычислении количества информации в элементах изображения, позволяет определить координаты движущегося объекта с точностью, превышающей разностный метод. Разработанный метод требует небольших вычислительных ресурсов, что делает возможным его применение при обработке данных в реальном масштабе времени.

Литература

- [1] *Yokoyma M., Poggio T.* A contour-based moving object detection and tracking // 14th Conference (International) on Computer and Networks Proceedings. Washington, DC, USA: IEEE Computer Society, 2005. P. 271–276.
- [2] *Яне Б.* Цифровая обработка изображений. — М.: Техносфера, 2007. 584 с.
- [3] *Алпатов Б. А., Бабаян П. В., Балашов О. Е., Степашикин А. И.* Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. — М.: Радиотехника, 2008. 176 с.
- [4] *Лукьяница А. А., Шишкин А. Г.* Цифровая обработка видеоизображений. — М.: Ай-Эс-Эс Пресс, 2009. 518 с.
- [5] *Петров Е. П., Медведева Е. В., Метелев А. П.* Метод синтеза математических моделей видеоизображений на основе многомерных цепей Маркова // *Нелинейный мир*, 2011. № 4. С. 213–231.
- [6] *Медведева Е. В., Карлушин К. А.* Метод выделения движущихся объектов в видеопотоке и оценка точности определения их координат // *Цифровая обработка сигналов*, 2014. № 3. С. 47–50.
- [7] *Medvedeva E. V., Kurbatova E. E.* A two-stage image preprocessing algorithm // *Pattern Recogn. Image Anal.*, 2011. Vol. 21. No. 2. P. 297–301.
- [8] *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. — М.: Техносфера, 2012. 1104 с.

- [9] Ester M., Kriegel H.-P., Sander J., Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise // 2nd Conference (International) on Knowledge Discovery and Data Mining (KDD-96) Proceedings. AAAI Press, 1996. P. 226–231.

Поступила в редакцию 09.07.2015

References

- [1] Yokoyama, M., and T. Poggio. 2005. A contour-based moving object detection and tracking. *14th Conference (International) on Computer and Networks Proceedings*. Washington, DC: IEEE Computer Society. 271–276.
- [2] Jahne, B. 2005. *Digital image processing: Concepts, algorithms, and scientific applications*. Berlin–Heidelberg: Springer-Verlag. 584 p.
- [3] Alpatov, B. A., P. V. Babayan, O. E. Balashov, and A. I. Stepashkin. 2008. *Methods of automatic detection and tracking of objects*. Moscow: Radiotechnica. 176 p.
- [4] Lukjanica, A. A., and A. G. Shishkin. 2009. *Digital video image processing*. Moscow: ISS Press. 518 p.
- [5] Petrov, E. P., E. V. Medvedeva, and A. P. Metelyov. 2011. Method of synthesis of video images mathematical models based on multidimensional Markov chains. *Nonlinear World* 4:213–231.
- [6] Medvedeva, E. V., and K. A. Karlushin. 2014. A method of moving objects detection in video stream and estimation accuracy of objects coordinates. *Digital Signal Processing* 3:47–50.
- [7] Medvedeva, E. V., and E. E. Kurbatova. 2011. A two-stage image preprocessing algorithm. *Pattern Recogn. Image Anal.* 21(2):297–301.
- [8] Gonzalez R., Woods R. 2012. *Digital image processing*. Moscow: Technosphaera. 1104 p.
- [9] Ester, M., H.-P. Kriegel, J. Sander, and X. Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *2nd Conference (International) on Knowledge Discovery and Data Mining (KDD-96) Proceedings*. AAAI Press. 226–231.

Received July 9, 2015

О моделях нейронов агрегирующего типа*

З. М. Шибзухов^{1,2}, Д. Ю. Чередников²

szport@gmail.com

¹Московский педагогический государственный университет, Москва

²Институт прикладной математики и автоматизации, Нальчик

Описан новый класс моделей искусственных нейронов агрегирующего типа. Модели агрегирующих нейронов строятся на основе следующих принципов: (1) все вклады синапсов суммируются при помощи агрегирующей операции; (2) вклады простых синапсов, которые образуют сложный синапс или синаптический кластер, преобразуются также при помощи некоторой другой агрегирующей операции. Они охватывают большую часть моделей искусственных нейронов функционального типа. Для класса агрегирующих нейронов, обобщающих модель $\Sigma\Pi$ -нейрона, показано, что они могут быть корректно обучены по конечным наборам прецедентов.

Ключевые слова: *нейронная сеть; модель нейрона; агрегирующая операция*

DOI: 10.21469/22233792.1.12.06

About models of neurons of aggregation type*

Z. M. Shibzukhov^{1,2} and D. Y. Cherednikov²

¹Moscow Pedagogical State University, 1 M. Pirogovskaya st., Moscow, Russia

²Institute of Applied Mathematics and Automatization, 89-A Shortanova, Nalchik, Russia

A new class of models of artificial neurons is described in this work. These models are based on the following principles: (i) contributions of synapses are summed with the help of certain aggregation operation; and (ii) contribution of complex synapse or synaptic cluster is computed with the help of another aggregation operation on the set of simple synapses. These models include a big part of the known functional models of neurons. For a class of the aggregating neurons generalizing model $\Sigma\Pi$ -neuron, it is shown that they can be correctly trained on the final sets of precedents.

Keywords: *neural networks; neuron model; aggregation operation*

DOI: 10.21469/22233792.1.12.06

В одном из подходов *искусственная нейронная сеть* (далее просто нейронная сеть) рассматривается как направленный граф, в узлах которого находятся *нейроны*. Каждый нейрон в сети имеет (1) несколько входных каналов от других нейронов или от входов сети и (2) только один выходной канал для передачи на выход нейронной сети или, разветвляясь, на входы других нейронов.

Значительные классы нейронных сетей — нейронные сети *прямого распространения*. Они разделяются на слои, так что (1) нейроны, принадлежащие одному слою, не связаны друг с другом по входам и выходам; (2) выходы нейронов в слое поступают на вход нейронов из другого слоя или на выход сети. Слои пронумерованы, так что на вход нейронов могут поступать сигналы из слоя только с меньшим номером или от входов сети.

Нейроны головного мозга являются сложными его элементами. Они имеют разветвленную *дендритную систему*, в которую сигналы поступают от других нейронов или входов

*Работа выполнена при поддержке гранта РФФИ № 15-01-03381а.

нейронной сети при помощи *синапсов*. Синапсы могут быть как простые, когда в их образовании участвует единственный вход, так и сложные, когда в их образовании участвуют одновременно несколько входов. Также могут иметь место пространственно локализованные *синаптические кластеры*, которые образуют зоны дендритной системы, где обработка информации ведется независимо от других зон. При этом внутри кластера синапсы могут также оказывать влияние друг на друга.

Для описания преобразования сигналов в нейроне вводится понятие *суммарного потенциала нейрона*, который затем преобразуется в выход нейрона. Входы нейрона преобразуются посредством синапсов и вносят определенный вклад в увеличение или уменьшение суммарного потенциала нейрона. Дендритная система нейрона *агрегирует* вклады синапсов и формирует суммарный потенциал нейрона. На его основе генерируется выходной сигнал нейрона.

Существуют разные типы моделей нейрона, например стохастические модели, когда процесс преобразования информации в нейроне рассматривается как случайный процесс. В детерминированных моделях нейрона его выход и его входы связаны при помощи детерминированной функциональной зависимости. Здесь рассматриваются *функциональные модели* нейронов, в которых значение на выходе является детерминированной функцией от значений на входах нейрона. Различные функциональные модели нейрона получаются в зависимости от того (1) как моделируются простые и сложные синапсы; (2) какие модели используются для описания синаптических кластеров; (3) как агрегируются вклады синапсов.

Большая часть теоретических исследований проводилась для нейронных сетей, построенных на основе классических нейронов вида:

$$y = \sigma(w_0 + w_1x_1 + \dots + w_nx_n),$$

где σ — сигмоидальная функция. Был получен ряд теоретических результатов, которые характеризуют способности таких нейронных сетей по аппроксимации зависимостей. Первый из них — это аппроксимационный вариант теоремы Колмогорова [1, 2], в соответствии с которой любая непрерывная функция может быть аппроксимирована с любой наперед заданной точностью $\varepsilon > 0$ функциями одной переменной следующим образом:

$$f(x_1, \dots, x_n) \approx \sum_{q=1}^M d_q \sigma \left(\sum_{p=1}^N a_{qp} \sigma (b_{qp} x_{i(qp)} + c_{qp}) + u_q \right),$$

однако N и M растут экспоненциально при $\varepsilon \rightarrow 0$.

Позже были получены интересные результаты [3], из которых следует, что любая непрерывная функция может быть аппроксимирована с любой наперед заданной точностью $\varepsilon > 0$ следующим образом:

$$f(x_1, \dots, x_n) \approx \sum_{q=1}^{4n+3} d_q \sigma \left(\sum_{p=1}^{2n+1} c_{qp} \sigma (\mathbf{w}_{pq} \cdot \mathbf{x} + \theta_{pq}) \right),$$

где σ — строго монотонная сигмоидальная бесконечно непрерывно дифференцируемая функция. Структура нейронной сети фиксирована для любой непрерывной функции f . Однако σ — катастрофически сложно устроенная функция и поэтому трудно вычисляемая (т. е. сложность вычислений экспоненциально быстро растет по мере увеличения точности вычисления значения функции). Можно было бы сделать заключение, что этот результат ставит под сомнение *практическую* ценность искусственных нейронных сетей

с *ограниченной структурной сложностью* (т. е. когда число нейронов в каждом слое нейронной сети зависит только от n и не зависит от ε) как универсального инструмента для аппроксимации зависимостей. Однако нужно принять во внимание тот факт, что нейронная сеть построена на основе предельно упрощенной модели нейрона, которая не отражает многих особенностей преобразования сигналов в природных нейронах.

Для того чтобы преодолеть или, по крайней мере, ослабить проблему экспоненциального роста структурной сложности нейронных сетей, необходимо использовать более адекватные модели искусственных нейронов, которые лучше отражают особенности преобразования информации в природных нейронах и имеют лучшие способности по аппроксимации зависимостей (например, модель $\Sigma\Pi$ -нейрона, радиального нейрона и др.).

В настоящей работе предлагаются функциональные модели нейронов, которые основаны на применении *агрегирующих функций* [4, 5] для вычисления суммарного потенциала нейрона, вклада сложных синапсов и синаптических кластеров. Эти модели обобщают классические модели искусственных нейронов (в том числе $\Sigma\Pi$ -нейрона), которые лежат в основе моделирования искусственных нейронных сетей. Диапазон классических моделей нейрона весьма ограничен, что накладывает ограничения на виды нейросетевых моделей, которые можно реально использовать для восстановления неизвестных зависимостей. Однако на основе аппарата *теории агрегирующих функций* можно построить широкий спектр моделей нейронов, которые, с одной стороны, существенно расширяют «номенклатуру» способов преобразования информации в искусственных нейронных сетях, а с другой стороны, лучше отражают процессы преобразования информации в природных нейронах. Кроме того, в рамках многослойной архитектуры применение аппарата теории агрегирующих функций также может позволить строить многослойные сети, содержащие модельно однородные алгоритмы в пределах одного слоя, но при этом модели алгоритмов в разных слоях могут быть принципиально отличными. Типичные примеры — это модель $\Sigma\Pi$ -нейрона, модель сети со скрытым слоем из радиальных нейронов.

При обучении нейронных сетей на основе классических моделей нейронов часто приходится сталкиваться с проблемой переобучения. Одной из причин переобучения является то, что используемая модель преобразования информации является неадекватной реальному методу преобразования информации. Для ее преодоления успешно применяются композиции модулей из отдельных нейронных сетей, так же как и в методах типа *boosting* или *bagging*.

Продвигаемая в настоящей работе идея состоит в том, что агрегирующие нейроны могут использоваться в качестве базовых алгоритмов в таких композициях вместо нейронных сетей, т. е. теперь композиция может строиться не из нейронных сетей, а из агрегирующих нейронов. Это возможно из-за того, что агрегирующие нейроны могут обладать способностями по аппроксимации, которые не уступают перцептронным нейронным сетям. Так, для одного класса агрегирующих нейронов, которые обобщают $\Sigma\Pi$ -нейроны, показывается, что класс таких агрегирующих нейронов является корректным в том смысле, что для любой конечной обучающей выборки можно построить агрегирующий нейрон, который будет давать правильные ответы на входах из своей обучающей выборки. Значимость корректности модели алгоритма продемонстрирована с помощью комбинаторного подхода к оценке обобщающей способности алгоритмов [6].

Сначала опишем общую схему функциональной модели нейрона, учитывающей сложные процессы обработки информации в реальных нейронах.

1 Схема функциональной модели нейрона

Пусть сегмент $\mathbf{Y} \subset \mathbb{R}$ (например, $[0, 1]$ или $[-1, 1]$) — множество значений, передаваемых между нейронами; сегмент $\mathbf{U} \subseteq \mathbb{R}$ — множество значений полного или суммарного потенциала нейрона; $\text{syn}: \mathbf{Y} \rightarrow \mathbf{U}$ — функция преобразования *простого синапса*. Классическая модель простого синапса линейная, т. е. $\text{syn}(x, w) = wx$. При построении нелинейных моделей простых синапсов можно исходить из естественного предположения, что функция преобразования простого синапса ограниченная и монотонная. Например, если $\mathbf{X} = [0, 1]$, то $\text{syn}(x, w) = \varphi(x - a, w)$, где $\varphi(s, w)$ — монотонно неубывающая по s , строго монотонно возрастающая по w и $\varphi(s, w) = 0 \Leftrightarrow s \leq 0$.

Сложный синапс рассматривается как композиция простых синапсов. Вклад сложного синапса получается в результате *нелинейного агрегирования* вкладов от простых синапсов:

$$u_k = \text{Agg}_S\{u_i : i \in \mathbf{i}_k\},$$

где u_k — вклад k -го сложного синапса ($k = n + 1, \dots, N$); u_i — вклад i -го простого синапса ($i < n$); $\mathbf{i}_k \subseteq \{1, \dots, n\}$ — набор индексов простых синапсов, входящих в k -й сложный синапс; Agg_S — агрегирующая функция для вычисления вкладов сложных синапсов.

Синаптический кластер рассматривается как композиция синапсов (простых или сложных):

$$s_l = \text{Agg}_C\{u_j : j \in \mathbf{j}_l\},$$

s_l — вклад l -го синаптического кластера ($l = 1, \dots, m$); $\mathbf{j}_l \subset \{1, \dots, n, \dots, N\}$ — набор индексов синапсов (сложных и простых), входящих в l -й синаптический кластер; Agg_C — агрегирующая функция для вычисления вкладов синаптических кластеров, которая может отличаться от Agg_S , однако если кластеры состоят из простых синапсов, то агрегирующие функции Agg_C и Agg_S , как правило, совпадают.

Суммарный потенциал нейрона получается в результате *агрегирования* вкладов от синаптических кластеров:

$$s = \text{Agg}\{s_1, \dots, s_m\},$$

где s_1, \dots, s_m — вклады синаптических кластеров; Agg — агрегирующая функция для вычисления суммарного потенциала нейрона, которая отличается от Agg_S , но может совпадать с Agg_C . Это хорошо иллюстрирует модель $\Sigma\Pi$ -нейрона, которая рассматривается ниже. Функция выхода $\text{out}: \mathbf{U} \rightarrow \mathbf{Y}$ задает закон преобразования суммарного потенциала нейрона в его выход.

2 Модели нейрона с простыми синапсами

Классический нейрон имеет только простые синапсы и реализует преобразование

$$y = \text{out}\left(\theta + \sum_{i=1}^n w_i x_i\right),$$

где w_i — вес i -го синапса. В этой модели слагаемое $w_i x_i$ соответствует преобразованию в i -м синапсе, суммарный потенциал нейрона есть арифметическая сумма вкладов простых синапсов и смещения θ .

Эта модель является частным случаем более общей модели нейрона с простыми синапсами:

$$y = \text{out}\left(\theta + \sum_{i=1}^n u_i\right), \quad (1)$$

где $u_i = \text{syn}(x, w, \dots)$ — нелинейная скалярная функция преобразования простого синапса, зависящая от *синаптического веса* w и, возможно, других параметров.

Естественным обобщением арифметического суммирования, которое сохраняет все его привычные алгебраические свойства, является так называемое g -суммирование [7, 8]:

$$\left(\theta + \sum_{i=1}^n u_i\right)_g = g^{-1}\left(g(\theta) + \sum_{i=1}^n g(u_i)\right),$$

где g — монотонная обратимая функция. Оно является точным аналогом взвешенного арифметического суммирования. Теорема Акцеля [9] объясняет, почему выделяется g -суммирование: всякая непрерывная и ассоциативная агрегирующая функция на сегменте представляет собой некоторое g -суммирование. Например:

| | |
|--|--|
| $g(u)$ | $\left(\theta + \sum_{i=1}^n u_i\right)_g$ |
| $u^{(p)} = \text{sign } u u ^p$ | $\left(\theta^{(p)} + \sum_{i=1}^n u_i^{(p)}\right)^{\langle 1/p \rangle}$ |
| e^u | $\ln\left(e^\theta + \sum_{i=1}^n e^{u_i}\right)$ |
| $\begin{cases} \ln u, & u > 0; \\ 0, & u = 0; \\ \ln(-u) + i\pi, & u < 0. \end{cases}$ | $\theta \prod_{i=1}^n u_i$ |

Модель классического нейрона, в котором вместо обычного суммирования используется g -суммирование, принимает следующий вид:

$$y = \widetilde{\text{out}}\left(g(\theta) + \sum_{i=1}^n g(u_i)\right),$$

где $\widetilde{\text{out}}(s) = \text{out}(g^{-1}(s))$; $u_i = \text{syn}(x_i, w_i)$.

Классическую модель нейрона с простыми синапсами можно обобщить до модели нейрона агрегирующего типа с простыми синапсами вида:

$$y = \text{out} \circ \text{Agg}\{\theta, u_1, \dots, u_n\},$$

где Agg — симметричная агрегирующая функция.

Классическая модель нейрона с простыми синапсами по форме идентична обобщенной линейной модели, которая имеет статистическую природу [10]. В ней функция связи соответствует функции выхода и отражает связь между математическим ожиданием целевой переменной и взвешенной линейной комбинацией предикторов. В линейной модели с нелинейным преобразованием признаков [11], наоборот, учитывается функция распределения значений входных предикторов. Объединение этих двух моделей дает обобщенную аддитивную статистическую модель, которая похожа на модель нейрона с простыми нелинейными синапсами (1).

3 Модель $\Sigma\Pi$ -нейрона

Примером нейрона со сложными синапсами или синаптическими кластерами является $\Sigma\Pi$ -нейрон. Эта модель нейрона впервые была предложена в [12] и использовалась как модель, которая отражает локальное взаимодействие входов в сложных синапсах

или простых синапсов в синаптических кластерах [13]. В работах [14, 15] было показано, что модель $\Sigma\Pi$ -нейрона более адекватно соответствует процессам обработки информации с учетом того, что вклады сложных синапсов суммируются арифметически. Достоинством этой модели является ее относительная простота и выразительность. Например, в модели $\Sigma\Pi$ -нейрона с функцией выхода, принимающей значения в $\{0, 1\}$, можно представить любую логическую функцию. Сплайны для аппроксимации непрерывных функций многих переменных можно рассматривать как $\Sigma\Pi$ -нейрон, на вход которого подаются сигналы, предварительно преобразованные при помощи скалярных усеченно-степенных функций. В работе [16] показано, что адекватная функциональная модель обучения нейронов коры головного мозга получается в случае, когда на вход $\Sigma\Pi$ -нейрона подаются сигналы, предварительно преобразованные при помощи радиальных функций. Более подробный анализ моделей со сложными синапсами и синаптическими кластерами можно найти в [17].

$\Sigma\Pi$ -нейрон моделирует сложные синапсы или синаптические кластеры при помощи агрегирующей функции произведения:

$$S(u_1, \dots, u_m) = \prod_{i=1}^m u_i,$$

где u_1, \dots, u_m — вклады простых синапсов, входящих в сложный синапс или синаптический кластер. Суммарный потенциал по-прежнему есть арифметическая сумма вкладов синапсов и смещения θ :

$$y = \text{out} \left(\theta + \sum_{k=1}^m \prod_{i \in \mathbf{i}_k} u_{ki} \right),$$

где $\mathbf{i}_k \subseteq \{1, \dots, n\}$; u_{ki} — вклад i -го простого синапса в k -м сложном синапсе или синаптическом кластере. Если простые синапсы линейные ($u = wx$), то функция преобразования имеет вид:

$$y = \text{out} \left(\theta + \sum_{k=1}^m w_k \prod_{i \in \mathbf{i}_k} x_i \right), \quad (2)$$

где $w_k = \prod_{i \in \mathbf{i}_k} w_{ki}$. Такая модель $\Sigma\Pi$ -нейрона является корректной для класса логических функций, т. е. для любой частично определенной логической функции $f: \mathbf{X} \rightarrow \{0, 1\}$, где $\mathbf{X} \subseteq \{0, 1\}^n$, существует $\Sigma\Pi$ -нейрон вида (2), который вычисляет значения функции f на \mathbf{X} .

Модель $\Sigma\Pi$ -нейрона с простыми синапсами вида $u = w\varphi(x - a)$ имеет вид:

$$y = \text{out} \left(\theta + \sum_{k=1}^m w_k \prod_{i \in \mathbf{i}_k} g_i(x_i - a_{ki}) \right),$$

где $g_i(x) = 0 \Leftrightarrow x \leq 0$. В [18] доказано, что эта модель при некоторых условиях общего вида способна корректно представлять произвольные дискретные функции, определенные на конечных подмножествах \mathbb{R}^n , при этом конструктивный процесс обучения сопровождается процессом минимизации сложности так, чтобы в синаптическом кластере участвовало минимально возможное число простых синапсов. Важным свойством конструктивной процедуры обучения $\Sigma\Pi$ -нейрона является то, что она позволяет строить множество различных $\Sigma\Pi$ -нейрона по одной обучающей выборке, которые выдают на ней правильные ответы. Это дает возможность уверенно использовать их в качестве базовых алгоритмов в процедурах типа бустинга, баггинга или взвешенного голосования, так как всегда можно

построить $\Sigma\Pi$ -нейрон, который не ошибается *более чем на половине* выборки, используя при этом в качестве обучающей *не более половины* выборки.

4 Модель нейрона с агрегирующими сложными синапсами

Модель $\Sigma\Pi$ -нейрона является частным случаем следующей модели нейрона со сложными синапсами:

$$y = \text{out} \left(\theta + \sum_{k=1}^m u_k \right).$$

Здесь u_k — вклад k -го агрегирующего сложного синапса или синаптического кластера, состоящего из простых синапсов:

$$u_k = \text{Agg}_{\Pi} \{u_i : i \in \mathbf{i}_k\},$$

где Agg_{Π} — агрегирующая функция, которая удовлетворяет требованию:

$$\text{Agg}_{\Pi} \{u_i : i \in \mathbf{i}\} = 0 \Leftrightarrow \exists i \in \mathbf{i} : u_i = 0.$$

Значительный класс таких агрегирующих функций — это *квазимультипликативные* функции:

$$\Pi_h \{u_i : i \in \mathbf{i}\} = h^{-1} \left(\prod_{i \in \mathbf{i}} h(u_i) \right),$$

где h — обратимая монотонная функция.

Таким образом, модель нейрона с агрегирующими ложными синапсами или синаптическими кластерами из простых синапсов имеет вид:

$$y = \text{out} \left(\theta + \sum_{k=1}^m \text{Agg}_{\Pi} \{u_{ki} : i \in \mathbf{i}_k\} \right). \quad (3)$$

Способности такой модели отражает следующее утверждение. Пусть out — *корректная функция выхода*, т. е. для любого $y \in \mathbf{Y}$ найдется значение s такое, что $y = \text{out}(s)$.

Пусть функция простого синапса $\text{syn}(x - a, w)$, такая что

- 1) $\text{syn}(x - a, w) = 0$ только при $x \leq a$;
- 2) для любого w функция $\text{syn}_w(x) = \text{syn}(x - a, w)$ — монотонная;
- 3) для любого x функция $\text{syn}_x(w) = \text{syn}(x - a, w)$ — строго монотонная;
- 4) для любого x и любого u существует w такая, что $u = \text{syn}_x(w)$.

Типичный пример: $\text{syn}(x - a, w) = w\varphi(x - a)$, где φ — монотонно неубывающая функция и $\varphi(x) = 0 \Leftrightarrow x \leq 0$ или $\varphi(x) = 0 \Leftrightarrow x < 0$.

Теорема 1. *Для любого конечного и непротиворечивого набора прецедентов¹ $\langle \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} \rangle$ можно построить агрегирующий нейрон*

$$\text{agn}(\mathbf{x}) = \text{out} \left(\theta + \sum_{k=1}^m \text{Agg}_{\Pi} \{u_{ki} : i \in \mathbf{i}_k\} \right),$$

такой, что $y = \text{agn}(\mathbf{x})$ для любой пары $\langle \mathbf{x}, y \rangle$ из $\langle \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} \rangle$.

Для краткости приведем только схему конструктивного доказательства, которое следует схемам доказательств аналогичных утверждений в [18] для $\Sigma\Pi$ -нейронов. Набор прецедентов $\langle \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} \rangle$ переупорядочивается так, что для любой пары $1 \leq j < k \leq N$ всегда

¹Набор прецедентов $\langle \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} \rangle$ является непротиворечивым, когда для любой пары $\langle \mathbf{x}', y' \rangle$ и $\langle \mathbf{x}'', y'' \rangle$ из $\langle \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} \rangle$ верно условие: если $\mathbf{x}' = \mathbf{x}''$, то $y' = y''$.

найдется индекс $i = i(j, k)$, такой, что $x_{ji} \leq x_{ki}$ (\neq -упорядоченность). При помощи процедуры исключения индексов, несущественных для порядка примеров в $\tilde{\mathbf{X}}$, строится последовательность $\{\mathbf{i}_k : k = \overline{1, N}\}$ ($N = |\tilde{\mathbf{X}}|$), которая содержит минимальные наборы индексов (т. е. из любого набора \mathbf{i}_k нельзя исключить какой-либо индекс так, чтобы не нарушилось свойство \neq -упорядоченности). В силу свойств агрегирующей функции Agg_{Π} последовательность функций $\{\Phi_k(\mathbf{x})\}$, где

$$\Phi_k(\mathbf{x}) = \text{Agg}_{\Pi} \{ \text{syn}(x_i - \nu(x_{ki}), w_{ki}) : i \in \mathbf{i}_k \},$$

является треугольно упорядоченной, а именно:

- (1) $\Phi_j(\mathbf{x}_k) = 0$ для любой пары $j < k$;
- (2) $\Phi_k(\mathbf{x}_k) \neq 0$.

Здесь $\nu(x_{ki})$ — предыдущее значение относительно x_{ki} в вариационном ряду значений x_{1i}, \dots, x_{Ni} .

