

Министерство образования Республики Беларусь
учреждение образования
«Белорусский государственный университет информатики
и радиоэлектроники»

ТЕЛЕКОММУНИКАЦИИ: СЕТИ И ТЕХНОЛОГИИ,
АЛГЕБРАИЧЕСКОЕ КОДИРОВАНИЕ
И БЕЗОПАСНОСТЬ ДАННЫХ

МАТЕРИАЛЫ МЕЖДУНАРОДНОГО НАУЧНО-ТЕХНИЧЕСКОГО СЕМИНАРА
(Минск, ноябрь – декабрь 2019 г.)

TELECOMMUNICATIONS: NETWORKS AND TECHNOLOGIES,
ALGEBRAIC CODING AND DATA SECURITY

Минск, 2019

УДК 654:004.056
ББК 32.88+32.973.202
Т31

Руководитель семинара В.Ю. Цветков

Редакционная коллегия:

М.Н. Бобов, А.А. Борискевич, Т.В. Борботько, В.Ф. Голиков
В.А. Лабунов, Л.М. Лыньков, В.К. Конопелько, Л.А. Шичко

Т31 **Телекоммуникации:** сети и технологии, алгебраическое кодирование и безопасность данных = Telecommunications: Networks and Technologies, Algebraic Coding and Data Security : материалы междунар. науч.-техн. семинара (Республика Беларусь, Минск, ноябрь – декабрь 2019 г.).– Минск : БГУИР, 2019. – 128 с. : ил.
ISBN 978-985-543-553-3.

Сборник содержит статьи, тематика которых посвящена научно-теоретическим разработкам в области сетей телекоммуникаций, информационной безопасности, алгебраического кодирования и обработки изображений.

Предназначен для научных сотрудников в области телекоммуникаций, преподавателей, аспирантов, магистрантов и студентов технических вузов.

Научное издание

Корректор *В.В. Чепикова*

Ответственный за выпуск *В.Ю. Цветков*

Компьютерный дизайн и верстка *Е.Г. Макейчик*

Подписано в печать 03.12.2019. Формат 60×84 1/8. Бумага офсетная. Гарнитура «Таймс».
Отпечатано на ризографе. Усл. печ. л. 15,35. Уч.-изд. л. 13,2. Тираж 60 экз. Заказ 376.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники».

Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий №1/338 от 24.03.2014,

№2/133 от 07.04.2014, №3/615 от 07.04.2014.

Ул. П. Бровки, 6, 220013, г. Минск

ISBN 978-985-488-834-7

© УО «Белорусский государственный университет информатики и радиоэлектроники», 2019

СОДЕРЖАНИЕ

Jun Ma, V.Yu. Tsviatkou, V.K. Konopelko Fast fully parallel one-subiteration thinning algorithm	5
Ren Xunhuan, V.K. Konopelko Classification of point forms and establish a library for coding the skeleton of the binary images	11
В.Н. Shaya, U.A. Vishniakou Model and structure of multiagent system for collection and processing sound information	18
А.Т. Нгуен, Т.Х. Доан, В.Ю. Цветков Волновое выращивание областей локальных экстремумов изображений	22
Н.С. Филиппов, С.Б. Саломатин Цифровая подпись веб-сервиса протокола TLS на основе эллиптических кривых и облачной платформы Негоку	29
А.А. Дудкин, А.А. Воронов, В.В. Ганченко, А.В., Инутин, Е.Е. Марушко Алгоритмическое обеспечение для оптического контроля критических размеров полупроводниковых пластин СБИС	34
А.В. Курилович, В.Ю. Цветков, В.К. Конопелько Структурный алгоритм текстурной сегментации полутоновых изображений	39
В.А. Аксенов, С.В. Смоляк, М.Ю. Хоменок Вертикальная секторизация для совместной работы узкополосного интернета вещей и мобильных сервисов 3G	43
В.А. Вишняков, М.М. Гондаг Саз Интеллектуальное управление выбором для доступа пользователей к мобильным приложениям	49
С. Б. Саломатин, А.Э. Алексеенко Корректирующие коды на основе биномиальных фрактальных структур	53
В.В. Рабцевич, В.Ю. Цветков Градиентно-волновое выращивание областей АСМ-изображений	58
М.Г. Моздурани Шираз, В.А. Вишняков Интеллектуальный классификатор обнаружения вторжений на базе генетических алгоритмов	64
Н.Н. Сергеев, В.Н. Урядов Возможность несанкционированного доступа к информации, передаваемой от абонента на другой абонентской розетке в пассивной оптической сети, при использовании отраженного сигнала от ответвителя	70
Э.Б. Липкович, А.А. Павловский Распределительная сеть интернет-вещания по технологии OTT	74
С.И. Наливайко, В.А. Вишняков Алгоритм анализа изображений в видеоаналитике	78
В.В. Новицкий, В.Ю. Цветков Метод реверсивного кодирования нуль-деревьев	82
С.А. Лукашевич, В.Н. Урядов, Т.В. Полуян Прямое фотодетектирование и когерентный прием с различными видами модуляции в открытых системах оптической связи со стационарными космическими объектами	88
А.Л. Калоша, А.В. Саскевич, М.В. Стержанов Детектирование объектов на изображениях при помощи фреймворка TensorFlow	95
Б.Д.С. Садик, В.Ю. Цветков, М.Н. Бобов Сжатие изображений без потерь на основе комбинированного эффективного кодирования битовых плоскостей	98
Э.Б. Липкович, А.А. Павловский, А.А. Серченя Оценка эффективности каналов распределительной системы стандарта DVB-C2	106
Н.И. Листопад, О.Н. Виничук Разработка веб-приложения для обработки цифровых изображений	114
Д.А. Белянков Возможности искусственного интеллекта и машинного обучения в телекоммуникационных сетях	118

CONTENTS

Jun Ma, V.Yu. Tsviatkou, V.K. Konopelko Fast fully parallel one-subiteration thinning algorithm	5
Ren Xunhuan, V.K. Konopelko Classification of point forms and establish a library for coding the skeleton of the binary images	11
B.H. Shaya, U.A. Vishniakou Model and structure of multiagent system for collection and processing sound information	18
A.T. Nguyen, T.H. Doan, V.Yu. Tsviatkou Wave growing of local extreme regions of images	22
N.S. Filippov, S.B. Salomatin Digital signature of the TLS protocol web service based on the elliptic curves and Heroku cloud platform	29
A.A. Doudkin, A.A. Voronov, V.V. Ganchenko, A.V. Inutin, E.E. Marushko Algorithmic support for optical inspection of critical dimensions on VLSI semiconductor wafers	34
A.V. Kurylovich, V.Yu. Tsviatkou, V.K. Konopelko Search of local extremums of half-tone images based on central symmetric scanning	39
V.A. Aksyonov, S.V. Smolyak, M.Yu. Homenok Vertical sectorization for joint operation of the narrow-band internet of things and 3G mobile services	43
U.A. Visniakou, M.M. Ghondagh Sas Intelligent management of choice for user access to mobile applications	49
S.B. Salomatin, A.E. Alekseenko Corrective codes based on binomial fractal structures	53
V.V. Rabtceвич, V.Yu. Tsviatkou Gradient-wave growing segmentation algorithm of AFM-images	58
M.M. Mosdurany Shiras, U.A. Vishniakou Intelligent classifier of intrusion detection on bases of genetic algorithms	64
N.N. Sergeev, V.N. Uryadov Possibility of unauthorized access to information transferred from subscriber to other subscriber socket in a passive optical network, when using a reflected signal from a splitter	70
E.B. Lipkovich, A.A. Paulouski OTT technology distribution network	74
S.I. Nalivaika, V.A. Vishnyakou Image analysis algorithm in video analytics	78
V.U. Navitski, V.Yu. Tsviatkou Method of reversive zero-tree coding	82
S.A. Lukashevich, V.N. Urjadov, T.V. Poluyan Direct photodetection and coherent reception with different modulation techniques in open optical communication systems with stationary space objects	88
A.L. Kalosha, A.V. Saskevich, M.V. Sterjanov Objects detection and recognition on images using TensorFlow	95
B.J.S. Sadiq, V.Yu. Tsviatkou, M.N. Bobov Lossless image compression based on the combined efficient coding of bit planes	98
E.B. Lipkovich, A.A. Paulouski, A.A. Serchenya Evaluation of efficiency of DVB-C2 distribution system channels	106
N.I. Listopad, O.N. Vinichuk Development of a web application for processing digital images	114
D.A. Beliankou Possibilities of artificial intelligence and machine learning in telecommunication networks	118

УДК 621.391

FAST FULLY PARALLEL ONE-SUBITERATION THINNING ALGORITHM

JUN MA, V.Yu. TSVIATKOU, V.K. KONOPELKO

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus**Submitted 30 October 2019*

Abstract. Fast thinning algorithm with one-subiteration was proposed. Results of implementing the proposed algorithm, on a variety of binary images and comparison with ZS algorithm and OPTA algorithm show better results in terms of thinning speed, thinning rate and visual quality.

Keywords: Thinning algorithm; Binary images; Thinning Rate; Thinning speed.

Introduction

Thinning algorithm is generally deemed as a fundamental preprocessing algorithm that plays a critical role in many pattern recognition applications. Thinning algorithm tend to remove edge contour layer by layer and only leave the skeleton which has same topology with the original image. There is a large amount of literatures on thinning algorithms [1–6].

The mainly structure of the classification of the thinning algorithm as shown in Fig. 1.

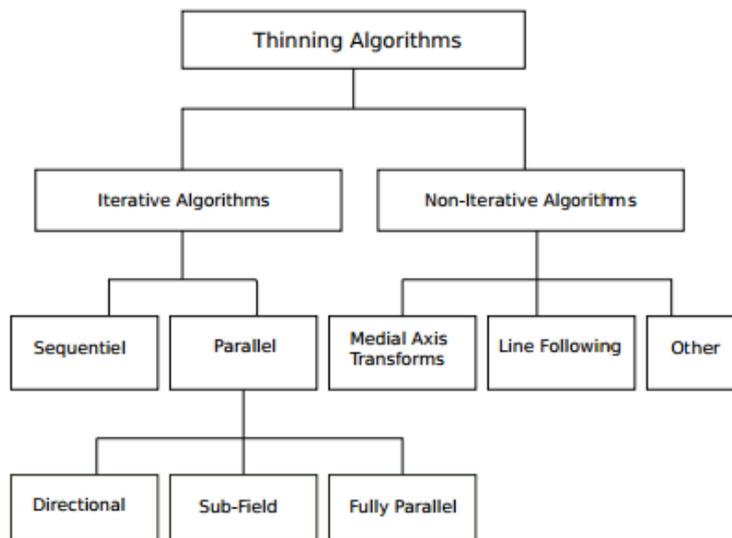


Fig. 1. Classification of the thinning algorithm

Thinning algorithm can be divided into iterative algorithm and non-iterative algorithm at first. Iterative algorithm works on the pixel by pixel based on thinning, however, non-iterative algorithm is just opposite to iterative. Non-iterative algorithm can further divide in medial axis transforms, line following and others.

According to the approach of examine pixels, iterative thinning algorithm can be classified as sequential or parallel [3]. In a sequential algorithm, each deletion pixel in the n -th iteration depend on all the operations that have been performed, however, in a parallel algorithm, the deletion of pixels in the n -th iteration only depend on the result that remains after the $(n-1)$ -th; therefore, all pixels can be examined independently in a parallel manner in each iteration.

Parallel thinning algorithms can be further classified in two categories according to used approach. There are mainly three parallel methods: the directional approach, the subfield approach and the fully parallel approach. The directional approach consists of breaking the thinning iteration into several sub – iteration of thinning of which based on a combination of direction. The subfield approach consists of breaking down the figure into sub – field according to some criterion such as the parity of pixels. As for the full parallel approach is a method without any sub – iteration, in this way all the pixels are examined by the same criterion.

A good thinning algorithm should be satisfied with the following requirements [7]. First of all, the output of the thinning algorithm should maintain the original topology and geometric properties of the input image, and at the same time have one-pixel width. Second, the skeleton of the output should be located at the medial line of the original image. Besides, thinning algorithm should not be sensitive to the boundary noise. And the last, thinning algorithm should be proven to be high efficiency, which means it should process a picture as fast as possible.

Basic notations

The binary image I is denoted by a matrix \mathbf{M} with the size $R \times C$, where $P(i, j)$ represents the binary value of the pixel (i, j) . If $P(i, j) = 1$, then the pixel (i, j) is black and if $P(i, j) = 0$, it is white. The black pixels build up the foreground of I , and the white pixels form the background which is the complement of the foreground.

The set of four pixels which called 4 – neighbourhood of the pixel p is noted as $N_4(p)$ as shown in Fig. 2.

There are 4 diagonals pixels that represent the diagonal neighborhood of the pixel p noted $N_D(p)$ as shown in the Fig. 3.

	$(i-1, j)$	
$(i, j-1)$	$P(i, j)$	$(i, j+1)$
	$(i+1, j)$	

Fig. 2. The 4 – Neighborhood of a pixel P

$(i-1, j-1)$		$(i-1, j+1)$
	$P(i, j)$	
$(i+1, j-1)$		$(i+1, j+1)$

Fig. 3. The D – Neighborhood of a pixel P

The $N_4(p)$ and $N_D(p)$ together formed the concept of the 8-neighborhood, which noted as $N_8(p)$. This set of points define on the other hand the 8-connectivity.

Two pixels can seem as connected if they appear in one another's 8-neighborhood.

In general, a non-zero pixel can be classified as 4 special points, which are break pixel, edge pixel, end pixel and simple pixel respectively. Break pixel: which is a non-zeros pixel for which the deletion would break the connectivity of the original pattern. Edge pixel, also called as boundary or border pixel, has at least one zero 4-neighbor pixel. And end point has at most one non-zero 8-neighbor pixel. As for simple pixel is an edge pixel whose removal from the object does not change the topology.

Here according to the criterion, which used to describe a good thinning algorithm, it is reasonable to deduced that an algorithm can be a good algorithm if and only if following requirement can be satisfied.

1. No object in I can be completely deleted by the algorithm in any iteration;
2. It does not connect any originally disjoint objects in image I , nor does it disconnect an existing hole or create any new hole in I ;
3. It does not disconnect any object in I ;
4. It does not connect any hole of I to another distinct hole or the background of I where a hole may be defined as a background region surrounding by a connected border of foreground pixels.

Thinning algorithm

The scan window used to find out all the deletable candidate points is expanded from the general window whose size is 3×3 in the fast fully parallels algorithm proposed in this paper, as is shown in Fig. 4.

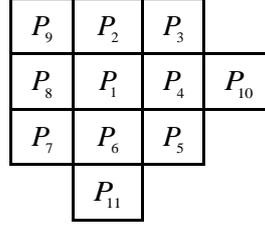


Fig. 4. Scan window

Our algorithm for extracting the skeleton of a picture consists of removing all the contour points of the picture except those points that belong to the skeleton. In each iteration, the contour point P_1 is deleted from the digital pattern if it satisfies the following conditions at the same time:

- (a) $2 \leq B(P_1) \leq 6$,
- (b) $A(P_1) = 1$,
- (c) $P_4 = 0$ or $P_8 + P_{10} > 0$,
- (d) $P_6 = 0$ or $P_2 + P_{11} > 0$,

where $A(P_1)$ is the number of 01 pattern in the ordered set P_2, P_3, \dots, P_9 hat are the eight neighbors of P_1 , and $B(P_1)$ is the number of nonzero neighbors of P_1 , that is,

$$B(P_1) = \sum_{i=2}^9 P_i.$$

By condition (a), the end points of a skeleton lines are preserved. In another word, it means there is only one 8-connected component in the neighborhood of P_1 , which implies there is deletable when P_1 is a boundary pixel and the deletion will not break the connectivity of the 8 neighborhood. Besides, condition (b) prevents the deletion of those points that lie between the end points of skeleton line. Also, condition (c) and (d) can preserve the excessive erosion of the 2×2 pattern.

After iteratively remove the boundary and left the skeleton of the original image by applying operations mentioned above, an extra delete procedure is introduced to ensure the single thick of the skeleton. The additional procedure only performed once, so it will not increase the cost of the time dramatically. In this procedure, a normal 3×3 scan window is used to examine the skeleton pixel by pixel. If one of the following conditions is satisfied, current pixel can be removed

- (a') $P_4 \times P_6 = 1$ and $P_9 = 0$,
- (b') $P_6 \times P_8 = 1$ and $P_3 = 0$,
- (c') $P_2 \times P_4 = 1$ and $P_7 = 0$,
- (d') $P_2 \times P_8 = 1$ and $P_5 = 0$.

After all the pixels have been processed, the whole thinning algorithm is finished.

Experiments and compare

In this section, well known thinning algorithm ZS [6], OPTA [8] and our algorithm were implemented in Matlab2018 and subsequently tested them on a variety of binary images which involves patterns from basic shapes to animal shapes.

In order to evaluate our new algorithm and these famous thinning algorithms, comparisons are done using the following performance measures:

Thinning rate (TR): Or the degree of thinness of the image as it was proposed in [9] by the following formula:

$$TR = 1 - TM1 \div TM2,$$

where $TM1$ – stands for total triangle count of the thinned image given by:

$$TM1 = \sum_{i=0}^n \sum_{j=0}^m TC(P(i, j)),$$

where $P(i, j)$ is a foreground pixel with coordinates i, j . TC is a function which counts the number of the black triangles which can be created from $P(i, j)$ and its neighboring pixels given by the following formula:

$$TC(P_i) = P_i \times [(P_8 \times P_9) + (P_2 \times P_9) + (P_2 \times P_3) + (P_3 \times P_4)].$$

$TM2$ represents the largest number of black triangles that an image could have computed. This number is computed as follows:

$$TM2 = 4 \times [\max(m, n) - 1]^2,$$

where m and n denote the number of the row and number of the column of the picture respectively.

When $TR = 1$ the image is perfectly thinned, but $TR = 0$ means that the image is not thinned at all [10].

Thinning Speed (TS): measures the number of pixels thinned per time unit (second) given by:

$$TS = DP \div ET,$$

$$DP = OP - SP,$$

where DP or (deleted points) is the number of black points deleted during thinning, OP is the number of black points of the image before thinning. SP or (skeleton points) are the number of the remaining points after applying the thinning algorithm or the number of pixels in the skeleton and ET denote the real execution time.

In the following, we report both visual and numerical results of ZS , $OPTA$ and our new algorithm on test image.

First test image consists of 2×2 square, 135° lines and 45° lines on which the application of the ZS algorithm and $OPTA$ algorithm will cause some distortion, such as excessive error or the existence of redundant pixels which cannot ensure that the skeleton is precisely one pixel thick, which mentioned in some literatures [5, 11–13]. The application results are shown in Fig. 5.

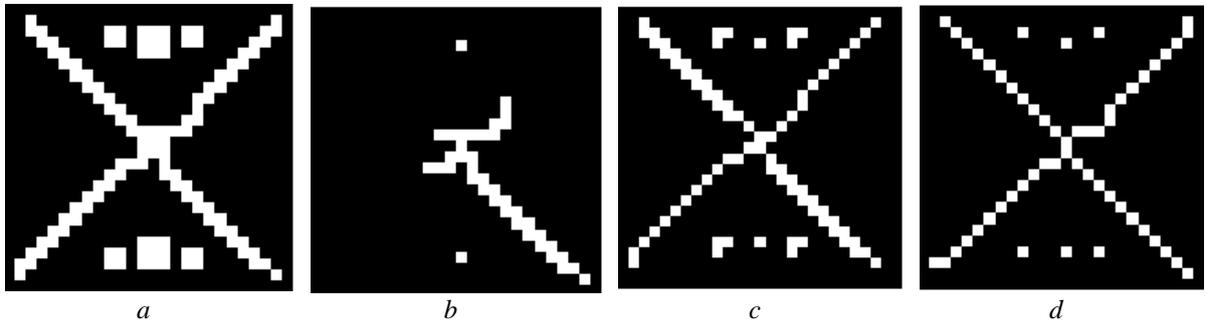


Fig. 5. Input image and its skeletons: a – original image; b – ZS ; c – $OPTA$; d – our algorithm

In Fig. 5, ZS algorithm both confronts serious excessive error in diagonal and in 2×2 square, what's more, it also suffered the redundant pixels. The only problem of the $OPTA$ is the existence of the redundant pixels. Our algorithm deals successfully with these patterns.

Next, more images from the dataset $kimia99$ have been introduced to test, the results of the different thinning algorithm are shown below in Fig. 6.