Затем при помощи однопроходной процедуры находятся веса w_{ki} , так что по построению $y_k = \text{agn}(\mathbf{x}_k)$ для любого $k = \overline{1, N}$. Эта процедура рекуррентная, так что в процессе обучения строится последовательность $\{\text{agn}_k(\mathbf{x}) = \text{out}(\text{SP}_k(\mathbf{x}))\}$ такая, что для любого $j = \overline{1, k}$: $y_j = \text{agn}_k(\mathbf{x}_j)$, при этом

$$\text{SP}_k(\mathbf{x}) = \begin{cases} \text{SP}_{k-1}(\mathbf{x}), & \text{если } y_k = \text{agn}_{k-1}(\mathbf{x}_{k-1}); \\ \text{SP}_{k-1}(\mathbf{x}) + w_k \Phi_k(\mathbf{x}_k), & \text{если } y_k \neq \text{agn}_{k-1}(\mathbf{x}_{k-1}), \end{cases}$$

где

$$\text{SP}_k(\mathbf{x}) = \theta + \sum_{j=1}^k \text{Agg}_{\Pi} \{ u_{ji} : i \in \mathbf{i}_j \}, \quad u_{ji} = \text{syn}(x_i - a_{ji}, w_{ji});$$

w_k — решение уравнения

$$y_k = \text{out}(\text{SP}_{k-1}(\mathbf{x}_k) + w_k \Phi_k(\mathbf{x}_k)).$$

Таким образом, получается корректный агрегирующий нейрон по построению.

Помимо квазимультимпликативных функций примером агрегирующих функций сложных синапсов или синаптических кластеров из простых синапсов также могут служить квазиаддитивные функции

$$u_k = \Sigma_h \{ u_i : i \in \mathbf{i}_k \} = h^{-1} \left(\sum_{i \in \mathbf{i}_k} h(u_{ki}) \right),$$

где h — обратимая монотонная функция, например:

| $h(u)$ | Суммарный потенциал нейрона |
|-----------|---|
| $u^{(p)}$ | $\theta + \sum_{k=1}^m \left(\sum_{i \in \mathbf{i}_k} u_{ki}^{(p)} \right)^{\langle 1/p \rangle}$ |
| e^{pu} | $\theta + \sum_{k=1}^m \frac{1}{p} \ln \left(\sum_{i \in \mathbf{i}_k} e^{pu_{ki}} \right)$ |

5 Двуслойная модель нейрона со сложными синапсами

В общем случае модель (3) агрегирующего нейрона можно обобщить, если заменить обычное суммирование вкладов сложных синапсов на симметричную агрегирующую функцию :

$$\begin{aligned} y &= \text{out} \circ \text{Agg}_{\Sigma} \{ \theta, s_1, \dots, s_m \}; \\ s_k &= \text{Agg}_{\Pi} \{ u_{ki} : i \in \dot{i}_k \}, \quad k = \overline{1, m}, \end{aligned} \quad (4)$$

где Agg_{Σ} — агрегирующая функция для вычисления суммарного потенциала нейрона, такая что

- (1) если $s_{k'} = 0$, то $\text{Agg}_{\Sigma} \{ s_k : k \in \mathbf{k} \} = \text{Agg}_{\Sigma} \{ s_k : k \in \mathbf{k}' \}$, где $\mathbf{k}' = \mathbf{k} \setminus \{ k' \}$;
- (2) для любого $\text{Agg}_{\Sigma} \{ s_1, \dots, s_M \}$, любого S найдется s_{M+1} такая, что

$$\text{Agg}_{\Sigma} \{ s_1, \dots, s_M, s_{M+1} \} = S.$$

Примерами такой агрегирующей функции являются квазиаддитивные функции вида:

$$\text{Agg}_{\Sigma} \{ s_1, \dots, s_M \} = h^{-1} \left(\sum_{k=1}^M w_k h(s_k) \right).$$

В этих предположениях и предположениях Теоремы 1 в части простых синапсов имеет место следующая теорема.

Теорема 2. *Для любого конечного и непротиворечивого набора прецедентов $\langle \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} \rangle$ можно построить агрегирующий нейрон $\text{agn}(\mathbf{x})$ вида (4) такой, что $y = \text{agn}(\mathbf{x})$ для любой пары $\langle \mathbf{x}, y \rangle$ из $\langle \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} \rangle$.*

Доказательство этой теоремы идентично доказательству Теоремы 1. Особенность заключается только в том, что агрегирующая операция Agg_{Σ} , по приведенному выше определению, обладает всеми необходимыми свойствами, которыми обладает операция обычного суммирования и которые непосредственно и используются при доказательстве Теоремы 1 и поэтому формальная ее замена на Agg_{Σ} обеспечивает истинность Теоремы 2.

6 Заключение

В настоящей работе описаны новые классы моделей нейронов агрегирующего типа, которые включают в себя большинство функциональных моделей нейронов, в том числе со сложными синапсами (нейроны перцептронного типа, сигма-пи нейроны, радиальные нейроны). Они существенно расширяет «номенклатуру» типов моделей нейронов, на основе которых можно строить нейронные сети. Для одного подкласса таких моделей, которые обобщают $\Sigma\Pi$ -нейроны, доказана способность по корректному представлению функций, определенных на конечном множестве.

В дальнейшем предстоит теоретическое и экспериментальное исследование возможностей моделей агрегирующих нейронов по аппроксимации различных типов зависимостей. Поскольку модели $\Sigma\Pi$ -нейронов и их обобщения показали хорошие способности по представлению многих типов зависимостей [18–20], то таких же способностей можно ожидать и от моделей агрегирующих нейронов в общем случае.

Также предстоит еще исследовать эффект от применения различных классов агрегирующих функций как для агрегирования вкладов в сложных синапсах и синаптических кластерах, так и для агрегирования всех вкладов в суммарный потенциал нейрона на качество и сложность аппроксимации зависимостей.

Литература

- [1] *Kurkova V.* Kolmogorov's theorem is relevant // *Neural Comput.*, 1991. Vol. 3. P. 617–622.
- [2] *Kurkova V.* Kolmogorov's theorem and multilayer neural networks // *Neural Networks*, 1992. Vol. 5. P. 501–506.
- [3] *Maiorov V., Pinkus A.* Lower bounds for approximation by MLP neural networks // *Neurocomputing*, 1999. Vol. 25. P. 81–91.
- [4] *Mesiar R., Komornikova M., Kolesarova A., Calvo T.* Aggregation functions: A revision // *Fuzzy sets and their extensions: Representation, aggregation and models* / Eds. H. Bustince, F. Herrera, J. Montero. — Berlin–Heidelberg: Springer, 2008. С. 121–144.
- [5] *Grabich M., Marichal J.-L., Pap E.* Aggregation functions. — *Encyclopedia of mathematics and its applications ser.* — Cambridge University Press, 2009. No. 127. 482 p.
- [6] *Воронцов К. В.* Комбинаторный подход к оценке качества обучаемых алгоритмов // *Математические вопросы кибернетики* / Под ред. О. Б. Лупанова. — М.: Физматлит, 2004. Т. 13. С. 5–36.
- [7] *Pap E.* *g*-calculus. *Univ. u Novom Sadu, Zb. Rad. Prirod.-Mat. Fak., Ser. Mat.*, 1993. Vol. 23. No. 1. P. 145–150.
- [8] *Pap E.* Applications of the generated pseudo-analysis on nonlinear partial differential equations. Vienna, 2004. Preprint ESI 1448. 22 p.
- [9] *Aczél J.* Lectures on functional equations and their applications. — New York, NY, USA: Academic Press, 1966. P. 253–273.
- [10] *McCullagh P., Nelder J.* Generalized linear models. — 2nd ed. — Boca Raton: Chapman and Hall/CRC, 1989. 532 p.
- [11] *Hasti T., Tibshirani R.* Generalized additive models // *Stat. Sci.*, 1986. Vol. 1. No. 3. P. 297–318.
- [12] *Feldman J. A., Ballard D. H.* Connectionist models and their properties // *Cognitive Sci.*, 1982. No. 6. P. 205–254.
- [13] *Rumelhart D. E., Hinton G., Williams R.* Learning internal representation by error propagation // *Parallel distributed processing. Vol. 1: Foundations.* — Cambridge, MA, USA: MIT Press, 1986. P. 318–362.
- [14] *Mel B. W.* The sigma-pi column: A model of associative learning in cerebral neocortex. Pasadena, CA, USA: California Institute of Technology, 1990. Cns Memo No. 6: Technical Report.
- [15] *Mel B. W.* The sigma-pi model neuron: Roles of the dendritic tree in associative learning // *Soc. NeuroSci. Abstr.*, 1990. Vol. 16. P. 205.4.
- [16] *Mel B. W., Koch C.* Sigma-pi learning: on radial basis functions and cortical associative learning // *Advances in neural information processing systems* / Ed. D. S. Touretzky. — San Mateo, CA, USA: Morgan Kaufmann, 2000. Vol. 2. P. 474–481.
- [17] *Mel B. W.* Why have dendrites? A computational perspective // *Dendrites* / Eds. G. Stuart, N. Spruston, M. Hausser. — 2nd ed. — Oxford University Press, 2007. P. 271–289.
- [18] *Шибзухов З. М.* Конструктивные методы обучения СП-нейронных сетей. — М: МАИК Наука, 2006. 159 с.
- [19] *Шибзухов З. М.* Рекуррентный метод конструктивного обучения некоторых сетей алгебраических СП-нейронов и СП-нейромодулей // *Журнал вычислительной математики и математической физики*, 2003. № 43. С. 1298–1310.
- [20] *Шибзухов З. М.* О некоторых конструктивных и корректных классах алгебраических СП-алгоритмов // *Докл. РАН*, 2010. Т. 432. С. 465–468.

Поступила в редакцию 29.05.2014

References

- [1] Kurkova, V. 1991. Kolmogorov's theorem is relevant. *Neural Comput.* 3:617–622.
- [2] Kurkova, V. 1992. Kolmogorov's theorem and multilayer neural networks. *Neural Networks* 5:501–506.
- [3] Maiorov, V., and A. Pinkus. 1999. Lower bounds for approximation by MLP neural networks. *Neurocomputing* 25:81–91.
- [4] Mesiar, R., M. Komornikova, A. Kolesarova, and T. Calvo. 2008. Aggregation functions: A revision. *Fuzzy sets and their extensions: Representation, aggregation and models*. Eds. Bustince, H., F. Herrera, and J. Montero. Berlin–Heidelberg: Springer. 121–144.
- [5] Grabich, M., J.-L. Marichal, and E. Pap. 2009. *Aggregation functions*. Encyclopedia of mathematics and its applications ser. Cambridge University Press. No. 127. P. 482 p.
- [6] Vorontsov, K. V. 2004. Kombinatornyy podkhod k otsenke kachestva obuchaemykh algoritmov. *Matematicheskie voprosy kibernetiki*. Ed. O. B. Lupanov. Moscow: Fizmatlit. 13:5–36.
- [7] Pap, E. 1993. g -calculus. *Univ. u Novom Sadu, Zb. Rad. Prirod.-Mat. Fak., Ser. Mat.* 23(1):145–150.
- [8] Pap, E. 2004. Application of the generated pseudo-analysis on nonlinear partial differential equations. Vienna. Preprint ESI 1448. 22 p.
- [9] Aczél, J. 1966. Lectures on functional equations and their applications. New York, NY: Academic Press. 253–273.
- [10] McCullagh, P., and J. Nelder. 1989. *Generalized linear models*. 2nd ed. Boca Raton: Chapman and Hall/CRC. 532 p.
- [11] Hasti, T., and R. Tibshirani. 1986. Generalized additive models. *Stat. Sci.* 1(3):297–318.
- [12] Feldman, J. A., and D. H. Ballard. 1982. Connectionist models and their properties. *Cognitive Sci.* 6:205–254.
- [13] Rumelhart, D. E., G. Hinton, and R. Williams. 1986. Learning internal representation by error propagation. *Parallel distributed processing. Vol. 1: Foundations*. Cambridge, MA: MIT Press. 318–362.
- [14] Mel, B. W. 1990. The sigma-pi column: A model of associative learning in cerebral neocortex. Pasadena, CA: California Institute of Technology. Cns Memo No. 6: Technical Report.
- [15] Mel, B. W. 1990. The sigma-pi model neuron: Roles of the dendritic tree in associative learning. *Soc. NeuroSci. Abstr.* 16:205.4.
- [16] Mel, B. W., and C. Koch. 2000. Sigma-pi learning: On radial basis functions and cortical associative learning. *Advances in neural information processing systems*. Ed. D. S. Touretzky. San Mateo, CA: Morgan Kaufmann. 2:474–481.
- [17] Mel, B. W. 2007. Why have dendrites? A computational perspective. *Dendrites*. 2nd ed. Eds. G. Stuart, N. Spruston, and M. Hausser. Oxford University Press. 271–289.
- [18] Shibzuхов, Z. M. 2006. Konstruktivnie metody obucheniya $\Sigma\Pi$ -neyronnykh setey. Moscow: MAIK Nauka. 159 p.
- [19] Shibzuхов, Z. M. 2003. Rekurrentnyy metod konstruktivnogo obucheniya nekotorykh setey algebraicheskikh $\Sigma\Pi$ -neyronov i $\Sigma\Pi$ -neyromoduley. *Zhurnal Vychislitel'noy Matematiki i Matematicheskoy Fiziki* 43:1298–1310.
- [20] Shibzuхов, Z. M. 2010. O nekotorykh konstruktivnykh i korrektnykh klassakh algebraicheskikh $\Sigma\Pi$ -algoritmov. *Dokl. RAN* 432:465–468.

Received May 29, 2014

Метод повышения разрешения космических изображений с использованием априорной информации в векторной форме для сохранения границ*

А. Н. Трекин^{1,2}, И. А. Матвеев³, А. Б. Мурьнин^{1,3}, В. Г. Бочкарева^{1,2}
alexey.trekin@gmail.com

¹Научно-исследовательский институт аэрокосмического мониторинга «АЭРОКОСМОС»

²Московский физико-технический институт (Государственный университет)

³Вычислительный центр им. А. А. Дородницына Российской академии наук

Разработан метод повышения разрешения изображений, использующий векторную информацию о геометрических свойствах объектов, содержащихся на изображениях. Априорная информация о контрастных границах разнородных областей позволяет сохранить резкий переход яркости при повышении разрешения. Произведена апробация метода на наборе космических изображений с использованием карты водоемов в векторной форме.

Ключевые слова: дистанционное зондирование Земли; повышение разрешения; обработка изображений

DOI: 10.21469/22233792.1.12.07

A method for upsampling of remote sensing images using vector data for preserving edges*

A. N. Trekin^{1,2}, I. A. Matveev³, A. B. Murynin^{1,3}, and V. G. Bochkareva^{1,2}

¹Institute for Scientific Research of Aerospace Monitoring “AEROCOSMOS”, 4 Gorokhovskii per., Moscow, Russia

²Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Russia

³Dorodnicyn Computing Centre of the Russian Academy of Sciences, 40 Vavilova st., Moscow, Russia

A method for image upsampling was developed. The method makes use of vector data about geometry of objects contained in image. *A priori* information about high-contrast boundaries helps to preserve sharp illuminance change from blurring due to upsampling procedure. The developed method was tested over a set of remote sensing images and vector map of water bodies.

Keywords: remote sensing; superresolution; image processing

DOI: 10.21469/22233792.1.12.07

1 Введение

Пространственное разрешение, т.е. частота дискретизации, с которой непрерывный сигнал переводится в дискретное изображение, — одно из ограничений, влияющих на информативность данных дистанционного зондирования [1, 2]. Оно играет решающую роль

*Работа выполнена при финансовой поддержке РФФИ (проекты №№ 14-01-00348, 13-05-12019 и 14-05-91759). Результаты получены в рамках государственного задания Минобрнауки России № 2015/Р8.

при поиске объектов, обработке изображений спектральными методами [3, 4]. Моделирование полей электромагнитных сигналов, поступающих на вход аппаратуры дистанционного зондирования Земли (ДЗЗ), позволяет определить ее возможности и задать оптимальные характеристики [5], однако на практике возможности создания систем ДЗЗ с высоким пространственным разрешением ограничиваются множеством факторов, что приводит к необходимости увеличивать частоту дискретизации имеющихся космических изображений. Существует множество методов повышения разрешения изображений, многие из которых применяются в ДЗЗ. Эти методы можно подразделить по числу используемых изображений на однокадровые (в том числе линейные [6], нелинейные градиентные [7] и фрактальные [8]) и многокадровые (сверхразрешение [9], паншарпенинг). По способу восполнения недостающей информации при формировании изображений высокого разрешения можно разделить методы на использующие интерполяцию [6, 10], повышающие разрешающую способность мультиспектральных изображений путем синтеза информации в спектральных каналах и основанные на использовании априорной информации об объектах. Заслуживает внимания класс методов, основанных на минимизации потоков [11–13].

В любом случае при повышении разрешения приходится иметь дело с недостатком информации, и это сказывается на качестве восстановленного изображения. Как правило, при этом проявляются следующие дефекты: эффект размытия, алиасинг (ступенчатость контуров) и эффект Гиббса (ложное оконтуривание) [14].

В задачах мониторинга стационарных объектов по их изображениям, в частности при космическом мониторинге участков поверхности Земли, возникает необходимость четкого прослеживания границ этих объектов с целью точного определения их расположения, измерения геометрических размеров и площадей. Такая задача возникает при мониторинге урбанизированных территорий, береговых линий, ледовой обстановки, природных пожаров и территорий, пройденных огнем. При использовании стандартных методов повышения разрешающей способности, не учитывающих информацию о границах объектов, происходит сглаживание, в том числе там, где не должно быть плавного перехода яркости, а существует резкая граница.

Существуют различные методы избежать размытия границ, включающие предварительное выделение перепадов яркости [15, 16], использование нелинейной диффузии [17], агентный подход [18]. В таких методах непосредственно перед интерполяцией проводится предварительная обработка, выделяющая границы. Данный подход имеет немало преимуществ: он не требует дополнительной информации и хорошо сохраняет границы. К его недостаткам относится значительная вычислительная сложность, связанная с необходимостью предварительной обработки изображения, и вероятность ошибки при выделении границ: пропуска реальной границы или выделения ее там, где в действительности происходит плавное, но быстрое изменение яркости.

Для объектов, не изменяющих свои геометрические свойства длительное время, часто существует их векторное представление, полученное тем или иным способом. В данной статье предложен метод, использующий для повышения четкости границ при повышении разрешения априорную информацию, т. е. уже имеющиеся векторные карты, содержащие границы объектов.

2 Постановка задачи

Основы данного метода аналогичны тем, что заданы в работе [19]. Обычное пиксельное изображение не позволяет сделать вывод, где проходят резкие границы между объектами, а где — плавный переход яркости, что приводит к размытию границ при попытке повысить

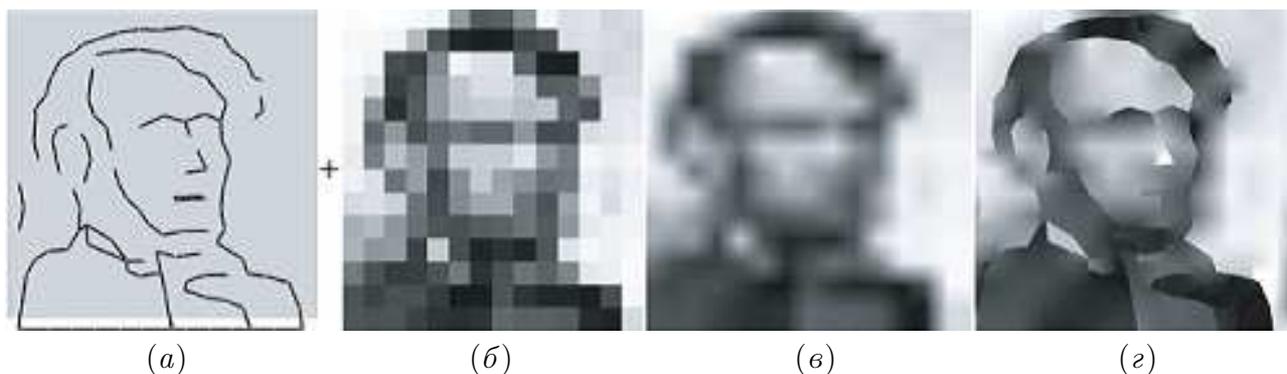


Рис. 1 Границы биксельного изображения (а); пиксельная матрица низкого разрешения (б); пиксельное изображение повышенного разрешения (билинейная интерполяция) (в) и биксельное изображение повышенного разрешения (з) [19]

разрешение изображения (рис. 1). В данной статье предлагается хранение изображения не в виде двумерной матрицы пикселей, а в виде так называемых бикселей (bixels). Такое изображение состоит из обычного растра, задающего значения яркости в вершинах равномерной сетки, и описания границ, проходящих вблизи каждой вершины, с субпиксельной точностью. Границы могут быть представлены в виде ломаных или кривых и обозначают линии, по которым проходит резкий перепад яркости, обычно — границы объектов, разделение объекта и фона. Таким образом, появляется возможность сохранить резкие перепады яркости, соответствующие границам, при повышении разрешения в любое количество раз, оставив в прочих местах плавные переходы, соответствующие интерполяционным алгоритмам. Для отображения таких изображений на экране или в печати используется пиксельное представление, поэтому существует алгоритм перевода изображения из бикселей в чисто растровый вид. Так как при этом могут получаться абсолютно резкие границы, для лучшего восприятия производится сглаживание (antialiasing). В зависимости от конфигурации близлежащих границ интерполяция яркости в выбранных координатах производится по одной из десяти функций билинейной интерполяции по точкам базового изображения. При этом для интерполяции, с помощью которой получают пиксели результирующего растрового изображения, используются только те значения яркостей, что не отделены от результирующего пикселя линией границы.

Таким образом, на параметры границы должны быть наложены некоторые ограничения, позволяющие восстановить яркость в любой точке. Для этого область изображения делится на квадратные плитки с вершинами в центрах пикселей:

- каждая плитка может содержать не более одной опорной точки линии, представляющей границу;
- каждая сторона плитки может пересекаться не более чем с одной линией границы;
- граница не может проходить через угол плитки.

При использовании космических изображений не имеет смысла требовать хранения в особом виде, но существует возможность использовать дополнительную информацию о поверхности, изображение которой требуется улучшить. В качестве такой информации доступны различные карты поверхности, созданные ранее по другим данным дистанционного зондирования, а также с помощью наземных измерений — карты подстилающей поверхности, граница суши и воды, карта дорожной сети и др. Предполагается использование таких карт в качестве априорной информации в векторной форме, например в формате шейпфайла [20], содержащего полигоны.

Сформулируем условия задачи.

На плоскости (x, y) задано изображение — функция яркости $I(x_i, y_j)$, заданная на регулярной сетке

$$x_i = id; \quad y_j = jd, \quad i \in \{1, \dots, h\}; \quad j \in \{1, \dots, w\},$$

где d — размер ячейки. Точки изображения, представленные своими координатами и значением яркости, будем называть также пикселями.

Дополнительно задано деление множества определения функции на области P_k , $k \in \{1, \dots, K\}$. Каждая область представляет собой отдельный объект, обладающий резкой границей с другими объектами и фоном, так что при операциях с изображением значения яркости внутри и снаружи объекта не должны смешиваться.

На практике области P_k , $k \in \{1, \dots, K - 1\}$, задаются в виде полигонов, а область P_K представляет собой фон — множество точек, не вошедших ни в одну другую область. Каждый полигон представляет собой плоский многоугольник, заданный последовательностью точек $p_l(x, y)$, $l \in \{1, \dots, n\}$.

Требуется получить пропорционально увеличенное изображение повышенного разрешения:

$$I(x_p, y_q), \quad x_p = pb, \quad y_q = qb, \quad p \in \{1, \dots, H\}, \quad q \in \{1, \dots, W\}, \quad b < d,$$

причем яркости точек, принадлежащих к определенному полигону, должны быть вычислены с использованием известных значений функции только в тех точках, которые лежат внутри данного полигона:

$$\forall (x, y) \in P : I(x, y) = f(I(x_{k_1}, y_{k_1}), x_{k_1}, y_{k_1}, \dots, I(x_{k_m}, y_{k_m}), x_{k_m}, y_{k_m}) : (x_{k_i}, y_{k_i}) \in P \\ \forall i \in \{1, \dots, m\}.$$

Также для обеспечения линейной сложности алгоритм должен быть локальным, т. е.

$$\exists L : \forall (x, y) : I(x, y) = f(I(x_{k_1}, y_{k_1}), \dots, I(x_{k_m}, y_{k_m}), x_{k_m}, y_{k_m}) : \\ \forall i \in \{1, \dots, m\} : \rho((x, y), (x_{k_i}, y_{k_i})) < L, \quad (1)$$

где $\rho((x_1, y_1), (x_2, y_2))$ — расстояние между точками (x_1, y_1) и (x_2, y_2) .

По аналогии с ограничением на описание границ в определении бикселей, необходимо задать условия на полигоны:

- каждый полигон должен содержать как минимум один пиксель исходного изображения, чтобы не было неопределенных значений:

$$\forall P \exists i \in \{1, \dots, h\}, j \in \{1, \dots, w\} : (x_i, y_j) \in P; \quad (2)$$

- полигоны не должны иметь пересечений.

$$\forall i, j : P_i \cap P_j = \emptyset.$$

Для каждой точки нового изображения (x_p, y_q) решается задача интерполяции функции $I(x, y)$ по известным точкам. Как базовый метод повышения разрешающей способности используется один из известных локальных методов, не требующих дополнительной информации — билинейная или бикубическая интерполяция [10].

В случае, когда все пиксели исходного изображения, влияющие на данные пиксель нового, согласно (1), расположены в одной области и условие (2) выполняется автоматически, можно считать, что в окрестности не проходит границы, и использовать базовый метод интерполяции.

В том случае, когда пиксели исходного изображения, влияющие на данный пиксель нового изображения, лежат в разных областях, для интерполяции нового пикселя используются только те пиксели исходного изображения, которые лежат в той же области, что и данный пиксель изображения повышенного разрешения.

Таким образом, вблизи заданных границ объектов невозможно построение регулярной прямоугольной сетки, поэтому требуется применение методов, работающих на нерегулярных сетках. В статье [21] описывается такой метод, позволяющий интерполировать значения, определенные на нерегулярной сетке гладкой функцией. В данном методе используется разное количество узлов в зависимости от расстояния их до искомой точки, направления и плотности узлов в данной области. Полученная функция соответствует значениям, заданным в узлах, является гладкой и непрерывной. Значение функции в точке выражается следующим образом:

$$f(P) = \begin{cases} \frac{|\sum_{D_i \in C'} w_i (f_i + \Delta f_i)|}{|\sum_{D_i \in C'} w_i|}, & \text{если } \forall D_i : d_i > \varepsilon; \\ f_i & \text{если } \exists D_i : d_i \leq \varepsilon, \end{cases},$$

где P — искомая точка, для которой производится интерполяция; D_i — узлы нерегулярной сетки, участвующие в интерполяции; f_i — значение функции в узле D_i ; d_i — расстояние между искомой точкой и узлом D_i ; ε — малое число, введенное для учета погрешностей округления при компьютерном расчете; Δf_i — параметр, характеризующий первую производную функции вблизи узла сетки; при этом предполагается, что только одна точка может быть настолько близка к искомой, так как исходные точки выбираются из узлов регулярной сетки.

Множество C' — те узлы нерегулярной сетки, которые используются при расчете интерполяции, их не менее 4 и не более 10. В него входят ближайшие к искомой точке узлы, входящие в радиус r , выбираемый так, чтобы в среднем в него входили 7 точек:

$$r = \sqrt{\frac{7A}{N\pi}},$$

где A — площадь выпуклой оболочки исходного множества точек; N — их количество, либо 4 ближайших узла, если в радиус r входят не более 3, либо 10 ближайших узлов, если в радиус r входят более 10. Соответственно r' — расстояние от P , в которое входят $|C'|$ точек.

Весовая функция, определяющая влияние каждого узла сетки на интерполируемое значение, задается следующим образом:

$$w_i = s_i^2(1 + t_i),$$

где s_i — член, учитывающий расстояние до узла D_i , а t_i — член, учитывающий направление.

Весовая функция зависит от расстояния таким образом, чтобы гладким образом сходиться к нулю при удалении, имея вблизи вид обратного квадрата:

$$s_i = \begin{cases} \frac{1}{d}, & 0 < d \leq \frac{r'}{3}; \\ \frac{27}{4r'} \left(\frac{d}{r'} - 1 \right)^2, & \frac{r'}{3} < d \leq r'; \\ 0, & r' < d \end{cases}$$

Зависимость весовой функции от направления введена для того, чтобы учитывать «затенение» более удаленных точек на одном и том же направлении от искомой:

$$t_i = \frac{\sum_{D_j \in C'} s_j |1 - \cos(D_i P D_j)|}{\sum_{D_j \in C'} s_j}.$$

В этом члене для точки D_i учитываются углы $D_i P D_j$ для всех j , и он тем меньше, чем меньше эти углы. Таким образом, точки, лежащие рядом друг с другом в одном направлении от точки, для которой проводится интерполяция, будут иметь пониженный весовой коэффициент по сравнению с точками, лежащими на менее плотных направлениях.

Несмотря на то что есть возможность использовать данный метод для всего изображения как вблизи границ, так и в отдалении от них, в рассматриваемом случае это не представляется целесообразным, так как при удалении от границы возможна интерполяция по регулярной сетке (например, бикубическая), которая дает хорошие результаты и менее вычислительно затратна.

3 Описание алгоритма

В качестве базового метода интерполяции выбрана бикубическая интерполяция, так как она позволяет производить достаточно качественное повышение разрешения и при этом является относительно простым, локальным, одноступенчатым методом.

Таким образом, базовым методом интерполяции можно воспользоваться, когда 16 пикселей исходного изображения, ближайших к точке, для которой необходимо вычислить яркость, находятся в одной области. В случае когда часть этих пикселей лежит в той же области, что и искомая точка, а часть отделена границей, существует множество возможностей вычисления и часть из них требует экспериментального сравнения.

В случае если не все 16 заданных точек лежат в данной области, применяется вычисление значения яркости методом Шепарда с выбором исходных точек из пересечения искомой области с ближайшими 16 пикселями.

Назовем квадрат, образованный центрами четырех соседних пикселей изображения низкого разрешения, ячейкой, а окружающие его 16 пикселей исходного изображения окрестностью. На рис. 2 представлена ячейка (заштрихованная область) и ее окрестность.

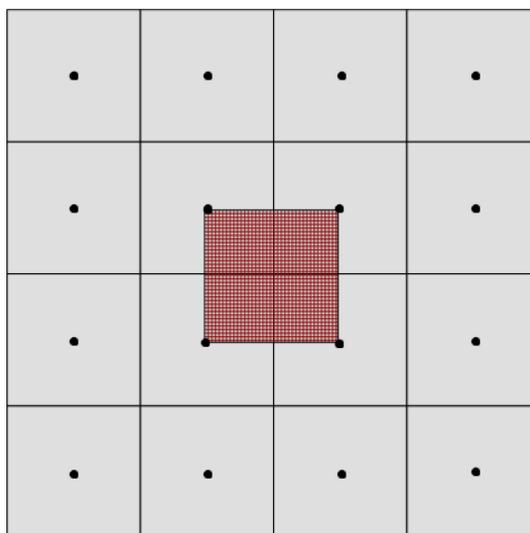


Рис. 2 Ячейка изображения и 16 пикселей, используемых для интерполяции в ее пределах

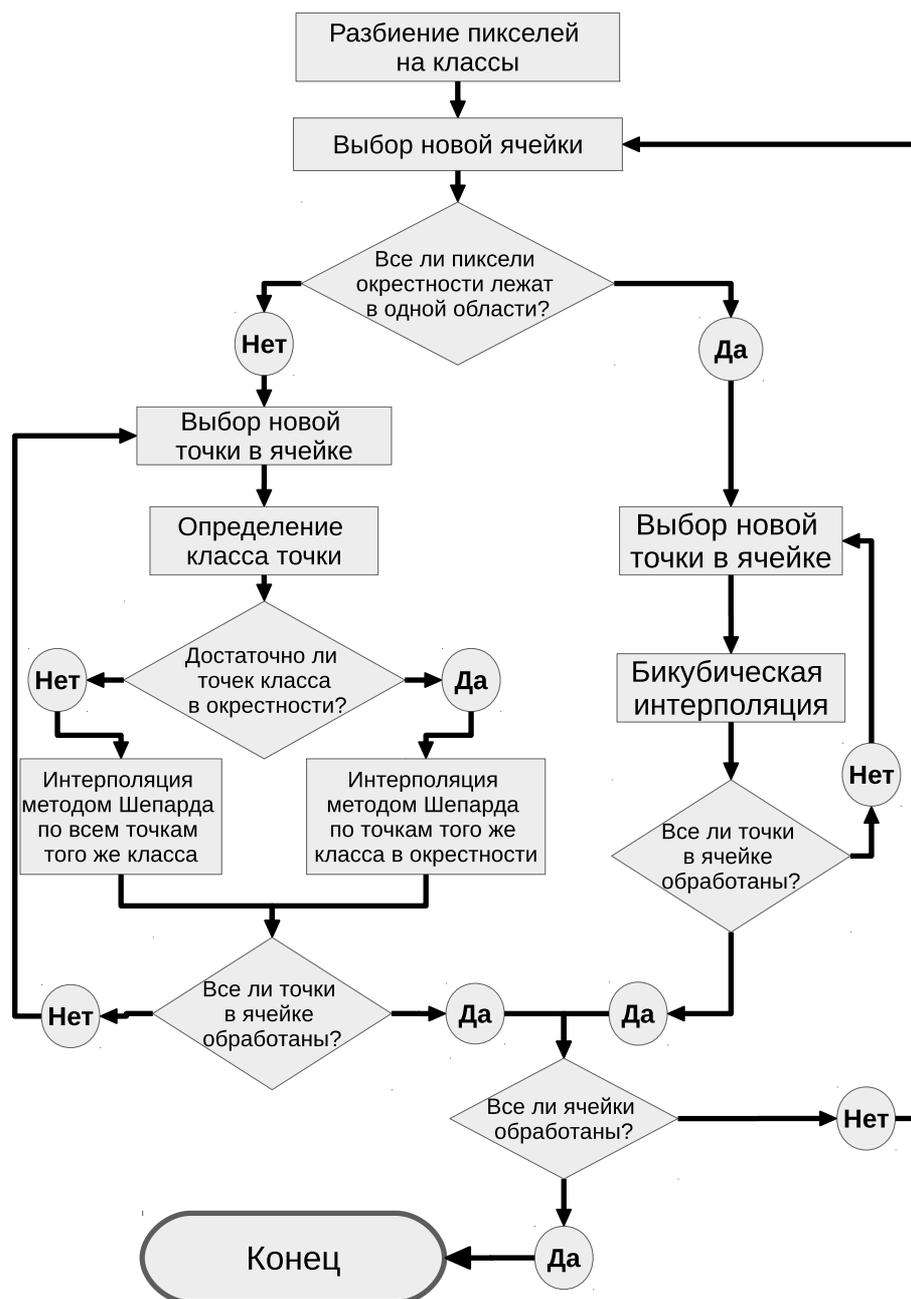


Рис. 3 Блок-схема работы алгоритма

Значения интерполяционной функции внутри ячейки зависят только от значений яркости в окрестности, причем для всех точек внутри ячейки метод интерполяции один и тот же: либо бикубическая интерполяция, если вся окрестность лежит в одной области, либо метод Шепарда для пикселей, лежащих в той же области, что и точка, для которой вычисляется функция. Поэтому повышение разрешения состоит из повторения идентичных шагов для каждой ячейки исходного изображения и каждого класса объектов, содержащихся в шейпфайле.

Алгоритм обработки изображения в виде блок-схемы представлен на рис. 3.

Области с контрастными границами задаются в одном или нескольких векторных файлах в виде полигонов. Для определения, какой области принадлежит определенная точка,

производится проверка по каждому шейпфайлу, лежит ли точка внутри какого-либо из его полигонов. Так как такую проверку необходимо делать для каждой точки, целесообразно использовать метод индексации векторного файла для ускорения поиска нужного полигона [22]. В зависимости от того, лежат ли точки внутри какого-либо из полигонов, каждой из них присваивается номер соответствующей области. Если все точки окрестности лежат в одной области, то к каждой точке, для которой требуется рассчитать значение, т. е. каждой точке изображения повышенного разрешения, лежащей внутри данной ячейки, применяется бикубическая интерполяция.

Если же в окрестности ячейки находятся пиксели, принадлежащие разным областям, то для каждого пикселя изображения повышенного разрешения тем же образом определяется область, к которой он принадлежит. После этого искомое значение вычисляется интерполяцией по методу Шепарда на основании точек окрестности, принадлежащих той же области. Ограничения, наложенные на полигоны, обеспечивают наличие хотя бы одного пикселя исходного изображения внутри каждого полигона, поэтому интерполяция возможна для любой точки.

4 Экспериментальная проверка

Проверка разработанных методов и алгоритмов возможна как на реальных, так и на модельных данных [1, 2, 23], в данном случае исследование на реальных данных может показать более репрезентативные результаты. Для экспериментальной проверки были использованы космические изображения Landsat 5 с разрешением 30 м в надире и карта водоемов России в векторной форме. Для лучшей контрастности водоемов был выбран спектральный канал изображения в коротковолновом инфракрасном диапазоне (длина волны 1,55–1,75 мкм). Для тестирования алгоритма квадратный фрагмент изображения Landsat размером 500×500 точек загружался в 4 раза, а затем проводилось повышение разрешения исследуемым методом и для сравнения — повышение разрешения методом бикубической интерполяции. Метод бикубической интерполяции был выбран для сравнения, так как он используется как базовый метод во всех точках, удаленных от границ объектов, и полученная разница будет непосредственно представлять то отличие, которое достигнуто использованием информации о границах объектов. Также для сравнения использован метод Шепарда для всего изображения. Также результат работы разработанного алгоритма сравнивался с увеличением всего изображения методом Шепарда в вариантах с применением векторных данных и без применения.

Алгоритм Шепарда использован в реализации Burkardt [24], бикубическая интерполяция и программа для экспериментальной проверки реализована на языке С. Исходные изображения Landsat получены с помощью онлайн-сервиса Earthexplorer [25].

Пример исходных данных и результата работы приведен на рис. 4. Видно, что при повышении разрешения методом бикубической интерполяции (см. рис. 4, б) границы реки размываются, а при использовании априорной информации в векторной форме — сохраняют резкость.

Результаты сравнительного исследования собраны в табл. 1. Чтобы оценить качество повышения разрешения, полученные результаты сравниваются с эталоном — изображением, из которого загрузлением получено исходное изображение. Для этого применяются различные количественные оценки, в рамках данного исследования были выбраны две: среднеквадратичная ошибка (RMSE — root mean square error) и индекс структурного сходства (SSIM — structural similarity) [26, 27].

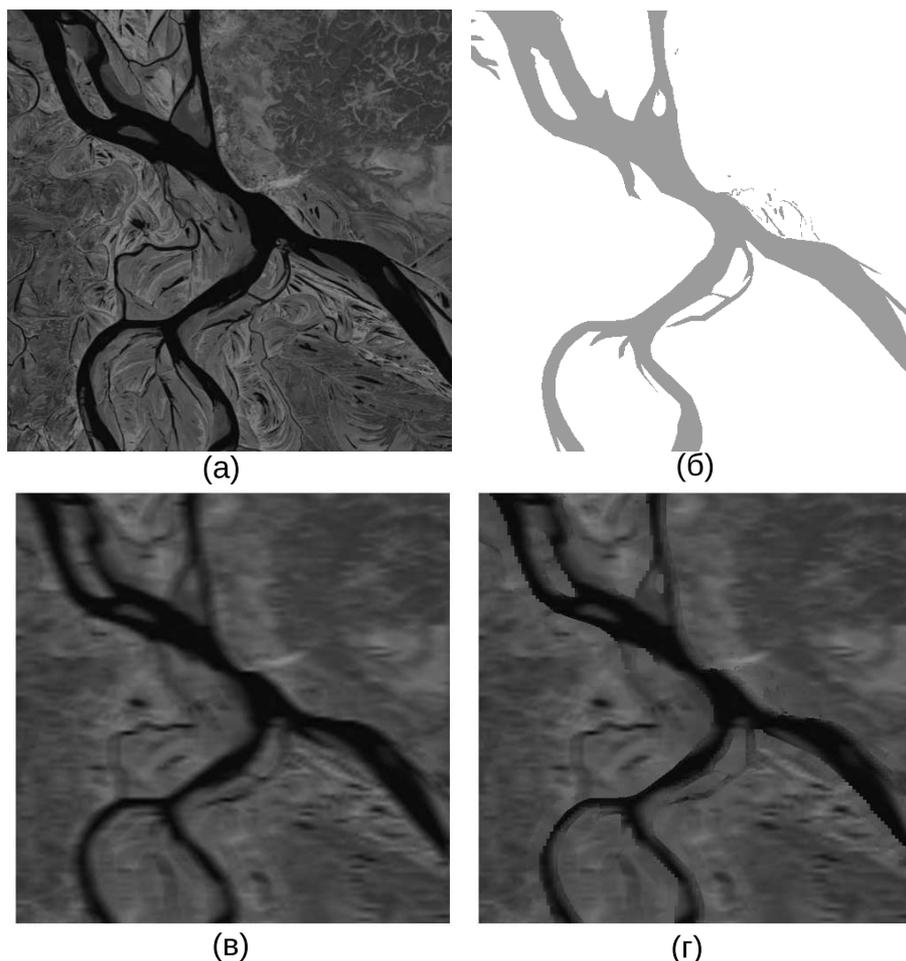


Рис. 4 Эталонное изображение (а); векторные данные — карта водоемов (б); результат бикубической интерполяции (в) и результат разработанного метода с использованием априорной информации в векторной форме (г)

Таблица 1 Результат экспериментальной проверки

| Метод | RMSE | MSSIM |
|---------------------------|---------------|----------------|
| Бикубическая интерполяция | 433,4 | 0,45844 |
| Метод Шепарда | 483,88 | 0,42699 |
| Метод Шепарда с границами | 386,71 | 0,45002 |
| Данный метод | 366,37 | 0,47107 |

Среднеквадратичная ошибка определяет изменение яркости пикселей по формуле [26]:

$$\text{RMSE}(I_1, I_2) = \sqrt{\frac{\sum_{i=1}^h \sum_{j=1}^w (I_1(i, j) - I_2(i, j))^2}{hw}},$$

где $I_1(i, j)$ и $I_2(i, j)$ — функции яркости сравниваемых изображений.

При измерении на однородной области изображения среднеквадратичная ошибка позволяет достаточно хорошо оценить спектральное качество изображения. Среднеквадратичная ошибка должна быть близка к нулю. Индекс структурного сходства характерен

тем, что учитывает «восприятие ошибки», благодаря учету структурного изменения информации. Идея заключается в том, что пиксели имеют сильную взаимосвязь, особенно когда они близки пространственно. Данные зависимости несут важную информацию о структуре объектов и о сцене в целом. Для двух квадратных окон $a \subset I_1$ и $b \subset I_2$, имеющих одинаковый размер [27]:

$$\text{SSIM}_{\text{local}}(a, b) = \frac{(2\mu_a\mu_b + c_1)(2\sigma_{ab}^2 + c_2)}{(\mu_a^2 + \mu_b^2 + c_1)(\sigma_a^2 + \sigma_b^2 + c_2)},$$

где μ_a — среднее значение яркости изображения I_1 в окне a ; μ_b — среднее значение яркости изображения I_2 в окне b ; σ_a^2 — дисперсия яркости a ; σ_b^2 — дисперсия b ; σ_{ab}^2 — ковариация a и b . Константы c_1 и c_2 вычисляются как $c_1 = (0,01L)^2$ и $c_2 = (0,03L)^2$, где L — динамический диапазон значений яркости пикселей.

Метрика SSIM рассчитана на различные размеры окна, как правило, это квадратный участок изображения с линейным размером 8 пикселей. Она принимает значения от -1 до 1 , причем значение 1 соответствует абсолютно идентичным участкам, когда $\mu_a = \mu_b$, $\sigma_a^2 = \sigma_b^2 = \sigma_{ab}^2$, при появлении различий $2\mu_a\mu_b < \mu_a^2 + \mu_b^2$ и $2\sigma_{ab}^2 < \sigma_a^2 + \sigma_b^2$ индекс уменьшается вплоть до значения -1 .

Для оценки качества всего изображения рассчитывается среднее значение SSIM по всевозможным положениям окна на изображениях:

$$\text{MSSIM} = \sum_{x \in I_1, y \in I_2} \text{SSIM}_{\text{local}}(x, y).$$

Полученный индекс, являясь усреднением SSIM, лежит в пределах от -1 до 1 , и чем он больше, тем более схожи изображения.

Время работы разработанного алгоритма зависит от концентрации границ, заданных векторными данными. В случае отсутствия границ время работы соответствует бикубической интерполяции, в случае наличия их в окрестности любой ячейки — методу Шепарда. На тестовых данных время работы бикубической интерполяции, разработанного алгоритма и метода Шепарда соотносится как $1 : 5 : 60$.

Полученные результаты показывают, что применение векторной информации при использовании бикубической интерполяции в качестве базового метода повышения разрешения позволяет улучшить результат. Разработанный метод, сочетая в себе подходы, примененные в базовых методах, в рассмотренных условиях превосходит их оба по качеству повышения разрешения изображения, измеренному с помощью двух различных критериев.

Следует заметить, что повышение качества требует точного позиционирования векторных данных относительно растровых, иначе неизбежно появление артефактов — перепадов яркости там, где их быть не должно, и соответственно снижаются численные характеристики качества повышения разрешения.

При этом улучшение в меньшей степени касается восприятия схожести изображения человеком, что показывают близкие значения индекса структурного сходства, и в большей степени реальной, математической попиксельной близости к изображению, которая выражается методом среднеквадратичного отклонения. Таким образом, данный метод может быть применен для повышения разрешения изображений с целью последующей компьютерной обработки методами, использующими значения яркости — сегментации, классификации и др.

5 Заключение

Разработан метод повышения пространственного разрешения изображений с использованием априорной информации о границах объектов, выраженной в векторной форме. Метод применим для улучшения космических изображений, причем в качестве векторной информации о границах объектов доступны различные карты местности, предварительно составленные по данным высокого разрешения, наземным наблюдениям и т. п.

Разработанный метод использует разные подходы к повышению разрешения вблизи известных контрастных границ и вдали от них, что позволяет комбинировать эффективные алгоритмы для оптимального повышения разрешения с сохранением четких, резких границ там, где это необходимо, и гладких, непрерывных решений в прочих местах. Для повышения разрешения вблизи границ использован метод Шепарда интерполяции по нерегулярной сетке, а в удалении от них — метод бикубической интерполяции.

Качество повышения разрешения исследовано на космических изображениях Landsat с использованием векторной информации в виде карты водоемов. Для сравнения качества повышения разрешения использованы критерии среднеквадратичного отклонения и среднее значение индекса структурного сходства, которые показали, что использование векторной информации позволяет улучшить результат по сравнению с бикубической интерполяцией. Приведенные результаты показывают возможность применения разработанного метода для повышения разрешающей способности аппаратуры, устанавливаемой на космических аппаратах ДЗЗ, в том числе в инфракрасном и микроволновом диапазонах спектра электромагнитных волн.

Разработанный метод может быть использован для космического мониторинга урбанизированных территорий, береговых линий, ледовой обстановки, природных пожаров и территорий, пройденных огнем, предупреждения и оценки последствий чрезвычайных ситуаций, создания и актуализации картографической продукции, решения других задач дистанционного зондирования, в том числе с применением мультиспектральных и гиперспектральных данных, содержащих значительное количество спектральной информации, однако зачастую имеющих недостаточное пространственное разрешение [28]. В качестве развития данного подхода предполагается применение других, более совершенных методов интерполяции на регулярной сетке и вблизи границ с учетом доли площади пикселя, попадающего в выделенную область.

Литература

- [1] *Бондур В. Г.* Методы моделирования полей излучения на входе аэрокосмических систем дистанционного зондирования // Исслед. Земли из космоса, 2000. № 5. С. 16–27.
- [2] *Бондур В. Г., Аржениенко Н. И., Линник В. Н., Титова И. Л.* Моделирование многоспектральных аэрокосмических изображений динамических полей яркости // Исслед. Земли из космоса, 2003. № 2. С. 3–17.
- [3] *Савин А. И., Бондур В. Г.* Научные основы создания и диверсификации глобальных аэрокосмических систем // Оптика атмосферы и океана, 2000. Т. 13. № 1. С. 46–62.
- [4] *Бондур В. Г.* Аэрокосмические методы в современной океанологии // Новые идеи в океанологии. Т. 1.: Физика. Химия. Биология. — М.: Наука, 2004. С. 55–117.
- [5] *Бондур В. Г., Савин А. И.* Принципы моделирования полей сигналов на входе аппаратуры ДЗ аэрокосмических систем мониторинга окружающей среды // Исслед. Земли из космоса, 1995. № 4. С. 24–33.
- [6] *Turkowski K.* Filters for common resampling tasks // Graphics gems. — Academic Press, Inc., 1990. P. 147–165.

- [7] Sun J., Xu Z., Shum H.-Y. Image super-resolution using gradient profile prior // IEEE Conference on Computer Vision and Pattern Recognition (CVPR2008) Proceedings, 2008. P. 1–8.
- [8] Price J. R., Hayes M. H. Resampling and reconstruction with fractal interpolation functions // IEEE Signal Proc. Let., 1998. Vol. 5. No. 9. P. 228–230.
- [9] Маркелов К. С. Модель повышения информативности цифровых изображений на базе метода суперразрешения // Инженерный вестник, 2013. № 3.
- [10] Ярославский Л. П. Введение в цифровую обработку изображений. — М.: Советское радио, 1979. 312 с.
- [11] Миронов А. А., Цурков В. И. Транспортные и сетевые задачи с минимаксным критерием // Журнал вычислительной математики и математической физики, 1995. Т. 35. № 1.
- [12] Миронов А. А., Цурков В. И. Транспортные задачи с минимаксным критерием // Докл. Акад. наук СССР, 1996. Т. 346. № 2.
- [13] Тизик А. П., Цурков В. И. Метод последовательной модификации функционала для решения транспортной задачи // Автоматика и телемеханика, 2012. № 1. С. 148–158.
- [14] Nasonov A. V., Krylov A. S. Finding areas of typical artifacts of image enhancement methods // Pattern Recognition Image Anal., 2011. Vol. 21. No. 2. P. 316–318.
- [15] Цурков В. И., Ковков Д. В. Способ удаления шума в изображении. Пат. на изобр. RUS 2316816, 25.08.2005.
- [16] Fattal R. Upsampling via imposed edges statistics // ACM SIGGRAPH Proceedings, 2007. Vol. 26. No. 3. Article No. 95.
- [17] Tsurkov V. I. An analytical model of edge protection under noise suppression by anisotropic diffusion // Int. J. Comput. Syst. Sci., 2000. Vol. 39. No. 3. P. 437–440.
- [18] Чебурашкин И. Ф., Цурков В. И. Синтез дискретных логических устройств обработки информации на основе теории агентов // Мехатроника, автоматизация, управление, 2011. № 3. С. 27–34.
- [19] Tumblin J., Chouldhury P. Bixels: Picture samples with sharp embedded boundaries // Eurographics Symposium on Rendering, 2004.
- [20] ESRI. <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>.
- [21] Shepard D. A two-dimensional interpolation function for irregularly-spaced data // ACM National Conference Proceedings, 1968. P. 517–524.
- [22] Бондур В. Г., Мурынин А. Б., Матвеев И. А., Трекин А. Н., Юдин И. А. Метод вычислительной оптимизации в задаче сопоставления растровой и векторной информации при анализе спутниковых данных // Современные проблемы дистанционного зондирования, 2013. Т. 10. № 4. С. 98–106.
- [23] Бондур В. Г. Моделирование двумерных случайных полей яркости на входе аэрокосмической аппаратуры методом фазового спектра // Исслед. Земли из космоса, 2000. № 5. С. 28–44.
- [24] Shepard Interpolation of 2D Data. http://people.sc.fsu.edu/~jburkardt/c_src/shepard_interp_2d/shepard_interp_2d.html.
- [25] EarthExplorer. <http://earthexplorer.usgs.gov/>.
- [26] Wald L., Ranchin T., Mangolini M. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images // Photogramm. Eng. Remote Sens., 1997. Vol. 63. P. 691–699.
- [27] Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. Image quality assessment: From error visibility to structural similarity // IEEE Trans. Image Proc., 2004. Vol. 13. No. 4. P. 600–612.

- [28] *Бондур В. Г.* Современные подходы к обработке больших потоков гиперспектральной и многоспектральной аэрокосмической информации // Исслед. Земли из космоса, 2014. № 1. С. 4–16.

Поступила в редакцию 17.06.2015

References

- [1] Bondur, V. 2000. Metody modelirovaniya poley izlucheniya na vkhode aerokosmicheskikh sistem distantsionnogo zondirovaniya [Methods for modeling of radiation fields at input of remote sensing aerospace systems]. *Issledovanie Zemli iz Kosmosa* 5:16–27. (In Russian.)
- [2] Bondur, V., N. Arzhenenko, V. Linnik, and I. Titova. 2003. Modelirovanie mul'tispektral'nykh aerokosmicheskikh izobrazheniy dinamicheskikh poley yarkosti [Modeling of multispectral aerospace images of dynamic brightness fields]. *Issledovanie Zemli iz Kosmosa* 2:3–17. (In Russian.)
- [3] Savin, A., and V. Bondur. 2000. Nauchnye osnovy sozdaniya i diversifikatsii globalnykh aerokosmicheskikh sistem. *Optika Atmosfery i Okeana* [Optics of Atmosphere and Ocean] 13(1):46–62. (In Russian.)
- [4] Bondur, V. 2004. Aerokosmicheskie metody v sovremennoy okeanologii [Aerospace methods in modern oceanology]. *Noviye idei v okeanologii. T. 1: Fizika. Khimiya. Biologiya* [New ideas in oceanology. Vol. 1: Physics. Chemistry. Biology]. Moscow: Nauka. 55–117. (In Russian.)
- [5] Bondur, V., and A. Savin. 1995. Printsipy modelirovaniya poley signalov na vkhode apparatury DZ aerokosmicheskikh sistem monitoringa okruzhayushey sredy. *Issledovanie Zemli iz Kosmosa* 4:24–33. (In Russian.)
- [6] Turkowski, K. 1990. Filters for common resampling tasks. *Graphics gems*. Academic Press, Inc. 147–165.
- [7] Sun, J., Z. Xu, and H.-Y. Shum. 2008. Image super-resolution using gradient profile prior. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR2008) Proceedings*. 1–8.
- [8] Price, J., and M. Hayes. 1998. Resampling and reconstruction with fractal interpolation functions. *IEEE Signal Proc. Let.* 5(9):228–230.
- [9] Markelov, K. 2013. Model' povysheniya informativnosti izobrazheniy na baze metoda super-razresheniya [A model of image comprehension based on superresolution]. *Inzhenernyy Vestnik*. 3. (In Russian.)
- [10] Yaroslavskiy, L. 1979. *Vvedenie v tsifrovuyu obrabotku izobrazheniy* [Introduction to digital image processing]. Moscow: Sovetskoye radio. 312 p. (In Russian.)
- [11] Mironov, A., and V. Tsurkov. 1995. Transportnye i setevye zadachi s minimaksnym kriteriem [Transportation and roadmap problems with min-max criterion]. *Zhurnal Vychislitel'noy Matematiki i Matematicheskoy Fiziki* [J. Comput. Math. Math. Phys.] 35(1). (In Russian.)
- [12] Mironov, A., and V. Tsurkov. 1996. Transportnye i setevye zadachi s minimaksnym kriteriem [Transportation problems with min-max criterion]. *Dokl. Akad. Nauk USSR* 346(2). (In Russian.)
- [13] Tizik, A., and V. Tsurkov. 2012. Metod posledovatel'noy modifikatsii funktsionala dlya resheniya transportnoy zadachi [A method for serial modification of functional for solving the transportation problem]. *Avtomatika i Telemekhanika* [Automatics and Telemechanics] 1:148–158. (In Russian.)
- [14] Nasonov, A., and A. Krylov. 2011. Finding areas of typical artifacts of image enhancement methods. *Pattern Recognition Image Anal.* 21(2):316–318.
- [15] Tsurkov, V., and D. Kovkov. 25.08.2005. Sposob udaleniya shuma v izobrazhenii [A method for noise elimination in images]. Invention Patent RUS 2316816. (In Russian.)