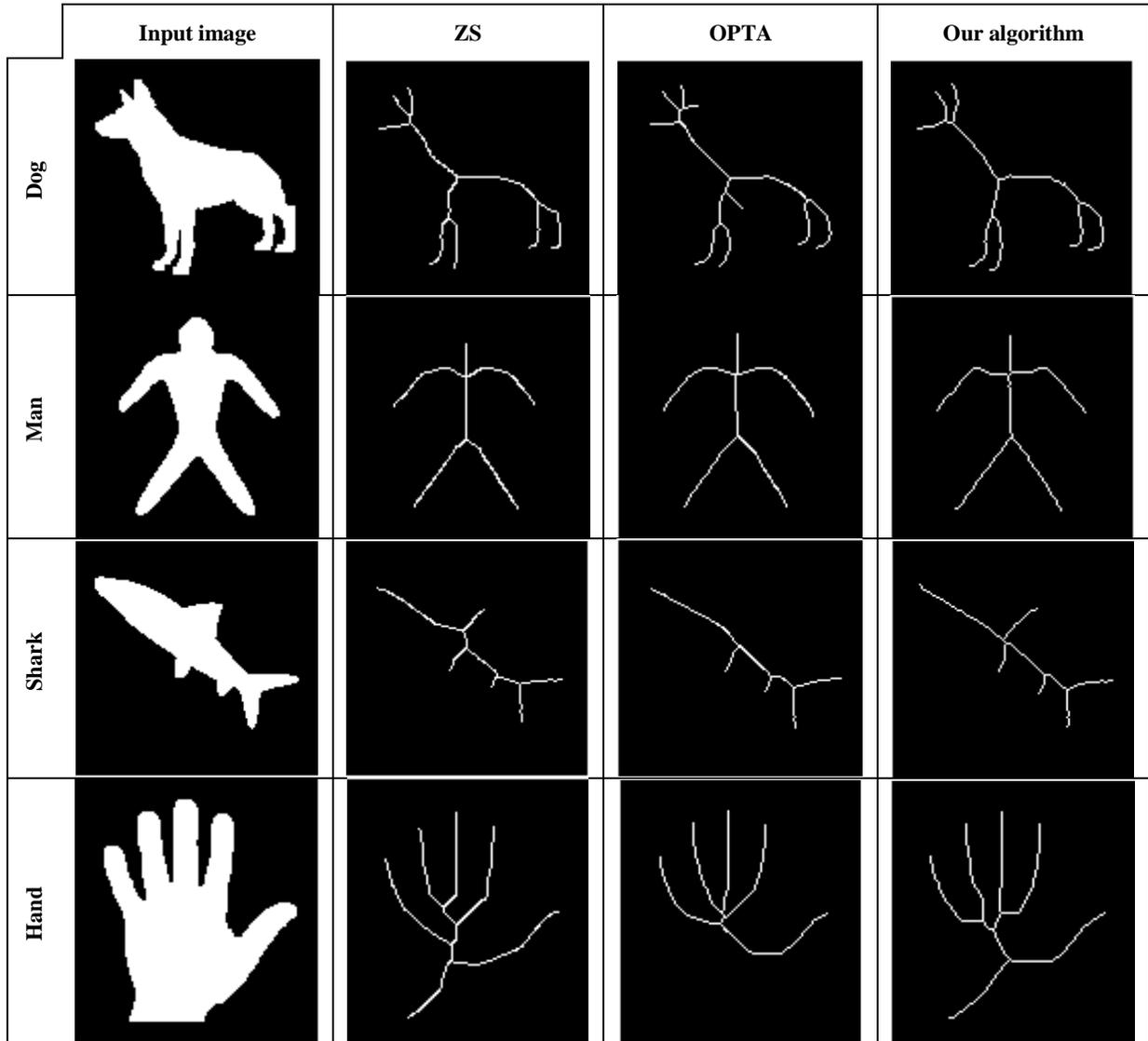


Fig. 6. Input images and its skeletons

Tabl. 1 gives a summary of performance details for all skeletons extracted from the binary image above.

Table 1. Comparison results of thinning performance for ZS, OPTA and our algorithm

Images	Object points	ZS			OPTA			Our algorithm		
		$TR1$	$TS1$ (p/s)	$SP1$	$TR2$	$TS2$ (p/s)	$SP2$	$TR3$	$TS3$ (p/s)	$SP3$
Dogs	3345	0,9991	426954	262	0,9996	191358	253	0,9999	601456	251
People	2699	0,9986	644012	232	0,9992	273963	227	1	909690	203
Shark	2402	0,9991	613532	187	0,9993	267551	167	1	810920	169
Hand	5270	0,9987	442326	333	0,9999	179904	250	1	641850	303

The performance evaluation confirms that our algorithm is better than ZS and OPTA in terms of: thinness, since our algorithm produces thinner skeletons of one width thick such $TR3 > \max(TR, TR2)$ in all test cases. Second, our results without excessive in diagonal lines. Moreover, indicator of the thinning speed much higher than both ZS and OPTA, with about 1,5 times and 3 times higher than the ZS algorithm and OPTA algorithm respectively. Finally, our algorithm produces better visual quality (middle of the shape, preserving the original shape).

Conclusion

In this paper we have proposed a new fully fast parallel one-subiteration thinning algorithm. After compare with the famous thinning algorithm ZS and OPTA, the results of performance evaluation and experiments show that new algorithm is faster and produces better skeleton without any excessive erosion.

References

1. Lam L., [et. al.] // Pattern Analysis and Machine Intelligence, IEEE Transactions. 1992. Vol. 14. P. 869–885.
2. Jagna. A. // International Journal of Advanced Research in Computer and Communication Engineering. 2014. Vol. 3. P. 8309–8311.
3. Brener N.E., Weian Deng S.S.I. // International Journal of High Performance Computing Application. 2000. Vol. 14. P. 65–81.
4. Rosenfeld A., Kak A. Digital Picture Processing. Elsevier. 1976.
5. Tarabek P. // 7th IEEE International Symposium on Applied Computational Intelligence and Informatics. 2012. P. 75–79.
6. Zhang. T.Y., Suan C.Y. // Commun. ACM. 1984. Vol. 27. P. 236–239.
7. Hastings E. // Pattern analysis and Machine Intelligence, IEEE Transactions. 1992. Vol. 14. P. 869–885.
8. Chen C.S., Tsai W.H. // Pattern Recognition Lett. 1990. P. 471–477.
9. Quek C., Ng R.G.S. // IEEE Transcation on System Man and Cybernetics. 1994.
10. Tarabek P. // Journal of Information. Control and Management Systems. 2008. Vol. 6. P. 125–132.
11. Abdulla W.H., Saleh A.O.M., Morad A.H. // Pattern Recognition Letters 7. 1988. P. 13–18.
12. Juan H.S. // Pattern Recoginition Letters 10. 1989. P. 77–80.
13. Boudaoud L.B., Sider A., Tari A. // Vis Comput. 2018. P. 689–706.

UDC 621.391.14

CLASSIFICATION OF POINT FORMS AND ESTABLISH A LIBRARY FOR CODING THE SKELETON OF THE BINARY IMAGES

REN XUNHUAN, V.K. KONOPELKO

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus**Submitted 30 October 2019*

Abstract. In this paper, the methods for classifying points which are allowed to determining the types of two-dimensional vectors and can be used to code the skeleton of the binary image were analyzed. On the basis of [3–5], our research focused on the matrix which is t by t , in which there are only t elements are one. These elements can be called as point (pixel). An improved algorithm for classifying points is proposed. The result of the preprocessing point is a compressed library of point patterns.

Keywords: pattern recognition, square matrix, transposed group, classifying points.

Introduction

The field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such as classifying the data into different categories. In 1872, Felix Klein in his famous "The Erlangen program" got the conclusion that in geometric space, if a set of objects can be classified, they must have some common features which can be extracted by implementing some transformation on these objects. Here given an example, two different matrices are equal if they both can translate to the same matrix though some elementary transformations of matrix respectively [1, 2].

In this paper, our research focused on the matrix which is t by t , in which there are only t elements are one. These elements can be called as point (pixel). The work is based on the research of the process of calculating the pattern point of the library is a complex calculate process, before processing the information, we need to reduce the pattern of point library in the initial stage.

The rest of the paper is organized as follows. In Section 2, a brief introduction of the classical method for classification numbers (points) will be presented. And in section 3, an improved algorithm which used to generate the patterns library from $t = 2:7$ has proposed. The conclusion will be given in Section 4.

The method for classification points

In [3–6] an exhaustive method for generating complete point vectors was proposed.

In statistics, where classification is often done with logistic regression or a similar procedure. The most primitive method to generate all possible point patterns through an exhaustive search, the point pattern can be equivalent to a matrix which is t by t and the set of possible patterns are represented by $A_t = C_t^t$. As can be seen from A_t , when $t = 2$, we need to analyze 6 possible patterns, then when $t = 7$ – about 90 million, as t increases, the amount of calculation increases exponentially. Therefore, in the following work, without changing the complete library patterns, we will reduce the computational complexity by reducing the total number of patterns.

Algorithm generating point patterns

The number of point patterns depends on the number of random permutation characters. When there exist two points – $t=2$, the set of point locations in the table consists of 6 different combinations. Fig. 1 shows these point patterns.

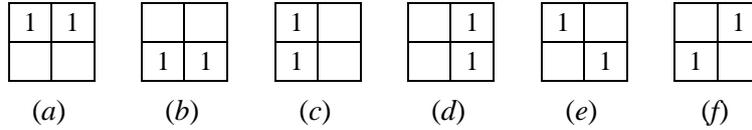


Fig. 1. The set of all patterns of the point at $t=2$

According to the definition, the two matrices are equivalent if they can be transformed into each other through the transformation. Select any one of the equivalent matrices as a representative, indicating this type of point pattern. Rules can be expressed as follows.

The rule for creating the library of point patterns:

1. Calculate the total number for each row (r) and column (c) and sort the results by size.
2. Calculation of rank for each matrix.
3. Calculating the number of intersections by row and column.
4. Calculating the sum and difference of the intersection coordinates.

At first we reduce all possible patterns according to the above rules, and then calculate their transposition, and finally get the total pattern library of corresponding point t , from the Fig. 1, the subset of $\{a;b;c;d\}$ fall into the same category, expressed as 2.1, the subset of $\{e;f\}$, as 2.2. Finally, calculate transposition of 2.1 and 2.2 we get the library of $t=2$ (Fig. 2).

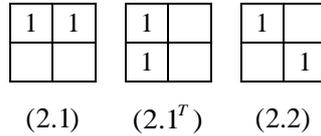


Fig. 2. The set of generating patterns when $t=2$

Based on the reduced point pattern of $t=2$ adding a point, we generate the point pattern of $t=3$, and so on, until the generation of the point patterns for $t=7$, finally we get a $t=2:7$ reduced point pattern library (Fig. 3).

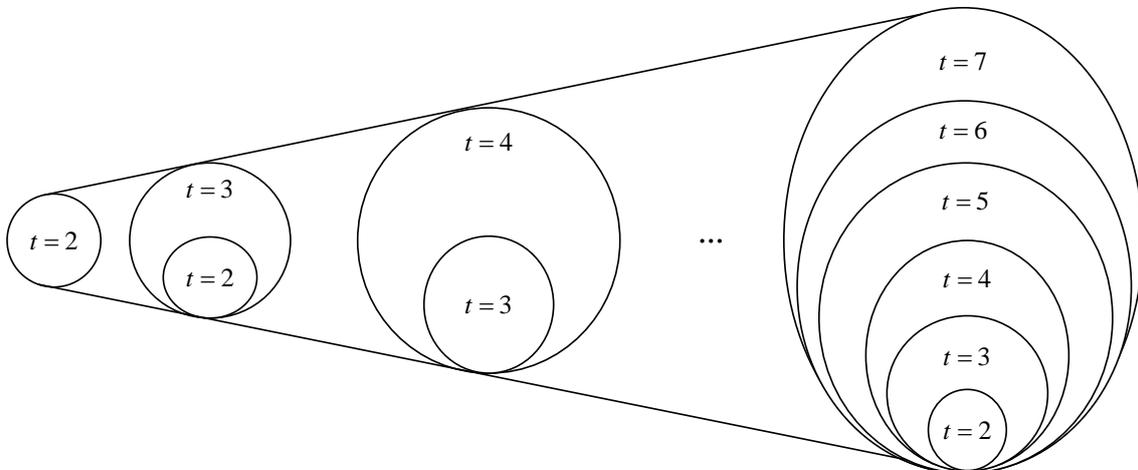


Fig. 3. Structure of the generated point patterns from $t=2:7$

In Fig. 4 and Fig. 5 show the process of $t=2$ generates the total point patterns for $t=3$.

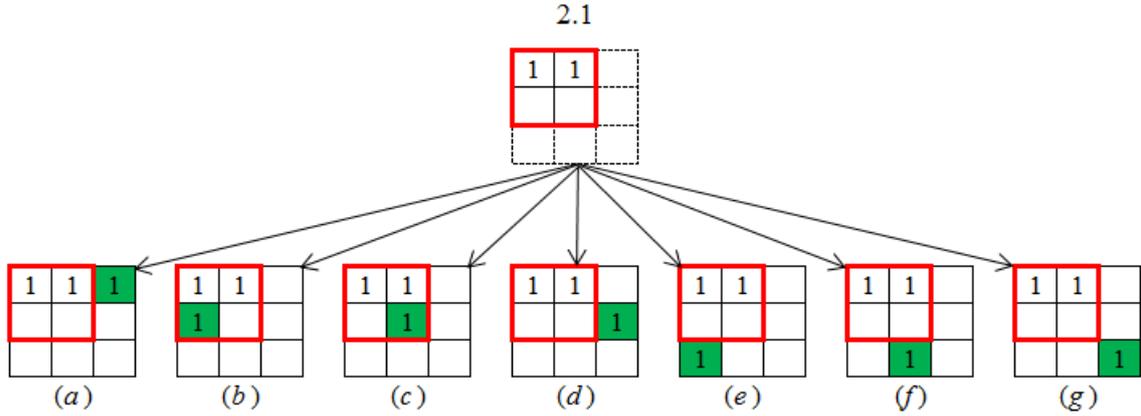


Fig. 4. From point patterns 2.1 obtain a subset point patterns of $t = 3$

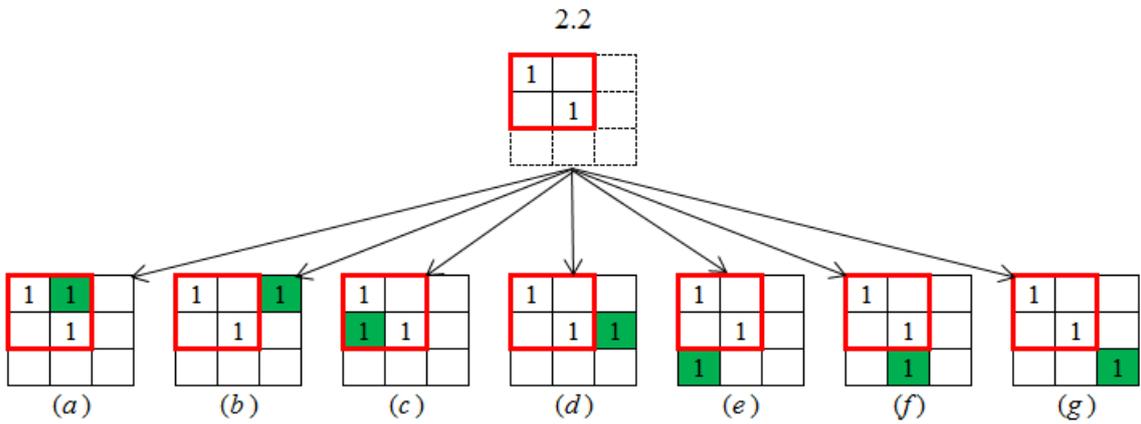


Fig. 5. From point patterns 2.1 obtain a subset of $t = 3$

According to the above rules, from Fig. 4 and Fig. 5 form four different patterns A_{com} . Compute the row (string) and the column which has "1" and record the string as s_i and the column as c_j . Here, $s_1 \geq s_2 \geq \dots \geq s_i \dots \geq s_n; c_1 \geq c_2 \geq \dots \geq c_i \dots \geq c_n$. In the set of 2.1 $\{b; c; e; f\}$ are the same as the set of 2.2 $\{a; c\}$, they match through a transposition of rows and columns. From the above Fig. we obtain 4 patterns for $t=3$, all of the reduced pattern set denoted by A_{red} , then from the set A_{red} can derive A_{red}^T , as a result we obtain compressed point patterns $A_{pat} = A_{red} \cup A_{red}^T$, when $t=3$, $A_{pat3} = A_3 \cup A_3^T$, as shown in Fig. 6.

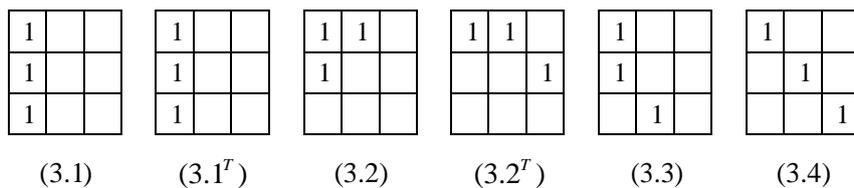


Fig. 6. The total pattern set when $t = 3$

In Fig. 6 it is shown the set of $t=3$, the transposition of 3.2 has not changed, it is still itself. We only need to continue to keep it in the library.

The association of the reduced point pattern from $t=2$ to $t=4$ is listed in the Tabl. 1.

Table 1. The association of the reduced point pattern from $t=2$ to $t=4$

The reduced basis point pattern of t	The obtained point pattern of $t+1$		
2.1	3.1	3.2	3.3
2.2	3.4	–	–
3.1	4.1	4.2	4.3
3.2	4.4	4.5	4.6
3.3	4.7	4.8	4.9
3.4	4.10	–	–

The structure of the reduced point pattern set from 3.1 to 5.11 is shown in Fig. 7.

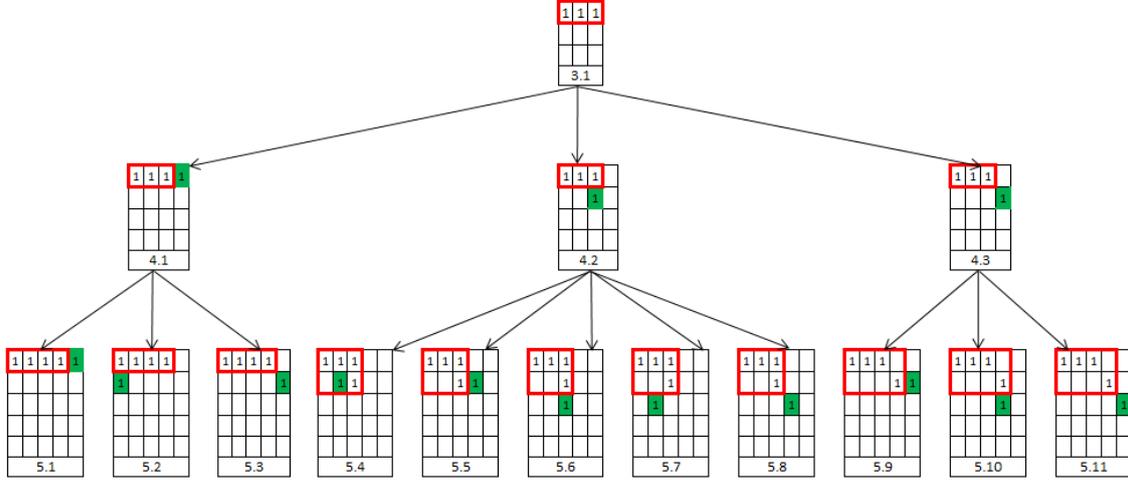


Fig. 7. From 3.1 obtain the reduced subset of $t=5$

Based on these definitions and guidelines, algorithm has been constructed as listed below.

Algorithm Generate point patterns

Input: t

Output: A_{pat}

Initialization: $t = 2$

While $t \leq N$ program do

Step 1: On the basis of matrix $t \times t$, we obtain the general total sets

$A_{tot} - (t+1) \times (t+1)$.

Step 2: according the rules reduce the set of A_{tot} and get the patterns of A_{red} .

Step 3: calculate the transpose of set A_{red}^T .

Step 4: calculate the set of point pattern – $A_{pat}, A_{pat} = A_{red} \cup A_{red}^T$.

Step 5: $t = t + 1$.

End while

Return A_{pat}

Then we will compare the performance of these methods in the computation experiment.

The computation experiment

The asymptotic growth of the original algorithm is faster than our improved algorithm (Tabl. 2), which means that the computational complexity of our new method is lower than the previous method.

Table 2. **Performance comparison of the methods of forming the total point patterns**

Method of generating the total patterns	The total number of generating patterns					
	2	3	4	5	6	7
Search algorithm	6	84	1820	53130	1947792	85900584
Fast algorithm	2	14	52	210	620	2193

The numbers of the feature point pattern extracted from all original point patterns are far less than the numbers of the original error, which will help to decrease the computational complexity (Tabl. 3).

Table 3. **The number of point patterns depending on the matrix of t**

The number of patterns of square matrix t	The reduced number of generating patterns					
	2	3	4	5	6	7
Reduced patterns A_{red}	2	4	10	20	51	119
Patterns set A_{pat}	3	6	16	34	90	214

Due to the limited space, reduced set A_{red} from $t=4$ to $t=5$ are listed in the Tabl. 4, those patterns whose transpose is equal to themselves are marked by the blue shadow.