- [16] Fattal, R. 2007. Upsampling via imposed edges statistics. *ACM SIGGRAPH Proceedings*. 26(3). Article No. 95.
- [17] Tsurkov, V. 2000. An analytical model of edge protection under noise suppression by anisotropic diffusion. *Int. J. Comput. Syst. Sci.* 39(3):437–440.
- [18] Cheburakhin, I., and V. Tsurkov. 2011. Sintez diskretnykh logicheskikh ustroystv obrabotki informatsii na osnove teorii agentov [Synthesis of discrete logical devices for agent theory-based information processing]. *Mekhatronika Avtomatizatsiya Upravlenie* 3:27–34. (In Russian.)
- [19] Tumblin, J., and P. Choudhury. 2004. Bixels: Picture samples with sharp embedded boundaries. *Eurographics Symposium on Rendering*.
- [20] ESRI. Available at: <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf> (accessed November 25, 2015).
- [21] Shepard, D. 1968. A two-dimensional interpolation function for irregularly-spaced data. *ACM National Conference Proceedings*. 517–524.
- [22] Bondur, V., A. Murynin, I. Matveev, A. Trekin, and I. Yudin. 2013. Metod vichislitel'noy optimizatsii v zadache sopostavleniya rastrovoy i vektornoy informatsii pri analize sputnikovykh dannykh [A computational optimization method for raster and vector information matching for satellite data analysis]. *Sovremennye Problemy Distantionnogo Zondirovaniya Zemli iz Kosmosa* 10(4):98–106. (In Russian.)
- [23] Bondur, V. 2000. Modelirovanie sluchaynykh poley yarkosty na vkhode aerokosmicheskoy apparatury metodom fazovogo spektra [Modeling of brightness random fields at input of aerospace instrumentation using phase spectrum method]. *Issledovanie Zemli iz Kosmosa* 5:28–44. (In Russian.)
- [24] Shepard Interpolation of 2D Data. Available at: http://people.sc.fsu.edu/~jburkardt/c_src/shepard_interp_2d/shepard_interp_2d.html (accessed November 25, 2015).
- [25] EarthExplorer. Available at: <http://earthexplorer.usgs.gov/> (accessed November 25, 2015).
- [26] Wald, L., T. Ranchin, and M. Mangolini. 1997. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogramm. Eng. Remote Sens.* 63:691–699.
- [27] Wang, Z., A. Bovik, H. Sheikh, and E. Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Proc.* 13(4):600–612.
- [28] Bondur, V. 2014. Sovremennye podkhody k obrabotke bol'shikh potokov giperspektral'noy i multispektral'noy aerokosmicheskoy informatsii [Modern approaches to processing of large hyperspectral and multispectral aerospace data flows]. *Issledovanie Zemli iz Kosmosa* 1:4–16. (In Russian.)

Received June 17, 2015

Восстановление пропущенных данных на ультразвуковых дефектограммах рельсового пути на основе локально-глобального выравнивания*

Д. Н. Чеprasов¹, А. А. Маленичев¹, В. В. Сулимова¹, О. В. Красоткина¹,
В. В. Моттль², А. А. Марков³

tcheprasov.mitya@yandex.ru; malenichev@mail.ru; vsulimova@yandex.ru,
o.v.krasotkina@yandex.ru; vmottl@yandex.ru; amarkovspb@gmail.com

¹Тульский государственный университет, Тула, Россия

²Московский физико-технический институт, Москва, Россия

³ОАО «Радиоавионика», Санкт-Петербург, Россия

Статья посвящена актуальной проблеме автоматического восстановления пропущенных данных при ультразвуковом контроле рельсового пути, например из-за плохих погодных условий. Предлагаемый подход основывается на извлечении пропущенных данных из дефектограммы, соответствующей предыдущему проходу по тому же участку рельсового пути. В данной работе авторы усовершенствуют предложенный ими ранее метод, делая его более точным и существенно более быстрым. Разработанная трехконная модель болтового соединения совместно с процедурой локально-глобального выравнивания и специальной мерой несходства элементов дефектограмм позволяют достаточно быстро обнаружить местоположение искомой области на предыдущей дефектограмме с точностью около 3,5 см, что обеспечивает возможность успешного восстановления по ней пропущенных данных.

Ключевые слова: неразрушающий контроль рельсов; ультразвуковая дефектограмма; пропущенные данные; сегментация сигнала; локально-глобальное выравнивание

DOI: 10.21469/22233792.1.12.08

Recovering missing data on ultrasonic rail defectograms via semiglobal warping*

D. N. Cheprasov¹, A. A. Malenichev¹, V. V. Sulimova¹, O. V. Krasotkina¹,
V. V. Mottl², and A. A. Markov³

¹Tula State University, 92 Lenina pr., Tula, Russia

²Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Russia

³Radioavionica Corp., 4B Troitskiy pr., Saint Petersburg, Russia

Background: The paper deals with the actual problem of automatic recovering missing data on ultrasonic rail defectograms, occurring due to, for example, bad weather conditions.

Methods: The proposed approach is based on retrieving missing data for current ultrasonic inspection from ultrasonic defectogram of previous inspection. In this work, the authors update their previous method, making it more accurate and appreciably more fast one. A special 3-windows model for fast localization of bolt-on joint areas and semiglobal warping procedure have been proposed based on special dissimilarity measure of defectogram elements for more precise definition of bolt-on joint area positions.

* Работа выполнена при поддержке РФФИ, проекты №№ 15-37-50751, 15-07-08967 и 14-07-00661.

Results: The experiments show that the proposed approach allows for finding place of the area of interest with accuracy about 3.5 cm. So, it possesses good possibility to recover missing data from the previous defectogram.

Keywords: *nondestructive rail inspection; ultrasonic defectogram; missing data; signal segmentation; semiglobal warping*

DOI: 10.21469/22233792.1.12.08

1 Введение

В процессе эксплуатации железнодорожного пути в рельсах под воздействием подвижного состава, природных и других факторов образуются дефекты и повреждения, угрожающие безопасности движения поездов. В связи с этим, во избежание чрезвычайных ситуаций, необходимо своевременное выявление и замена рельсов, содержащих дефекты [1, 2].

Большинство дефектов располагаются внутри рельса и оказываются не видны при обычном осмотре, в связи с чем для их выявления в настоящее время в российской железнодорожной отрасли активно используются управляемые операторами средства ультразвукового контроля — съемные двухниточные дефектоскопы (рис. 1), осуществляющие регистрацию сигналов ультразвукового контроля на дефектограмме [3–5].

Однако, хотя существует специальное программное обеспечение для сбора, просмотра и интерпретации ультразвуковых дефектограмм [6, 7], ежегодно все-таки случается около 40–60 изломов рельсов [8], причем каждый из таких изломов может иметь чрезвычайно серьезные последствия.

Одной из основных причин обнаружения дефектов являются так называемые непроконтролируемые области — участки рельсового пути, на которых ультразвуковой сигнал не был зарегистрирован [9]. Такие области могут появляться, например, из-за специфического повреждения поверхности рельсов или неровного их соединения [9, 10], высокой скорости движения дефектоскопа и(или) временного плохого акустического контакта приемника ультразвукового сигнала с поверхностью исследуемого участка рельсового пути.

При этом наиболее частой причиной появления непроконтролируемых участков являются неблагоприятные погодные условия [9, 10], о чем, в частности, говорит существенное увеличение количества таких участков в холодные месяцы года [4]. Более того, статистика показывает, что лишь около 15% непроконтролируемых участков также являются непроконтролируемыми и на предыдущих проходах дефектоскопом по тому же участку



Рис. 1 Современный съемный двухниточный дефектоскоп АВИКОН-14

рельсового пути [4], что, с учетом достаточно частого контроля, осуществляемого 2–4 раза в месяц [11], дает возможность восстановления пропущенных данных на основе анализа дефектограммы, зарегистрированной во время предыдущего прохода по тому же участку.

Однако в большинстве работ проблема наличия непроконтролированных участков даже не поднимается [1, 3, 6, 7, 12, 13]. На практике же в настоящее время данная проблема обычно решается путем повторного исследования оператором непроконтролированного участка [2], что, естественно, приводит к дополнительным временным и материальным затратам. Иногда осуществляется ручной поиск соответствующего участка на дефектограмме предыдущего прохода [14], однако очевидно, что гораздо более предпочтительным было бы привлечение средств автоматизации [15].

Вообще говоря, избежать непроконтролированных областей позволяет применение магнитодинамических (МД) методов [16]. Однако МД-устройства имеют свои специфические особенности по сравнению с ультразвуковыми: возможность обнаружения дефектов только в верхней части головки рельсов; значительные габариты и массу намагничивающей системы и, в связи с этим, возможность реализации метода в основном на автомотрисах и вагонах-дефектоскопах. В [17] предложена технология совместного использования данных магнитодинамических и ультразвуковых методов для формирования диагностической карты рельсового пути. Целью же данной работы является решение проблемы наличия непроконтролированных областей, оставаясь в рамках исключительно ультразвукового контроля, без привлечения дополнительных технических средств.

Хотя известно множество методов восстановления пропущенных данных во временных рядах [18–20], однако следует отметить, что обычно в таких задачах требуется восстановить пропущенные данные на основе анализа остальной части того же сигнала, в результате чего соответствующие методы основаны на интерполяции пропущенных данных тем или иным способом.

Особенность же задачи восстановления пропущенных данных на ультразвуковой дефектограмме заключается в возможности привлечения дополнительной информации с другой дефектограммы, что делает нецелесообразным применение традиционных методов.

Фактически в данном случае задача восстановления пропущенных данных сводится к задаче поиска соответствующей области на дефектограмме с другого прохода. При этом следует обратить внимание, что, несмотря на то что при регистрации ультразвукового сигнала фиксируются координаты дефектоскопа, данная задача не является тривиальной, поскольку регистрируемые координаты являются весьма приблизительными и при разных проходах для одного и того же объекта могут отличаться на несколько метров (что соответствует более чем тысяче отсчетов дефектограммы). Более того, важная особенность данной задачи, отличающая ее от задач поиска некоторой области в сигнале, заключается в том, что сама искомая область не известна.

Следует отметить, что, несмотря на глобальную тенденцию к переходу на бесстыковые рельсовые пути, в настоящее время все еще остается достаточно большое количество болтовых соединений. Более того, проведенный анализ показывает, что именно в зонах болтовых соединений наиболее часто возникают весьма опасные изломы, появление которых связывается с наличием непроконтролированных областей [9].

В связи с этим в рамках данной работы, вслед за [21], предлагается решать задачу восстановления пропущенных данных на дефектограмме на основе выделения характерных элементов, которыми могут служить зоны болтовых соединений рельсового пути, т. е. производить частичную разметку дефектограммы.

Вообще говоря, известно множество различных методов разметки (сегментации) речевых [22], биомедицинских [23] и других сигналов [24–26], однако многие из них имеют очень высокую вычислительную сложность [23, 27]. В работах [14, 21] предложен способ осуществления разметки сигналов, ориентированный именно на задачу выделения зон болтовых соединений на ультразвуковых дефектограммах, за счет чего он является существенно более простым по сравнению с методами разметки общего вида. Тем не менее на практике оказывается, что и он имеет низкую вычислительную эффективность, поскольку требует прохождения скользящим окном по дефектограмме и осуществления парного выравнивания каждого фрагмента, вырезанного окном, с некоторым эталонным фрагментом, соответствующим типовому болтовому соединению. Данное обстоятельство делает невозможным практическое применение данного подхода в реальных условиях.

В связи с этим в данной работе предлагается новый подход, развивающий предложенный авторами ранее [21]. Предлагаемый подход тоже основан на прохождении скользящим окном по дефектограмме для поиска опорных зон болтовых соединений, однако требует лишь однократного выравнивания фрагментов дефектограмм для каждого болтового соединения за счет двух основных нововведений: во-первых, применения специальной трехконной модели для быстрой примерной локализации зоны болтового соединения и, во-вторых, применения специальной процедуры парного выравнивания фрагментов дефектограмм для уточнения ее положения.

2 Регистрация ультразвуковой дефектограммы в виде многокомпонентного дискретного сигнала

Наиболее перспективным способом представления сигналов, регистрируемых многоканальным дефектоскопом, является так называемая *B*-развертка — представление сигнала в координатной плоскости «время распространения ультразвуковых колебаний в рельсе — координата контролируемой нити пути» [4]. Однако, как было отмечено выше, координаты являются условными и лишь приближенно отражают местоположение объектов.

При ультразвуковой дефектоскопии рельсового пути, как правило, используются несколько каналов — ультразвуковых излучателей/приемников, посылающих зондирующие импульсы заданной амплитуды под некоторым углом к поверхности катания рельса (различные каналы отличаются друг от друга углом излучения/приема сигнала). Регистрация сигнала приемником происходит лишь в том случае, если в течение некоторого времени от момента излучения на приемник поступает отраженный сигнал, обладающий достаточно большой амплитудой.

Чаще всего при хорошем акустическом контакте ультразвукового излучателя и приемника, а также наличии отражающей поверхности, расположенной под углом 90° к направлению распространения ультразвука, происходит однократное отражение сигнала, и дефектоскоп регистрирует однократно отраженные от плоскости трещины импульсы. Однако, когда дефект или какой-либо конструктивный отражатель находится близко к поверхности рельса, ультразвуковой сигнал не успевает затухнуть, и могут наблюдаться случаи множественных переотражений, в результате которых будет зарегистрировано несколько групп (пачек) импульсов, принятых с разной задержкой относительно момента излучения. В случае отсутствия отражения ввиду плохого акустического контакта или отражения от дефекта, направленного к излучателю под углом, значительно отличающимся от 90° , импульс не будет зарегистрирован вовсе [15].

При наличии дефектов или каких-либо конструктивных отражателей (например, болтов при болтовом соединении рельсов) на дефектограмме появляются линии разной фор-

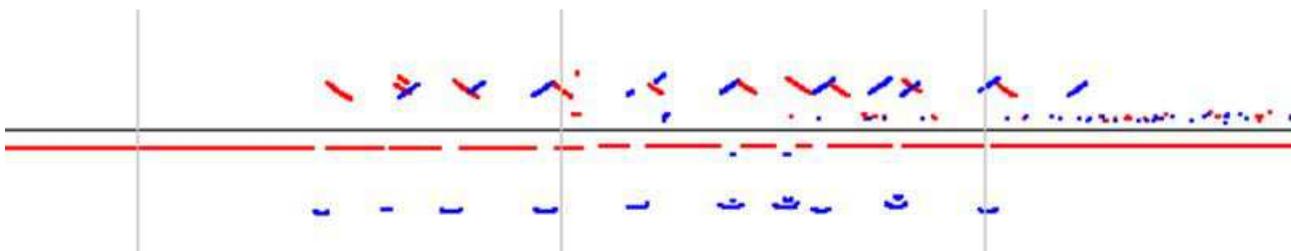


Рис. 2 Пример фрагмента ультразвуковой дефектограммы рельсового пути, содержащего зону болтового соединения

мы, ориентации и продолжительности в зависимости от вида объекта, оказавшегося на пути распространения ультразвукового сигнала (рис. 5б).

Согласно данному способу представления каждый элемент дефектограммы по отдельному каналу представляет собой импульсный сигнал в пространстве «задержка»–«амплитуда» и может содержать некоторое количество $n = 0, 1, 2, \dots$ импульсов, характеризующихся своей задержкой τ_i и амплитудой a_i , $i = 1, \dots, n$.

Для удобства дальнейших рассуждений элемент дефектограммы, не содержащий импульсов, будем рассматривать как сигнал, состоящий из одного элемента $n = 1$ с нулевой задержкой $\tau_1 = 0$ и нулевой амплитудой $a_1 = 0$. С учетом этого каждый элемент дефектограммы оказывается представлен двухкомпонентным сигналом длины n , составленным из пар $x = \{(\tau_i, a_i), i = 1, \dots, n\}$.

Для объединения информации, получаемой по разным ультразвуковым каналам в данной работе, так же как и в [21], каждый элемент дефектограммы рассматривается как m -мерный вектор

$$\mathbf{x} = [x_1, \dots, x_m]^T, \quad x_k = \{(\tau_i, a_i) \in R^2, i = 1, \dots, n\}, \quad k = 1, \dots, m, \quad (1)$$

а дефектограмма, соответственно, как m -компонентный дискретный сигнал $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_X})$ некоторой длины N_X .

3 Задача восстановления пропущенных данных на дефектограмме

3.1 Постановка задачи

Как было сказано выше, под пропущенными данными в рассматриваемом случае понимается отсутствие на зарегистрированной дефектограмме сигнала на некотором участке пути, причиной чего могут быть, например, неблагоприятные погодные условия.

В рамках данной работы задача восстановления пропущенных данных на дефектограмме формулируется как задача поиска пропущенной области на дефектограмме, соответствующей предыдущему проходу по тому же участку рельсового пути.

3.2 Основной принцип восстановления пропущенных данных

Поскольку на искомом участке дефектограммы вид сигнала неизвестен, то для его поиска в данной работе используется принцип, предложенный в [21], основывающийся на поиске на дефектограммах опорных областей, в качестве которых используются зоны болтовых соединений.



Рис. 3 Графическая иллюстрация принципа восстановления пропущенных данных

Согласно данному принципу, проиллюстрированному на рис. 57, поиск искомой области на дефектограмме предыдущего прохода осуществляется в несколько этапов:

1. Поиск фрагмента дефектограммы, соответствующего ближайшей зоне болтового соединения, слева от искомого фрагмента.
2. Поиск фрагмента дефектограммы, соответствующего ближайшей зоне болтового соединения, справа от искомого фрагмента.
3. Поиск конкретных, найденных на первых двух этапах «левой» и «правой» зон болтового соединения на дефектограмме, соответствующей предыдущему проходу.
4. Уточнение координат искомого фрагмента на дефектограмме предыдущего прохода.

Следует отметить, что, хотя в основу используемого в данной работе подхода положен принцип, предложенный в работе [21], реализация каждого конкретного этапа в нем существенно отличается от реализаций, применяемых в [21], а общая эффективность оказывается существенно более высокой.

3.3 Поиск опорных зон на дефектограмме текущего прохода

Поиск «левой» и «правой» зон болтового соединения предлагается осуществлять в два этапа:

1. Быстрая примерная локализация зоны болтового соединения на основе описанной ниже специально разработанной трехконной модели.
2. Уточнение координат искомой зоны болтового соединения на основе установления оптимальных парных соответствий между элементами фрагмента дефектограммы, найденного на первом этапе, с фрагментом дефектограммы, соответствующим типовой зоне болтового соединения. При этом в качестве критерия поиска оптимальных парных соответствий предлагается использовать специальный критерий локально-глобального выравнивания.

Результатом данного этапа являются фрагменты дефектограмм, соответствующие «левой» и «правой» опорным зонам болтового соединения на дефектограмме текущего прохода, координаты соответствующих рельсовых стыков, а также фрагмент дефектограммы текущего прохода, ограниченный найденными зонами болтовых соединений.

3.4 Поиск опорных зон на дефектограмме предыдущего прохода

Поиск опорных зон на предыдущем проходе предлагается осуществлять аналогично поиску опорных зон на текущем проходе, однако учитывая, что в данном случае требуется найти конкретные зоны болтовых соединений.

В связи с этим на втором этапе осуществляется локально-глобальное выравнивание не с фрагментом дефектограммы, соответствующим типовой зоне болтового соединения, а с конкретными фрагментами, соответствующими опорным «левой» и «правой» зонам болтовых соединений. При этом зона болтового соединения считается найденной, если вычисленное значение критерия оптимальности парного выравнивания фрагментов окажется выше заданного порога.

Кроме того, следует обратить внимание, что процедуру локализации зоны болтового соединения в данном случае достаточно применять не для всей дефектограммы предыдущего прохода, а лишь для некоторой окрестности координат «левого» и «правого» рельсовых стыков, определенных на предыдущем этапе.

3.5 Уточнение координат искомого фрагмента на дефектограмме предыдущего прохода

Уточнение положения искомого фрагмента предлагается осуществлять на основе предположения о том, что на протяжении участка дефектограммы между найденными «левой» и «правой» зонами болтовых соединений скорость движения дефектоскопа сильно не изменялась. В этом случае положение искомой области, содержащей пропуск, может быть определено на основе применения подтвердившегося на практике принципа пропорциональности.

В соответствии с данным принципом координата начала t_{start} и конца t_{end} искомой области относительно начала вырезанного фрагмента текущей дефектограммы, ограниченного опорными болтовыми соединениями (рис. 58), определяется следующим образом:

$$t_{start} = l_{cur} = l_{prev} \frac{L_{cur}}{L_{prev}} ; \quad t_{end} = l_{cur} + \Delta l .$$

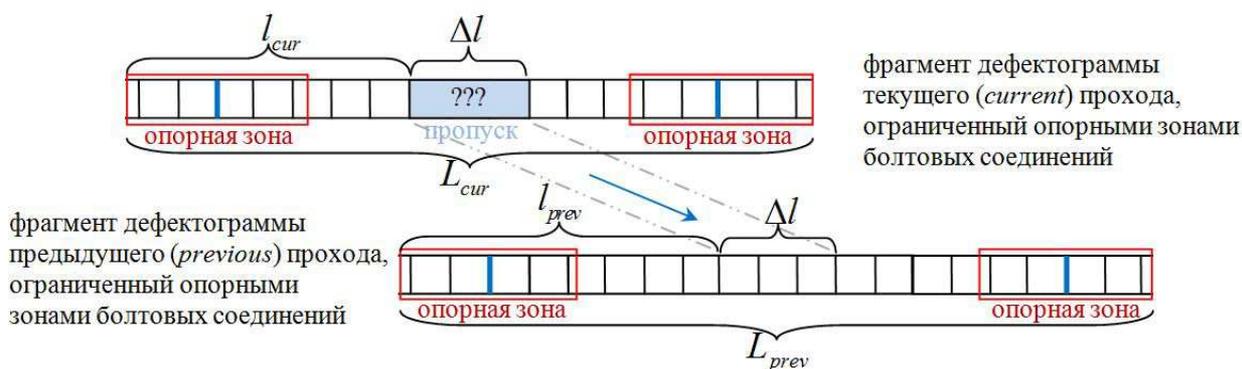


Рис. 4 Иллюстрация принципа пропорциональности

4 Трехконная модель для быстрой примерной локализации зоны болтового соединения

Для примерной локализации зоны болтового соединения в данной работе предлагается подход, основанный на прохождении скользящим окном некоторой ширины W вдоль дефектограммы с шагом ΔW , с применением специально разработанной трехконной модели для быстрого определения среди фрагментов дефектограммы, вырезанных скользящим окном, фрагмента, содержащего зону болтового соединения.

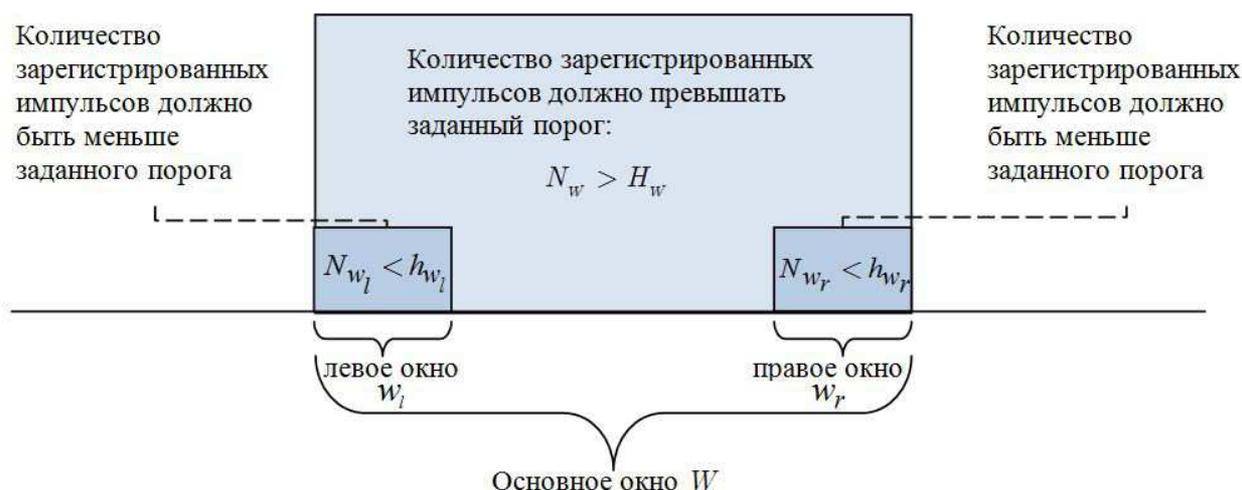


Рис. 5 Трехоконная модель для быстрой примерной локализации зоны болтового соединения

Основная идея использования трехоконной модели заключается в том, чтобы в дополнение к основному окну W использовать два дополнительных окна небольшого размера, расположенных по краям основного окна: левое w_l и правое w_r (рис. 59), а также в принятии решения о наличии зоны болтового соединения в окне W на основе информации о количестве зарегистрированных импульсов.

Предлагается считать, что окно W содержит зону болтового соединения только в том случае, если одновременно выполняются три условия:

- (1) количество импульсов, зарегистрированных в основном окне, больше заданного для него порогового значения: $N_W > H_W$;
- (2) количество импульсов, зарегистрированных в левом окне, меньше заданного для него порогового значения: $N_{w_l} < h_{w_l}$;
- (3) количество импульсов, зарегистрированных в правом окне, меньше заданного для него порогового значения: $N_{w_r} < h_{w_r}$.

В случае если сразу несколько подряд идущих окон подходят под заданную модель болтового соединения, из них может быть оставлено одно — центральное окно.