Table 4. **The library of reduced point patterns from $t=4$ to $t=5$**

$t=4$ and $t=5$																																																													
4.1	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td></td><td></td><td></td><td></td></tr></table>	1	1	1	1					4.2	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td></td></tr><tr><td>1</td><td></td><td></td><td></td></tr></table>	1	1	1		1				4.3	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td>1</td></tr></table>	1	1	1					1																																
1	1	1	1																																																										
1	1	1																																																											
1																																																													
1	1	1																																																											
			1																																																										
4.4	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td></td><td></td></tr><tr><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>	1	1			1	1							4.5	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td></td><td></td></tr><tr><td>1</td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>	1	1			1		1						4.6	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td></tr></table>	1	1				1					1																					
1	1																																																												
1	1																																																												
1	1																																																												
1		1																																																											
1	1																																																												
	1																																																												
		1																																																											
4.7	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td>1</td></tr><tr><td></td><td></td><td></td><td></td></tr></table>	1	1					1	1					4.8	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td>1</td><td></td></tr></table>	1	1					1				1		4.9	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td>1</td></tr></table>	1	1					1					1																				
1	1																																																												
		1	1																																																										
1	1																																																												
		1																																																											
		1																																																											
1	1																																																												
		1																																																											
			1																																																										
4.10	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td></td><td></td><td></td></tr><tr><td></td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td>1</td></tr></table>	1					1					1					1	5.1	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>	1	1	1	1	1																5.2	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td>1</td><td></td></tr><tr><td>1</td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>	1	1	1	1		1														
1																																																													
	1																																																												
		1																																																											
			1																																																										
1	1	1	1	1																																																									
1	1	1	1																																																										
1																																																													
5.3	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>1</td></tr></table>	1	1	1	1						1	5.4	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td></td><td></td></tr><tr><td>1</td><td>1</td><td></td><td></td><td></td></tr></table>	1	1	1			1	1				5.5	<table border="1" style="display: inline-table; text-align: center;"><tr><td>1</td><td>1</td><td>1</td><td></td><td></td></tr><tr><td>1</td><td></td><td></td><td>1</td><td></td></tr></table>	1	1	1			1			1																											
1	1	1	1																																																										
				1																																																									
1	1	1																																																											
1	1																																																												
1	1	1																																																											
1			1																																																										

Extension of the Tabl. 4

5.6		5.7		5.8	
5.9		5.10		5.11	
5.12		5.13		5.14	
5.15		5.16		5.17	
5.18		5.19		5.20	

From the Tabl. 4 select the patterns with unit-pixel connect in the Tabl. 5.

Table 5. The patterns with unit-pixel connect from $t=3$ to $t=5$

$t=3$					
3.1		3.3		3.4	
$t=4$ and $t=5$					
4.1		4.3		4.7	
4.8		4.9		4.10	

Extension of the Tabl. 5

5.1	<table border="1"><tr><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>	1	1	1	1	1						5.3	<table border="1"><tr><td>1</td><td>1</td><td>1</td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>1</td></tr></table>	1	1	1	1						1	5.9	<table border="1"><tr><td>1</td><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td></td><td>1</td><td>1</td></tr></table>	1	1	1						1	1																																													
1	1	1	1	1																																																																												
1	1	1	1																																																																													
				1																																																																												
1	1	1																																																																														
			1	1																																																																												
5.10	<table border="1"><tr><td>1</td><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td>1</td><td></td></tr></table>	1	1	1						1					1		5.11	<table border="1"><tr><td>1</td><td>1</td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>1</td></tr></table>	1	1	1						1						1	5.17	<table border="1"><tr><td>1</td><td>1</td><td></td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>1</td></tr></table>	1	1						1	1						1																														
1	1	1																																																																														
			1																																																																													
			1																																																																													
1	1	1																																																																														
			1																																																																													
				1																																																																												
1	1																																																																															
		1	1																																																																													
				1																																																																												
5.18	<table border="1"><tr><td>1</td><td>1</td><td></td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>	1	1						1					1						1							5.19	<table border="1"><tr><td>1</td><td>1</td><td></td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>1</td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>	1	1						1						1						1						5.20	<table border="1"><tr><td>1</td><td></td><td></td><td></td><td></td></tr><tr><td></td><td>1</td><td></td><td></td><td></td></tr><tr><td></td><td></td><td>1</td><td></td><td></td></tr><tr><td></td><td></td><td></td><td>1</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>1</td></tr></table>	1						1						1						1						1
1	1																																																																															
		1																																																																														
		1																																																																														
			1																																																																													
1	1																																																																															
		1																																																																														
			1																																																																													
				1																																																																												
1																																																																																
	1																																																																															
		1																																																																														
			1																																																																													
				1																																																																												

Conclusion

In this paper, we analyze methods for generating the library of point patterns, all the possible patterns are accessible in our library which implies coding the image in the future is feasible. The improved method has reduced the time of generating all the possible patterns. We select the patterns with unit-pixel connect in our library and in the next stage we will code the binary skeleton image by using parts of our library.

References

1. Klein F. // Comparative review of the latest geometric research (Erlangen program) /on the foundations of geometry: Collection of classical works on Lobachevsky's geometry and the development of its ideas. M.: 1956. P. 399–434.
2. Hardy G.Y, Wright E.M. An introduction to the theory of Numbers. Berlin. 1973.
3. Konopelko V.K., Lipnitsky V.A, Spichekova N.V. // Dokl. BGUIR. M.: BSUIR, 2010. P. 112–117.
4. Konopelko V.K., Lipnitsky V.A, Spichekova N.V. // Dokl. BGUIR. M.: BSUIR, 2010. P. 40–46.
5. Konopelko V.K., Lipnitsky V.A, Spichekova N.V. // Dokl. BGUIR. M.: BSUIR, 2011. P. 17–25.
6. Konopelko V.K., Smolyakova O.G. // Dokl. BGUIR. M.: BSUIR, 2008. P. 19–28.
7. Konopelko V.K., Smolyakova O.G. // Dokl. BGUIR. M.: BSUIR, 2008. P. 142–153.

УДК 004.89

MODEL AND STRUCTURE OF MULTIAGENT SYSTEM FOR COLLECTION AND PROCESSING SOUND INFORMATION

B.H. SHAYA, U.A. VISHNIAKOU

*Belarusian state university of informatics and radioelectronics, Republic of Belarus**Submitted 01 November 2019*

Abstract. A symbolic model of a multi-agent system (MAS) for collecting and processing sound information (CPSI) from the environment was proposed. The model includes agents of input and encoding of local sound information, database agents, knowledge base agents, agents of calculation of integral estimates of the sound situation, decision-making agent. On the basis of this model the structure of the MAC for CPSI consisting of preprocessors for audio input and encoding, wireless communication channels, server and operator console was developed

Keywords: sound preprocessor, databases and knowledge, server, multi-agent system.

Introduction

Collecting and processing the various types of sounds and their spatio-temporal effects becomes crucial. Using a model that embeds the representation of sounds properties on sources, and the rules that govern their propagation across the various surrounding mediums would help in both, tracking the historical variations and predicting the future changes of sounds properties along the spatio-temporal dimensions. In fact, such a model can represent levels of noise in a large urban space and help in studying noise pollution at various layers: inside a given building, in a specific public park or around the whole city. It will also help in predicting how spatiotemporal changes may affect the levels of noise pollution at any of these layers, for instance when a new building complex or a compound community take place in the city.

In fact, the amplitude and effect of sound waves vary considerably across the continuous spatio-temporal dimensions. For instance, the noise produced by a taking-off airplane is perceived by its neighborhood with varying amplitudes over time: it starts loud, and then decreases gradually while it is flying away. On the other hand, data modeling and especially spatio-temporal data is an important phase that gives the users a clear way to understand a case of study and to make a decision [1–3].

Existing approaches to sound processing

The most important classification methods for information processing use Hidden Markov Models (HMM), Gaussian Mixture Models (GMM) and Support Vector Machines (SVM), which will be discussed in below, although there are other useful methods that are summarized as follows [4]:

– *k*-Nearest Neighbours (*k*-NN): a simple algorithm that, given a testing pattern, uses the majority vote of the *k*-nearest training patterns to assign a class label. It is often described as a lazy algorithm, as all computation is deferred to testing, and hence can have a slow performance for a large number of training samples. For the case of the 1-NN, the method has a 100 % recall performance, which is unique.

– Dynamic Time Warping (DTW): the algorithm can find the similarity between two sequences, which may vary in time or speed. It works well with the *bag-of-frames* approach, as it can decode

the same word spoken at different speeds. However, it has largely been superseded by HMM for Automatic Speech Recognition (ASR).

– Artificial Neural Networks (ANN): the method, also referred to as a Multi-Layer Perceptron (MLP), is a computational model inspired by neurons in the brain. Given a sufficient number of hidden neurons, it is known as universal approximator, but is often criticized for being a black-box, as the function of each neuron in the network is hard to interpret. It also suffers from difficulty in training, as the most common method of back propagation is likely to get stuck in local minima.

There are many features that can be used to describe audio signals. The feature vector for the experiments consisted of features summarized in [5, 6].

Three different classification methods were investigated: k -Nearest Neighbours (k -NN) [7], Gaussian Mixture Models (GMM) [8], and Support Vector Machine (SVM) [9]. For k -NN, we used the Euclidean distance as the distance measure and the 1-nearest neighbor queries to obtain the results. As for GMM, we set the number of mixtures for both training and testing to 5. For the SVM classifiers, we used a 2-degree polynomial as its kernel with regularization parameter $C=10$ and the epsilon $\varepsilon = 1e^{-7}$, which controls the width of the ε -insensitive zone, which used to fit the training data, affecting the number of support vectors used.

Multiagent approach and offering model

Multi-agent systems (MASs) have been usually accepted and imbedded in deferent domains applications because of the benefits and advantages which they can offer. Some of them that offered by using Multi-agent systems (MASs) in large systems are [10]:

1. An enhance in the effectiveness and process speed due to parallel computation and asynchronous operation.
2. In case of system degradation (when one or more of the agents fail), common reliability and robustness of the system is acceptable.
3. Scalability and flexibility – agents can be added as required.
4. Reduced cost – individual agents cost much less than a centralized architecture.
5. Reusability – agents have a modular structure and they can be easily replaced in other systems or be upgraded more easily than a monolithic system.

Let's consider the authors' approach to the symbolic description of a multi-agent model for collecting and processing sound information (CPSI) from the environment, based on the concept of object algebra [11]. Let's present this model as a set of agents:

$$MAS_{CPSI} = (A_{IP}, A_{FC}, A_{IDB}, A_{CM}, A_{KB}, A_{DS}, A_{UI}),$$

where A_{IP} – agent of input and preprocessing of sound information, A_{FC} – filtering and classification agent, A_{IDB} – agent working with data base of sound information, A_{CM} – conceptual modeling agent, A_{KB} – agent working with knowledge base of sound information classification, A_{SD} – agent for of decision support (DS) making about sound situation, A_{UI} – agent of user interface. In general, this model works including several algorithms. In the following, the main ones will be discussed.

The first algorithm starts from work of many agents A_{IP} which input and collect local sound information in some region and send it to the A_{FC} agent which will also get information from the A_{CM} agent and then send it to the agent A_{DS} . It works with agents A_{IDB} , A_{KB} , A_{UI} (data and knowledge bases, DS and user interface). The second algorithm have the same cycle as the first one, but in this case if there is some sound information missing so the A_{DS} agent will get back to the A_{IP} agent to get necessary information, and then the cycle will complete as the first one to get a better result.

The third algorithm is: before the A_{CM} agent give information to the A_{FC} agent it will communicate with the A_{IP} agent and get the required information.

More details of working these agents will be writing in further author publications.

MAS Structure for CPSI

The structure of multi-agent system for Collection and Processing Sound Information (CPSI) is composed of several different components that works together to collect sounds from the environment to get a required output (Fig. 1). These components can be classified in two parts. The first part is detecting and collecting sounds processors from the environment (A_{ip} model agents), while the other part is software-hardware realization on Server ($A_{FC}, A_{IDB}, A_{CM}, A_{KB}, A_{DS}, A_{UI}$ model agents).

The detecting or collecting processor is composed of detectors or sound sensors (SD) so that the number of detectors will take in to consideration the zone or place that will be covered so that we will insure that we collect all the sound waves from a specified environment or zone.

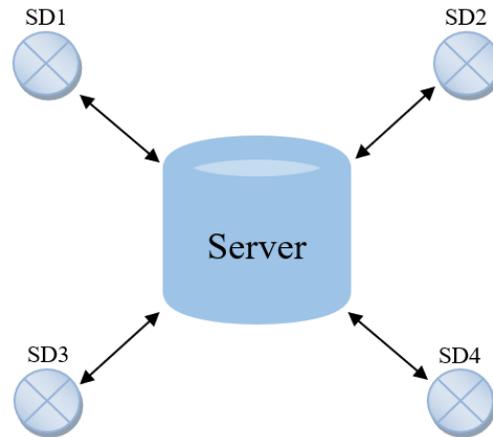


Fig. 1. MAS for Collection and Processing Sound Information (CPSI)

These sensors (SD) will be connected to the Server by the usage of wireless connection, so that we will have the flexibility to relocate the sensors depending on the zone that it is necessary to cover, and at the same time to insure that these detectors can be dynamic detectors not static, and also be flexible with different types of environment.

These components of MAS CPSI realize the model special algorithms which were given above. Common algorithm of MAS will allow specialist to select a specific sound length or sound range. If user wants to collect not only this but also avoid duplication in sound collecting so that if two or more detectors (SD) collect the same sound source from the environment and send it to the Server, it will drop all the common sound waves to decrease the data that will be studied and select the most clear data that was collected with the help of special methods and sensors that are provided by the second part of these agents as shown at Fig. 1.

Conclusion

Authors present the model and structure of MAS for CPSI. As a perspective of this work one will focus on each agent and will do each component of hardware-software implementation, a solution that would be practical for the end user. The detailed plan includes:

1. To create a hierarchy for conceptual representation and a complete set of pictograms to cover all noise types of sound.
2. In the transformation agent will focus on decision tree creation in order to swap between scales.
3. To regard the specific types of sound different types of detectors will propose the suggestion for the related sound detectors.
4. The last will be used as GUI graphical user interface that create automatic report for the end user.

References

1. Miralles A., Libourel T. // Encyclopedia of GIS. 2008. P. 700–705.
2. Moisc B., Gensel J., Davoine P., Martin H. // W2GIS. 2006. P. 146–157.
3. Parent C., Spaccapietra S., Zimányi E. // Encyclopedia of Database Systems. 2009. P. 2571–2576.
4. Pelekis N., Theodoulidis B., Kopanakis I., Theodoridis Y. // Knowledge Engineering Review. 2004. Vol. 19 (3). P. 235–274.
5. Visniakou U.A., Shaya B.H. // Science among us. 2019. Vol. 10 (26) P. 1–10.
6. Mitchell H.A., Schaefer P.A. // Int. Journal of Intelligent Systems, April 2001. P. 459–468.
7. Moore A.W. Clustering with gaussian mixtures. [Electronic resource]. URL: <http://www.autonlab.org/tutorials/gmm.html>. (date of access: 02.02.2019).
8. Boser B.E., Guyon I.M., Vapnik V.N. // 5th Annual ACM Workshop on COLT. Pittsburgh, PA. 1992. P. 144–152.
9. Stone P., Veloso M. // Autonomous Robotics. 2000. Vol. 8, № 3. P. 1–56.
10. Vishniakou U.A. Information security in corporate systems, e-Commerce and cloud computing: methods, models, software and hardware solutions. Monograph. / U.A. Vishniakou. – Minsk: Bestprint, 2016.

УДК 621.391

ВОЛНОВОЕ ВЫРАЩИВАНИЕ ОБЛАСТЕЙ ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ ИЗОБРАЖЕНИЙ

А.Т. НГУЕН, Т.Х. ДОАН, В.Ю. ЦВЕТКОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 10 ноября 2019*

Аннотация. Предложен алгоритм сегментации полутоновых изображений на основе квазипараллельного волнового выращивания областей локальных экстремумов с постепенным изменением порога яркости от значения экстремума. В отличие от известных алгоритмов сегментации предложенный алгоритм позволяет разделять области с плавными перепадами яркости, адаптироваться к ограниченному времени сегментации и контролировать число сегментов, достаточное для аппроксимации изображений и компактного описания их сегментов.

Ключевые слова: локальный экстремум, сегментация, волновое выращивание областей.

Введение

Сегментация является одной из базовых процедур цифровой обработки изображений, лежащих в основе их анализа, визуализации и объектно-ориентированного кодирования [1]. Сегментация используется для обнаружения местоположения объектов и границ в задачах визуализации медицинских изображений, поиска, классификации, распознавания и сопровождении объектов на изображениях. Сегментация обеспечивает присвоение номера каждому пикселю таким образом, что пиксели с одинаковыми номерами имеют общие визуальные характеристики. Это приводит к разделению изображения на области, которые соответствуют различным объектам или их частям [2, 3].

Точность сегментации определяет качество результатов последующей обработки. В некоторых случаях время сегментации может быть ограничено, либо необходимо контролировать число сегментов изображения [4]. К основным характеристикам методов сегментации изображений относятся время, ошибки и компактность представления результатов сегментации.

Время сегментации является основным фактором, определяющим эффективность методов сегментации. В зависимости от размера, сложности изображения, выбранного метода сегментации и вычислительных ресурсов оно составляет от долей секунды до нескольких минут.

Ошибки сегментации проявляются в точности и стабильности локализации областей при изменении условий видеорегистрации, приводящих к изменению яркости, контрастности, смещению и повороту изображения [5]. Основной причиной ошибок методов сегментации в реальных условиях является неравномерность освещения сцены, возникающая из-за нестабильности источника света, неравномерного распределения света по поверхности объекта (особенно крупного), невозможности оптической изоляции объекта от тени других объектов [6].

Учет компактности представления результатов сегментации [7] позволяет оценивать эффективность методов в плане требований к вычислительным ресурсам.

Сегментация на основе квантования по гистограмме не обеспечивает точное разделение областей из-за присвоения одинаковых номеров сегментам с одинаковой яркостью. Кроме того, известные методы сегментации, основанные на формировании областей с использованием водораздела [8–11], квантовании по гистограмме [12], разделении и слиянии областей

с использованием квадрата-дерева [13–15], выращивании областей [16–19], не эффективны для разделения областей с плавными перепадами яркости. Все рассмотренные методы не обеспечивают адаптацию к ограничениям на время сегментации и не позволяют контролировать число сегментов. В этой связи актуальна задача разработки метода сегментации изображений, учитывающего перечисленные недостатки.

Целью работы является разработка метода сегментации изображений, позволяющего разделять области с плавными перепадами яркости, адаптироваться к ограничениям на время сегментации и контролировать число сегментов для их компактного описания в задачах аппроксимации и компрессии изображений.

Метод сегментации изображений на основе встречно-волнового выращивания областей

Предлагается алгоритм сегментации полутоновых изображений на основе встречно-волнового выращивания областей локальных экстремумов. Сущность алгоритма состоит в выделении на изображении локальных экстремумов, постепенном присоединении к ним новых элементов с учетом их местоположения вокруг экстремумов, значения порога, пошагово-изменяемого от значения экстремума в противоположном направлении во избежание блокировки процесса волнового выращивания. Процесс встречно-волнового выращивания продолжается до тех пор, пока все области не будут выделены.

Отличие предложенного алгоритма от известного алгоритма выращивания областей заключается в квазипараллельном увеличении размеров выделенных областей вокруг максимумов и минимумов, что позволяет повысить точность сегментации изображений с плавными перепадами яркости. Предложенный алгоритм отличается от алгоритма волнового выращивания областей [20, 21] и регрессивного волнового выращивания областей [22, 23] выбором начальных точек роста областей в результате поиска локальных экстремумов изображения [24, 25], и увеличением их областей за счет постепенно присоединения к ним соседних элементов с учетом значений порогов, изменяемых постепенно на интервале между смежными локальными экстремумами противоположных знаков.

Алгоритм сегментации изображений на основе встречно-волнового выращивания областей экстремумов состоит из следующих шагов.

1. Инициализация. На данном шаге осуществляется буферизация изображения $I = \|i(y, x)\|_{(y=0, Y-1, x=0, X-1)}$ размером YX пикселей и определение значения порога $\Delta_T = [1, 255]$ и шаг изменения порога $\Delta_S = [0, 254]$, определяющих условие и скорость присоединения пикселя к выращиваемой области. Формируются матрица сегментации $SM = \|sm(y, x)\|_{(y=0, Y-1, x=0, X-1)}$, элементы которой являются результатом сегментного поиска или блочно-сегментного поиска локальных экстремумов изображения (положительные номера – максимумы, отрицательные номера – минимумы и нулевые номера – неэкстремальные пиксели), стеки координат экстремумов в качестве стеков начальных точек роста $SY = \|sy(k)\|_{(k=0, YX-1)}$ и $SX = \|sx(k)\|_{(k=0, YX-1)}$ ($SY = \|sy(k)\|_{(k=K, YX-1)} \leftarrow 0$, $SX = \|sx(k)\|_{(k=K, YX-1)} \leftarrow 0$) и стеки $SE = \|se(n)\|_{(n=0, N-1)}$ значений яркости экстремумов, где K – число экстремальных пикселей, N – число экстремумов изображения и $K \geq N$ [24, 25].

Инициализация счетчика сканирования текущих выращиваемых пикселей $k \leftarrow 0$.

Инициализация счетчика пошагового повышения значения порога $N_T \leftarrow 1$.

2. Начало цикла сегментации.

2.1 Начало цикла пошагового повышения значения порога яркости.

Если значение указателя счетчика N_T больше нуля ($N_T > 0$), то присваивается значение 0 счетчику N_T ($N_T \leftarrow 0$), переход на следующий шаг, иначе ($N_T = 0$) – переход на шаг 2.9.

2.2 Начало цикла сканирования текущих выращиваемых пикселей.

Если последний пиксель стека роста области не обрабатывается ($k \leq K$), то переход на следующий шаг, иначе ($k > K$) – переход на шаг 2.8.

2.3 На основе координат текущего выращиваемого пикселя (y_E, x_E) указателя k вычисляются координаты (y_N, x_N) текущего окрестного пикселя:

$$\begin{aligned} y_N &= y_E + j, \\ x_N &= x_E + i, \end{aligned} \quad (1)$$

где $(j \in \{-1, 0, 1\}) \wedge (i \in \{-1, 0, 1\}) \setminus (j = 0 \wedge i = 0)$, $y_E = sy(k)$, $x_E = sx(k)$.

Проверяется на ноль значение элемента матрицы сегментации, координаты которого соответствуют найденным координатам (y_N, x_N) окрестного пикселя. Если это значение не равно нулю ($sm(y_N, x_N) \neq 0$), то переход на следующий окрестный пиксель, иначе ($sm(y_N, x_N) = 0$) – переход на следующий шаг.

2.4 Абсолютное значение разности значений текущего выращиваемого экстремума $I_E = se(|sm(y_E, x_E)|)$ и текущего окрестного пикселя $i(y_N, x_N)$ сравниваются с заданным порогом Δ_T . Если абсолютное значение разности меньше или равно порогу ($|I_E - i(y_N, x_N)| \leq \Delta_T$ – окрестный пиксель должен быть присоединен к областям), то переход на следующий шаг, иначе ($|I_E - i(y_N, x_N)| > \Delta_T$) – переход на шаг 2.7.

2.5 Элементу матрицы сегментации, координаты которого соответствуют координатам (y_N, x_N) текущего окрестного пикселя, присваивается значение элемента матрицы сегментации, координаты которого соответствуют координатам (y_E, x_E) текущего выращиваемого пикселя $sm(y_N, x_N) \leftarrow sm(y_E, x_E)$.

2.6 Указатель текущего числа выращиваемых пикселей K стеков роста области увеличивается на единицу. В стеки роста области заносятся координаты текущего окрестного пикселя:

$$\begin{aligned} sy(K) &\leftarrow y_N, \\ sx(K) &\leftarrow x_N, \\ K &\leftarrow K + 1. \end{aligned} \quad (2)$$

2.7 Указатель N_T счетчика пошагового повышения значения порога яркости увеличивается на единицу. В стеки роста области заносятся координаты текущего окрестного пикселя:

$$\begin{aligned} N_T &\leftarrow N_T + 1, \\ sy(N_T) &\leftarrow y_N, \\ sx(N_T) &\leftarrow x_N. \end{aligned} \quad (3)$$

2.8 Проверка указателя счетчика сканирования текущих выращиваемых пикселей на суммарное число выращиваемых пикселей K . Если указатель меньше или равно K ($k \leq K$), то переход на шаг 2.2 и увеличивается указатель k на единицу ($k \leftarrow k + 1$), иначе ($k > K$) – переход на следующий шаг.

2.9 Проверка указателя счетчика пошагового повышения значения порога яркости на ноль. Если указатель больше нуля, то присваиваются единица указателю k ($k \leftarrow 1$), значение N_T указателю K ($K \leftarrow N_T$) и увеличивается порог Δ_T на шаг Δ_S ($\Delta_T \leftarrow \Delta_T + \Delta_S$), иначе ($N_T = 0$) – выход из алгоритма.

В результате выполнения данного алгоритма формируется матрица сегментации, значение каждого элемента которой указывает на номер сегмента, которому принадлежит пиксель сегментируемого изображения с соответствующими координатами. С каждым циклом размеры сегментов постепенно увеличиваются, в чем проявляется прогрессивный характер сегментации. Число полученных сегментов совпадает с числом локальных экстремумов

изображения, что позволяет контролировать это число, а также местоположение и размеры сегментов.

Оценка эффективности алгоритмов сегментации изображений

Для оценки эффективности методов сегментации на рис. 1 показаны четыре тестовых полутоновых изображения: Image1 и Image2 с плавно изменяемой яркостью, низкочастотное изображение Lena и высокочастотное изображение City размером 256×256 пикселей.



Рис. 1. Тестовые полутоновые изображения: *a* – Image1; *б* – Image2; *в* – Lena; *г* – City

На рис. 2–5 показаны результаты сегментации вышеуказанных полутоновых изображения с помощью предложенного метода, методов сегментации на основе выращивания областей, разделения и слияния областей, водораздела на основе градиентов.

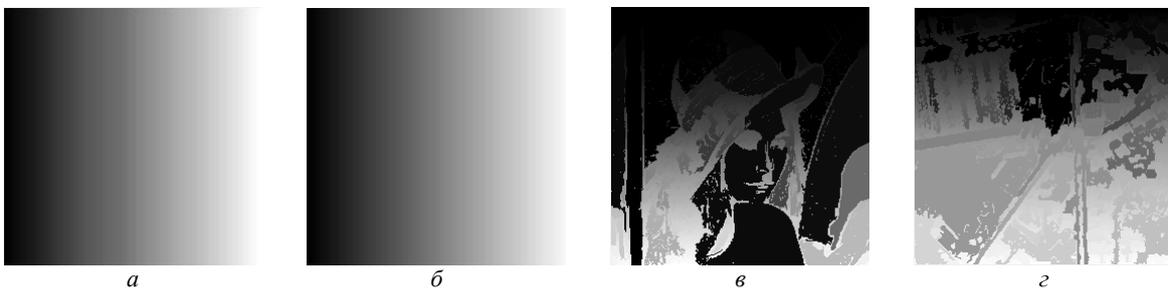


Рис. 2. Результат применения алгоритма выращивания областей: *a* – 152 сегментов ($\Delta_T = 0$);
б – 145 сегментов ($\Delta_T = 0$); *в* – 6988 сегментов ($\Delta_T = 5$); *г* – 9323 сегментов ($\Delta_T = 5$)

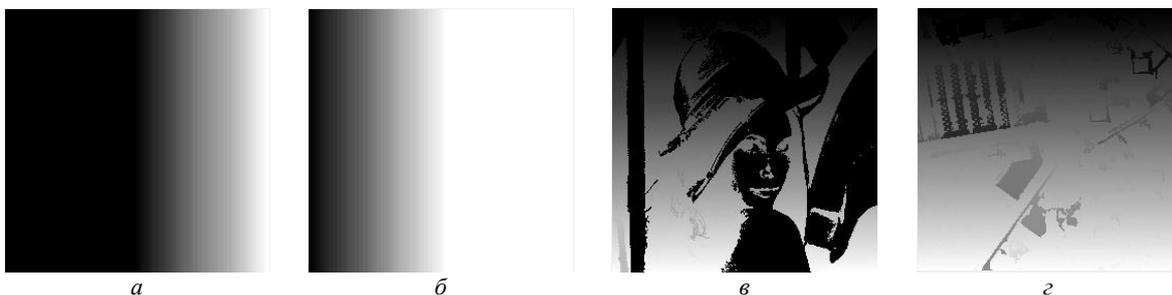


Рис. 3. Результат применения алгоритма разделения и слияния областей: *a* – 79 сегментов;
б – 77 сегментов; *в* – 25753 сегментов; *г* – 40081 сегментов

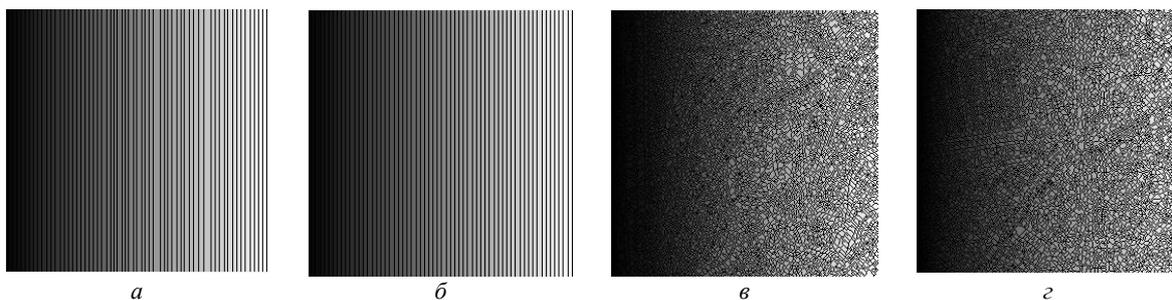


Рис. 4. Результат применения алгоритма водораздела на основе градиентов: *a* – 62 сегментов;
б – 62 сегментов; *в* – 8728 сегментов; *г* – 8003 сегментов

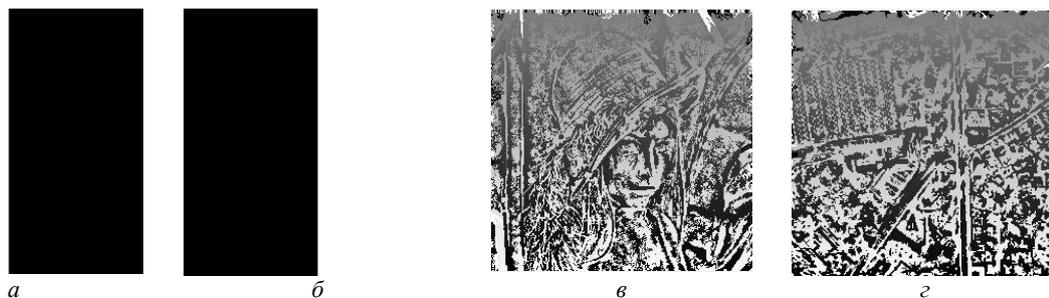


Рис. 5. Результат разработанного метода: *а* – 2 сегмента ($\Delta_T = 1, \Delta_S = 1$); *б* – 2 сегмента ($\Delta_T = 1, \Delta_S = 1$); *в* – 7507 сегментов ($\Delta_T = 1, \Delta_S = 1$); *г* – 4830 сегментов ($\Delta_T = 1, \Delta_S = 1$)

Из рис. 2–5 видно, что предложенный алгоритм обеспечивает стабильное положение границы и выделение 2-х областей при любом значении перепада яркости (рис. 5, *а, б*). Остальные алгоритмы при некоторых значениях перепада яркости показывают ошибку или избыточность сегментации, выделяя на изображении много избыточных областей. Таким образом, предложенный алгоритм обеспечивает повышение чувствительности сегментации к перепадам яркости, формы полученных сегментов которой не являются сложными и зависят от местоположений, форм и расстояний между экстремумами на изображении, что позволяют компактно их описать для последующей обработки. Такое компактное описание сегментов представляет интерес для распознавания и сжатия изображений.

Произведена оценка эффективности предложенного алгоритма и известных алгоритмов сегментации. В качестве показателей эффективности использованы время сегментации, стабильность границ и числа сегментов.

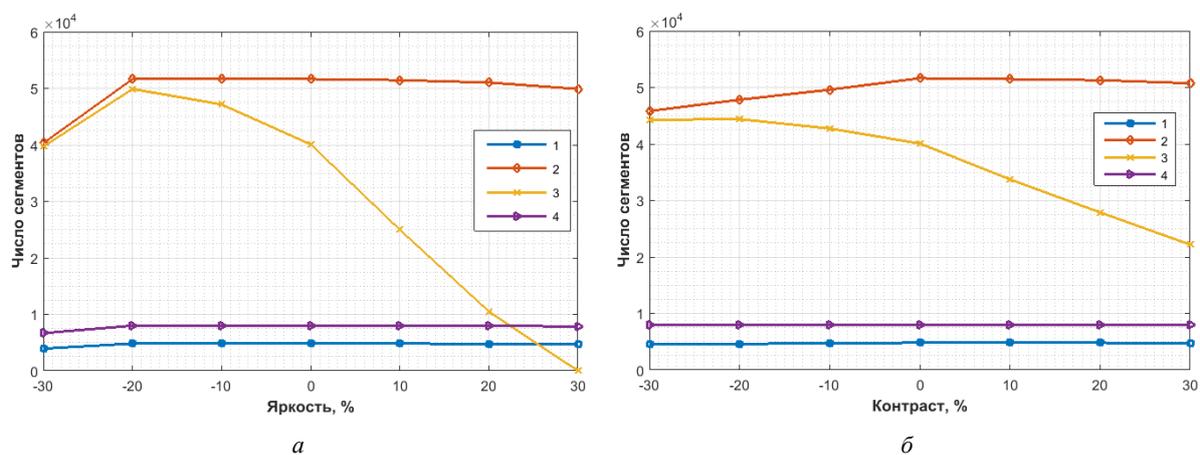


Рис. 6. Зависимости числа сегментов от изменения условий формирования изображения City (рис. 1, *г*): *а* – при изменении яркости; *б* – при изменении контраста (1 – предложенный метод; 2 – выращивания областей; 3 – разделения и слияния областей; 4 – водораздела на основе градиентов)

На рис. 6 приведены зависимости числа сегментов от изменения яркости и контраста, характеризующие устойчивость результатов сегментов. Устойчивость оценивается по отношению числа сегментов для базового изображения к числу сегментов для модифицированного изображения, подвергнутого изменению яркости и контраста. Установлено, что предложенный алгоритм выигрывает в стабильности числа сегментов при изменении яркости, контраста по сравнению с алгоритмами сегментации на основе выращивания областей, разделения и влияния областей, в уменьшении избыточности сегментов по сравнению с алгоритмом сегментации на основе водораздела.

В табл. 1 приведены результаты оценки времени сегментации для рассматриваемых алгоритмов в среде Matlab и Matlab ®Image Processing Toolbox (R2015a), в системе Intel Core i3 3,1 ГГц с 6 ГБ ОЗУ. Для этого эксперимента использованы четыре изображения размерами 128×128, 256×256, 512×512 и 1024×1024 пикселей.

Таблица 1. **Время сегментации тестовых изображений**

№	Алгоритмы сегментации	Язык программирования	Время сегментации, с			
			Lena 128×128	Barbara 256×256	France 512×512	Man 1024×1024
1	Предложенный	Matlab	0,067	0,247	0,937	3,374
2	Выращивания областей	Matlab	0,084	0,291	1,115	4,422
3	Разделения и слияния областей	C/C++	0,029	0,077	0,311	1,246
4	Водораздела на основе градиентов	C/C++	0,017	0,043	0,162	0,917

Из табл. 1 следует, что предложенный алгоритм выигрывает в скорости сегментации по сравнению с алгоритмом выращивания областей при любом размере изображений, но проигрывает в скорости сегментации по сравнению с алгоритмами разделения и слияния областей, водораздела на основе градиентов, написанными на языке программирования C/C++. Скорость обработки предложенного алгоритма улучшается в несколько раз при реализации на языке программирования C++.

Заключение

Для сегментации изображений предложен алгоритм на основе встречно-волнового выращивания областей локальных экстремумов. Сущность алгоритма заключается во встречном квазипараллельном выращивании областей вокруг локальных минимумов и максимумов, выбранных в качестве начальных точек роста. Это обеспечивает автоматическое разделение областей с плавным перепадом яркости, которые известные алгоритмы сегментируют с ошибками. Показано, что предложенный алгоритм позволяет четко выделять сегменты и контролировать их количество по сравнению с известными алгоритмами сегментации. Установлено, что предложенный алгоритм выигрывает в скорости сегментации по сравнению с алгоритмом выращивания областей при любых размерах изображений и в стабильности числа сегментов по сравнению с известными алгоритмами. Для уменьшения избыточности и повышения устойчивости сегментации к шумам в качестве предварительной обработки может быть использована низкочастотная фильтрация.

WAVE GROWING OF LOCAL EXTREME REGIONS OF IMAGES

A.T. NGUYEN, T.H. DOAN, V.Yu. TSVIATKOU

Abstract. An algorithm for the segmentation of grayscale images based on quasi-parallel wave growth of local extrema with a gradual change in the brightness threshold from the extremum value was proposed. Unlike well-known segmentation algorithms, the proposed algorithm allows to separate areas with smooth differences in brightness, adapt to a limited segmentation time and control the number of segments sufficient to approximate images and compactly describe their segments.

Keywords: local extrema, image segmentation, wave region growing, counter-wave region growing.

Список литературы

1. Milan S. [et al.] // Thomson press, west. 2008. P. 175–240.
2. Chijindu V.C., Inyama H.C., Uzedhe G. // African Journal of Computing & ICT. 2012. Vol. 5, № 5. P. 90–98.
3. Shivhare P., Gupta V. // International Journal of Engineering and Advanced Technology. 2015. Vol. 4. P. 153–157.
4. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в среде MATLAB. М. Техносфера. 2006.
5. Fabijańska A., Strzecha K., Sankowski D. // CADSM'2007, Polyana, Ukraine, 20–24 February, 2007. P. 439–441.
6. Chandrakala M., Devi P.D. // International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE). 2016. Vol. 5. P. 163–168.
7. Гладырева А. Ю. // Автоматика, Львів, Україна, 28–30 вересня 2011. 2011. С. 348–349.
8. Lalitha M., Kiruthiga M., Loganathan C. // International Journal of Science and Research (IJSR). 2013. Vol. 2, № 2. P. 348–358.
9. Gauch J.M. // IEEE transactions on image processing. 1999. Vol. 8, № 1. P. 69–79.
10. Khiyal M.S.H., Khan A., Bibi A. // Informing Science and Information Technology. 2009. Vol. 6. P. 876–886.
11. Arindrajit Seal, Arunava Das, Prasad Sen // International Journal of Computer Science and Information Technologies (IJCSIT). 2015. Vol. 6, № 3. P. 2295–2297.
12. Chang J.H., Fan K.Ch., Chang Y.L. // Image and Vision Computing. 2002. Vol. 20. P. 203–216.
13. Muhsin Z.F., [et. al.] // The Imaging Science Journal. 2014. Vol. 62, № 1. P. 56–62.
14. Xiaolin Wu // IEEE transactions on pattern analysis and machine intelligence. 1993. Vol. 15, № 8. P. 808–815.
15. Dass R., Priyanka, Devi S. // International Journal of Electronics & Communication Technology (IJECT). 2012. Vol. 3, Issue 1. P. 66–70.
16. Singh K.K., Singh A. // International Journal of Computer Science Issues. 2010. Vol. 7, № 5. P. 414–417.
17. Shih F. Y., Cheng S. // Image and Vision Computing. 2005. № 23. P. 877–886.
18. Sharma Ritu, Sharma Rajesh // International Journal of Innovative Research in Computer and Communication Engineering. 2014. Vol. 2, Issue 9. P. 5686–5692.
19. Mohd Saad N., [et. al.] // Proceedings of the International MultiConference of Engineers and Computer Scientists. Hong Kong, 14–16 March 2012. P. 674–677.
20. Альмияхи О.М., Конопелько В.К. // Телекоммуникации: сети и технологии, алгебраическое кодирование и безопасность данных: материалы международного научно-технического семинара. Минск, БГУИР. 2015. С. 61–67.
21. Альмияхи О.М., Цветков В.Ю., Конопелько В.К. // Докл. БГУИР. 2016 № 3 (97). С. 24–30.
22. Рабцевич В.В., Цветков В.Ю., Ловецкий М. Ю. // Мониторинг техногенных и природных объектов: сб. материалов междунар. научн. -техн. конф. Минск, БГУИР. 2017. С. 77–84.
23. Рабцевич В.В., Нгуен А.Т., Цветков В.Ю. // BIG DATA Advanced Analytics: collection of materials of the fourth international scientific and practical conference. Minsk, BSUIR. 2018. С. 373–377.
24. Nguyen A.T, Tsviatkou V.Yu. // International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE). 2019. Vol. 8, № 9. P. 1–10.
25. Цветков В.Ю., Нгуен А.Т. // Информатика. 2019. 16 (3). С. 23–36.

УДК 681.3

ЦИФРОВАЯ ПОДПИСЬ ВЕБ-СЕРВИСА ПРОТОКОЛА TLS НА ОСНОВЕ ЭЛЛИПТИЧЕСКИХ КРИВЫХ И ОБЛАЧНОЙ ПЛАТФОРМЫ HEROKU

Н.С. ФИЛИППОВ, С.Б. САЛОМАТИН

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 02 ноября 2019*

Аннотация. Рассмотрены криптографические алгоритмы цифровой подписи на эллиптических кривых в контексте «Интернет вещей», а также расширение возможностей стандартного протокола защиты транспортного уровня TLS на основе устройства *raspberry pi* и облачного протокола *Heroku*.

Ключевые слова: криптографическая защита информации, эллиптическая кривая, протокол TLS, интернет вещей, веб-сервис.

Введение

Сети «Интернет вещей» (IoT) используют симметричные и асимметричные криптографические алгоритмы для защиты информации [1–3]. В качестве сенсоров часто используются устройства типа *raspberry pi*, имеющие малое потребление энергии, ограниченную пропускную способность и вычислительную мощность.

Криптографические алгоритмы на эллиптических кривых протокола TLS реализуются достаточно сложным образом, что приводит к значительным накладным расходам по отношению к времени выполнения, энергопотреблению и затрудняет их использование в устройствах *raspberry pi*, большая часть ресурсов которых занята прикладным программным обеспечением, оставляя очень ограниченное пространство для протоколов безопасности.

Одним из путей решения задачи по расширению возможностей протокола TLS на эллиптических кривых в контексте сетей IoT – *raspberry pi* связан с привлечением облачных технологий и, в частности, платформы *Heroku*.

Веб-сервис цифровой подписи TLS IoT на базе алгоритма ECDSA

В работе рассматривается реализация веб-сервиса цифровой подписи на базе алгоритма *ECDSA (Elliptic Curve Digital Signature Algorithm)* с использованием языка программирования *Python* и облачной платформы *Heroku*.