5 Поиск оптимальных парных соответствий между элементами дефектограмм

5.1 Сравнение элементов дефектограмм

Очевидно, что поиск оптимальных парных соответствий между элементами дефектограмм должен основываться на сравнении составляющих их элементов. Причем от того, насколько удачно будет выбрана мера несходства элементов дефектограмм, во многом зависит и качество результата поиска оптимальных соответствий.

Поскольку каждый элемент дефектограммы по отдельному каналу представляет собой импульсный сигнал и может содержать некоторое количество импульсов (1), то прямое их сравнение оказывается невозможным.

В связи с этим в данной работе используется специальная математическая модель, предложенная авторами в [14], для представления элементов дефектограмм в виде, удобном для их дальнейшего сравнения.

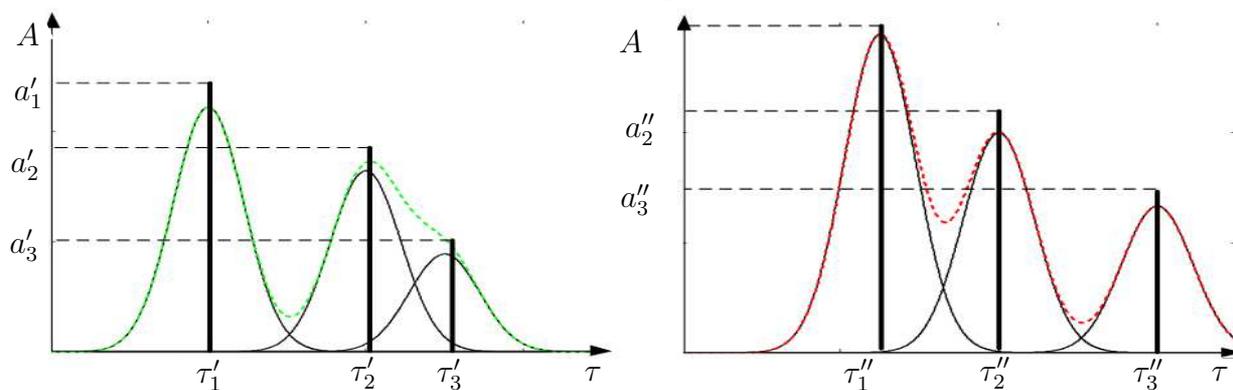


Рис. 6 Графическое представление модели описания элемента дефектограммы по одному каналу

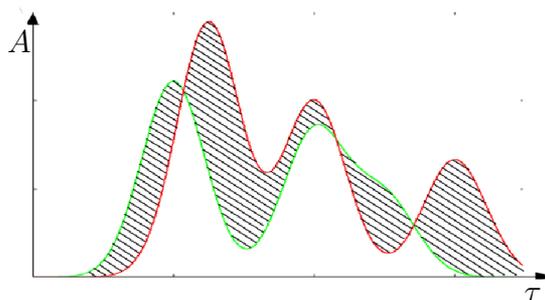


Рис. 7 Графическое представление меры несходства двух элементов дефектограмм

Согласно данной модели, каждый элемент дефектограммы по отдельному каналу $x = \{(\tau_i, a_i), i = 1, \dots, n\}$ описывается взвешенной суммой нормальных распределений с некоторой дисперсией σ и математическим ожиданием, равным задержке соответствующего импульса τ_i :

$$f(\tau|x) = \sum_{i=1}^n a_i \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\tau - \tau_i)^2\right). \tag{2}$$

При этом для точки, в которой не зарегистрировано импульсов, $f(x) = 0$, что не противоречит здравому смыслу.

Пусть $x' = \{(\tau'_i, a'_i) \in R^2, i = 1, \dots, n'\}$ и $x'' = \{(\tau''_i, a''_i) \in R^2, i = 1, \dots, n''\}$ — два элемента дефектограммы. Пример их графического представления согласно выбранной модели представлен на рис. 6.

Меру их несходства вслед за [21] будем вычислять как интеграл от квадрата разности представляющих их функций (2):

$$\tilde{r}(x', x'') = \sqrt{\int_{-\infty}^{\infty} [f(\tau|x') - f(\tau|x'')]^2 d\tau}. \tag{3}$$

Геометрически данная мера несходства может быть интерпретирована как площадь фигуры, ограниченной графиками распределений, описывающих сравниваемые элементы дефектограммы (рис. 7).

Удобство используемого представления заключается в том, что интеграл в (3) может быть вычислен аналитически. В результате мера несходства (3) может быть записана в следующем эквивалентном виде:

$$\begin{aligned} \tilde{r}(x', x'') &= \sqrt{\rho(x', x') + \rho(x'', x'') - 2\rho(x', x'')} ; \\ \rho(x', x'') &= \sum_{i=1}^{n'} \sum_{j=1}^{n''} a'_i a''_j \exp \left[- \left(\frac{\tau'_i - \tau''_j}{2\sigma} \right)^2 \right]. \end{aligned} \quad (4)$$

При этом несходство некоторого элемента дефектограммы $x' = \{(\tau'_i, a'_i) \in R^2, i = 1, \dots, n'\}$ с нулевым элементом $\varphi = (0, 0)$, не содержащим импульсов, может быть вычислено по еще более простой формуле:

$$\tilde{r}(x', \varphi) = \sqrt{\sum_{i=1}^{n'} \sum_{k=1}^{n'} a'_i a'_k \exp \left[- \left(\frac{\tau'_i - \tau'_k}{2\sigma} \right)^2 \right]} .$$

Очевидно, что одинаковые по форме сигналы будут иметь нулевую меру несходства. Наибольшее же значение меры несходства будет наблюдаться при сравнении сигнала с большим количеством высокоамплитудных импульсов и нулевого сигнала.

Совместное использование информации, полученной с разных каналов, в рамках данной работы осуществляется путем вычисления расширенной меры несходства элементов дефектограммы, представляющей собой линейную комбинацию частных мер несходства (4) с некоторыми неотрицательными весами c_i , в сумме составляющими единицу:

$$r(x', x'') = \sum_{i=1}^m c_i \tilde{r}(x'_i, x''_i), \quad \sum_{i=1}^m c_i = 1. \quad (5)$$

5.2 Постановка задачи поиска оптимальных парных соответствий

Пусть имеются два многокомпонентных дискретных сигнала, $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_X})$ и $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_{N_Y})$, представляющих согласно (1) фрагменты дефектограмм и имеющих в общем случае разную длину. Один из них — \mathbf{X} — примем за «базовый», а другой — \mathbf{Y} — за «ссылочный».

Пусть также на множестве элементов этих сигналов определена метрика, например согласно (5).

Требуется найти оптимальные парные соответствия элементов дефектограмм (оптимальное парное выравнивание соответствующих сигналов)

$$\hat{T} = \arg \min_T J(\mathbf{X}, \mathbf{Y}, T),$$

т. е. каждому элементу из базового сигнала сопоставить некоторый элемент из ссылочного, а результат запомнить в виде таблицы ссылок

$$T = (\theta_t, t = 1, \dots, N_X),$$

где $\theta_t \in \{1, N_Y\}$ — абсолютная ссылка (номер элемента в ссылочном сигнале, соответствующий элементу t в базовом).

Применительно к дискретным сигналам такую задачу принято называть задачей парного выравнивания (Dynamic Time Warping в англоязычной терминологии) [28].

5.3 Критерий локально-глобального выравнивания для поиска оптимальных парных соответствий

Очевидно, что решение задачи поиска оптимальных парных соответствий (парного выравнивания) фрагментов дефектограмм будет существенным образом зависеть от того, какой критерий парного выравнивания $J(\mathbf{X}, \mathbf{Y}, T)$ будет выбран для установления оптимальных парных соответствий.

Традиционным критерием парного выравнивания дискретных сигналов является критерий глобального выравнивания [14, 21, 28–30]. Особенностью глобального выравнивания является то, что при выравнивании обязательно ставятся в соответствие друг другу первые и последние элементы сравниваемых сигналов и ни один элемент не остается не связанным.

Однако такой способ хорошо подходит лишь для сравнения близких по длине фрагментов, что в работах [14, 21] обусловило необходимость прохождения вдоль дефектограммы скользящим окном фиксированной длины, равной длине зоны типового болтового соединения с последующим их выравниванием, и, соответственно, повлекло за собой необходимость выполнения большого объема вычислений.

В данной работе для поиска оптимальных парных соответствий между элементами дефектограмм впервые применяется критерий локально-глобального выравнивания, который в отличие от традиционного для сигналов критерия глобального выравнивания требует обязательного сопоставления всех элементов только одного из сравниваемых сигналов, во втором же сигнале произвольное число начальных и конечных элементов разрешается оставить несвязанными без наложения на это какого-либо штрафа:

$$\left. \begin{aligned}
 J(\mathbf{X}, \mathbf{Y}, T) &= r(\mathbf{x}_1, \mathbf{y}_{\theta_1}) + \sum_{t=2}^{N_X} \gamma(\theta_{t-1}, \theta_t); \\
 \gamma(\theta_{t-1}, \theta_t) &= \begin{cases} \beta + r(\mathbf{x}_t, \mathbf{y}_{\theta_t}), & \theta_t = \theta_{t-1}; \\ \beta |\theta_t - \theta_{t-1} - 1| + \sum_{j=\theta_{t-1}+1}^{\theta_t} r(\mathbf{x}_t, \mathbf{y}_j), & \theta_t > \theta_{t-1}; \\ \infty, & \theta_t < \theta_{t-1}, \end{cases}
 \end{aligned} \right\} \quad (6)$$

где β — штраф на непараллельные ссылки, соответствующие локальным сжатиям или растяжениям осей сигналов, представляющих фрагменты дефектограмм, при их выравнивании, а запись $r(\mathbf{x}_1, \mathbf{y}_{\theta_1})$ означает, что первый элемент базовой дефектограммы может

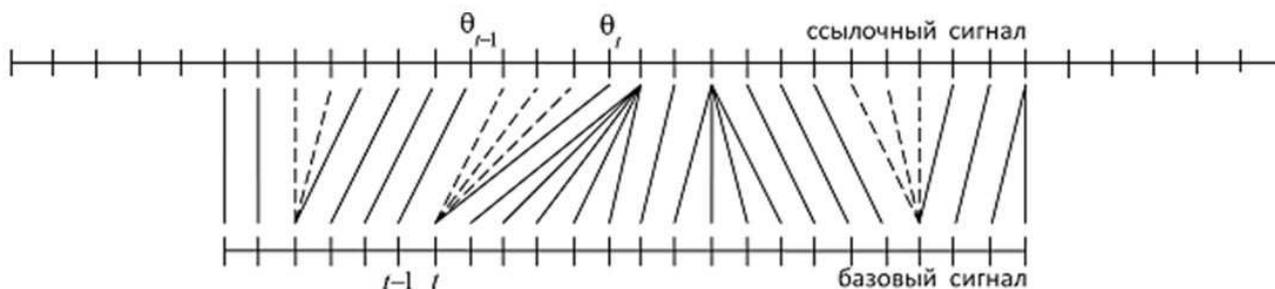


Рис. 8 Пример установления парных соответствий при локально-глобальном выравнивании

соответствовать любому элементу ссылочной дефектограммы. Схематично идея локально-глобального парного выравнивания фрагментов дефектограмм представлена на рис. 8.

Следует обратить внимание, что согласно критерию (6), в случае если $|\theta_t - \theta_{t-1}| > 1$, считается, что элемент t базового сигнала ставится в соответствие сразу нескольким элементам ссылочного сигнала $\theta_{t-1} + 1, \dots, \theta_t$.

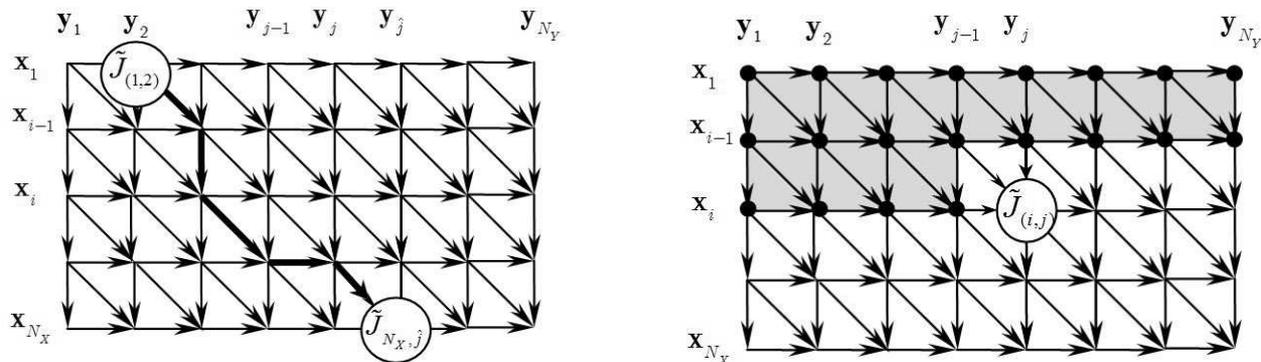
Самой важной особенностью критерия (6), эксплуатируемой в данной работе, является то, что он разрешает оставлять свободными начальные и конечные элементы ссылочного сигнала. Такая особенность делает его более адекватным по сравнению с критерием глобального выравнивания в случаях, когда требуется найти местоположение некоторого заданного небольшого фрагмента сигнала в более длинном, что как раз требуется в данной работе на стадии уточнения координат опорных зон болтового соединения.

5.4 Алгоритм локально-глобального выравнивания фрагментов дефектограмм

Следует обратить внимание, что критерий (6) является сепарабельной функцией, поскольку может быть представлен как сумма более простых функций, зависящих только от двух соседних переменных θ_{t-1} и θ_t . Поэтому его минимум может быть найден при помощи процедуры динамического программирования [31]. Однако более удобным, наглядным и эффективным является представление алгоритма при помощи графа парных соответствий, каждая вершина которого связывает определенный элемент базового сигнала с определенным элементом ссылочного сигнала, представляющих фрагменты дефектограмм.

Заметим, что любое локально-глобальное выравнивание фрагментов дефектограмм может быть представлено в виде пути на данном графе, начинающимся в одной из вершин верхней грани, проходящим последовательно по некоторым ребрам графа через его вершины и заканчивающимся в одной из вершин нижней грани (рис. 9). При этом перемещение по горизонтальному ребру соответствует локальному растяжению оси базового сигнала, перемещение по вертикальному ребру — локальному растяжению оси ссылочного сигнала.

Алгоритм поиска оптимальных парных соответствий при помощи графа парных соответствий заключается в последовательном прохождении по вершинам графа с координатами (i, j) , $i = 1, \dots, N_x, j = 1, \dots, N_y$, начиная с левой верхней и заканчивая правой



(а) Пример локально-глобального выравнивания

(б) Иллюстрация алгоритма поиска оптимального парного выравнивания

Рис. 9 Примеры применения графа парных соответствий для выравнивания дефектограмм

нижней, и вычислении в каждой вершине значения неполного критерия $\tilde{J}_{(i,j)}$:

$$\begin{aligned} \tilde{J}_{(1,j)} &= r(\mathbf{x}_1, \mathbf{y}_j), \quad j = 1, \dots, N_Y; \\ \tilde{J}_{(i,j)} &= r(\mathbf{x}_i, \mathbf{y}_j) + \min \begin{cases} \beta + \tilde{J}_{(i-1,j)}, \\ \tilde{J}_{(i-1,j-1)}, \\ \beta + \tilde{J}_{(i,j-1)}, \end{cases} \quad i = 2, \dots, N_X, \quad j = 1, \dots, N_Y. \end{aligned}$$

В качестве последней вершины оптимального пути выбирается вершина нижней грани, обеспечивающая минимальное значение неполного критерия:

$$\hat{j} = \arg \min_{j=1, \dots, N_Y} \tilde{J}_{(N_X, j)}.$$

Оптимальное значение критерия равно значению в этой вершине:

$$J(T, \mathbf{X}, \mathbf{Y}) = \tilde{J}_{(N_X, \hat{j})}.$$

Для восстановления траектории оптимального выравнивания, т.е. непосредственно таблицы соответствий $T = (\theta_t, t = 1, \dots, N_X)$, $\theta_t \in \{1, \dots, N_Y\}$ элементов сравниваемых фрагментов дефектограмм, на обратном ходе, последовательно перемещаясь в обратном направлении, начиная с вершины ($t = N_X, \theta_{N_X} = \hat{j}$), определяется оптимальное направление обратного движения — индекс вершины, движение из которой в данную вершину обеспечивает минимум неполного критерия.

В целом данный алгоритм очень похож на алгоритм локально-глобального выравнивания символьных последовательностей [32], но имеет и некоторые принципиальные отличия, связанные с тем, что элементы символьных последовательностей в отличие от элементов дефектограмм принадлежат некоторому конечному алфавиту, а естественным способом деформации при выравнивании является удаление, вставка и замена символов, а не сжатие и растяжение, как в случае сигналов, представляющих дефектограммы.

6 Экспериментальное исследование

6.1 Выбор параметров и исследование применимости трехконной модели для локализации зоны болтового соединения

Выбор параметров предложенной трехконной модели осуществлялся на основе анализа дефектограмм, включающих 20 зон болтовых соединений, а также участки между ними.

В результате было выявлено, что наиболее подходящим ультразвуковым каналом для применения трехконной модели является 6-й ультразвуковой канал наклонного (под углом 42°) ввода ультразвуковых колебаний двухниточного дефектоскопа АВИКОН-11. Также были выбраны следующие параметры модели: ширина основного окна $W = 2000$; ширина левого и правого дополнительных окон $w_l = w_r = w = 500$; пороговое количество ультразвуковых импульсов в основном окне $N_W = 160$; пороговое количество импульсов в левом и правом окнах $N_{w_l} = N_{w_r} = N_w = 50$; шаг скольжения основного окна $\Delta W = 50$.

На рис. 10 представлен пример зависимостей количества зарегистрированных импульсов по 6-му каналу в левом (зеленый), основном (синий) и правом (красный) окнах от положения основного окна для выбранных значений параметров. Пример демонстрирует, что при выбранных значениях параметров трехконная модель позволяет уверенно выделить фрагменты дефектограмм, соответствующие зонам болтовых соединений.

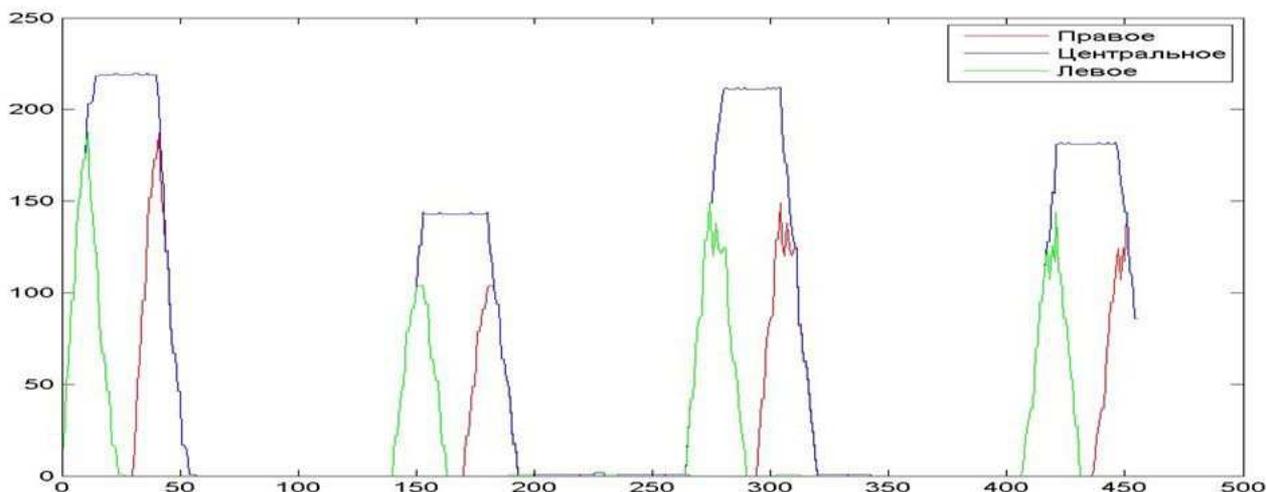


Рис. 10 График зависимости количества зарегистрированных импульсов от положения центрального окна для выбранных значений параметров

Проверка адекватности построенной модели осуществлялась на 60 фрагментах дефектограмм, не включающих фрагменты, использовавшиеся на этапе выбора параметров. Каждый из 60 анализируемых фрагментов дефектограмм содержит две зоны болтового соединения и три участка рельсового пути вне зоны болтового соединения. Таким образом, всего было проанализировано 120 зон болтовых соединений и 180 неболтовых соединений, что соответствует рельсовому пути общей протяженностью около 9 км.

Корректность работы разработанной процедуры локализации зон болтовых соединений оценивалась путем сравнения результатов ее работы с экспертной оценкой.

В результате все зоны болтовых соединений, размеченные экспертом, были локализованы верно — каждая зона болтового соединения была вырезана единственным окном, причем не было ни одного случая обрезания зоны болтового соединения (т. е. выбора фрагмента, дефектограммы, содержащего лишь ее часть), а также не было зафиксировано ни одного случая ложного срабатывания (т. е. выделения фрагмента, не содержащего зону болтового соединения).

6.2 Экспериментальное исследование точности выделения зон болтовых соединений

Целью данной части исследования являлась оценка точности выделения зон болтовых соединений на дефектограмме при помощи применения предложенной в данной работе трехконной модели совместно с процедурой локально-глобального выравнивания фрагментов дефектограмм.

На данном этапе экспериментального исследования использовался участок дефектограммы, содержащий 28 зон болтовых соединений, на котором экспертами были отмечены центры, соответствующие непосредственно рельсовым стыкам. Кроме того, дополнительно был выбран эталонный фрагмент дефектограммы, содержащий типовую зону болтового соединения, на котором также экспертом был отмечен рельсовый стык.

Согласно предложенному подходу сначала осуществлялась примерная локализация каждой из 28 зон болтовых соединений при помощи разработанной трехконной модели с параметрами, выбранными в предыдущем разделе.

Таблица 1 Ошибка определения координат рельсового стыка (см)

| Канал | min | max | mean | std | median |
|-----------|-------|--------|--------|--------|--------|
| 5 | 0,000 | 4,2000 | 1,6929 | 0,9622 | 1,7000 |
| 6 | 0,000 | 29,600 | 3,9857 | 7,0077 | 1,2000 |
| 7 | 0,000 | 83,000 | 10,414 | 21,474 | 1,3000 |
| 5 + 6 + 7 | 0,000 | 5,2000 | 1,2357 | 1,3065 | 0,600 |

Таблица 2 Количество и процент зон болтовых соединений с заданной ошибкой определения координаты рельсового стыка

| Канал | Ошибка сегментации (отклонение от отметки эксперта) | | | |
|-----------|---|-------------|-------------|-------------|
| | 0 см | < 1 см | < 3 см | < 6 см |
| 5 | 0 (0%) | 8 (28,57%) | 26 (92,85%) | 28 (100%) |
| 6 | 2 (7,14%) | 14 (50%) | 22 (78,57%) | 28 (100%) |
| 7 | 1 (3,57%) | 13 (46,43%) | 20 (71,43%) | 22 (78,57%) |
| 5 + 6 + 7 | 1 (3,57%) | 16 (57,14%) | 26 (92,85%) | 28 (100%) |

В результате применения процедуры локализации зон болтовых соединений было получено 28 фрагментов дефектограмм длиной 2000 отсчетов (4 м), каждый из которых содержит одну зону болтового соединения.

Далее производилось уточнение положения каждой зоны болтового соединения при помощи процедуры локально-глобального выравнивания каждого из найденных фрагментов с выбранным эталонным фрагментом, соответствующим типовому болтовому соединению.

Центр каждой найденной зоны болтового соединения (положение рельсового стыка) определялся как отсчет, поставленный в результате парного выравнивания в соответствие отмеченному экспертом рельсовому стыку на эталонном фрагменте.

Описанные вычисления производились 4 раза: по трем различным ультразвуковым каналам отдельно (5-, 6- и 7-му каналам) и по всем каналам одновременно. Для каждого из четырех указанных случаев были вычислены отклонения найденных координат рельсовых стыков для каждого из 28 зон болтовых соединений от координат стыков, отмеченных экспертом. Статистики данных отклонений приведены в табл. 1.

В табл. 2 приведены количества и доли зон болтовых соединений, найденные рельсовые стыки которых отклоняются от отмеченных экспертами на значение из указанного диапазона.

Как видно из табл. 1 и 2, использование информации, полученной по разным каналам, позволяет улучшить точность определения координат рельсового стыка. В результате, предложенный подход позволил достаточно точно произвести выделение зон болтовых соединений с ошибкой, не превышающей 6 см, причем в 92,85% случаев ошибка не превышает 3 см, что является достаточной точностью при анализе рельсового пути.

6.3 Экспериментальное исследование алгоритма восстановления пропущенных данных

Поскольку задача восстановления пропущенных данных сформулирована в данной работе как задача поиска на дефектограмме предыдущего прохода области, соответствующей пропуску данных на дефектограмме текущего прохода, то возможность успешного

Таблица 3 Ошибки определения координат начала и конца искомой области на предыдущем проходе (см)

| Проход | Положение пропуска относительно зоны болтового соединения | | | | | |
|---|---|-----------------|-----------------|---------------|--------------------------|----------------|
| | Вне зоны | | В середине | | В левой или правой части | |
| | Начало | Конец | Начало | Конец | Начало | Конец |
| 1 | 2,6 | 1,6 | 2,2 | 3,4 | 3,2 | 4,4 |
| 2 | 1,8 | 2,8 | 2,6 | 3,8 | 3,4 | 3,8 |
| 3 | 3,8 | 4,2 | 3,0 | 4,2 | 2,8 | 4,2 |
| 4 | 2,2 | 2,8 | 3,4 | 1,8 | 3,6 | 4,4 |
| 5 | 1,2 | 0,4 | 3,4 | 3,0 | 1,8 | 0,8 |
| 6 | 3,0 | 3,8 | 3,8 | 4,2 | 3,2 | 4,4 |
| 7 | 3,8 | 4,2 | 3,6 | 1,6 | 3,8 | 4,2 |
| 8 | 6,4 | 5,2 | 4,6 | 2,2 | 3,2 | 4,4 |
| 9 | 3,8 | 4,6 | 5,0 | 3,8 | 3,8 | 4,2 |
| 10 | 1,8 | 4,2 | 3,4 | 9,0 | 5,6 | 6,2 |
| Среднее | $3,04 \pm 1,5$ | $3,38 \pm 1,48$ | $3,49 \pm 0,83$ | $3,7 \pm 2,1$ | $3,44 \pm 0,95$ | $4,1 \pm 1,33$ |
| Среднее по всем экспериментам: $3,53 \pm 1,4$ | | | | | | |

восстановления данных (при условии отсутствия пропуска на дефектограмме предыдущего прохода) определяется точностью определения координат интересующей области на дефектограмме предыдущего прохода. В связи с этим в данной части исследования был проведен эксперимент по определению на дефектограмме предыдущего прохода координат области, соответствующей пропуску, зная его координаты на дефектограмме с текущего прохода.

Для этого отдельно рассматривались три случая локализации пропуска данных:

- 1) пропуск данных находится вне зоны болтового соединения;
- 2) пропуск данных находится в начале или конце зоны болтового соединения;
- 3) пропуск данных находится в середине зоны болтового соединения.

Для каждого типа локализации пропуска данных было рассмотрено 10 различных фрагментов дефектограмм.