В качестве устройства сети «Интернет вещей» рассматривается *raspberry pi zero w* (рис. 1) со следующими техническими характеристиками [1]:

- однокристальная система *SoC Broadcom BCM2835*;
- процессор 32-битный 1-ядерный *ARMv6Z ARM1176JZF-S* с тактовой частотой 1 ГГц;
- графический 2-ядерный сопроцессор *Video Core IV Multimedia*;
- ОЗУ 512 Мб *LPDDR2 SDRAM*.
- WIFI 802.11n + Bluetooth 4.1 Low Energy (BLE).



Рис. 1. *Raspberry pi zero w*

В общем случае «Интернет вещей» состоит из беспроводных сенсорных узлов, которые собирают информацию и отправляют ее на ближайшую базовую станцию, называемую «шлюзом», которая, в свою очередь, связывается с общим или частным облаком [2].

Криптосистемы с открытым ключом работают медленнее, чем симметричные криптосистемы. Одна из наиболее распространенных асимметричных криптосистем – *RSA*.

Известно, что у данных алгоритмов имеются ряд недостатков, такие как большие размеры ключей, а также большие затраты времени для генерации ключей и цифровой подписи. Для использования криптографии в устройствах, память и вычислительная мощность которых сильно ограничены, необходимы более эффективные алгоритмы [3]. В качестве такого алгоритма в работе используется *ECDSA*.

Алгоритм криптографической защиты *ECDSA TLS* на основе эллиптических кривых и устройства *raspberry pi*

Эллиптическая кривая – это множество точек, описываемое уравнением

$$y^2 = x^3 + ax + b \pmod{p}.$$

Данные кривые обладают следующими характеристиками:

1. Простое p , задающее размер конечного поля;
2. Коэффициенты a и b уравнения эллиптической кривой;
3. Базовая точка G , генерирующая подгруппу.
4. Порядок n подгруппы.
5. Кофактор h подгруппы.

Примером эллиптической кривой, используемая в современных протоколах защиты информации, таких как *TLS* является кривая *secp256k1* [4], представленная на рис. 2 и имеющая следующие параметры:

1. $p=0xfffefffffc2f$;
2. $a=0$;
3. $b=7$;
4. $g=(0x79be667ef9dcbbac55a06295ce870b07029bfcdb2dce28d959f2815b16f81798, 0x483ada7726a3c4655da4fbfc0e1108a8fd17b448a68554199c47d08ffb10d4b8)$;
5. $n=0xffffffffffffffffffffffffffffffebaaedce6af48a03bbfd25e8cd0364141$;
6. $h=1$.

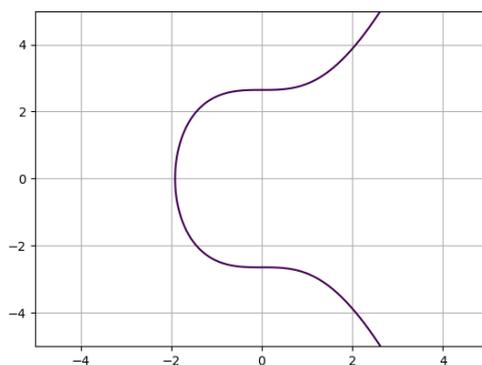


Рис. 2. График эллиптической кривой *secp256k1*

В качестве эксперимента выберем несколько эллиптических кривых, предоставляемых *OpenSSL*. *OpenSSL* – криптографическая библиотека с открытым исходным кодом, широко известна из-за расширения *SSL/TLS*, используемого в веб-протоколе *HTTPS*. Алгоритм цифровой подписи [5] был реализован на языке *Python*. Общая функция генерации цифровой подписи представлена на рис. 3.

```
def sign_message(private_key, message):
    z = hash_message(message)
    r = 0
    s = 0
    while not r or not s:
        k = random.randrange(1, curve.n)
        x, y = scalar_mult(k, curve.g)
        r = x % curve.n
        s = ((z + r * private_key) * inverse_mod(k, curve.n)) % curve.n
    return (r, s)
```

Рис. 3. Функция генерации цифровой подписи

Полученные результаты времени, необходимого *raspberry pi* для генерации цифровой подписи, представлены в табл. 1.

Таблица 1. Результаты измерений времени генерации цифровой подписи на *raspberry pi*

Эллиптическая кривая	Время генерации подписи, с
Secp192k1	0,78
Secp192r1	0,73
Secp224r1	0,92
Secp256k1	1,20
Secp256r1	1,32
Secp384r1	2,91

Алгоритм криптографической защиты на основе эллиптических кривых, устройства *raspberry pi* и платформы *Heroku*

Для увеличения производительности данный алгоритм перенесен на облачную платформу *Heroku* [6] в виде веб-приложения.

Heroku – облачная *PaaS (Platform as a Service)* платформа, поддерживающая ряд языков программирования.

PaaS – модель предоставления облачных вычислений, при которой потребитель получает доступ к использованию информационно-технологических платформ: операционных систем, систем управления базами данных, связующему программному обеспечению, средствам разработки и тестирования, размещенным у облачного провайдера.

Для работы приложения были созданы 2 таблицы в базе данных платформы *Heroku*:

– для хранения информации о каждом устройстве в сети, которому необходим доступ к сервису;

– для хранения используемых эллиптических кривых и их параметров.

Схема базы данных приложения представлена на рис. 4.

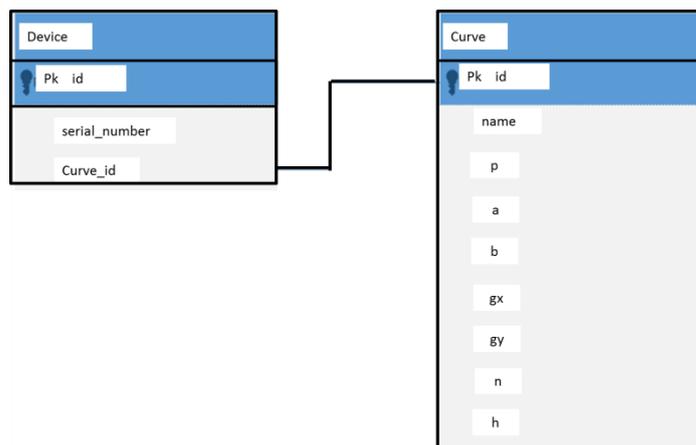


Рис. 4. Схема сущностей в базе данных

Результаты эксперимента для анализа затраченного времени для генерации цифровой на платформе *Heroku* приведены в табл. 2.

Таблица 2. Результаты измерений времени генерации цифровой подписи с использованием веб-сервиса

Эллиптическая кривая	Время генерации подписи, с
secp192k1	0,16
secp192r1	0,15
secp224r1	0,34
secp256k1	0,56
secp256r1	0,58
secp384r1	1,05

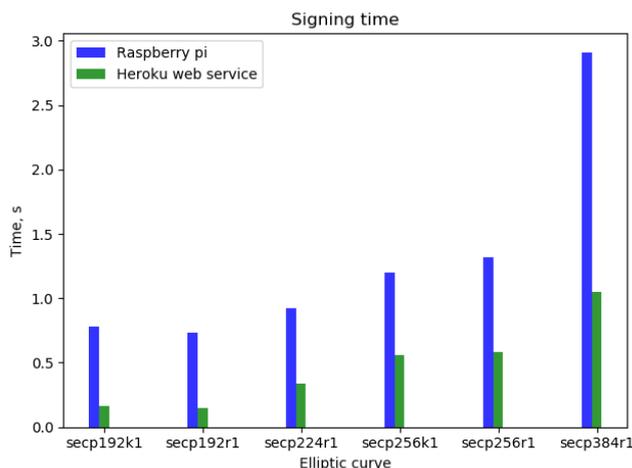


Рис. 5. Сравнение времени генерации цифровой подписи на *raspberrypi* и веб-сервисе

Заключение

Веб-приложение на базе платформы *Heroku*, позволило освободить часть ресурсов *raspberrypi* за счет переноса алгоритма генерации цифровой подписи на облачный веб-сервис, тем самым увеличив производительность устройства. Благодаря встроенной базе данных, данное приложение позволяет гибко изменять используемую эллиптическую кривую и ее параметры для каждого устройства в сети, что не позволяет стандартная реализация протокола TLS [7].

DIGITAL SIGNATURE OF THE TLS PROTOCOL WEB SERVICE BASED ON THE ELLIPTIC CURVES AND HEROKU CLOUD PLATFORM

N.S. FILIPPOV, S.B. SALOMATIN

Abstract. The cryptographic algorithms of digital signature on elliptic curves in the context of the Internet of Things were considered, as well as the expansion of the capabilities of the standard TLS transport layer protection protocol based on the *raspberrypi* device and *Heroku* cloud protocol.

Keywords: cryptographic information protection, elliptic curve, TLS protocol, Internet of things, web service IoT.

Список литературы

1. Why Is IoT? [Electronic resource]. URL: <https://www.oracle.com/internet-of-things/what-is-iot.html> (date of access: 02.11.2019).
2. Raspberry Pi Zero W [Electronic resource]. URL: <https://www.raspberrypi.org/products/raspberry-pi-zero-w/> (date of access: 02.11.2019).
3. Nils Gura, [et. al.] // Workshop on Cryptographic Hardware and Embedded Systems, LNCS 3156. 2004. P. 119–132.
4. SEC 2: Recommended Elliptic Curve Domain Parameters. Daniel R. L. Brown. – NIST, 2010.
5. Johnson D., Menezes A. // International Journal of Information Security. 2001. Vol. 1. P. 36–63.
6. Learn about building, deploying, and managing your apps on Heroku [Electronic resource]. URL: <https://devcenter.heroku.com/> (date of access: 02.11.2019).
7. Ключи, шифры, сообщения: как работает TLS [Электронный ресурс]. URL: <https://tls.dxdt.ru/tls.html#crypto-params> (дата доступа: 02.11.2019).

УДК 004.021

АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ОПТИЧЕСКОГО КОНТРОЛЯ КРИТИЧЕСКИХ РАЗМЕРОВ ПОЛУПРОВОДНИКОВЫХ ПЛАСТИН СБИС

А.А. ДУДКИН, А.А. ВОРОНОВ, В.В. ГАНЧЕНКО,
А.В. ИНУТИН, Е.Е. МАРУШКО

*Государственное научное учреждение «Объединенный институт проблем информатики
Национальной академии наук Беларуси», Республика Беларусь*

Поступила в редакцию 01 октября 2019

Аннотация. Приводится описание алгоритмического обеспечения программного комплекса управления оборудованием контроля критических размеров полупроводниковых пластин СБИС на базе систем компьютерного зрения.

Ключевые слова: Полупроводниковая пластина, СБИС, оптический контроль, машинное зрение, цифровое изображение.

Введение

Актуальной задачей при разработке и выпуске конкурентоспособных наукоемких изделий микроэлектроники является создание современной научно-технической и производственно-технологической базы производства сверхбольших интегральных микросхем (СБИС), в том числе оптико-механического, контрольно-измерительного и сборочного оборудования.

В связи с переходом на субмикронные нормы проектирования и усложнением самих СБИС возникает необходимость решения задач обработки, хранения, приема и передачи больших объемов данных с использованием современных интерфейсов, для чего требуется разработка оригинальных подходов, методов и алгоритмов цифровой обработки изображений и приборов контроля технологических процессов литографии, компьютерных методов и алгоритмов анализа получаемых данных, которые в совокупности позволят обеспечить качественную отработку процессов фотолитографии при изготовлении СБИС по субмикронным нормам.

Объектами контроля на топологических слоях СБИС являются контактная площадка, контактное окно, диффузионный или металлизированный проводник, другие элементы топологии. С точки зрения фотолитографии они создаются как комбинация двух областей: прямоугольник и линия (промежуток).

Области структур создаются локальным введением в подложку примесей (посредством диффузии из газовой фазы или ионным легированием), осуществляемым через маску (обычно из пленки формируемую при помощи фотолитографии). Последовательно проводя процессы окисления, фотолитографии и введения примесей, можно получить легированную область любой требуемой конфигурации, а также внутри области с одним типом проводимости (уровнем концентрации примеси) создать другую область с другим типом проводимости. Наличие на одной стороне пластины выходов всех областей позволяет осуществить их коммутацию в соответствии с заданной схемой при помощи пленочных металлических проводников, формируемых также с помощью методов фотолитографии.

Основной параметр, характеризующий совершенство и качество фотолитографии, — это воспроизводимое и стабильно получаемое разрешение, или минимальный размер элемента. При этом топологическая норма определяется как полуширина линии и пространства между линиями в регулярных плотно упакованных полосчатых структурах. Тенденции в развитии микроэлектроники определяются законом Мура.

Требования к функциям алгоритмического обеспечения контроля

Алгоритмы должны обеспечить возможность выполнения следующих функций:

- предобработки изображений с учетом конструкторско-технологических ограничений (КТО);
- обработки и анализа изображений с поддержкой оборудования видеозахвата сторонних производителей;
- анализа изображений для контроля конструкторско-технологических ограничений;
- хранения и доступа к данным с возможностью импорта и экспорта данных в различных форматах;
- синтеза программы для автоматического режима работы контроля;
- управления механизмами сторонних производителей;
- визуализации данных.

Основные алгоритмы предварительной обработки

Алгоритмы предварительной обработки изображений слоев СБИС с учетом КТО позволяют исправить их дисторсию и геометрические искажения, повысить их информативность, удалить шум, уточнить положение кадров слоя, улучшить изображения.

Отметим особо разработанные оригинальные алгоритмы:

- повышения информативности снимков слоев полупроводниковых микросхем с учетом КТО, основанный на объединении объектов интереса с нескольких экземпляров кристалла;
- семантической фильтрации изображений с учетом КТО, основанный на использовании дескрипторов формы сегментов изображения: общей площади сегмента, геометрического центра сегмента, длины главной диагонали сегмента, размеров аппроксимирующего прямоугольника для сегмента;
- предобработки изображений с учетом КТО на основе мягкой морфологии, условной эрозии и дилатации;
- сшивки изображений с учетом КТО;
- улучшения изображений с учетом КТО, позволяющий выровнять неоднородности освещения отдельных кадров и перепадов яркости на их границе.

Алгоритмы автофокусировки предназначены для количественной оценки функции резкости с использованием различных критериев на конкретных сериях изображений, непрерывного автоматического определения фокусного расстояния для съемочной системы по изображениям топологии СБИС и определения неравномерности освещенности, что обеспечивает требуемой точности идентификации объектов контроля.

Алгоритмы анализа

Операция автоматического контроля является одной из самых ответственных при производстве изделий микроэлектроники. Так, оборудование для автоматического контроля оригиналов топологии является самым дорогим во всей технологической цепочке кристалльного производства, что обусловлено с одной стороны, высокой сложностью, с другой стороны – высокой ценой возможного пропуска дефектов.

На сегодняшний момент отсутствует методика, позволяющая на основании анализа характеристик контролируемых объектов и условий контроля выбрать метод контроля, поэтому предлагается комплекс алгоритмов, реализующий следующие основные методы контроля:

- метод контроля проектных норм, который заключается в проверке элементов топологии на соответствие КТО, детальное описание которых содержится в руководствах и стандартах (SEMI [1] и др.);
- метод сравнения с эталоном, который основан на использовании моделей с заранее определенными информативными признаками в качестве эталона (изображений, получаемых

с идеальной топологии или с САПР), с которыми непосредственно сравнивается тестовое изображение фотошаблона или топологического слоя СБИС и фиксируются различия;

– метод морфологических описаний, который предполагает анализ критериев корректности элементов рисунка топологии как списка признаков или образов в комбинации с некоторыми базовыми критериями правил проектирования.

Эта группа методов позволяет идентифицировать локальные дефекты топологии, такие как прокол, островок, вырыв, выступ, а также дефекты, связанные с нарушением предельных КТО, например, допусков на толщину проводника, расстояние между проводниками и т.п.

Каждый из названных методов имеет как свои достоинства, так и недостатки. Техника сравнения проста и эффективна с точки зрения вычислений, ограничена необходимостью иметь образец, а также требует точного совмещения изображений. Морфологический подход работает с большим числом форм и требуется постоянная доработка алгоритма контроля. Применение процесса верификации правил проектирования непосредственно к изображению образов является затратным по времени и требует больших вычислительных мощностей. Кроме того, проверка правил проектирования приводит к большому числу ложных дефектов. Поэтому важно комбинирование различных методов контроля для достижения лучшего результата.

Разработанные перечисленные выше оригинальные алгоритмы обеспечивают:

– контроль критических размеров на базе системы получения видеоизображения с поддержкой минимальных элементов размером 350–500 нм и повторяемостью не хуже 2 нм;

– определение неровности края на всем участке измерений;

– определение размеров периодических решеток;

– измерение с проходом по фокусировке для повышения повторяемости и точности контроля.

Основную функциональность алгоритмического обеспечения опишем диаграммой прецедентов или вариантов использования. Прецеденты можно условно разделить на основные и вспомогательные.

Основными прецедентами использования являются:

– загрузка объекта исследования (полупроводниковой пластины или фотошаблона) – подготовка объекта исследования для дальнейшей работы (предварительная ориентация в пространстве, перемещение в рабочую зону);

– выгрузка объекта исследования – удаление объекта исследования из рабочей зоны в хранилище (контейнер, кассету);

– инициализация установки и базирование механизмов – загрузка в установку данных, описывающих исходное состояние оборудования для решения конкретной задачи, и установка механизмов в соответствующее состояние/положение;

– управление перемещениями координатного стола – формирование обобщенных команд для управления движением координатного стола;

– управление механизмами – подготовка команд управления оборудованием (метакоманд) и соответствующих параметров;

– формирование управляющих команд – преобразование метакоманд в формат, требуемых тем либо иным микроконтроллером управления оборудованием;

– создание программы контроля и измерений для автоматического режима – формирование списка управляющих воздействий с соответствующими параметрами и их сохранение в виде файла либо записи в базе данных;

– создание карты-структуры объекта (объектом может служить как пластина, так и шаблон) – подготовка описания структуры объекта для дальнейшего его исследования (кадры, модули);

– сохранение результатов контроля и измерений – формирование структур, предназначенных для дальнейшего хранения в базе данных и последующее сохранение полученного блока данных с использованием СУБД.

Вспомогательные прецеденты требуются для идентификации.

Для эффективной реализации приведенных выше функций был разработан программный комплекс, структура которого показана на рис. 1.

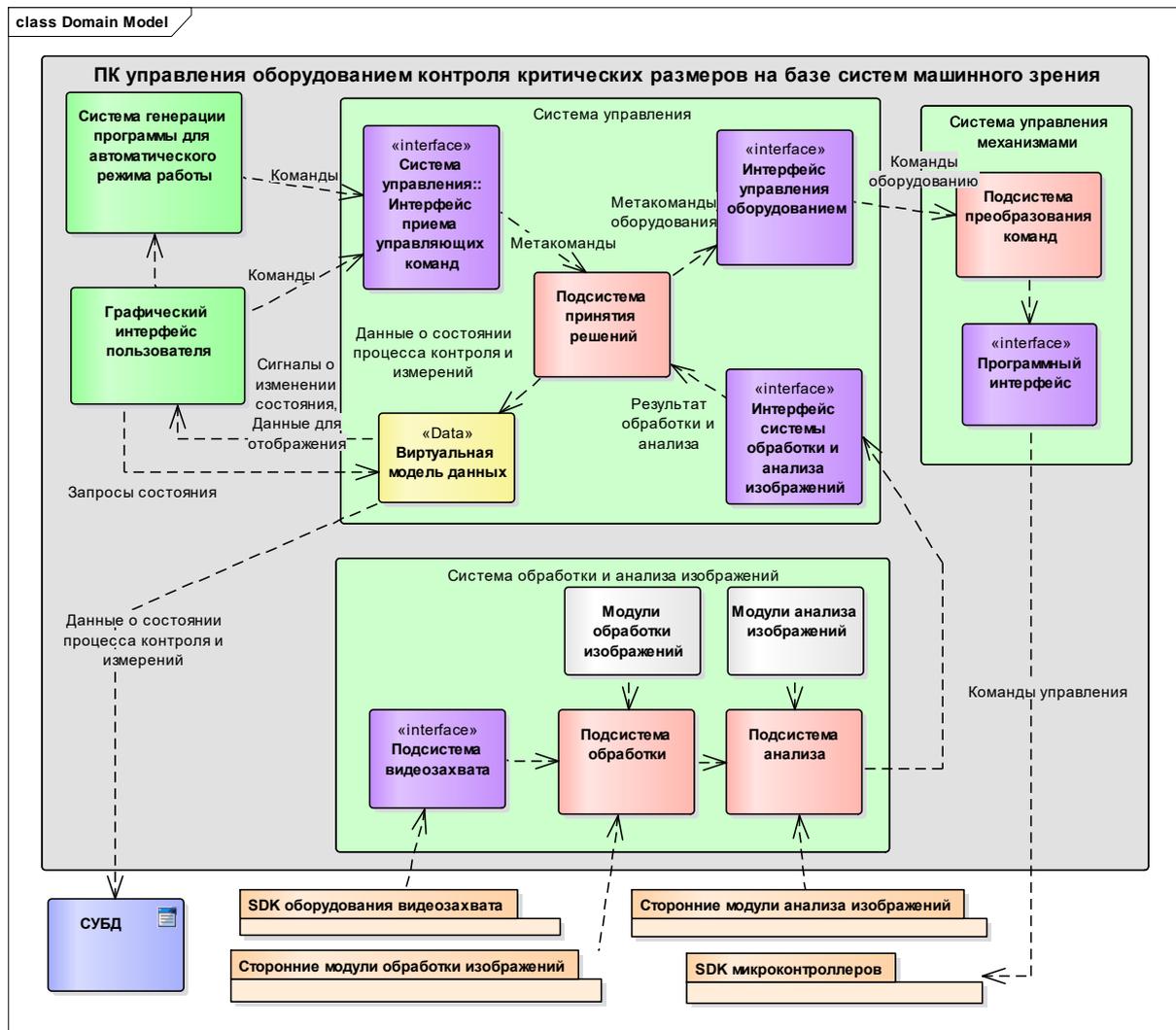


Рис. 1. Структурная схема программного комплекса контроля

Пользователь может самостоятельно компоновать цепочки простых алгоритмов для получения более сложных. При этом реализована возможность подключения внешних модулей, реализующих ту или иную функциональность. Архитектурно и функционально обеспечена возможность хранения параметров работы, а также программ автоматической обработки, в базе данных программного комплекса. При реализации встроенных алгоритмов обработки и анализа изображений использованы средства для повышения производительности путем распараллеливания: использование векторных инструкций процессора, использование стандарта OpenCV для обработки нескольких изображений одновременно и возможность конвейерной обработки для обеспечения равномерной загрузки вычислительных ресурсов при обработке больших объемов данных.