Для определения координат участка дефектограммы предыдущего прохода, соответствующего области пропущенных данным дефектограммы текущего прохода, применялся предложенный в данной работе подход, описанный в разд. 3.

В табл. 3 приведены полученные результаты.

Как видно из табл. 3, предложенный подход позволяет определить координаты начала и конца области, соответствующей пропущенным данным на дефектограмме предыдущего прохода с точностью около 3,5 см, при этом точность почти не зависит от того, попадает пропуск на зону болтового соединения или располагается вне ее. Однако точность определения положения пропуска вне зоны в среднем немного выше по сравнению с точностью определения положения пропуска, расположенного в зоне болтового соединения. Предположительной причиной этого является то, что в данном случае вырезаемый в соответствии с предложенным подходом фрагмент дефектограммы, ограниченный опорными зонами болтовых соединений, имеет двойную длину, поскольку зона болтового соединения, содержащая пропуск, не может являться опорной и находится внутри вырезаемого фрагмента. Используемый принцип пропорциональности при этом дает менее точные ре-

зультаты, поскольку на больших участках пути выше вероятность изменения скорости движения дефектоскопа.

Кроме того, следует отметить, что точность определения координаты конца интересующей нас зоны в среднем несколько ниже точности определения координаты начала соответствующей зоны. Предположительной причиной этого является то, что данная координата определяется исходя из предположения, что участок, соответствующий пропуску на дефектограмме предыдущего прохода, имеет ту же длину, что и пропуск на текущем проходе. Однако если скорость движения дефектоскопа была неодинаковой на первом и втором проходах, то и длина соответствующей области тоже может измениться, поэтому, возможно, применение принципа пропорциональности и для определения конца области может дать более точные результаты.

В целом следует отметить, что достигнутая точность является адекватной рассматриваемой прикладной задаче, что позволяет использовать предлагаемый подход для автоматического восстановления пропущенных данных на дефектограмме.

7 Заключение

В рамках данной работы предложен принцип восстановления пропущенных данных на ультразвуковой дефектограмме текущего прохода на основе анализа дефектограммы, полученной при предыдущем проходе дефектоскопом по тому же участку рельсового пути, а именно: на основе поиска опорных зон болтовых соединений на дефектограммах текущего и затем предыдущего проходов.

Данный подход является развитием предложенного ранее [14], делая его более точным и существенно более эффективным с вычислительной точки зрения.

Повышение эффективности стало возможным благодаря применению разработанной трехконной модели зоны болтового соединения и идентификации с ее помощью зоны болтового соединения не по результатам парного выравнивания, а на основе анализа количества зарегистрированных импульсов.

Предложенная трехконная модель позволяет осуществить быстрое, уверенное, но достаточно грубое определение местоположения зоны болтового соединения. Для уточнения координат зоны болтового соединения в рамках данной работы предложена локально-глобальная процедура парного выравнивания фрагментов дефектограмм, существенно отличающаяся от традиционной для сигналов процедуры глобального выравнивания тем, что ориентирована именно на поиск местоположения короткого фрагмента в длинном, а не просто на вычисление несходства сигналов. Хотя в данном случае тоже применяется процедура выравнивания фрагментов дефектограмм, следует отметить, что она применяется лишь однократно для каждой зоны болтового соединения, что обеспечивает существенно большую эффективность по сравнению с [21].

Еще большего повышения эффективности можно достичь за счет распространения принципа быстрой оптимизации критерия глобального выравнивания [33, 34] на процедуру локального выравнивания, а также за счет привлечения средств параллельной обработки данных. В частности, существенного повышения производительности можно достичь, осуществив параллельное вычисление несходства элементов дефектограмм. Кроме того, отдельные этапы предложенного подхода также обеспечивают естественную возможность применения параллельных вычислений, например одновременно может выполняться поиск левой и правой опорных зон болтовых соединений.

Достигнутая точность определения координат искомой области на дефектограмме предыдущего прохода ($3,5 \pm 1,4$) является адекватной рассматриваемой прикладной за-

даче, что позволяет использовать предлагаемый подход для автоматического восстановления пропущенных данных на ультразвуковой дефектограмме рельсового пути.

Литература

- [1] *Jimenez-Redondo N., Bosso N., Zeni L., Minardo A., Shubert F., Heinicke F., Simrothhubert A.* Automated and cost effective maintenance for railway (ACEM-Rail) // *Procedia Social Behavioral Sci.*, 2012. Vol. 48. P. 1058–1067.
- [2] *Марков А. А., Шпагин Д. А.* Ультразвуковая дефектоскопия рельсов. — 2-е изд. перераб. и доп. — СПб: Образование — Культура, 2013. 284 с.
- [3] *Jemec V., Grum J.* Automated non-destructive testing and measurement systems for rails // 10th European Conference on Non-Destructive Testing. — Moscow, 2010. www.ndt.net/article/ecndt2010/reports/1_10_42.pdf.
- [4] *Марков А. А., Козьяков А. Б., Кузнецова Е. А., Шпагин Д. А.* Утраченные и новые технологии контроля рельсов // *Путь и путевое хозяйство*, 2013. № 8. С. 2–9.
- [5] *Марков А. А., Кузнецова Е. А.* Дефектоскопия рельсов. Формирование и анализ сигналов. Кн. 2. Расшифровка дефектограмм. — СПб: УльтраПринт, 2014. 332 с.
- [6] *Шилов М. Н.* Методическое, алгоритмическое и программное обеспечение регистрации и анализа дефектограмм при ультразвуковом контроле рельсов. Дисс. ... канд. техн. наук. СПб, 2007.
- [7] *Heckel T., Thomas H., Kreuzbruck M., Ruhe S.* High speed non-destructive rail testing with advanced ultrasound and eddy-current testing techniques // *Indian National Seminar and Exhibition on Non-Destructive Evaluation — NDE*, 2009.
- [8] *Марков А. А.* Современное состояние, проблемы и направления развития диагностики рельсов и элементов инфраструктуры // *Путь и путевое хозяйство*, 2015. № 4. С. 19–24.
- [9] *Марков А. А., Гараева В. С.* Об акустическом контакте в зоне болтовых стыков // *Путь и путевое хозяйство*, 2008. № 12. С. 15–17.
- [10] *Марков А. А., Козьяков А. Б., Кузнецова Е. А.* Расшифровка дефектограмм ультразвукового контроля рельсов. — СПб, 2006. 206 с.
- [11] *Бирюзов В. П.* Об основных направлениях развития систем диагностики и мониторинга путевого хозяйства до 2025 г. // *Путь и путевое хозяйство*, 2015. № 4. С. 4–9.
- [12] *Jarmulak J., Kerchoffs E.* Case-based reasoning for interpretation of data from non-destructive testing // *Eng. Appl. Artificial Intell.*, 2001. Vol. 14. No. 4. P. 401–417.
- [13] *Papaelias M., Roberts C., Davis C.* A review on nondestructive evaluation of rails: State-of-the-art and future development // *Proc. IMechE, Part F, J. Rail Rapid Transit*, 2008. Vol. 222. P. 367–384.
- [14] *Маленичев А. А., Красоткина О. В., Сулимова В. В., Моттль В. В., Марков А. А.* Применение процедуры парного выравнивания для разметки стыков на ультразвуковой дефектограмме рельсового пути // *Известия ТулГУ. Сер. Технические науки*, 2013. Вып. 9. Ч. 1. С. 115–128.
- [15] *Федоренко Д. В.* Проблемы автоматизации расшифровки сигналов многоканального ультразвукового контроля рельсов // *Радиоэлектронные комплексы многоцелевого назначения: сб. научн. тр. Юбилейный выпуск. 1991–2011.* — СПб.: Изд-во Политехн. ун-та, 2011. С. 117–120.
- [16] *Марков А. А., Антипов А. Г.* Магнитодинамический метод контроля рельсов // *В мире НК*, 2012. № 3(57). С. 66–71.

- [17] Марков А. А., Кузнецова Е. А., Антипов А. Г., Вережкин А. Ю. Способ диагностики рельсового пути. Патент на изобретение № 2521095. Оpubл. 27.06.2014.
- [18] Candes J., Romberg J., Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information // *IEEE Trans. Inform. Theory*, 2006. No. 52. P. 489–509.
- [19] Candes J., Tao T. Near-optimal signal recovery from random projections: Universal encoding strategies? // *IEEE Trans. Inform. Theory*, 2006. Vol. 52. No. 12. P. 5406–5425.
- [20] Cai J.-F., Shen Z., Ye G.-B. Approximation of frame based missing data recovery // *Appl. Comput. Harmonic Anal.*, 2011. Vol. 31. № 2. P. 185–204.
- [21] Malenichev A., Sulimova V., Krasotkina O., Markov A., Mottl V. Automatic analysis of ultrasonic railway defectograms with uncontrolled regions // *Machine learning and data mining in pattern recognition. — Lecture notes in computer science ser. — Springer*, 2014. Vol. 8556. P. 315–327.
- [22] Wang D., Lu L., Zhang H. J. Speech segmentation without speech recognition // *Conference (International) on Multimedia and Expo (ICME '03)*, 2003. Vol. 1. P. 405–408.
- [23] Kehagias A., Nidelkou E., Petridis V. A dynamic programming segmentation procedure for hydrological and environmental time series // *Stochastic Environmental Res. Risk Assessment*, 2006. P. 77–94.
- [24] Lavielle M. Optimal segmentation of random processes // *IEEE Trans. Signal Proc.*, 1998. Vol. 46. No. 5. C. 365–373.
- [25] Fearnhead P. Exact Bayesian curve fitting and signal segmentation // *IEEE Trans. Signal Proc.*, 2005. Vol. 53. No. 6. P. 2160–2166. doi: 10.1109/TSP.2005.847844.
- [26] Sasan M., Bayan S. A nonlinear variational method for signal segmentation and reconstruction using level set algorithm // *Signal Proc.*, 2006. Vol. 86. Iss. 11. P. 3496–3504.
- [27] Mico P., Mora M., Cuesta-Frau D., Aboy M. Automatic segmentation of long-term ECG signals corrupted with broadband noise based on sample entropy // *Computer Methods Programs Biomedicine*, 2010. Vol. 98. Iss. 2. P. 118–129.
- [28] Martens R., Claesen L. On-line signature verification by dynamic time-warping // *13th IEEE Conference (International) on Pattern Recognition Proceedings*, 1996. Vol. 1. P. 38–42.
- [29] Вербовой В. В. Метрики для сравнения звуковых сигналов с учетом особенностей человеческого уха // *Компьютерная графика и мультимедиа*, 2005. № 3. С. 2–10.
- [30] Wang, X., Mueen A., Ding H., Trajcevski G., Scheuermann P., Keogh E. Experimental comparison of representation methods and distance measures for time series data // *Data Mining Knowledge Discovery*, 2013. Vol. 26. No. 2. P. 275–309.
- [31] Беллман Р., Калаба Р. Динамическое программирование и современная теория управления. — М.: Наука, 1969. 118 с.
- [32] Jones N. C., Pevzner P. A. An introduction to bioinformatics algorithms. The MIT Press, 2004. 435 p.
- [33] Salvador S., Chan P. FastDTW: Toward accurate dynamic time warping in linear time and space // *KDD Workshop on Mining Temporal and Sequential Data*, 2004. P. 70–80.
- [34] Al-Naymat G., Chawla S., Taheri J. SparseDTW: A novel approach to speed up dynamic time warping // *The 2009 Australasian Data Mining*, 2009. Vol. 101. P. 117–127.

Поступила в редакцию 14.06.2015

References

- [1] Jimenez-Redondo, N., N. Bosso, L. Zeni, A. Minardo, F. Shubert, F. Heinicke, and A. Simrothhubert. 2012. Automated and cost effective maintenance for railway (ACEM-Rail). *Procedia Social Behavioral Sci.* 48:1058–1067.

- [2] Markov, A. A., and D. A. Shpagin. 2013. *Ul'trazvukovaya defektoskopiya rel'sov*. 2nd ed. SPb: Obrazovanie — Kul'tura. 284 p. (In Russian.)
- [3] Jemec, V., and J. Grum. 2010. Automated non-destructive testing and measurement systems for rails. *10th European Conference on Non-Destructive Testing*. Moscow. Available at: www.ndt.net/article/ecndt2010/reports/1_10_42.pdf.
- [4] Markov, A. A., A. B. Koz'jakov, E. A. Kuznecova, and D. A. Shpagin. 2013. Utrachennye i novye tekhnologii kontrolya rel'sov. *Put' i putevoe hozyaystvo* 8:2–9. (In Russian.)
- [5] Markov, A. A., and E. A. Kuznecova. 2014. *Defektoskopiya rel'sov. Formirovanie i analiz signalov. Book 2. Rasshifrovka defektogramm*. SPb: Ul'traPrint. 332 p. (In Russian.)
- [6] Shilov, M. N. 2007. Metodicheskoe, algoritmicheskoe i programmnoe obespechenie registratsii i analiza defektogramm pri ul'trazvukovom kontrole rel'sov. PhD Diss. SPb. (In Russian.)
- [7] Heckel, T., H. Thomas, M. Kreutzbruck, and S. Ruhe. 2009. High speed non-destructive rail testing with advanced ultrasound and eddy-current testing techniques. *Indian National Seminar and Exhibition on Non-Destructive Evaluation — NDE*.
- [8] Markov, A. A. 2015. Sovremennoe sostoyanie, problemy i napravleniya razvitiya diagnostiki rel'sov i elementov infrastruktury. *Put' i putevoe hozyaystvo* 4:19–24. (In Russian.)
- [9] Markov, A. A., and V. S. Garaeva. 2008. Ob akusticheskom kontakte v zone boltovykh stykov. *Put' i putevoe hozyaystvo* 12:15–17. (In Russian.)
- [10] Markov, A. A., A. B. Koz'jakov, and E. A. Kuznecova. 2006. *Rasshifrovka defektogramm ul'trazvukovogo kontrolya rel'sov*. SPb. 206 p. (In Russian.)
- [11] Birjuzov, V. P. 2015. Ob osnovnykh napravleniyakh razvitiya sistem diagnostiki i monitoringa putevogo hozyaystva do 2025 g. *Put' i putevoe hozyaystvo* 4:4–9. (In Russian.)
- [12] Jarmulak, J., and E. Kerchoffs. 2001. Case-based reasoning for interpretation of data from non-destructive testing. *Eng. Appl. Artificial Intell.* 14(4):401–417.
- [13] Papaalias, M., C. Roberts, and C. Davis. 2008. A review on nondestructive evaluation of rails: State-of-the-art and future development. *Proc. IMechE, Part F, J. Rail Rapid Transit* 222:367–384.
- [14] Malenichev, A. A., O. V. Krasotkina, V. V. Sulimova, V. V. Mottl', and A. A. Markov. 2013. Primenenie procedury parnogo vyravnivaniya dlya razmetki stykov na ul'trazvukovoy defektogramme rel'sovogo puti. *Izvestiya TulGU, ser. Tehnicheskie nauki* 9(1):115–128. (In Russian.)
- [15] Fedorenko, D. V. 2011. Problemy avtomatizatsii rasshifrovki signalov mnogokanal'nogo ul'trazvukovogo kontrolya rel'sov. *Radioelektronnye kompleksy mnogotselevogo naznacheniya: Sb. nauchn. tr. Jubileyny vypusk. 1991—2011*. SPb.: Izd-vo Politehn. un-ta. 117–120. (In Russian.)
- [16] Markov, A. A., and A. G. Antipov. 2012. Magnitodinamicheskii metod kontrolya rel'sov. *V mire NK* 3(57):66–71. (In Russian.)
- [17] Markov, A. A., E. A. Kuznecova, A. G. Antipov, and A. Yu. Verevkin. 2014. Sposob diagnostiki rel'sovogo puti. Patent na izobretenie No. 2521095. Opubl. 27.06.2014. (In Russian.)
- [18] Candes, J., J. Romberg, and T. Tao. 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory* 52:489–509.
- [19] Candes, J., and T. Tao. 2006. Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inform. Theory* 52(12):5406–5425.
- [20] Cai, J.-F., Z. Shen, and G.-B. Ye. 2011. Approximation of frame based missing data recovery. *Appl. Comput. Harmonic Anal.* 31(2):185–204.

- [21] Malenichev, A., V. Sulimova, O. Krasotkina, A. Markov, and V. Mottl. 2014. Automatic analysis of ultrasonic railway defectograms with uncontrolled regions. *Machine learning and data mining in pattern recognition*. Lecture notes in computer science ser. Springer. 8556:315–327.
- [22] Wang, D., L. Lu, and H. J. Zhang. 2003. Speech segmentation without speech recognition. *Conference (International) on Multimedia and Expo (ICME '03)*. 1:405–408.
- [23] Kehagias, A., E. Nidelkou, and V. Petridis. 2006. A dynamic programming segmentation procedure for hydrological and environmental time series // *Stochastic Environmental Res. Risk Assessment* 77–94.
- [24] Lavielle, M. 1998. Optimal segmentation of random processes. *IEEE Trans. Signal Proc.* 46(5):365–373.
- [25] Fearnhead, P. Exact Bayesian curve fitting and signal segmentation. 2005. *IEEE Trans. Signal Proc.* 53(6):2160–2166. doi: 10.1109/TSP.2005.847844.
- [26] Sasan, M., and S. Bayan. 2006. A nonlinear variational method for signal segmentation and reconstruction using level set algorithm. *Signal Proc.* 86(11):3496–3504.
- [27] Mico, P., M. Mora, D. Cuesta-Frau, and M. Aboy. 2010. Automatic segmentation of long-term ECG signals corrupted with broadband noise based on sample entropy. *Computer Methods Programs Biomedicine* 98(2):118–129.
- [28] Martens, R., and L. Claesen. 1996. On-line signature verification by dynamic time-warping. *13th IEEE Conference (International) on Pattern Recognition Proceedings*. 38–42.
- [29] Verbovoj, V. V. 2005. *Metriki dlya sravneniya zvukovykh signalov s uchetom osobennostey chelovecheskogo uha. Komp'yuternaya grafika i mul'timedia* 3:2–10. (In Russian.)
- [30] Wang, X., A. Mueen, H. Ding, G. Trajcevski, P. Scheuermann, and E. Keogh. 2013. Experimental comparison of representation methods and distance measures for time series data. *Data Mining Knowledge Discovery* 26(2):275–309.
- [31] Bellman, R., and R. Kalaba. 1969. *Dinamicheskoe programmirovaniye i sovremennaya teoriya upravleniya*. Moscow: Nauka. 118 p. (In Russian.)
- [32] Jones, N. C., and P. A. Pevzner. 2004. *An introduction to bioinformatics algorithms*. The MIT Press. 435 p.
- [33] Salvador, S., and P. Chan. 2004. FastDTW: Toward accurate dynamic time warping in linear time and space. *KDD Workshop on Mining Temporal and Sequential Data*. 70–80.
- [34] Al-Naymat, G., S. Chawla, and J. Taheri. 2009. SparseDTW: A novel approach to speed up dynamic time warping. *The 2009 Australasian Data Mining* 101:117–127.

Received June 14, 2015

Автоматизация метода компьютерной окулографии для исследований центральной нервной системы на основе пассивного анализа видеоизображения

*Е. А. Новиков¹, И. А. Ваколюк², Р. Д. Ахапкин¹, И. А. Варчак¹,
И. Г. Шалагинова², Д. А. Швайко², Е. А. Буденкова²*

eugen@novikov.de

¹Институт прикладной математики и информационных технологий БФУ им. И. Канта, Калининград, Россия

²Химико-биологический институт БФУ им. И. Канта, Калининград, Россия

Данные, описывающие изменения положения центров зрачков во времени, называются окулограммами. Анализ окулограмм позволяет судить о функциональном состоянии определенных отделов головного мозга, вовлеченных в процесс программирования и регуляции движений глаз. Данная статья рассматривает новый вариант автоматизации процесса регистрации и последующего анализа произвольных и спровоцированных движений глаз, что в общем случае называется компьютерной окулографией. В основном, системы компьютерной окулографии опираются на активное инфракрасное сканирование глаз при жесткой фиксации головы, что является дорогостоящим и не доступным для простых пользователей. Однако благодаря развитию технологий регистрации цифровых изображений и общему росту вычислительной мощности персональных компьютеров и портативных устройств методы пассивного сканирования изображений начинают набирать популярность. Предлагаемый в данной статье метод рассчитан на использование вместе со стандартными цифровыми камерами. Метод может быть применен для получения и анализа качественных окулограмм на основе видеоизображения, полученного с частотой не менее 30 кадр/с.

Ключевые слова: *окулография; саккады; eyetracking; диагностика; ЦНС*

DOI: 10.21469/22233792.1.12.09

Automation method of computer oculography for research of the central nervous system based on passive video analysis

*E. A. Novikov¹, I. A. Vakoliuk², R. D. Akhapiin¹, I. A. Varchak¹,
I. G. Shalanginova², D. A. Shvaiko², and E. A. Budenkova²*

¹Immanuel Kant Baltic Federal University, Institute of Applied Mathematics and Information Technologies, 14 A. Nevskogo st., Kalinigrad, Russia

²Institute of Chemistry and Biology, 14 A. Nevskogo st., Kalinigrad, Russia

Data of changing pupil center position in time are called oculogram. The oculogram allows to define functional condition of the brain divisions involved in the programming process and regulation of eye movements. This article describes new version of registration automation for process and subsequent analysis of arbitrary and provoked eye movements that in general is called computer oculography. Basically, computer oculography relies on an active infrared eye tracking by the rigid fixation of the head. Such methods are quite costly and unpopular decision. However, due to development of digital images registration technologies and general increase in computing power of personal computers and portable devices, the methods of passive scanning images start to gain popularity. The proposed method relies on standard digital

cameras. This method can be applied to analyze qualitative oculogram based on video images taken from at least 30 frame/s.

Keywords: *oculography; saccade; eyetracking; diagnosis; CNS*

DOI: 10.21469/22233792.1.12.09

1 Введение

Возможность познать тайны мозга через движения глаз привела научную мысль к необходимости регистрации глазодвигательных реакций. Начиная с XIX в. были разработаны и применялись на практике различные методы записи движений глаз: механический метод, метод электроокулографии, инфракрасный метод и, наконец, метод анализа видеоизображения глаза.

Известные на данный момент методы могут полностью удовлетворить разнообразные потребности исследователей [1].

Из-за своей очевидной простоты в использовании, сравнительно легкой настройки и опоре на быстро развивающуюся индустрию оптических и электронных устройств обработки изображений методы компьютерной окулографии, основанные на видеоанализе, становятся наиболее популярными.

В данной статье рассматривается наиболее дешевый в эксплуатации метод, подразумевающий использование стандартного и широкодоступного оборудования от встроенных web-камер переносных гаджетов до подключаемых к персональному компьютеру высокоскоростных устройств получения видеоизображения человеческого лица. Данный метод основан на видеоанализе изображения (среднего — 1280×720 точек и низкого разрешения — 640×480 точек) лица пациента и нацелен на эффективный и точный поиск положения центров зрачков. Глаза человека постоянно совершают мелкие движения, взгляд перескакивает с объекта на объект. Так мозг сканирует окружающее пространство, создавая постоянно обновляющуюся карту мира. Эти быстрые движения глаз, переводящие взор на новый объект интереса, называются саккадами. Саккадическая система, объединяющая в себе различные структуры мозга, участвующие в контроле саккадических движений глаз, работает без постоянного осознанного контроля [2].

Предлагаемый метод представляет собой многоуровневую схему обработки и анализа изображения с целью построения графиков движений глаз, называемых окулограммами. Видеорегистрация глазодвигательных реакций осуществляется параллельно с демонстрацией на экране заранее подготовленной видеопрезентации — последовательности стимулов — точек, появляющихся в переменном темпе и в различных местах.

2 Постановка задачи

2.1 Синдром дефицита внимания и гиперактивности

Предлагаемый метод изначально был задуман как инструмент для изучения и диагностики синдрома дефицита внимания и гиперактивности (СДВГ).

Диагностирование СДВГ основано на сравнении имеющейся симптоматической картины с диагностическими критериями МКБ-10 (международной классификации болезней 10-го пересмотра) и заключается в сборе анамнеза, опросе и тестировании родителей [3, 4]. Диагноз СДВГ ставит невролог, опираясь не на физиологически обусловленные методы, а только лишь на опыт и наблюдение.

Существующие на сегодняшний день методы коррекции (биологическая обратная связь, комплексная поведенческая терапия) позволяют нивелировать проявления этого расстройства на ранней стадии до минимума, что возвращает человеку возможность полноценной жизнедеятельности [5].

Многолетние работы зарубежных и отечественных исследователей [6–10] позволили выявить функциональные связи между определенными параметрами саккад, когнитивными функциями и теми или иными структурами мозга.

Разрабатываемый авторами метод основан на регистрации окулограмм и анализе нарушений в выполнении глазодвигательных реакций, определяемых через вычисление таких значений, как величина латентного периода саккадических реакций, точность выполнения саккад, наличие и степень выраженности саккадических движений. Это позволяет оценить функциональное состояние систем мозга, обеспечивающих реализацию когнитивных функций, страдающих при СДВГ. Клинические исследования, проводимые сотрудниками лаборатории физиологии человека и регуляции когнитивных функций БФУ им. И. Канта, подтверждают возможность применения окулографического тестирования в качестве физиологически обоснованного элемента диагностики психопатологий [11–13].

Накопление базы глазодвигательных профилей по разным отклонениям позволит осуществлять диагностику состояния центральной нервной системы (ЦНС) путем сравнения паттернов окулограмм.

Таким образом, с помощью видеоокулографии существует возможность исследовать функциональное состояние мозга человека без непосредственного физического контакта (рис. 1 и 2). Интеграция видеоокулографической методики с современными математиче-

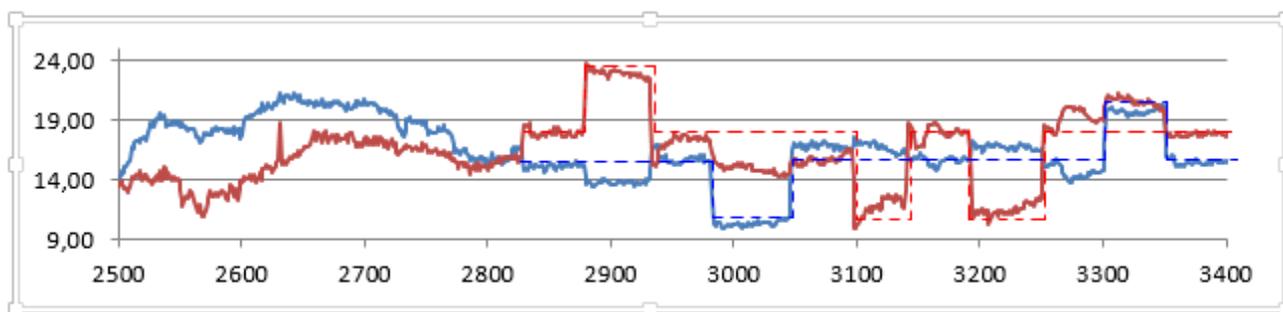


Рис. 1 Видеоокулограмма ребенка без СДВГ, возраст 6 лет: горизонтальная ось — временная разверстка (цена деления 20 мс), вертикальная ось — амплитуда движений глаз (пиксель), два графика описывают перемещения центра одного глаза по осям X и Y в координатах видеоизображения

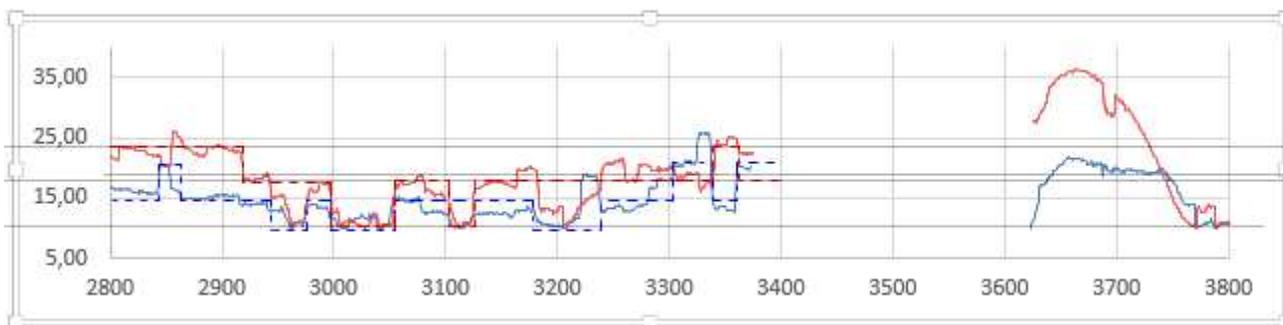


Рис. 2 Видеоокулограмма ребенка с СДВГ, возраст 6 лет: горизонтальная ось — временная разверстка (цена деления 20 мс), вертикальная ось — амплитуда движений глаз (пиксель)

скими методами обработки видеопотока и анализа данных предоставляет удобный инструмент для диагностики ряда нейропатологий, например СДВГ.

3 Решение

В практической реализации обработка и анализ видеозаписи вызванных движений глаз состоит из двух основных этапов — это тест-презентация стимульного материала и анализ материалов видеofиксации.

3.1 Построение тест-презентации

Для решения задачи регистрации глазодвигательных реакций испытуемого при отслеживании предъявляемых зрительных стимулов авторы разработали систему проектирования видеопрезентаций с псевдослучайным возникновением зрительных стимулов — презентационных точек.

Видеопрезентация состоит из появления, движения (если задан соответствующий параметр) и исчезновения презентационных точек на квадратной области экрана.

Ограничивая количество расположений презентационных точек до пяти (S, W, E, N, C), получаем конечное число возможных комбинаций появлений и исчезновений. Таким образом, не только упрощается создание видеопрезентации, но и уменьшается количество базовых движений глаз до 25, что позволит быстро и объективно сравнивать данные разных исследований, проведенных различными специалистами.

Для проведения испытаний специалист может использовать одну из включенных в систему тест-презентаций или создать собственную.

Одновременно с предъявлением стимульного материала той же машиной в параллельном потоке производится видеозапись лица испытуемого.

3.2 Предварительная подготовка изображения

Чтобы нивелировать влияние внешних факторов на результаты испытаний, необходимо провести некоторую подготовку видеозаписи лица испытуемого, полученной в ходе просмотра им тест-презентации.

Основная задача, которая стоит перед исследователями, — это нахождение области лица, глаз и центров зрачков. На этапе предварительной обработки необходимо сделать эти области как можно более четко различимыми, а само изображение как можно более «чистым».

В общем случае основной проблемой видеопотока, снятого обычной камерой, является наличие на изображении «шума», связанного как с внутренними факторами — высокой температурой светочувствительной матрицы, так и с внешними — плохой, неравномерной освещенностью объекта. Поэтому первое, что нужно сделать, — это избавиться от шума. В рассматриваемом случае лучше всего работает Гауссовское размытие.

На втором этапе проводят коррекцию освещенности объекта на экране методом наложения «пустого» изображения (рис. 3). «Пустое» изображение строится путем расчета линейной регрессии яркости пикселей по обеим осям. Метод линейной регрессии является наиболее быстрым для применения в реальном времени. В ходе собственных экспериментов выяснилось, что его применение улучшает результат поиска зрачков от 5% до 15% в зависимости от уровня неравномерности освещения лица на изображении. Описание меры точности находится в п. 3.4.

Таким образом авторы пытаются добиться как можно более высокого качества изображения и наиболее отчетливого выделения необходимых областей без потери точности передачи информации.

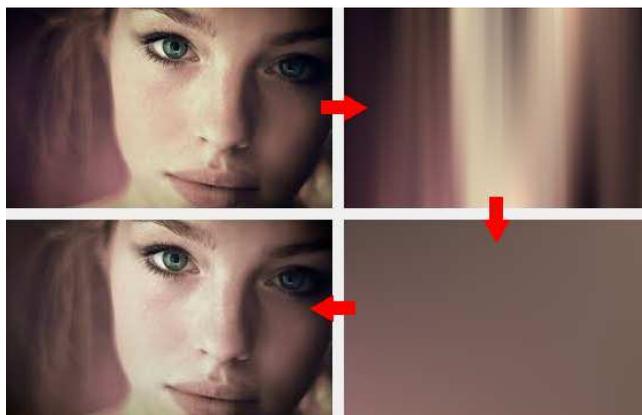


Рис. 3 Коррекция освещенности путем наложения «пустого» изображения

3.3 Определение положения лица

После подготовки изображения можно приступить к поиску положения лица в кадре. Для этого используем метод, основанный на известных каскадах классификаторов Хаара [14].

Каскад Хаара представляет собой некоторый набор примитивов, для каждого из которых считается его свертка с изображением. Пол Виола и Майкл Джонс при создании своего метода Виолы–Джонса усовершенствовали и адаптировали идею использования вейвлетов Хаара в своих алгоритмах.

На данном этапе разработки системы используем стандартную выборку классификаторов. Однако в будущем необходимо будет создать собственную библиотеку с учетом реальных условий, в которых будет эксплуатироваться представленная система. При тщательном подходе к созданию собственной выборки классификаторов можно добиться высокой эффективности определения положения лица — до 95% — при частоте ложных срабатываний около 0,001.

3.4 Определение центров зрачков

После того как определено положение лица в кадре, можно приступить к поиску самого зрачка испытуемого. Но прежде нужно выделить области его глаз (рис. 4). Для этого используем простой эмпирический подход, допуская, что местоположение глаз относительно центра лица у всех людей примерно одинаково.

Далее в качестве изображения будем использовать область правого или левого глаза. Рассмотрим изображение глаза как двумерную матрицу чисел, где каждое число соответствует значению яркости одного из элементов изображения (пиксела). Известно, что гео-

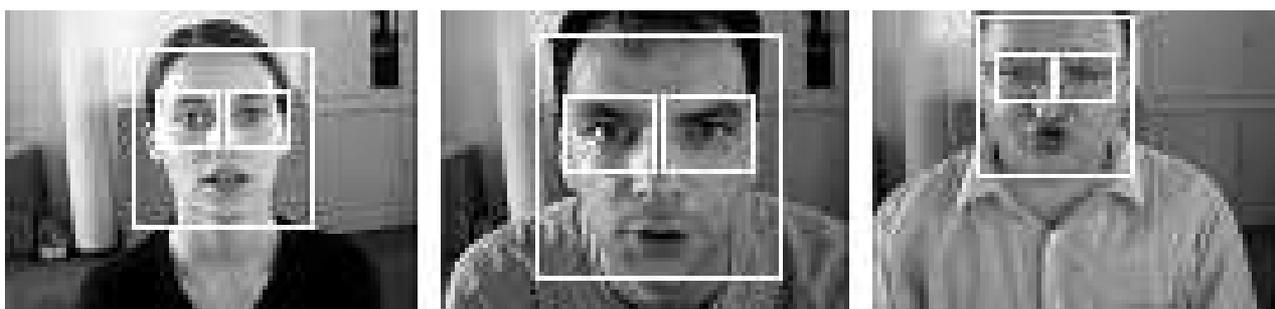


Рис. 4 Эмпирическое определение области глаз

Таблица 1 Сравнение эффективности существующих методов определения положения глаз на примере общедоступной базы образцов лиц BioID (BioID Face Database)

| Метод | $e \leq 0,05$ | $e \leq 0,10$ | $e \leq 0,15$ | $e \leq 0,20$ | $e \leq 0,25$ |
|------------------------|---------------|---------------|---------------|---------------|---------------|
| Asadifard, 2010 | 47,0% | 86,0% | 89,0% | 93,0% | 96,0% |
| Valenti и Gevers, 2008 | 72,2% | 82,1% | 86,2% | 93,8% | 96,4% |
| Turkan, 2007 | 18,6% | 73,7% | 94,2% | 98,7% | 99,6% |
| Campadelli, 2006 | 62,0% | 85,2% | 87,6% | 91,6% | 96,1% |
| Niu, 2006 | 75,0% | 93,0% | 95,8% | 96,4% | 97,0% |
| Asteriadis, 2006 | 44,0% | 81,7% | 92,6% | 96,0% | 97,4% |
| Hamouz, 2005 | 58,6% | 75,0% | 80,8% | 87,6% | 91,0% |
| Jesorsky, 2001 | 38,0% | 78,8% | 84,7% | 87,2% | 91,8% |
| Метод авторов | 81,0% | 92,5% | 94,7% | 96,6% | 98,0% |

метрический центр круглых объектов может быть определен через анализ векторного поля градиентов матрицы изображения. Применение данного принципа было предложено для определения центров зрачков с учетом факта значительного контраста, возникающего на переходе между склерой и зрачком глаза [15]. Таким образом используется ориентация каждого вектора градиента, чтобы нарисовать линии через все изображение, точки пересечения этих линий подсчитываются и аккумулируются в счетчиках, связанных с соответствующими точками изображения. Счетчик, набравший максимальное количество пересечений векторов, соответствует центру глаза на изображении.

На практике для получения градиентного изображения можно использовать фильтр Собеля. В каждой точке изображения результатом оператора Собеля является либо соответствующий вектор градиента, либо норма этого вектора.

В качестве меры точности метода определения центров глаз оценивают нормализованную ошибку, которая указывает на максимальную из двух погрешностей, полученных для обоих глаз. Эта мера была введена Есорским [16] и определяется как

$$e \leq \frac{1}{d} \max \{e_l, e_r\},$$

где e_l , e_r — дистанции между найденными и корректными центрами левого и правого глаз. При анализе качества метода поиска глаз данная мера должна иметь следующие характеристики:

- (1) $e \leq 0,25 \approx$ расстояние между центром и боковыми краями глаза;
- (2) $e \leq 0,10 \approx$ внешний диаметр радужки;
- (3) $e \leq 0,05 \approx$ диаметр зрачка.

Таким образом, метод поиска глаз должен обеспечивать качественный результат на уровне $e \leq 0,05$. Ошибка на уровне 0,25 и более может привести к отклонению найденного зрачка за пределами реального глаза, что недопустимо для качественного определения центров глаз. Поэтому при сравнении с существующими методами будем ориентироваться на результаты с ошибкой $e \ll 0,25$ (табл. 1).

Данный алгоритм хорошо подходит для определения положения зрачка в отдельном кадре, однако если речь идет об обработке видеопотока, то такое решение является достаточно затратным и недостаточно стабильным, поэтому для улучшения результативности используем описанный метод в комбинации с дополнительным методом покадрового отслеживания перемещений указанного объекта. Суть использования метода состоит

в том, что после получения устойчивой позиции центра зрачка в одном кадре (накопленная ошибка между предыдущими 5 кадрами $e_{\text{com}} \leq 0,25$) все последующие кадры будут проанализированы на предмет перемещения заданного объекта (зрачка) авторским методом Новикова–Падалко, что значительно повышает стабильность работы айтрекера и увеличивает его быстродействие. Авторский метод Новикова–Падалко разработан для отслеживания любого заданного объекта в видеопотоке. Данный метод реализует новый подход к задаче распознавания образов и основывается на применении комбинации классических методов прямого преобразования Радона к матрице изображения, одномерного преобразования Фурье к полученным интегральным проекциям и статистического анализа интегральных коэффициентов Фурье, рассматриваемых в качестве основных дескрипторов объектов изображения.

3.5 Построение и анализ окулограмм

Для построения окулограммы необходимо секвенировать полученную видеозапись и для каждого кадра выполнить следующий алгоритм:

- (1) найти зрачок левого глаза (см. п. 3.4);
- (2) найти зрачок правого глаза (аналогично);
- (3) определить точку отсчета движения глаз на изображении — в данном случае центральной точкой отсчета можно считать середину детектированного прямоугольника лица (рис. 5);

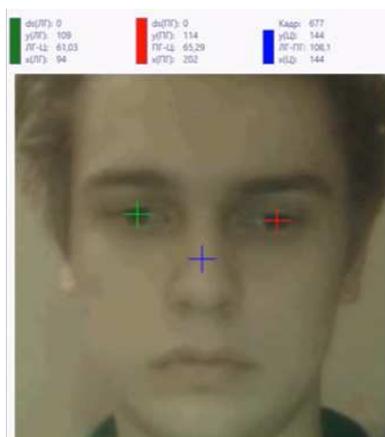


Рис. 5 Найденные зрачки и центр кадра

- (4) рассчитать расстояние между центральной статичной точкой и центром левого глаза;
- (5) рассчитать расстояние между центральной статичной точкой и центром правого глаза.

После выполнения данного алгоритма для каждого кадра получаем окулограмму (см. рис. 1), описывающую горизонтальные и вертикальные перемещения центра зрачка испытуемого.

Последующий объективный анализ окулограммы является завершающим этапом. Специалист (исследователь или медик) сможет выявить заболевание, связанное с нарушением функций ЦНС, по следующим ключевым параметрам: латентный период саккады, дисметричная саккада, гиперметричная саккада, гипометричная саккада, коррекционная саккада, латентный период коррекционной саккады, мультисаккада, дрейф. В настоящий момент уже собраны некоторые экспериментальные данные, подтверждающие эту гипотезу [11, 12]. Основной целью разработки описываемого метода является системное накоп-

ление базы статистических данных результатов соответствующих исследований, подтверждающих концепцию прямой зависимости глазодвигательных реакций человека от состояния его ЦНС.

3.6 Перспективы

Данный метод анализа полученных окулограмм имеет хорошие перспективы. Видеокамеры, широко распространенные и используемые в современных ноутбуках, планшетах и телефонах, дают частоту генерации около 30 кадр/с. Данного количества достаточно для построения окулограмм и объективного статистического анализа. Но при частоте ≥ 50 кадр/с открывается большой спектр дополнительных возможностей. Например, становится доступна регистрация микродвижений глаз.

Различный перечень характеристик глазодвигательных реакций зависит и от ситуационного психического состояния человека. Алкоголь, психотропные и наркотические вещества, усталость по-разному отражаются на характеристиках саккадической активности, что позволяет идентифицировать текущее состояние ЦНС исследуемого пациента [8].

4 Заключение

Представленный метод автоматического вычисления и анализа окулограмм предназначен для исследования состояния ЦНС. Метод опирается на пассивный видеоанализ и нацелен на массовое внедрение и повсеместное использование как на уровне профессионального медицинского и научно-исследовательского инструмента, так и на уровне повседневной проверки состояния человека.

Литература

- [1] *Kenneth H., Marcus N., Richard A., Richard D., Halszka J., Joost W.* Eye tracking: A comprehensive guide to methods and measures. — Oxford: Oxford University Press, 2011. 560 p.
- [2] *Schall J. D., Purcell B. A., Heitz R. P., Logan G. D., Palmeri T. J.* Neural mechanisms of saccade target selection: Gated accumulator model of visual-motor cascade // *Eur. J. Neurosci.*, 2011. Vol. 33. P. 1991–2002.
- [3] *Заваденко Н. Н.* Синдром дефицита внимания с гиперактивностью: как помочь ребенку // *Школа здоровья*, 2007. № 3. С. 16–23.
- [4] *Чутко Л. С., Юрьева Р. Г., Сурушжина С. Ю., Нижишена И. С., Яковенко И. С., Анисимова Т. И., Айтбеков К. А.* Принципы оказания помощи детям с синдромом дефицита внимания с гиперактивностью // *Журнал неврологии и психиатрии*, 2008. № 11. С. 76–79.
- [5] *Заваденко Н. Н.* Синдром дефицита внимания и гиперактивности: современные принципы диагностики и лечения // *Вопросы современной педиатрии*, 2014. Т. 13. № 4. С. 54–62.
- [6] *Ярбус А. Л.* Роль движений глаз в процессе зрения. — М.: Наука, 1965. 173 с.
- [7] *Just M. A., Carpenter P. A.* A theory of reading: From fixations to comprehension // *Psychological Rev.*, 1980. No. 87. P. 329–354.
- [8] *Шульговский В. В.* Психофизиология пространственного зрительного внимания у человека // *Соросовский образовательный журнал*, 2004. Т. 8. № 1. С. 17–23.
- [9] *Барабанчиков В. А., Белопольский В. И.* Стабильность видимого мира. — М.: Институт психологии РАН, 2008. 303 с.
- [10] *Hakvoort Schwerdtfeger R. M., Alahyane N., Brien D. C., Coe B. C., Stroman P. W., Munoz D. P.* Preparatory network is impaired in adults with Attention-Deficit/Hyperactivity Disorder during saccade tasks // *NeuroImage Clinical*, 2013. No. 2. P. 63–78.

- [11] Ваколюк И. А., Швайко Д. А., Иванова А. И., Голубицкий В. В. Опыт применения метода видеоокулографии для количественной оценки глазодвигательных реакций у людей, страдающих от алкогольной зависимости // 2-я Междунар. научн.-практич. конф. «Высокие технологии, фундаментальные и прикладные исследования в физиологии и медицине», 2011. Т. 2. С. 7–9.
- [12] Буденкова Е. А. Опыт применения видеоокулографии (ВОГ) при синдроме дефицита внимания с гиперактивностью (СДВГ) у детей младшего дошкольного возраста // Региональная научн.-практич. конф. «Перспективы развития Калининградской области» БФУ им. И. Канта, 2012. С. 53–57.
- [13] Шалагинова И. Г., Ваколюк И. А. Абсолютные и относительные показатели глазодвигательных реакций у пациентов с тревожными расстройствами // 6-я Междунар. конф. по когнитивной науке: Тезисы докладов, 2014. 17 с.
- [14] Viola P., Jones M. Robust real-time face detection // CVPR, 2004. Vol. 2. No. 57. P. 137–154.
- [15] Kothari R., Mitchell J. Detection of eye locations in unconstrained visual images // IEEE ICIP, 1996. Vol. 3. P. 519–522.
- [16] Jesorsky O., Kirchberg K., Frischholz R. 2001. Robust face detection using the Hausdorff distance // Audio- and video-based biometric person authentication / Eds. J. Bigun, F. Smeraldi. — Lecture notes in computer science ser. — Berlin–Heidelberg: Springer. Vol. 2091. P. 90–95.

Поступила в редакцию 24.07.2015

References

- [1] Kenneth, H., N. Marcus, A. Richard, D. Richard, J. Halszka, and W. Joost. 2011. *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press, 2011. 560 p.
- [2] Schall, J. D., B. A. Purcell, R. P. Heitz, G. D. Logan, and T. J. Palmeri. 2011. Neural mechanisms of saccade target selection: Gated accumulator model of visual-motor cascade. *Eur. J. Neurosci.* 33:1991–2002.
- [3] Zavadenko N. N. 2007. Sindrom defitsita vnimaniya s giperaktivnost'yu: Kak pomoch' rebenku [Attention-Deficit/Hyperactivity Disorder: How to help a child]. *Shkola Zdorov'ya* [Health School]. 3:16–23. (In Russian.)
- [4] Chutko, L. S., R. G. Yur'eva, S. Yu. Surushkina, I. S. Nikishena, I. S. Yakovenko, T. I. Anisimova, and K. A. Aytbekov. 2008. Printsipy okazaniya pomoshchi detyam s sindromom defitsita vnimaniya s giperaktivnost'yu [Principles to help children with Attention-Deficit/Hyperactivity Disorder]. *Zhurnal Nevrologii i Psikhiatrii* [J. Neurology and Psychiatry] 11:76–79. (In Russian.)
- [5] Zavadenko, N. N. 2014. Sindrom defitsita vnimaniya i giperaktivnosti: Sovremennyye printsipy diagnostiki i lecheniya [Attention-Deficit/Hyperactivity Disorder: Modern principles of diagnosis and treatment]. *Voprosy Sovremennoy Pediatrii* [Contemporary Pediatrics] 13(4):54–62. (In Russian.)
- [6] Yarbus, A. L. 1965. *Rol' dvizheniy glaz v protsesse zreniya* [The role of eye movements in vision]. Moscow: Nauka, 1965. 173 p. (In Russian.)
- [7] Just, M. A., and P. A. Carpenter. 1980. A theory of reading: From fixations to comprehension. *Psychological Rev.* 87:329–354.
- [8] Shul'govskiy, V. V. 2004. Psikhofiziologiya prostranstvennogo zritel'nogo vnimaniya u cheloveka [Psychophysiology spatial visual attention in humans]. *Sorosovskiy Obrazovatel'nyy Zhurnal* [Soros Educational J.] 8(1):17–23. (In Russian.)
- [9] Barabanshchikov, V. A., and V. I. Belopol'skiy. 2008. *Stabil'nost' vidimogo mira* [The stability of the visible world]. Moscow: Psychology Institute of RAS. 303 p. (In Russian.)

- [10] Hakvoort Schwerdtfeger, R. M., N. Alahyane, D. C. Brien, B. C. Coe, P. W. Stroman, and D. P. Munoz. 2013. Preparatory network is impaired in adults with Attention-Deficit/Hyperactivity Disorder during saccade tasks. *NeuroImage Clinical* 2:63–78.
- [11] Vakolyuk, I. A., D. A. Shvayko, A. I. Ivanova, and V. V. Golubitskiy. 2011. Opyt primeneniya metoda videookulografii dlya kolichestvennoy otsenki glazodvigatel'nykh reaktsiy u lyudey, stradayushchikh ot alkogol'noy zavisimosti [Experience of applying the method to quantify videooculography oculomotor reactions in people suffering from alcohol addiction]. *2nd Scientifically-Practical Conference (International) "The High Technologies Fundamental and Applied Researches in Physiology and Medicine" Proceedings*. St. Petersburg. 2:7–9. (In Russian.)
- [12] Budenkova, E. A. 2012. Opyt primeneniya videookulografii (VOG) pri sindrome defitsita vnimaniya s giperaktivnost'yu (SDVG) u detey mladshego doshkol'nogo vozrasta [Experience with videooculography (VOG) with Attention-Deficit/Hyperactivity Disorder (ADHD) in children of preschool age]. *Regional'naya nauchn.-praktich. konf. «Perspektivy razvitiya Kaliningradskoy oblasti»* [Regional Scientific and Practical Conference "Perspectives of Kalinigrad Region Development"]. Kaliningrad: Immanuel Kant Baltic Federal Univesity. 53–57. (In Russian.)
- [13] Shalaginova, I. G., and I. A. Vakolyuk. 2014. Absolyutnye i otnositel'nye pokazateli glazodvigatel'nykh reaktsiy u patsientov s trevozhnymi rasstroystvami [Absolute and relative indicators of oculomotor reactions in patients with anxiety disorders]. *6th Conference (International) on Cognitive Science Abstracts*. 17 p. (In Russian.)
- [14] Viola, P., and M. Jones. 2004. Robust real-time face detection. *CVPR* 2(5):137–154.
- [15] Kothari, R., and J. Mitchell. 1996. Detection of eye locations in unconstrained visual images. *IEEE ICIP* 3:519–522.
- [16] Jesorsky, O., K. Kirchberg, and R. Frischholz. 2001. Robust face detection using the Hausdorff distance. *Audio- and video-based biometric person authentication*. Eds. J. Bigun and F. Smeraldi. Lecture notes in computer science ser. Berlin–Heidelberg: Springer. 2001:90–95.

Received July 24, 2015

Метод быстрого сжатия изображений без потерь

Е. П. Петров, Н. Л. Харина, П. Н. Сухих

eppetrov@mail.ru

ФГБОУ ВО «Вятский государственный университет», Киров, Россия

Предложен метод сжатия цифровых изображений, отличающийся простотой реализации и отсутствием вычислительных операций на этапе предсказания. Метод включает процедуру предсказания пикселей изображения на основе теории условных марковских процессов и процедуру кодирования. Представлены результаты апробации алгоритма на реальных изображениях различных классов.

Ключевые слова: *цепь Маркова; цифровое изображение; метод сжатия; алгоритм предсказания*

DOI: 10.21469/22233792.1.12.10

Fast lossless image compression method

E. P. Petrov, N. L. Kharina, and P. N. Sukhikh

Vyatka State University, 36 Moskovskaya st., Kirov, Russia

A suggested digital image compression method is characterized by simplicity of implementation and the lack of computing operations at the stage of prediction.

Background: Miniature space watch facilities (small satellites) cannot provide a continuous data transmission because of the severely constrained requirements for energy resources usage efficiency that produces the necessity of new energy efficient low-cost digital images compression methods, which would not be inferior to the known multidigit digital images compression methods with high resolution but surpass them.

Methods: The algorithm consists of the following procedures, splitting the digital images into binary images, predicting of each element of binary images being based on the theory of the conditional Markov processes with discrete states, and coding using any known algorithm (here, Huffman method is used).

Results: To prove the efficiency of the proposed method, the compression of the Earth surface space pictures (group A) and photos (group B) is done. In each groupm, there were 50 one-type images. The known lossless compression algorithms such as PNG, JPEG-LS, JPEG 2000, BMF, Qlic, and ImageZero, are used as analogs. The obtained research results indicate that the proposed method performance and compression ratio comparable with analogues.

Concluding Remarks: The suggested method has the following advantages: the capability of simultaneous binary images processing, the capability of digital images processing with digit capacity, and the lack of computational operations.

Keywords: *Markov chain; digital image; compression method; prediction algorithm*

DOI: 10.21469/22233792.1.12.10

1 Введение

В настоящее время использование цифровых изображений в различных сферах деятельности, таких как медицина, геодезия, картография, системы мониторинга и т.п.,

привело к необходимости использования различных методов сжатия изображений. Существующие методы сжатия, такие как JPEG-LS, JPEG2000, успешно справляются с этой задачей [1, 2]. Однако существует ряд приложений, где использование этих методов не всегда возможно из-за ограничений на вычислительные и энергетические ресурсы, например системы дистанционного зондирования Земли, формирующие на борту космического аппарата огромные объемы информационных потоков, значительная часть которых связана с передачей изображений. В настоящее время для кодирования изображений используются различные алгоритмы, базирующиеся на следующих основных методах: ВТС (block truncation coding — усеченное блочное кодирование), ДИКМ (дифференциальная импульсно-кодовая модуляция), DCT (discrete cosine transform — дискретное косинусное преобразование), DWT (discrete wavelet transform — дискретное вейвлет-преобразование) [1, 2].

Необходимость сжатия на борту космического аппарата изображений обусловлена ростом генерируемых современными космическими системами дистанционного зондирования информационных потоков (до нескольких сотен мегабит в секунду). Это связано с увеличением как пространственного, так и спектрального разрешения съемочной аппаратуры при сохранении широкой полосы обзора, сравнительно невысокой пропускной способности радиоканалов и ограниченности частотного диапазона при передаче изображений на наземные приемные станции. В большинстве современных систем дистанционного зондирования используется непрерывный режим непосредственной передачи или хранения изображений на бортовых записывающих устройствах. Большинство современных спутников обладают энергетическими возможностями для обеспечения непрерывного режима передачи информации, что требует больших вычислительных ресурсов на борту космического аппарата.