Программный комплекс совместим с оборудованием контроля, выпускаемым ОАО «КБТЭМ-ОМО» [2].

ALGORITHMIC SUPPORT FOR OPTICAL INSPECTION OF CRITICAL DIMENSIONS ON VLSI SEMICONDUCTOR WAFERS

A.A. DOUDKIN, A.A. VORONOV, V.V. GANCHENKO,
A.V. INUTIN, E.E. MARUSHKO

Abstract. The description of algorithmic maintenance of the program complex of control equipment of the critical dimensions of semiconductor wafer plates on the basis of computer vision systems was given.

Keywords: Semiconductor chip, VLSI, optical inspection, computer vision, digital image.

Список литературы

1. SEMI International Standards [Electronic resource]. URL: <http://www1.semi.org/en/Standards>. (date of access: 27.03.2019).
2. Достанко А.П., Аваков С.М., Агеев О.А. и др. Технологические комплексы интегрированных процессов производства изделий электроники. Минск: Беларуская Навука, 2016.

УДК 621.391

СТРУКТУРНЫЙ АЛГОРИТМ ТЕКСТУРНОЙ СЕГМЕНТАЦИИ ПОЛУТОНОВЫХ ИЗОБРАЖЕНИЙ

А.В. КУРИЛОВИЧ, В.Ю. ЦВЕТКОВ, В.К. КОНОПЕЛЬКО

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 11 ноября 2019*

Аннотация. Предложен структурный алгоритм текстурной сегментации полутоновых изображений на основе сегментации, составления дескрипторов сегментов и их дальнейшей фильтрации по заданным критериям. Сущность алгоритма состоит в сегментации исходного изображения, определении основных параметров сегментов (средняя яркость, количество пикселей, форм-фактор, координаты центра), группировке сегментов по совпадению одного или нескольких параметров с учетом допустимых отклонений и фильтрации образованных групп по критериям текстурных объектов.

Ключевые слова: текстурная сегментация изображения, дескриптор сегмента.

Введение

В связи с постоянным ростом разрешения устройств получения изображений [1], на формируемых с их помощью изображениях увеличивается количество текстурных областей. Из-за высокочастотного характера текстур, коэффициенты сжатия содержащих их изображений значительно меньше коэффициентов сжатия изображений, содержащих преимущественно однородные по яркости области и малое число резких перепадов яркости [2]. В связи с этим актуальной является задача повышения эффективности сжатия текстурных изображений, решение которой требует, прежде всего, локализации текстурных областей. Это, в свою очередь, делает актуальной задачу текстурной сегментации. Если в результате текстурного анализа необходимо определить тип (класс) текстуры или различить несколько текстур [3, 4], то в результате сегментации необходимо найти границу разделения текстурных областей [5]. Несмотря на данное различие, для текстурного анализа и текстурной сегментации используются общие методы [6–11]. Среди них – энергетические карты [11, 12], которые широко используются в настоящее время в текстурном анализе и могут применяться для текстурной сегментации. Их недостатком является высокая вычислительная сложность и разделение простых текстур с высокой степенью однородности. В ряде работ [13] предлагается использовать контурные элементы изображений для текстурного анализа. Однако данные подходы основаны на небольшом наборе контурных примитивов и достаточно грубой статистической оценке, что приводит к значительным ошибкам разделения сложных текстур.

Структурный алгоритм текстурной сегментации на основе группировки и фильтрации дескрипторов сегментов

Предлагается структурный алгоритм текстурной сегментации на основе группировки и фильтрации дескрипторов сегментов. Согласно данному алгоритму, исходное полутоновое изображение подвергается обработке известным методом волновой сегментации, в результате которой находятся две матрицы одинакового размера, совпадающего с размером исходного изображения. Mz – матрица яркостей пикселей исходного полутонового изображения и Ms – матрица номеров сегментов. На основе матриц Mz и Ms находятся элементы матрицы дескрипторов сегментов Md . Количество строк в матрице дескрипторов сегментов Md

соответствует количеству различных чисел в матрице номеров сегментов, то есть количеству построенных сегментов, после обработки изображения методом волновой сегментации. Для изображения (рис. 1) количество сегментов составляет 4913. Каждая строка матрицы Md содержит дескриптор заданного сегмента.



Рис. 1. Обработка полутонового изображения методом волновой сегментации:
a – исходное изображение; *б* – результат сегментации

Дескриптор сегмента состоит из следующих компонентов: номер сегмента; координата Y центра сегмента; координата X центра сегмента; количество пикселей в сегменте; средняя яркость пикселей сегмента; форм-фактор сегмента. Первые 20 строк матрицы Md из 4913 представлены в табл. 1.

Таблица 1. Дескрипторы сегментов

№ сегмента	Координата Y центра	Координата X центра	Количество пикселей	Средняя яркость	Форм-фактор
2	5	10	6	204	1,00
3	9	1	4	221	1,00
4	20	16	3	170	1,00
5	3	13	8	197	1,00
6	20	8	15	185	0,63
7	7	7	23	213	0,77
8	8	2	3	217	1,00
9	93	205	4	183	1,00
10	96	192	5	145	0,63
11	20	11	4	207	1,00
12	24	10	7	223	0,78
13	82	186	16	204	0,41
14	4	16	15	195	0,36
15	9	3	3	224	1,00
16	9	10	14	221	0,70

На основе матрицы дескрипторов сегментов Md строится матрица Mdd , содержащая 4913 строк и 7 столбцов. В столбцах матрицы Mdd содержатся номера групп, в которые объединяются сегменты по результатам совпадения одного или нескольких параметров дескрипторов с заданным допустимым отклонением. Заданные отклонения составляют: по количеству пикселей – 2; по яркости – 10; по форм-фактору – 20 %. Так, в первом столбце содержится номер группы для каждого сегмента, в которой количество пикселей в сегменте различается не более чем на 2. Во втором столбце объединение в группы происходит по совпадению критерия средней яркости, в третьем – форм-фактору. В остальных столбцах формируются группы одновременного совпадения по паре критериев, а в последнем одновременно трех критериев.

На рис. 2 приведены исходное изображение и соответствующие ему центры сегментов. Алгоритм предполагает их распределение по битовым плоскостям в зависимости от значений дескрипторов соответствующих им сегментов. Примеры таких битовых плоскостей приведены на рис. 3.

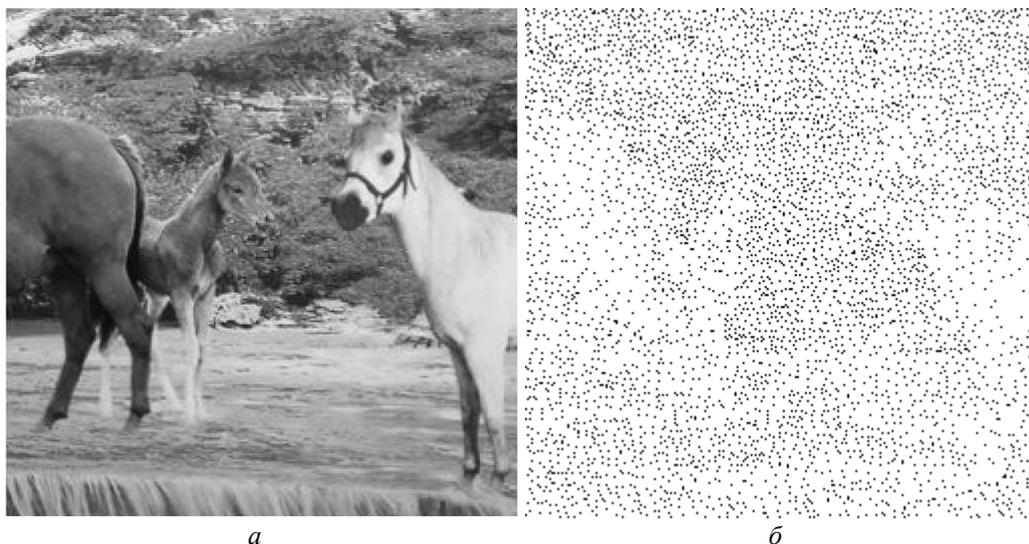


Рис. 2. Обработка дескрипторов сегментов: *а* – исходное изображение; *б* – центры сегментов

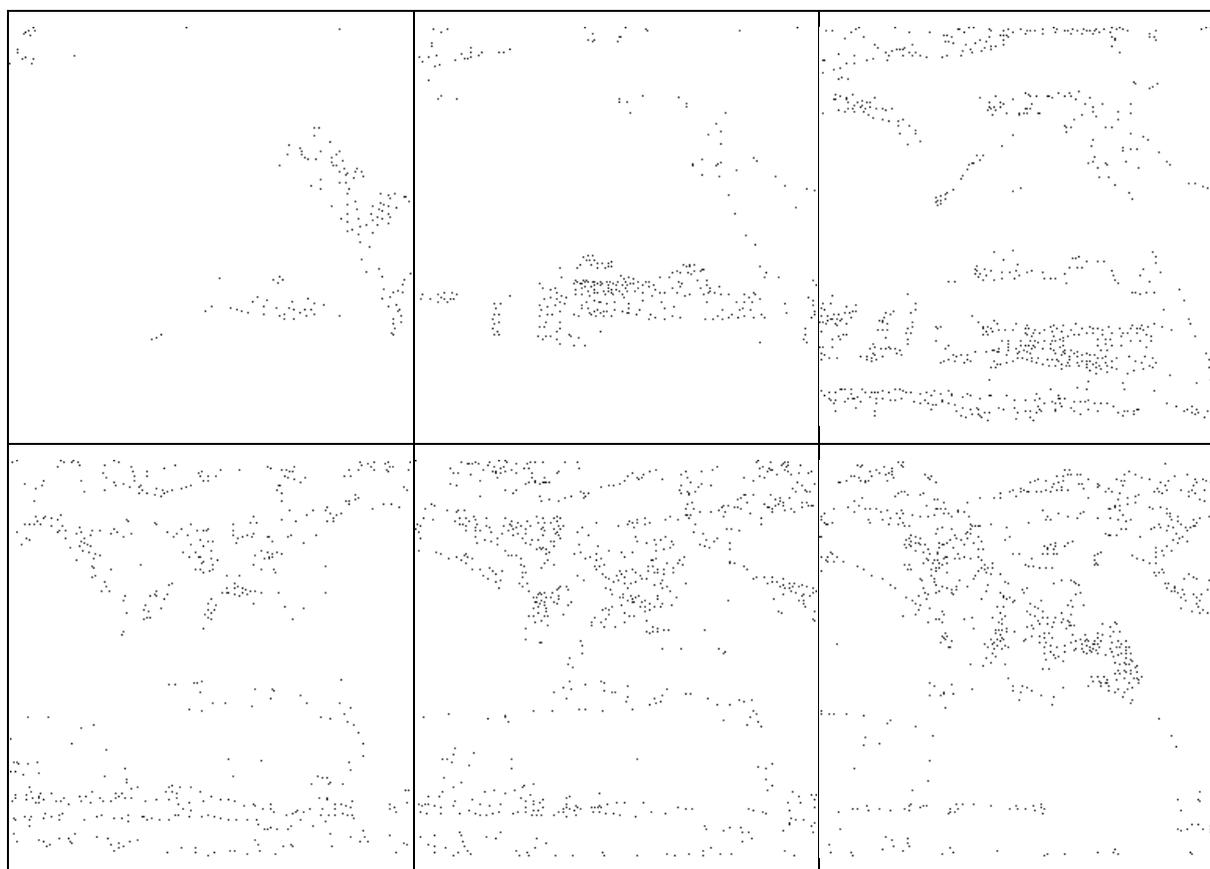


Рис. 3. Примеры битовых плоскостей с центрами сегментов

Дальнейшая обработка связана с оценкой расстояний между центрами сегментов в каждой битовой плоскости, объединением в области сегментов, центры которых находятся примерно на одинаковом расстоянии друг относительно друга, объединением битовых плоскостей, в результате чего определяются текстурные области.

Заключение

Предложен структурный алгоритм текстурной сегментации полутоновых изображений на основе вычисления дескрипторов сегментов, группировки сегментов по совпадению одного или нескольких параметров дескриптора и дальнейшей фильтрации по критериям соответствия текстурным объектам.

SEARCH OF LOCAL EXTREMUMS OF HALF-TONE IMAGES BASED ON CENTRAL SYMMETRIC SCANNING

A.V. KURYLOVICH, V.Yu. TSVIATKOU, V.K. KANAPELKA

Abstract. Structural algorithm for textural segmentation of halftone images based on segmentation, segmentation descriptors and their further filtering by specified criteria was proposed. The essence of the method consists segmentation of the original image, determination of the main parameters of segments (average brightness, number of pixels, form factor, center coordinates), grouping of segments by coincidence of one or several parameters taking into account permissible deviations and filtering of the formed groups according to the criteria of texture objects.

Keywords: texture image segmentation arch, segment descriptor.

Список литературы

1. Mastriani M. // Engineering and Technology International Science Index 48. 2010. № 4 (12). P. 485–497.
2. Ватолин Д., [и др.] Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М., 2003.
3. Materka A., Strzelecki M. // Technical university of lodz, institute of electronics .1998. № 11. P. 9–11.
4. Tuceryan M., Jain A.K. // The Handbook of Pattern Recognition and Computer Vision (2nd Edition). 1998. P. 207–248.
5. Bhosle V.V., Pawar V.P. // International Journal of Soft Computing and Engineering (IJSCE). 2013. Vol. 3. P. 69–74.
6. Grigorescu S.E., Petkov N., Kruizinga P. // IEEE Transactions On Image Processing. 2002. № 10. P. 1160–1167.
7. Reulke R., Lippok A. // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2008. Vol. 37. P. 615–620.
8. Clausi D.A. // Pattern Recognition. 2002. № 35. P. 1959–1972.
9. Liu X., Wang D. // IEEE Transactions On Image Processing. 2006. Vol. 15, № 10. P. 3066–3077.
10. Kokkinos L., Evangelopoulos G., Maragos P. // Actions On Pattern Analysis And Machine Intelligence. 2009. Vol. 31, № 1. P. 142–157.
11. Lee D-Ch., Shchenk T. // A Collection of Papers Presented At the XVII Congress of ISPRS. 1992. № 48. P. 75–80.
12. Ertuğrul Ö. // International Journal of Intelligent Information Systems. 2014. Vol. 3, № 2. P. 13–18.
13. Jithendra M., [et. al.] // International Journal of Computer. 2001. № 43 (1). P. 7–27.

УДК 654.16

ВЕРТИКАЛЬНАЯ СЕКТОРИЗАЦИЯ ДЛЯ СОВМЕСТНОЙ РАБОТЫ УЗКОПОЛОСНОГО ИНТЕРНЕТА ВЕЩЕЙ И МОБИЛЬНЫХ СЕРВИСОВ 3G

В.А. АКСЕНОВ, С.В. СМОЛЯК, М.Ю. ХОМЕНОК

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 30 октября 2019*

Аннотация. Проведен анализ возникающей у сотовых операторов проблемы радиопокрытия при совместном культивировании в диапазоне 900 МГц узкополосного интернета вещей NB-IoT и 3G на одной территории. Выполнены расчеты для оценки целесообразности вертикальной секторизации сот в такой ситуации.

Ключевые слова: зашумлением пилотных сигналов, U-900, NB-IoT, угол наклона ДН, вертикальная секторизация, антенна Huawei ADU4515R5v06.

Введение

Оператор сотовой связи А1 первым развернул в 2017 году сеть узкополосного интернета вещей NB-IoT в г. Минске и, далее, по территории РБ. За прошедшие годы оператором накоплен полезный опыт эксплуатации сети с этой технологией, выявлены ее положительные и проблемные места. На основе этого опыта ниже обсуждается вопрос об оптимизации радиопокрытия сети при совместном использовании в диапазоне 900 МГц систем WCDMA, известной так же под названием U-900 и NB-IoT.

Особенности организации радиопокрытия для U-900 и NB-IoT

В радиосетях WCDMA известен эффект «зашумлением пилотных сигналов», особых служебных сигналов, передаваемых только в нисходящем канале (DL) и служащих для выбора соты, ее индивидуализации, первичной синхронизации в соте и прочее. В описываемой системе передача пилотных сигналов происходит в рамках физического канала Common Pilot Channel (CPICH). Протокол «парковки» терминала на соту всегда начинается с поиска, приема и правильной интерпретации этих сигналов. К сожалению, интерференция из соседних сот, работающих на том же радиоканале и такими же по структуре пилотными сигналами, может сделать невозможным их корректный прием и подключение мобильного терминала к данной соте. Положение ухудшается для U-900, где из-за хорошего прохождения радиоволн диапазона 900 МГц образуются помехи не только от соседней соты, но и от других более далеких сот сети.

Опыт эксплуатации сотовых радиосетей показал, что эффективным средством борьбы с зашумлением пилотных сигналов может быть увеличение угла наклона диаграммы направленности (ДН) антенны базовой станции в вертикальной плоскости. Под углом наклона понимают угол между горизонталью и направлением максимального значения ДН в вертикальной плоскости.

Можно ожидать, что большой угол наклона для U-900 будет сильно снижать уровень дальней интерференции в радиосети и в определенной мере снижать интерференцию с соседней сотой, но, к сожалению, также будет снижаться уровень сигнала (и отношение сигнал/шум) в приграничной области соты, что ведет к потере ее пропускной способности.

Наклон может быть реализован механическим способом (Mechanical DownTilt, MDT), фиксированным электрическим (Electrical DownTilt, EDT), переменным электрическим (Variable Electrical Tilt, VET), электрическим с дистанционным управлением (Remote Electrical Tilt, RET)

и плавно регулируемым электрическим наклоном (Continuously Adjustable Electrical DownTilt, CAEDT) с точностью выше половины градуса [1].

В практике сотовой связи известны следующие подходы к выбору угла наклона главного лепестка ДН по вертикали [1, 2]. При геометрии соты с высотой подвеса антенны базовой станции H_{bts} , расстоянием (радиусом) R от антенны до границы соты, понимаемой, как граница по разделению обслуживаемого трафика, образуется угол θ_{geo} между горизонталью и линией (АВ) визирования антенны на границу соты (рис. 1). Значение угла θ_{geo} равно

$$\theta_{geo} = \arctan\left(\frac{H_{bts} - H_{ms}}{R}\right), \quad (1)$$

где H_{ms} – высота расположения антенны мобильного терминала, обычно принимаемая за 1,5 м (для упрощения на рис. 1 это значение принято за 0 м).

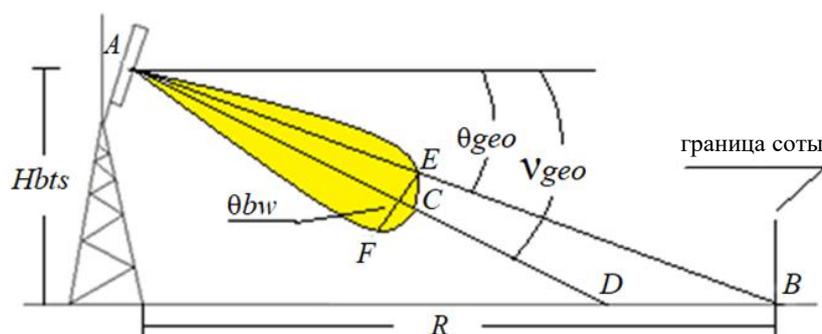


Рис. 1. Пояснения вариантов угла наклона лепестка ДН антенны в вертикальной плоскости

Выбор этого угла для наклона ДН имеет хорошее обоснование, т.к. при нем максимум лепестка ДН направлен в самый удаленный участок соты, на ее границу.

Для минимизации интерференции в соседнюю соту (соты) угол наклона, как правило, увеличивают. В [2] имеется вариант наклона на угол ν_{geo} , который больше θ_{geo} на половину ширины лепестка ДН θ_{bw} :

$$\nu_{geo} = \theta_{geo} + \frac{\theta_{bw}}{2}. \quad (2)$$

При этом на границу соты наводят линию (АЕВ), проходящую через точку Е в верхней части главного лепестка со спадом -3 дБ. Именно такой вариант используют для U-900 на сети оператора А1.