Наметившаяся в последнее время тенденция на «миниатюризацию» космических средств наблюдения (малые спутники) и увеличение разрядности передаваемых цифровых изображений привела к необходимости сокращения вычислительных ресурсов на реализацию алгоритмов сжатия изображений, что не позволяет реализовать вышеперечисленные методы. Таким образом, для малых космических аппаратов необходимы новые методы сжатия изображений, не уступающие по своей эффективности известным методам, но требующие меньших вычислительных ресурсов.

В данной работе предлагается один из возможных методов сжатия цифровых изображений (как панхроматических, так и цветных), удовлетворяющий указанным требованиям. Предложенный метод позволяет максимально удалить статистическую избыточность изображения, не имеет спектральных преобразований и вычислительных операций (таких, как умножение, сложение). Метод строится на процедуре предсказания на основе теории условных дискретно-значных марковских процессов. Сжатие осуществляется по битовым плоскостям, что позволяет обрабатывать их параллельно независимо друг от друга, при этом отсутствует ограничение на разрядность изображения.

2 Описание метода

Будем полагать, что g -разрядное изображение размером $m \times n$ пикселей является двумерным марковским процессом с несколькими состояниями $N = 2^g$ с вектором вероятностей начальных состояний $P = \|p_1, p_2, \dots, p_N\|^T$ и матрицами вероятностей переходов (МВП) ${}^1\Pi$ и ${}^2\Pi$ из M_i в соседнее M_j состояние ($i, j \in N$) по горизонтали и вертикали соответственно [3]:

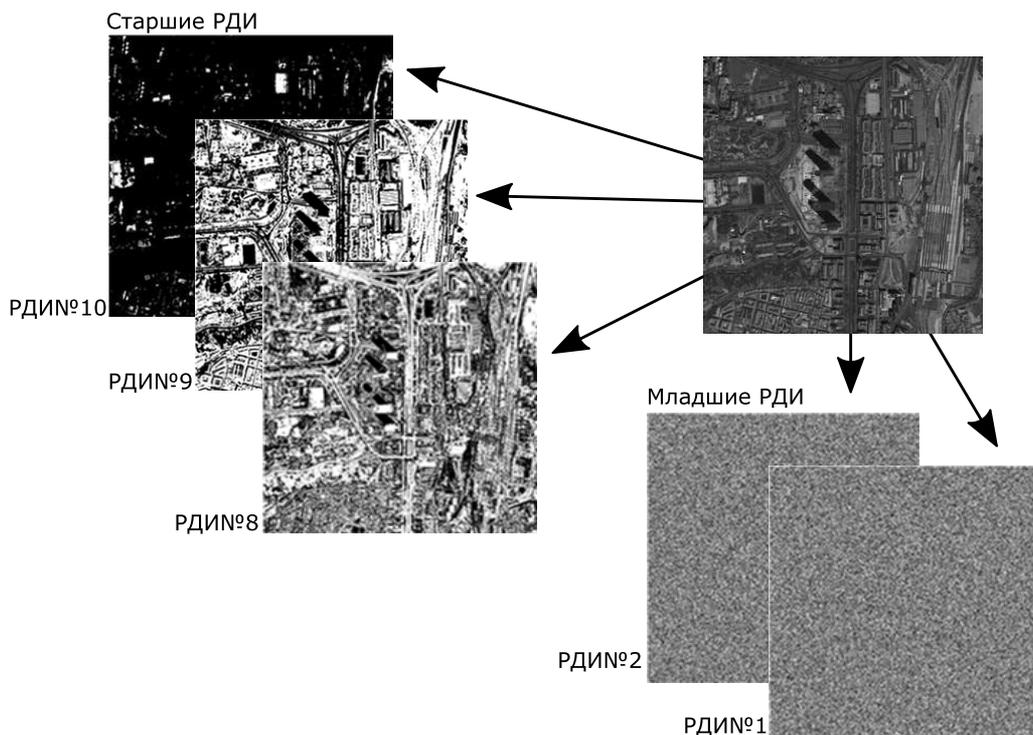


Рис. 1 Разделение панхроматического изображения на РДИ

$${}^1\Pi = \begin{vmatrix} {}^1\pi_{11} & {}^1\pi_{12} & \cdots & {}^1\pi_{1n} \\ {}^1\pi_{21} & {}^1\pi_{22} & \cdots & {}^1\pi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ {}^1\pi_{n1} & {}^1\pi_{n2} & \cdots & {}^1\pi_{nn} \end{vmatrix}; \quad {}^2\Pi = \begin{vmatrix} {}^2\pi_{11} & {}^2\pi_{12} & \cdots & {}^2\pi_{1n} \\ {}^2\pi_{21} & {}^2\pi_{22} & \cdots & {}^2\pi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ {}^2\pi_{n1} & {}^2\pi_{n2} & \cdots & {}^2\pi_{nn} \end{vmatrix}.$$

Разработка метода сжатия изображений в таком представлении требует больших вычислительных ресурсов, особенно для снимков с большим разрешением. Поэтому предлагается разделить g -разрядное цифровое изображение на g разрядных двоичных изображений (РДИ) (рис. 1).

Каждое РДИ представляет собой двумерный марковский процесс с двумя ($N = 2$) равновероятными состояниями ($p_1 = p_2$) и МВП ${}^1\Pi$ и ${}^2\Pi$ из M_i в соседнее M_j состояние ($i, j \in N$) по горизонтали и вертикали соответственно [3]:

$${}^1\Pi = \begin{vmatrix} {}^1\pi_{11} & {}^1\pi_{12} \\ {}^1\pi_{21} & {}^1\pi_{22} \end{vmatrix}; \quad {}^2\Pi = \begin{vmatrix} {}^2\pi_{11} & {}^2\pi_{12} \\ {}^2\pi_{21} & {}^2\pi_{22} \end{vmatrix}. \tag{1}$$

Элементы МВП (1) удовлетворяют условию нормировки

$$\sum_{j=1}^N {}^q\pi_{ij} = 1, i \in N, q = \overline{1, 2},$$

и стационарности

$$p_i = \sum_{j=1}^N p_j \pi_{ij}, i \in N.$$

Разрядное двоичное изображение представляет собой марковское случайное поле размером $m \times n$ на несимметричной полуплоскости, полученной классической разверткой с левого верхнего угла слева направо вниз (рис. 2).

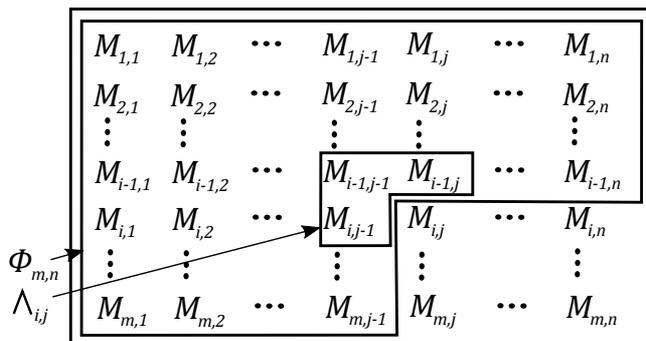


Рис. 2 Модель РДИ

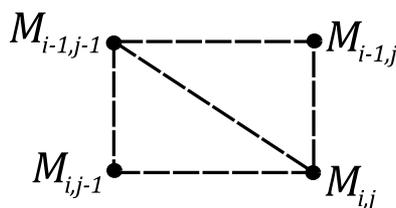


Рис. 3 Окрестность элемента РДИ

Состояние элемента $M_{i,j}$ ($i \in m, j \in n$) зависит только от известных ранее обработанных элементов $\Phi_{i,j}$, в частности от подмножества ближайших элементов $\Lambda_{i,j}$, называемого окрестностью элемента $M_{i,j}$. Лучшим образом удовлетворяющее условию каузальности является конфигурация окрестности (рис. 3) [3]:

$$\Lambda_{i,j} = \{M_{i,j-1}, M_{i-1,j}, M_{i-1,j-1}\}.$$

Вероятности перехода от состояний элементов окрестности $\Lambda_{i,j}$ к состоянию $M_{i,j}$ образуют МВП вида [3]:

$${}^3\Pi = \left\| \begin{matrix} \pi_{111} & \pi_{121} & \pi_{211} & \pi_{221} \\ \pi_{112} & \pi_{122} & \pi_{212} & \pi_{222} \end{matrix} \right\|^T. \tag{2}$$

Основой алгоритма является процедура предсказания каждого элемента РДИ на основе известной окрестности $\Lambda_{i,j}$ и МВП (2), используя критерий максимальной вероятности. В результате предсказания формируется РДИ по следующему правилу: если элемент предсказан верно, то 0, иначе 1. При восстановлении РДИ нулевые элементы заменяются результатами предсказания, а единичные — инверсией результата предсказания. В качестве примера на рис. 4, б приведен результат предсказания старшего разряда панхроматического цифрового изображения (см. рис. 1), белым показаны области, предсказанные неверно [4, 5]. Применение кода Грея поэлементно позволяет увеличить успешность предсказания.

Разрядное двоичное устройство после предсказания представляет собой разреженную матрицу, что позволяет эффективно применять методы энтропийного кодирования. В предложенном алгоритме применяется код Хаффмана. Данный метод можно использовать для работы с цветными изображениями путем обработки каждой цветовой компоненты в отдельности. Представление изображения в цветовом пространстве YCoCg повышает степень сжатия.

Формализованное описание алгоритма:

- 1) изображение преобразуется в цветовое пространство YCoCg;



(а) Исходный разряд



(б) Разряд после процедуры предсказания

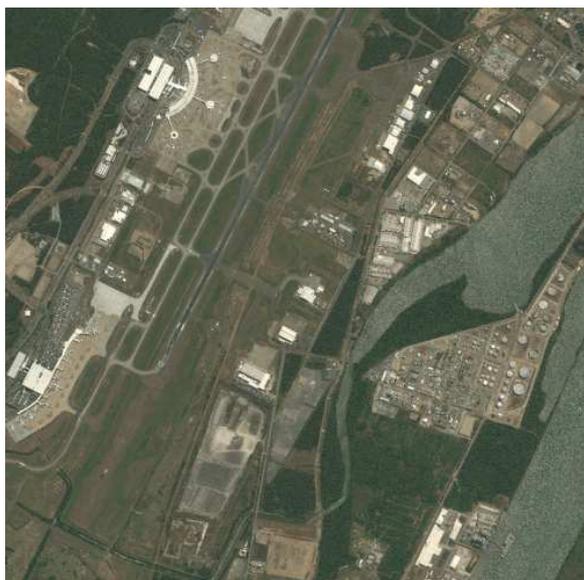
Рис. 4 Старший разряд цифрового изображения

- 2) код Грея попиксельно применяется к каждой цветовой компоненте;
- 3) каждая цветовая компонента разделяется на g РДИ;
- 4) для каждого РДИ вычисляются МВП:
 - по горизонтали $^1\Pi$;
 - по вертикали $^2\Pi$;
- 5) для каждого РДИ исходя из МВП $^1\Pi$ и $^2\Pi$ вычисляется МВП $^3\Pi$:
 - по горизонтали $^1\Pi$;
 - по вертикали $^2\Pi$;
- 6) для каждого РДИ выполняется процесс предсказания: прогнозируется элемент $\hat{M}_{i,j}$ на основе МВП $^3\Pi$ и окрестности $\wedge_{i,j}$:
 - если $M_{i,j} = \hat{M}_{i,j}$, то элемент предсказан верно;
 - если $M_{i,j} \neq \hat{M}_{i,j}$, то элемент предсказан неверно;
- 7) РДИ объединяются обратно в цветовые компоненты;
- 8) каждая цветовая компонента кодируется кодом Хаффмана;
- 9) полученные данные и служебная информация записываются в файл формата `mrk`.

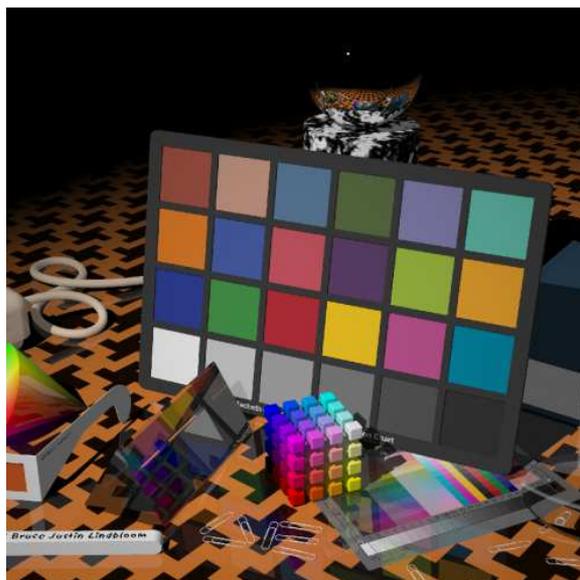
3 Результаты экспериментов

Для исследования эффективности предложенного алгоритма (MRK) произведено сжатие снимков поверхности Земли со спутника (рис. 5, а) [6] и базы изображений сайта `imagecompression.info` (рис. 5, б) [7]. Исходное разрешение спутниковых снимков 8192×8192 пикселей, изображения с сайта `imagecompression.info` имеют различное разрешение в диапазоне 2000×3000 – 5000×7000 пикселей, все изображения цветные (24 бита/пиксель). Исследование проводилось на компьютере, обладающем следующими характеристиками: процессор Intel Celeron G1820 — 2,7 ГГц, 8 ГБ оперативной памяти, SSD TOSHIBA 128 ГБ, операционная система Windows 7.

В качестве аналогов представлены известные алгоритмы сжатия без потерь PNG, JPEG-LS и JPEG 2000, а также современные разработки BMF [8], Qlic [9] и ImageZero [10]. На рис. 6 представлены графики зависимости среднего коэффициента сжатия в зависимо-



(а) Снимок со спутника



(б) База изображений сайта imagecompression.info

Рис. 5 Примеры снимков из выборки

сти от размера изображения. Для оценки быстродействия алгоритма на рис. 7 представлены зависимости скорости работы алгоритма от размера изображения. Оценки скорости получены при работе в однопоточном режиме. При работе в параллельном режиме скорость работы на современном четырехядерном процессоре увеличится в 3–4 раза.

4 Заключение

По результатам исследования видно, что предложенный метод можно отнести к классу методов с высоким быстродействием. Метод обладает потенциалом по увеличению производительности. Алгоритм построен на простых битовых операциях, тестирование произведено на компьютере, оперирующем байтами. Переход на оборудование, оперирующее битами, даст увеличение производительности в разы. Использование параллельного режима работы даст прирост производительности в 3–4 раза. Алгоритм имеет ряд преимуществ:

- сжатие каждого РДИ происходит независимо от других, это дает возможность увеличения быстродействия за счет параллельной обработки РДИ;
- алгоритм позволяет обрабатывать изображения любой разрядности;
- алгоритм не требователен к памяти и может работать на лету (для работы алгоритму необходимо хранить только две очередные строки изображения);
- алгоритм строится на простых вычислительных операциях, что снижает требования к элементной базе. Так, реализовав данный алгоритм на ПЛИС (программируемых логических интегральных схемах), можно существенно сократить энергетические затраты и повысить быстродействие, что актуально для использования в системах дистанционного зондирования Земли.

Литература

- [1] Гонсалес Р., Вудс Р. Цифровая обработка изображений / Пер. с англ. — М.: Техносфера, 2005. 1072 с. (*Gonzalez R., Woods R. Digital image processing. — 2nd ed. — Prentice Hall, 2002. 813 p.*)

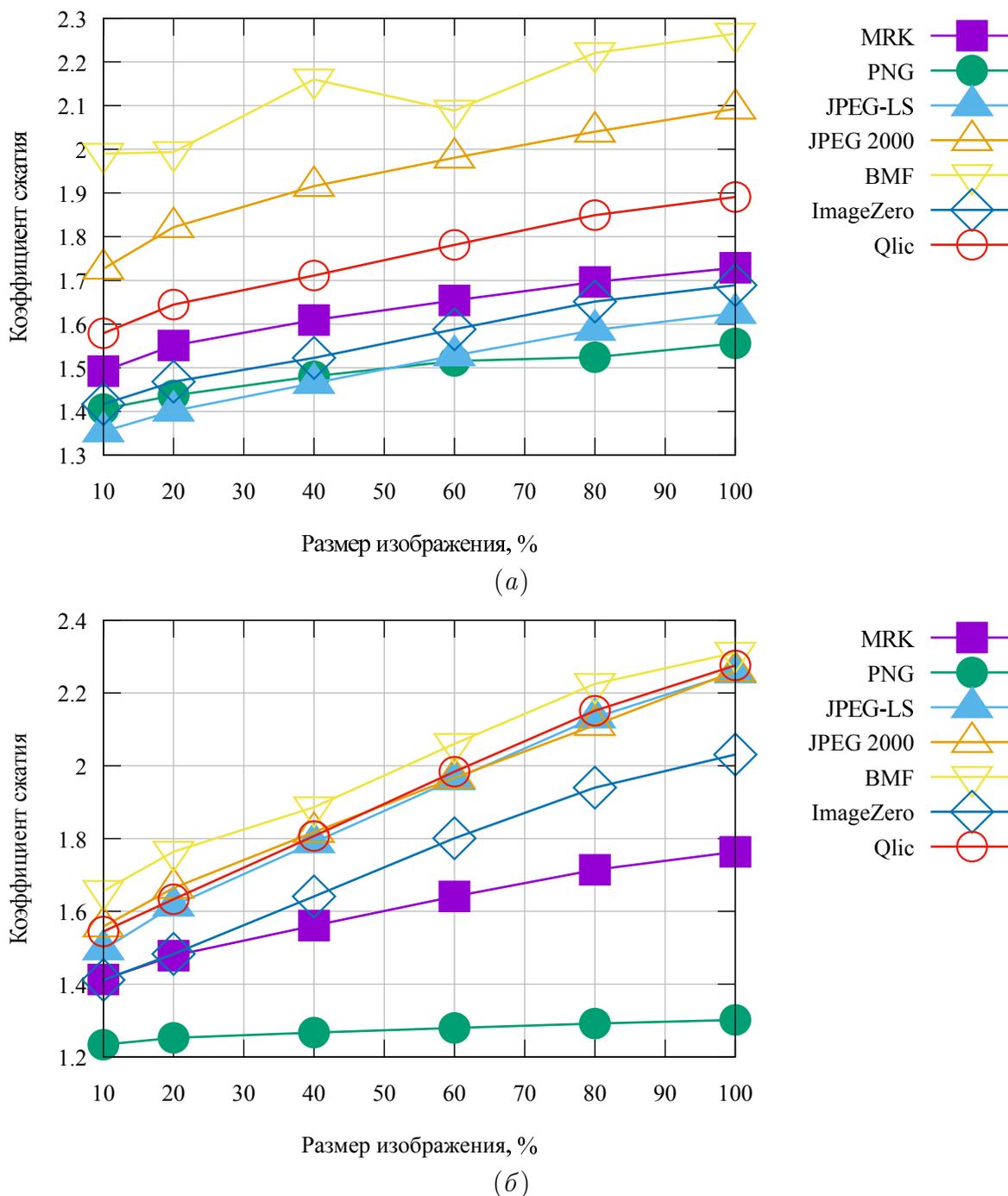


Рис. 6 Зависимость среднего коэффициента сжатия от размера изображения: (а) спутниковые снимки; (б) база изображений, сайт imagecompression.info

[2] Яне В. Цифровая обработка изображений / Пер. с англ. — М.: Техносфера, 2007. 584 с. (*Jahne V. Digital image processing. — 6th ed. — Springer-Verlag, 2005. 639 p.*)

[3] Петров Е. П., Медведева Е. В., Харина Н. Л. Математическая модель цифровых полутоновых изображений Земли из космоса // Мат-лы II Всеросс. НТК «Актуальные проблемы ракетно-космической техники». — Самара, 2011. С. 179–185.

[4] Петров Е. П., Харина Н. Л., Ржанникова Е. Д. Метод сжатия цифровых полутоновых изображений на основе цепей Маркова с несколькими состояниями // Мат-лы III Всеросс. НТК «Актуальные проблемы ракетно-космической техники». — Самара, 2013. С. 163–170.

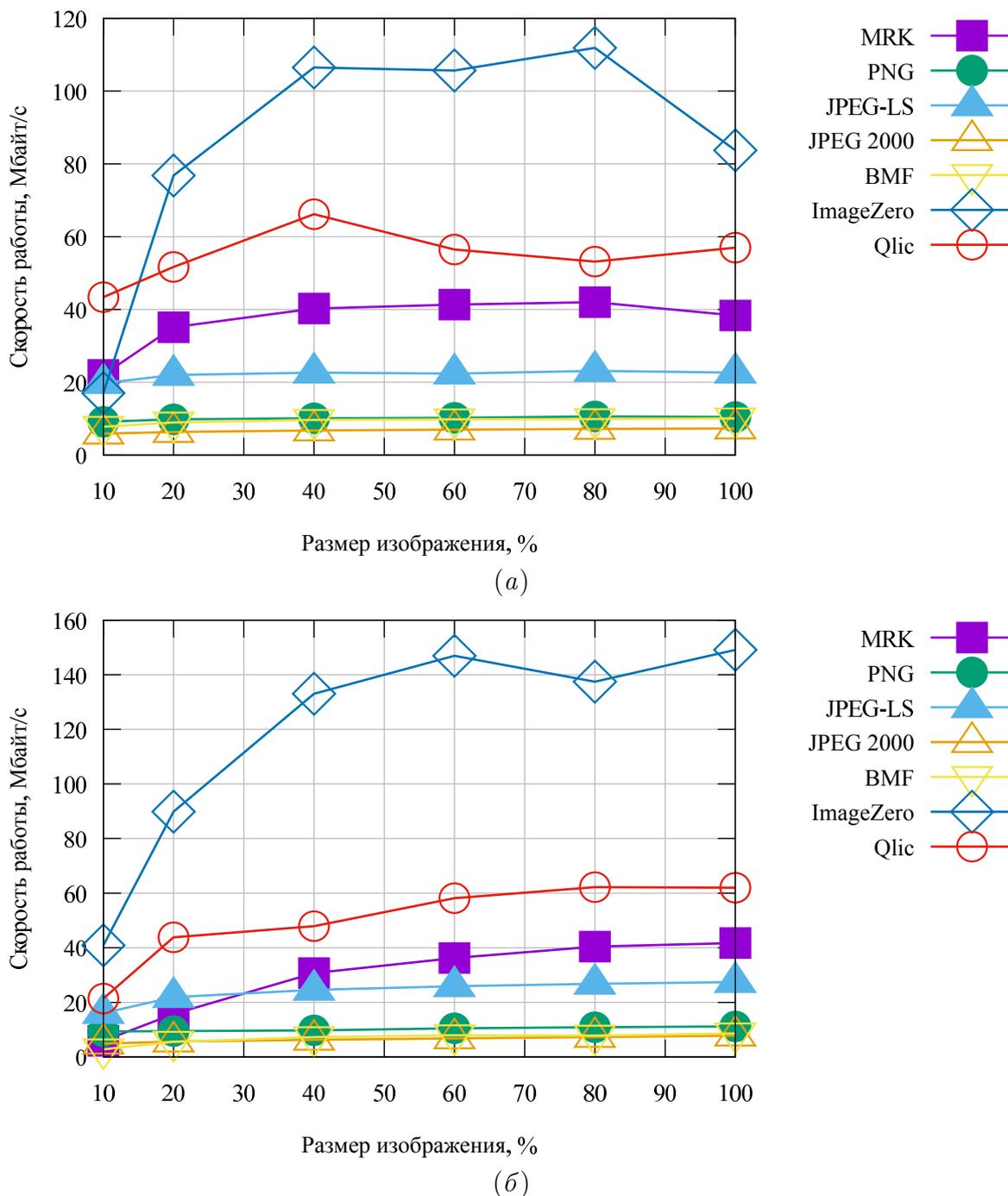


Рис. 7 Зависимость скорости работы алгоритма от размера изображения: (а) спутниковые снимки; (б) база изображений, сайт imagecompression.info

- [5] Петров Е. П., Харина Н. Л., Ржаникова Е. Д. Метод сжатия цифровых изображений без потерь минимальными вычислительными ресурсами // Тр. РНТОРЭС им. А. С. Попова. Сер. Цифровая обработка сигналов и ее применение. — М., 2015. С. 544–548.
- [6] DigitalGlobe <https://www.digitalglobe.com/resources/imagery-product-samples>.
- [7] The New Test Images — image compression benchmark. http://imagecompression.info/test_images.
- [8] BMF. <http://www.compression.ru/ds/>.

- [9] Rhatus A. QLIC — Quick Lossless Image Compressor. <http://www.imagecompression.info/gralic/qlic1d.zip>.
- [10] ImageZero — fast lossless color image compression. <http://imagezero.maxiom.de/>.

Поступила в редакцию 15.06.2015

References

- [1] Gonzalez, R., and R. Woods. 2002. *Digital image processing*. 2nd ed. Prentice Hall. 813 p.
- [2] Jahne, B. 2005. *Digital image processing*. 6th ed. Springer-Verlag. 639 p.
- [3] Petrov, E. P., E. V. Medvedeva, and N. L. Kharina. 2011. Matematicheskaya model' tsifrovyykh polutonovykh izobrazheniy Zemli iz kosmosa [Mathematical model of digital half-tone images of Earth from space]. *Mat-ly II Vseross. NTK "Aktual'nye Problemy Raketno-Kosmicheskoy Tekhniki"* [2nd Scientific-Technical Conference "Actual Problems of the Missile and Space Equipment Proceedings"]. Samara. 179–185. (In Russian.)
- [4] Petrov, E. P., N. L. Kharina, and E. D. Rzhanikova. 2013. Metod szhatiya tsifrovyykh polutonovykh izobrazheniy na osnove tsepey Markova s neskol'kimi sostoyaniyami [Method of digital half-tone images compression on the basis of Markov chains with several states]. *Mat-ly III Vseross. NTK "Aktual'nye Problemy Raketno-Kosmicheskoy Tekhniki"* [3rd Scientific-Technical Conference "Actual Problems of the Missile and Space Equipment" Proceedings]. Samara. 163–170. (In Russian.)
- [5] Petrov, E. P., N. L. Kharina, and E. D. Rzhanikova. 2015. Metod szhatiya tsifrovyykh izobrazheniy bez poter' minimal'nymi vychislitel'nymi resursami [Method of digital images compression without loss with minimum computing resources]. *Tr. RNTORES im. A. S. Popova. Ser. "Tsifrovaya obrabotka signalov i ee primeneniye"* [17th Conference (International) on Digital Signal Processing and its Applications Proceedings]. Moscow. 544–548. (In Russian.)
- [6] DigitalGlobe. Available at: <https://www.digitalglobe.com/resources/imagery-product-samples> (accessed November 24, 2015).
- [7] The New Test Images — image compression benchmark. Available at: http://imagecompression.info/test_images (accessed November 24, 2015).
- [8] BMF. Available at: <http://www.compression.ru/ds/> (accessed November 24, 2015).
- [9] Rhatus, A. QLIC — Quick Lossless Image Compressor. Available at: <http://www.imagecompression.info/gralic/qlic1d.zip> (accessed November 24, 2015).
- [10] ImageZero — fast lossless color image compression. Available at: <http://imagezero.maxiom.de/> (accessed November 24, 2015).

Received June 15, 2015