На основе моделирования сети для макросот WCDMA с радиус 1500 м и более, в [2] предложено соотношение для оптимального по пропускной способности угла наклона:

$$\nu_{opt} = 3 \left[\ln(H_{bts}) - R^{0.8} \right] \lg(\theta_{bw}). \quad (3)$$

Используя (2) и (3), можно получить зависимость соответствующих углов наклона от радиуса соты, показанную на рис. 2. При вычислениях использовались высоты подъема антенн 30 и 1,5 м, ширина ДН 10 градусов. Как видно, на наиболее актуальном для города интервале радиусов от 500 до 1500 м выражение (2) дает меньшее значение угла наклона, хотя отличие от (3) не превышает одного градуса.

Основная особенность узкополосного интернета вещей NB-IoT с точки зрения радиопокрытия – заметно больший радиус зоны обслуживания базовой станции данной технологии в сравнении с U-900. Это обусловлено следующими основными факторами.

Во-первых, терминал NB-IoT создает крайне низкий трафик, что позволяет базовой станции обслуживать тысячи терминалов в соте. Например, в сети NB-IoT оператора А1 был заложен показатель 10000 терминалов на типовой сайт из трех секторных сот.

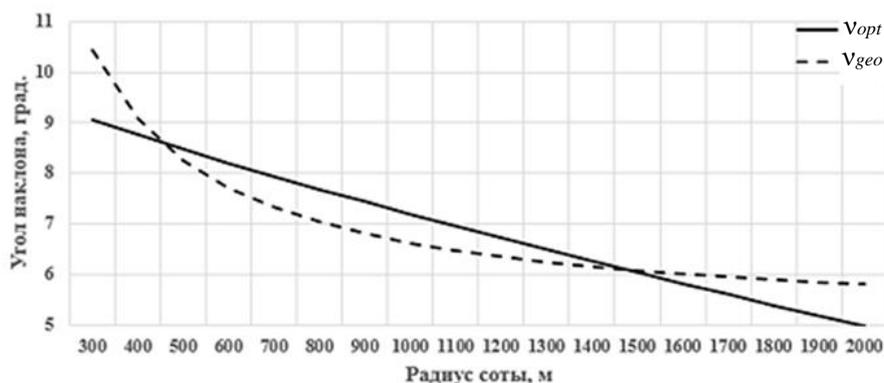


Рис. 2. Зависимость углов наклона ДН антенны по вертикали от радиуса соты

Во-вторых, эффективная чувствительность терминалов NB-IoT, закладываемая в расчет дальности по нисходящему/восходящему каналам (DL/UL), обычно составляет 130–138 дБм. Это на 20–28 дБ больше, чем у сотовых терминалов других технологий.

В-третьих, пиковая мощность излучения терминалов NB-IoT достигает 23 дБм, что сопоставимо с аналогичным показателем терминалов других технологий.

И, наконец, оператор сотовой связи платит за разрешение на эксплуатацию технологии NB-IoT на каждой задействованной базовой станции. Тот факт, что для сети с той же площадью покрытия можно использовать лишь отдельные существующие базовые станции, его чрезвычайно устраивает.

В результате на сегодняшний день в сети NB-IoT оператора А1 радиус сот составляет около 1200 м, а сот U-900 – только 500 м. Следовательно, по мере роста количества терминалов интернета вещей, зоны обслуживания придется постепенно сузить до размеров существующих сот 3G–4G. При современной динамике развития Интернета вещей специалисты оператора А1 считают, что это произойдет ориентировочно через 5 лет. Но этого может не произойти вообще, если нынешняя низкоплотная сеть NB-IoT будет плохо работать, чем опугнет от себя потенциальных потребителей.

Сказанное выше об оптимальном угле наклона ДН справедливо и для сот NB-IoT. Однако, если в (1–3) подставить в два с половиной раза больший радиус, то, естественно, получим заметно меньшие значения соответствующих углов. Таким образом, для оптимизации покрытия под U-900 главный лепесток ДН антенн желательно опустить вниз, а для NB-IoT его желательно поднять вверх.

Хорошим выходом из сложившегося технического противоречия – и к этому склоняются ведущие производители антенной техники – может быть, так называемая, «вертикальная секторизация» соты, т.е. использование панельных антенн с двумя самостоятельными главными лепестками ДН в вертикальной плоскости [3–7], имеющими разный наклон и передающими сигналы для разных стандартов сотовой связи, как показано на рис. 3.

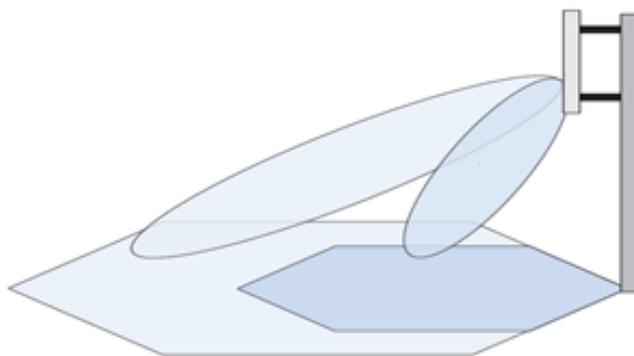


Рис. 3. Применение ДН антенны с двумя лепестками в вертикальной плоскости

В простейшем варианте это могут быть две антенны, каждая со своим EDT (или RET), помещенные в один корпус. Можно просто взять две отдельные панельные антенны с разными углами наклона методом MDT и установить их рядом.

Экономически такая антенная система будет на порядок дешевле smart-антенны с цифровым синтезом многолепестковой ДН. Остаются только два вопроса: на какие углы следует наклонять оба лепестка и как такую антенну (антенны) подключить к типовым базовым станциям.

Моделирование радиопокрытия при наклоне ДН антенны

Задачей моделирования было оценка целесообразности применения двухлепестковой антенны для вышеописанных целей. В расчетах использовались данные по 15 диаграммам направленности для антенны Huawei ADU4515R5v06 на частоте 960 МГц [8]. Указанная антенна может формировать два самостоятельно управляемых электрическим способом лепестка в вертикальной плоскости с шириной 12,36 градусов и общим коэффициентом усиления от 11,97 до 12,32 дБ в зависимости от угла наклона. Диаграммы соответствовали углам EDT от 0 до 14 градусов с шагом 1 градус.

Для сравнения были выполнены расчеты для простой антенны производителя "KATHREIN" с заметно меньшей шириной ДН по вертикали 9,5 градусов.

Для оценки уровня принимаемого сигнала по глубине соты при механическом наклоне ДН на угол V_{geo} использовались данные ДН с нулевым углом EDT и следующие геометрические построения, показанные на рис. 4.

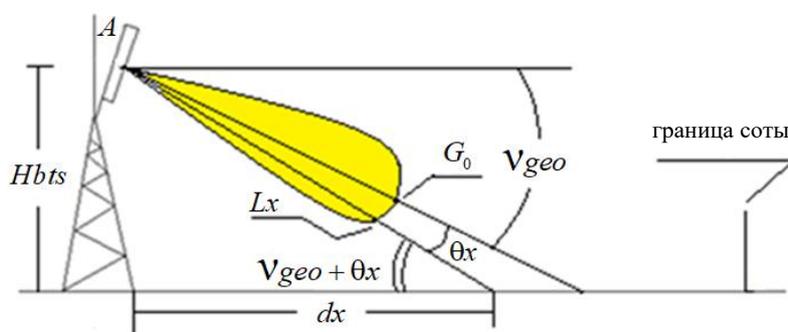


Рис. 4. Геометрическая модель для расчета уровня сигнала

Измеренные значения ДН предоставляются производителями антенн в виде табличных файлов с шагом по углу в 1 градус от 0 до 359 градусов. При этом максимальное значение ДН принимается за 0 дБ, а остальные точки имеют смысл потерь в дБ от коэффициента усиления антенны G_0 . При моделировании удобно брать угол θ_x по ДН и пересчитывать его в расстояние dx .

Тогда принимаемый на расстоянии dx сигнал будет равен

$$Pr(dx) = Pt + G_0 - Lx(\theta_x) - Lp(dx), \quad (4)$$

где $Pr(dx)$ – принимаемая мощность (дБм) на расстоянии dx (км) под углом вертикальной диаграммы направленности θ_x ; Pt – передоваемая антенной мощность (использовалась значение 44 дБм или 25 Вт); G_0 – коэффициент усиления антенны; $Lx(\theta_x)$ – значение потерь (дБ) в диаграмме направленности для угла θ_x , отсчитанного от направления максимального излучения; $Lp(dx)$ – потери (дБ) при распространении на расстояние dx под углом диаграммы направленности θ_x .

Для вычисления потерь при распространении использовалась известная модель Хата [9] для условий городской застройки. Для значений частоты 900 МГц, высоты подъема антенны базовой станции 30 м и антенны мобильного терминала 1,5 м, выражение для модели Хата преобразуется к виду

$$L_p(dx) = 126,44 + 35,22 \lg(dx). \quad (5)$$

Как видно из рис. 3, для угла передачи θx расстояние dx (в км) будет равно

$$dx = \frac{(H_{bts} - H_{ms}) \cdot 0,001}{\text{tg}(\sqrt{V_{geo}} \pm \theta x)}. \quad (6)$$

После подстановки (6) в (5) и преобразования логарифма отношения в разность логарифмов, получим

$$L_p(dx) \equiv L_p(\theta x) = 71,84 - 35,22 \lg(\text{tg}(\sqrt{V_{geo}} \pm \theta x)). \quad (7)$$

При моделировании электрического наклона использовались соответствующие ДН для разных углов EDT. Выше представленные соотношения сохраняют справедливость, но угол V_{geo} приравнивается нулю, а угол θx должен выбираться больше нуля.

На рис. 5 и 6 представлены результаты вычислений для механического наклона ДН на 7 и 9 градусов, которые близки V_{geo} для соты 1200 и 500 м соответственно, а также на максимальный угол 14 градусов. Аналогичные вычисления для электрического наклона ДН дали практически совпадающие результаты и здесь не приводятся.

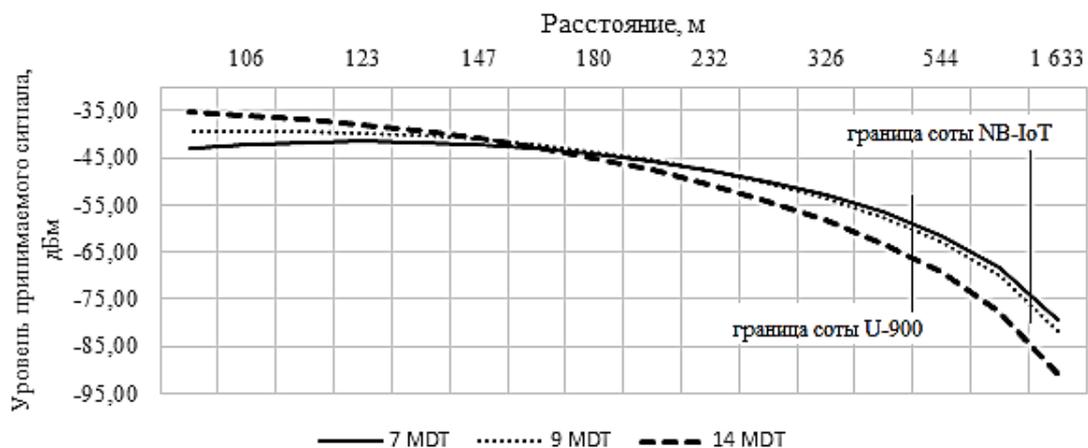


Рис. 5. Результаты распределения уровня принимаемого сигнала по дальности для антенны Huawei ADU4515R5v06 с шириной ДН по вертикали 12,36 градусов

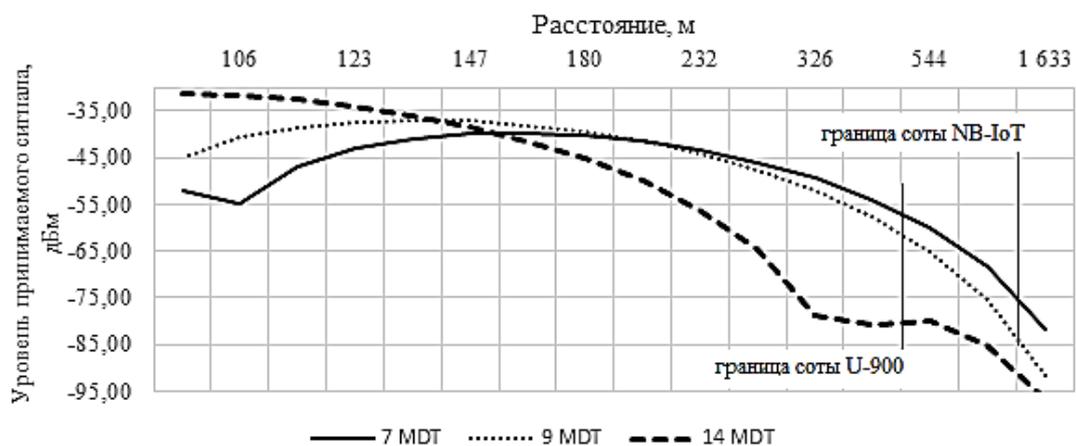


Рис. 6. Результаты распределения уровня принимаемого сигнала по дальности для антенны KATHREIN с шириной ДН по вертикали 9,5 градусов

Заклучение

Как следует из полученных результатов, вертикальная секторизация для решения поставленной проблемы имеет смысл только при использовании достаточно узких лепестков ДН в вертикальной плоскости (менее 10 градусов) или существенно разного наклона лепестков, численно близкого к ширине ДН. Применение широких лепестков (более 10 градусов) с незначительной разницей в наклоне не дает никаких преимуществ в сравнении с использованием однолепестковой антенны с той же шириной диаграммы.

VERTICAL SECTORIZATION FOR JOINT OPERATION OF THE NARROW-BAND INTERNET OF THINGS AND 3G MOBILE SERVICES

V.A. AKSYONOV, S.V. SMOLYAK, M.Yu. HOMENOK

Abstract. The analysis of the problem of radio coverage during joint cultivation in the 900 MHz band of the narrow-band Internet of things NB-IoT and 3G in one territory was made. Calculations to assess the appropriateness of the vertical sectorization of cells in such a situation were made.

Keywords: noise suppression of pilot signals, U-900, NB-IoT, beam down tilt angle, vertical sectorization, Huawei ADU4515R5v06 antenna.

Список литературы

1. Basic Antenna Principles for Mobile Communications. Dipl. Ing. Peter Scholz, KATHREIN-Werke KG [Electronic resource]. URL: <https://www.kathrein.de>. (date of access: 25.10.2019).
2. Niemela J., Isotalo T., Lempiainen J. Optimum Antenna Downtilt Angles for Macrocellular WCDMA network [Electronic resource]. URL: http://www.cs.tut.fi/tlt/RNG/publications/docs/topology/antenna_downtilt.pdf (date of access: 25.10.2019).
3. Dual vertical beam cellular array. United States Patent Application Publication. No.: US 2015/0236430A1 Pub. Date: Aug. 20, 2015.
4. Weiliang Xie, [et. al.] // IEEE Communications Magazine. 2016. P. 89–95.
5. Lei Song, Mugen Peng, Yan Li. Resource allocation for vertical sectorization in LTE-Advanced systems. Hindawi Publishing Corporation International Journal of Antennas and Propagation. Volume 2013, Article ID 456760, 7 pages. [Electronic resource]. URL: <http://dx.doi.org/10.1155/2013/456760>. (date of access: 25.10.2019).
6. Jorge Garrido-Holgado, José Luis Gómez-Tornero. LTE low-cost vertical sectorization through a leaky-wave antenna. Conference Paper April 2014. [Electronic resource]. URL: <https://www.researchgate.net/publication/286592323>. (date of access: 25.10.2019).
7. Mohamed Sanad, Noha Hassan. A Multiple Beam LTE Base Station Antenna with Simultaneous Vertical and Horizontal Sectorization. World Academy of Science, Engineering and Technology International Journal of Electronics and Communication Engineering Vol. 12, No.1, 2018. [Electronic resource]. URL: <https://publications.waset.org/10008490/pdf>. (date of access: 25.10.2019).
8. Huawei ADU4515R5v06 Data Sheet. [Electronic resource]. URL: <https://yadi.sk/d/S4ZM6vzEkDi3m> (date of access: 25.10.2019).
9. ETSI TR 101 362 V8.4.0 (2005-06) Technical Report "Radio Network Planning Aspects".

УДК 004.822:514

ИНТЕЛЛЕКТУАЛЬНОЕ УПРАВЛЕНИЕ ВЫБОРОМ ДЛЯ ДОСТУПА ПОЛЬЗОВАТЕЛЕЙ К МОБИЛЬНЫМ ПРИЛОЖЕНИЯМ

В.А. ВИШНЯКОВ, М.М. ГОНДАГ САЗ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 01 ноября 2019*

Аннотация. Предложена модель интеллектуального выбора метода аутентификации пользователей мобильных приложений, основанная на экспертном подходе. Разработана структура такой экспертной системы и приведены детали ее программной реализации.

Ключевые слова: метод аутентификации, интеллектуальный выбор, экспертная система.

Введение

Вопросы надежного доступа к инфокоммуникационным ресурсам – это основной момент защиты. Контроль доступа, системы шифрования, фаерволы, виртуальные частные сети – все основано на аутентификации пользователей и выдачи прав доступа к программам, с которым устанавливается соединение, если этого недостаточно, то упомянутые средства защиты не эффективны [1]. Данные исследований безопасности отражают факт, что более частыми являются атаки, с загрузкой паразитических программ или подключение атакующего [1, 2]. Использование аутентификации с многократным паролем не совсем надежно и плохо влияют на обеспечение информационной защиты локальных сетей и мобильных приложений. Для предотвращения предоставления атакующим возможности для нарушения конфиденциальности, доступности и целостности в мобильных приложениях, следует выполнять рекомендации по частому изменению многократного пароля и использовать варианты строгой аутентификации [1]. Но это требует больших вычислительных ресурсов и относительно дорого. Поэтому требуется реализация интеллектуального выбора вида доступа в зависимости от ряда факторов.

Контроль доступа в локальных сетях и облачных средах

Аутентификации пользователей – один из основных компонентов организации информационной безопасности в локальных сетях. В многочисленных работах рассматриваются средства аутентификации [1, 2]:

– применение многократных паролей. Этот подход используется часто, но является наиболее уязвимым – код пароля может быть прочитан атакующим;

– использование одноразовых паролей. Достоинством является невозможность их использования при перехвате, поскольку пароль уже не будет действовать. Хорошей реализацией этого подхода является схема S/Key;

– контрольные суммы используются при вычислении значения для проверки сообщений переменной длины, но, возникает проблема передачи их второй стороне. Решением является шифрование контрольной суммы или включение их в цифровую подпись.

– электронная подпись, созданная хешированием сообщения с дополнительной информацией (контрольной суммы, ключа), шифрованием полученного кода при помощи личного ключа отправителя. Пользователь может расшифровать это послание, используя открытый ключ, вычислив хэш от кода сообщения и сравнив его с расшифрованным. Если они совпадут, все хорошо.

- Аутентификация для организации облачных вычислений имеет особенности [3]:
- для удобства пользователей должны использоваться механизмы однократной удаленной аутентификации при доступе к различным облачным сервисам;
 - для обеспечения взаимодействия облачных сервисов с сервисом аутентификации должны использоваться распространенные протоколы, стандарты и модели контроля доступа;
 - должна обеспечиваться информационная безопасность сервисов аутентификации.

Способы и протоколы аутентификации в мобильных приложениях

Основные способы аутентификации, их применение и соответствующие протоколы приведены в табл. 1.

Таблица 1. Способы и протоколы аутентификация пользователей МП [4–6]

Способ	Основное применение	Протоколы
По паролю	Аутентификация клиентов	HTTP, Forms
По сертификатам	Аутентификация клиентов в безопасных приложениях; аутентификация сервисов	SSL/TLS
По одноразовым паролям	Дополнительная аутентификация клиентов (двухразовая проверка – TFA)	Forms
По ключам доступа	Аутентификация сервисов и приложений	SSL/TLS
По знакам (токенам)	Предоставляемая аутентификация клиентов; предоставляемая авторизация программ	SAML, WS-Federation, OAuth, OpenID Connect

Для выбора того или другого вида аутентификации, нужно понимать, как различия технологий при их использовании, так и затраты на реализацию. Выбрав ряд критериев, можно провести сравнительную оценку различных технологии аутентификации при том или ином решении. Критерии, по которым вычисляется интегральная оценка вида аутентификации, можно определить следующим образом: разновидность способа ввода кода; тип считывающего устройства; стойкость; затраты на разработку и внедрение; трудность использования для клиента; применимость удаленного доступа; множественность настроек; степень стойкости от настроек; срок хранения аутентифицирующей информации; вероятность ошибок. Данные положения были положены в основу работы системы по интеллектуальному выбору вида доступа пользователей к мобильным приложениям.

Интеллектуальный выбор аутентификации и идентификации

Естественно, возникает вопрос выбора той или иной модели аутентификации (МА) и модели идентификации (МИ). Рассмотрим детально модели для поддержки принятия решения по выбору МА и МИ. В [7] были приведены символичные описания видов моделей аутентификации клиента и было предложено выбирать вид аутентификации по ряду критериев для реализации обеспечения информационной безопасности.

Алгоритм аутентификации состоит из ряда выполняемых проверок двух видов: к первому относятся процедуры входа и идентификации нового пользователя ИС, ко второму – процедуры предъявления аутентификационных признаков (АП), протоколы обмена претендент – контролирующая сторона, проверка и принятие решения о результате прохождения клиентом процесса аутентификации. Модель классификации аутентификации – *MCA* представим тройкой:

$$MCA = \{A, W, C\},$$

где *A* – доступность (accessibility), *W* – целостность (wholeness), *C* – конфиденциальность (confidently). Детализацию модели классификации выразим таким образом:

$$A = \{GA, DA, CA, PA\}; W = \{WS, WRR, WAP, WPC\}; C = \{CPP, CAP, CPC\},$$

где *GA* – гарантия обработки запросов пользователей на аутентификацию, *DA* – разделение доступа пользователей, *CA* – контроль доступа, *PA* – персонификация доступа; *WS* –

целостность ПО, *WRR* – целостность учетных записей, *WAP* – целостность АП пользователя, *WPC* – целостность АП пользователя в облачной среде; *CPP* – конфиденциальность учетных записей, *CAP* – конфиденциальность АП пользователя, *CPC* – конфиденциальность АП пользователя в облачной среде.

МИ базируется на трехуровневом семифакторном подходе классификации идентификаторов:

- как характеристик принадлежности – универсальный (*U*), корпоративный (*C*); личный (*P*);
- для распознавания личности владельца – анонимный (*N*), персональный (*I*);
- доступа владельца к ресурсам – одноразовый (*O*) или многократный (*M*). Тогда

в электронном пространстве имеем разновидности идентификации, представленной семеркой МИ (*MCI*):

$$MCI = \{UNM, UPO, UPM, CNO, CNM, CPM, IPM\},$$

где следующие идентификаторы: *UNM* – универсальный анонимный многократный (пользователь Интернета), *UPO* – универсальный персональный одноразовый (генератор одноразовых паролей), *UPM* – универсальный персональный, содержащийся в реестре (электронный паспорт), *CNO* – корпоративный анонимный одноразовый (электронный билет), *CNM* – корпоративный анонимный многократный (банковская карта), *CPM* – корпоративный персональный многократный (пропуск – смарт-карта) анонимный многократный идентификатор, *IPM* – личный персональный многократный (биометрия на карте или сервере).

Интеллектуальный подход для выбора вариантов СИА базируется на составлении базы правил выбора, в качестве интеллектуального инструмента используются основанные на правилах, экспертные системы (ЭС). ЭС в базе знаний содержат описание классификационных правил СИА, соответствующие профилям легальных пользователей. Эксперт (с инженером по знаниям) формирует базу правил для выбора варианта СИА. Работа такой ЭС может базироваться как на экспертном подходе, так и в автоматическом режиме сервера. Для первого варианта организуется диалог, в ходе которого выявляются пожелания или требования пользователя, или администратора КИС. На основании результатов опроса формируется вариант СИА. В автоматическом режиме сервера вариант СИА формируется по профилям легальных клиентов.

Структура экспертной системы для выбора модели аутентификации

На базе критериев, перечисленных выше, будет происходить оценка пяти различных разновидностей аутентификации: использованием многократного пароля; одноразового пароля; код по персональному; графический пароль; биометрическая аутентификация.

Для подбора наиболее оптимального вида аутентификации эксперту нужно ввести следующую информацию:

- ряд вопросов $Q = \{q_1, \dots, q_N\}$, где q_1, \dots, q_N – вопросы пользователю;
- ряд ответов $A = \{a_1, \dots, a_G\}$, где a_1, \dots, a_G – ответы на вопросы Q ;
- несколько критериев $C = \{c_1, \dots, c_K\}$, где c_1, \dots, c_K варианты критериев.
- соответствие вопросов и критериев $W = A \times C$;
- виды аутентификации $S = \{s_1, \dots, s_M\}$;
- соответствие методов и критериев $R = S \times C$.

Ряд вопросов будет состоять из десяти заранее подготовленных вариантов, с двумя выборами ответа «ДА/НЕТ».

Разработана структура экспертной системы для выбора модели аутентификации, включающая интерфейс клиента, базу знаний по описанию видов аутентификации, процессор вывода лучшего варианта аутентификации. Клиент сам формирует желаемый вариант доступа, отвечая на ряд вопросов. Предусмотрен режим автоматического выбора вариантов доступа в зависимости от доступных ресурсов.

Процесс разработки физической структуры разрабатываемого программного продукта инструментами Visual Studio начинается с формирования пользовательского интерфейса

и разработки кода и различных рабочих форм с учетом эргономичности, эстетичности, минимализма. При разработке программы использовано два вида форм:

– формы, которые создаются при запуске программы и затем, при открытии или закрытии, просто прорисовываются либо скрываются;

– диалоговые окна, уведомляющие пользователя о произошедшем событии. При работе с такими окнами нельзя начать работу с какими-либо другими формами, пока это окно не будет закрыто.

События, использованные на форме FormTest:

– private void private void init_test() – метод инициализации формы, где объявляются основные переменные;

– private void private void load_questions() – обработчик загрузки вопросов для прохождения опроса;

– private void show_question() – обработчик события отображения вопроса и вариантов ответа;

– private void button_next_Click – обработчик события клика по кнопке далее, где сохраняется ответ пользователя;

– private void show_result() – обработчик события отображения результата.

При запуске программы пользователю показывается страница с вопросом и двумя вариантами ответа, отвечая на которые выбирается лучшая форма аутентификации.

Заключение

1. Представлены символьные модели на основе теории множеств для поддержки принятия решения по выбору видов идентификации и аутентификации пользователей мобильных приложений в конкретной ситуации, которые позволяют по ряду критериев выбрать лучший вариант. Для реализации данных моделей предложен экспертный подход.

2. Разработана структура экспертной системы для выбора модели идентификации и аутентификации, включающая интерфейс, базу знаний, процессор вывода. Клиент формирует желаемый вариант доступа, отвечая на ряд вопросов. Предусмотрен режим автоматического выбора вариантов доступа в зависимости от доступных ресурсов. Приведены детали программной реализации экспертной системы с использованием инструментария системы Visual Studio.

INTELLIGENT MANAGEMENT OF CHOICE FOR USER ACCESS TO MOBILE APPLICATIONS

U.A. VISNIAKOU, M.M. GHONDAGH SAS

Abstract. The model of intelligent choice of authentication method for mobile application user on expert approach was proposed. The structure of such expert system and its details of software realization was shown.

Keywords: authentication method, intelligent choice, expert system.

Список литературы

1. Смит Р.И. Аутентификация: от паролей до открытых ключей. М.: Вильямс, 2002.
2. Бобов М.Н., Конопелько В.К. Основы аутентификации в телекоммуникационных системах. Уч. пособие. Минск: БГУИР, 2009.
3. Вишняков В.А. Информационная безопасность в корпоративных системах, электронной коммерции и облачных вычислениях: методы, модели, программно-аппаратные решения. Монография. Минск: Бестпринт, 2016.
4. Вход в приложения для мобильных устройств или компьютеров [Электронный ресурс]. URL: <https://support.google.com/a/answer/1032419?hl=ru>. (дата доступа: 12.10.2019).
5. Выростков Д. Обзор способов и протоколов аутентификации в веб-приложениях [Электронный ресурс]. URL: <https://habrahabr.ru/company/dataart/blog/262817/>. (дата доступа: 12.10.2019).
6. Как обеспечить аутентификацию на мобильных устройствах [Электронный ресурс]. URL: <http://www.cnews.ru/reviews/security2014/articles/>. (дата доступа: 12.10.2019).
7. Вишняков В.А., М.М. Гондаг Саз. // Докл. БГУИР. 2017. № 1. С. 82–86.

УДК 681.3

КОРРЕКТИРУЮЩИЕ КОДЫ НА ОСНОВЕ БИНОМИАЛЬНЫХ ФРАКТАЛЬНЫХ СТРУКТУР

С.Б. САЛОМАТИН, А.Э. АЛЕКСЕЕНКО

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

Поступила в редакцию 30 октября 2019

Аннотация. Рассмотрена связь полярных кодов и кодов Рида-Маллера с фрактальными биномиальными матрицами Паскаля, что позволяет увеличить многообразие кодов для достижения пропускной способности канала.

Ключевые слова: корректирующие коды, фрактальные матрицы, спектральные преобразования, биномиальные коэффициенты, пропускная способность.

Введение

Полярные коды были первым семейством кодов, которое было доказано как способствующее пропускной способности [1]. Коды Рида-Мюллера могут быть полезны для достижения пропускной способности канала двоичного стирания. Метод использует кодовую симметрию последовательности кодов Рида-Мюллера при этом с увеличением длины блока и скоростью сходимости достигается пропускная способность канала [2]. Отличительной особенностью полярных кодов и кодов Рида-Маллера является простота процедур их построения, кодирования и декодирования, что делает их привлекательными для практического использования.

В работе рассматривается связь полярных кодов и кодов Рида-Маллера с фрактальными биномиальными матрицами Паскаля, что позволяет увеличить многообразие кодов для достижения пропускной способности.

Биномиальные фрактальные матрицы и преобразование Паскаля

Треугольник Паскаля является геометрическим расположением биномиальных коэффициентов в треугольнике [4, 5]

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 2 & 1 & 0 & \dots & 0 \\ 1 & 3 & 3 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{p-1}^0 & C_{p-1}^1 & C_{p-1}^2 & C_{p-1}^3 & \dots & C_{p-1}^{p-1} \end{bmatrix}.$$

Каждый элемент треугольника p_{ij} равен биномиальному коэффициенту a_{ij} ,

$$a_{ij} = C_i^j = \binom{i}{j} = \frac{i!}{j!(i-j)!},$$

где i, j – неотрицательные целые числа, $i \geq j$.

Матрица Паскаля представляет собой бесконечную матрицу $\mathbf{P} = [p_{ij}]$, элементами которой являются биномиальные коэффициенты, расположенные в виде двумерного массива. Биномиальные коэффициенты – это коэффициенты (положительные целые числа), возникающие при алгебраическом разложении неотрицательных целочисленных степеней бинома $(x + y)^n$. Степень n определяет строку матрицы Паскаля, в которой расположены соответствующие биномиальные коэффициенты. Для $y=1$ биномиальные коэффициенты могут быть идентифицированы как коэффициенты слагаемых x^k в полиномиальном разложении $(1 + x)^n$.

Основные свойства матрицы \mathbf{P} :

1. Элементы первого столбца равны 1.
2. Все нижние матрицы треугольные.
3. Сумма элементов каждой строки (кроме первой) равна нулю.
4. Все матрицы равны их обратным.

Дискретное преобразование Паскаля (ДПП)

Преобразование Паскаля может быть представлено в виде:

$$\mathbf{X} = \mathbf{P}\mathbf{x},$$

где \mathbf{X}, \mathbf{x} – векторы размером $N \times 1$ и $\mathbf{P} - (N \times N)$ – матрица преобразования Паскаля, элементы которой равны

$$p_{ij} = \begin{cases} (-1)^j a_{ij}, & i \geq j, \\ 0, & i < j. \end{cases}$$

Элементы матрицы преобразования Паскаль получают непосредственно из записей треугольника Паскаля путем чередования знаков столбцов. Матрица преобразования Паскаля для $N = 4$ приведены ниже

$$\mathbf{P}_4 = [p_{ij}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -3 & 3 & -1 \end{bmatrix}.$$

Преобразование Паскаля может быть представлено в виде:

$$X(k) = \sum_{n=0}^{N-1} p_{kn} x(n),$$

где $k = 0, 1, \dots, N - 1$.

Элементы матрицы преобразования Паскаля выше диагонали равны нулю, т.е. $p_{kn} = 0$ для $k > n$. Дискретное преобразование Паскаля может быть переписано следующим образом:

$$X(k) = \sum_{n=0}^k p_{kn} x(n), \quad n = 0, 1, \dots, N - 1,$$

или

$$X(k) = \sum_{n=0}^{N-1} (-1)^n \binom{k}{n} x(n).$$

Обратное ДПП может быть записано в матричной форме

$$\mathbf{x} = \mathbf{P}\mathbf{X},$$

или

$$x(n) = \sum_{k=0}^n p_{nk} X(k) = \sum_{k=0}^n (-1)^k \binom{n}{k} X(k), \quad k = 0, 1, 2, \dots, N-1.$$

Вычислительная сложность ДПП

Вычисление N -точечного ДПП вектора на основе матричного определения приводит к N^2 операций умножения и $N(N-1)$ операций сложения. Если принять во внимание, что верхние элементы треугольника матрицы преобразования равны нулю, а диагональные элементы равны 1 или -1 , то есть операций умножения не требуется, то вычислительная сложность составляет $N(N-1)/2$ операций умножения и $[N(N-3)+1]/2$ операций сложения.

Граф вычисления N -точечного ДПП показан на рис. 1. Преобразование включает $(N-1)$ этапов и имеет вид «бабочки».

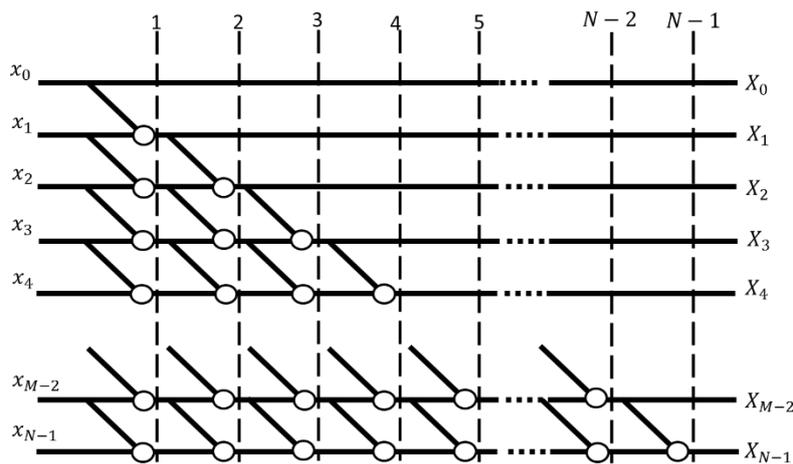


Рис. 1. Граф дискретного преобразования Паскаля

ДПП обладает низкой сложности и фактически представляет собой разновидность быстрого спектрального преобразования.

Структура матрицы Паскаля обладает свойствами фрактала (треугольник Серпинского (рис. 2)) и тесно связана с кодами Риды-Маллера (РМ) [2] и полярными кодами [1].

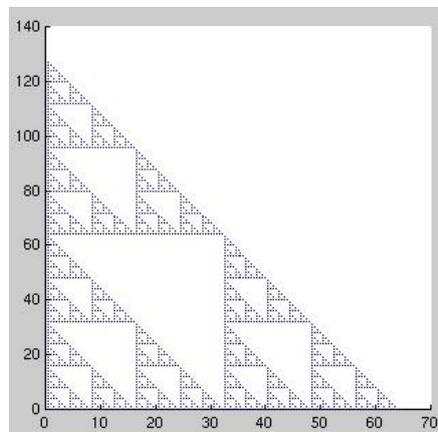


Рис. 2. Треугольник Серпинского

Генераторная матрица \mathbf{G} полярных кодов и РМ может быть получена просто с помощью операции по модулю 2 на матрице Паскаля.

На самом деле, можно создать бесконечно много фракталов, используя следующее выражение

$$\mathbf{G} = [\text{матрица Паскаля}] N \bmod q \text{ (для простого } q\text{)}.$$

Вышеуказанный метод не является уникальным. Можно построить любой фрактал, который исходит из матрицы Паскаля, используя степень Кронекера. Нужно только определить матрицу ядра \mathbf{F} . Это простая задача, поскольку можно получить матрицу ядра, используя

$$\mathbf{F} = [\text{матрица Паскаля}] q \bmod q \text{ (для простого } q\text{)}.$$

Матрицы Паскаля и полярные коды

Поляризация канала была первоначально предложена в [1] для дискретных каналов без памяти с двоичным входом как метод кодирования, который был использован для построения нового семейства кодов для передачи данных. Эти коды, называемые полярными кодами, могут достигать «симметричной пропускной способности» любого канала двоичного ввода, используя алгоритмы кодирования и декодирования малой сложности. Позднее было доказано, что эта конструкция кода может обеспечить пропускную способность любого дискретного канала без памяти с первичным входом [1, 2].

Матрица Паскаля троичного полярного кода

Матрица ядра для $q = 3$ имеет вид

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Матрица генератора является фракталом треугольника Серпинского. Линейное преобразование с матрицей ядра объединяет три идентичных канала в новый синтетический канал W_3 :

$$\{W, W, W\} \rightarrow \{(y_1^3; u_1), (y_1^3, u_1; u_2^3), (y_1^3, u_1^2, u_3)\},$$

где $\{u_i\}$ – входы, а $\{y_i\}$ – выходы канала.

Структура канала показана на рис. 3.

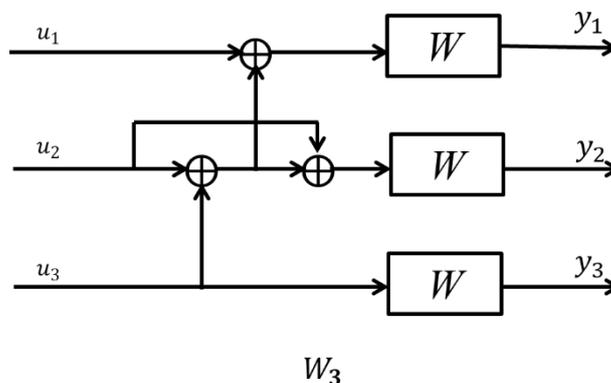


Рис. 3. Структура симметрично канала полярного кода

Сложность кодирования и декодирования троичного полярного кода оценивается как $O(N \log_3 N)$.

Заключение

Фрактальные биномиальные матрицы и ДПП являются перспективным инструментом для кодирования и обработки кодов и сигналов. Треугольная структура дискретного преобразования матрицы Паскаля производит полезное свойство локализации, что позволяет реализовать симметричную пропускную способность канала и уменьшить сложность кодирования и декодирования.

CORRECTIVE CODES BASED ON BINOMIAL FRACTAL STRUCTURES

S.B. SALOMATIN, A.E. ALEKSEENKO

Abstract. The connection of polar codes and Reed-Muller codes with fractal binomial Pascal matrices, which allows to increase the variety of codes to achieve channel throughput was considered.

Keywords: fractal matrix correction codes, spectral transformations, binomial coefficients, throughput.

Список литературы

1. Arikan E. // Proc. IEEE Inform. Theory Workshop. 2010. P. 1–5.
2. Park W. // IEEE Transactions on Information Theory. 2012. Vol. 59. P. 955–969.
3. Doan D., [et. al.] Demonstratio Mathematica, Vol. XLII, № 1, 2009.
4. Massopust P. Fractal function, fractal surfaces, and wavelets. Academic Press. 1994.

УДК 004.932

ГРАДИЕНТО-ВОЛНОВОЕ ВЫРАЩИВАНИЕ ОБЛАСТЕЙ АСМ-ИЗОБРАЖЕНИЙ

В.В. РАБЦЕВИЧ, В.Ю. ЦВЕТКОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 02 ноября 2019*

Аннотация. Предложен квазипараллельный градиентно-волновой алгоритм выращивания областей вокруг локальных максимумов в результате присоединения к ним соседних пикселей, которые не превышают пороговую скорость изменения градиента. Произведено сравнение с основными алгоритмами сегментации, применяемыми к изображениям атомно-силовой микроскопии (далее АСМ) по количеству сегментов, по однородности яркости сегмента и по контрасту между соседними сегментами. Показано, что разработанный алгоритм способен определить границы объектов, по выбранным критериям, лучше, чем известные алгоритмы.

Ключевые слова: АСМ, волновое выращивание, градиент, мера однородности, мера контраста.

Введение

Для различных алгоритмов сегментации, применяемых для обработки изображений, необходимы специально подобранные методы тестирования и оценки, которые учитывают требования к результату, используемый алгоритм и особенности обрабатываемых данных. При автоматическом анализе неорганических наноструктур, полученных с помощью атомно-силовой микроскопии, необходимо учитывать, что исходными данными являются матрицы, содержащие информацию о высоте размещения объектов на подложке, которую можно интерпретировать, как яркость. Сегментация является первым шагом в процессе обработке АСМ-изображений.

В работе [1] показано, что оценка алгоритмов сегментации зависит от: самого алгоритма; исходных данных; заданных критериев для объединения в сегменты; выбранных методов оценки. Сегментация является промежуточным этапом, используемым при обработке данных в специализированном программном обеспечении, для проведения сканирования и анализа. Основные подходы при сегментации АСМ-изображений можно выделить в три большие группы: пороговые алгоритмы (чувствительны к различным артефактам сканирования, например, при значительном наклоне подложки часть информации будет потеряна; применим для выделения отдельно расположенных объектов); алгоритмы водораздела (применимы как для материалов с ячеистой структурой так и для отдельно стоящих объектов, но способны построить линию водораздела только при наличии ярко выраженного контраста между областями) [2]; текстурные алгоритмы на основе нейросетей, которые анализируют все три канала АСМ-изображений (topography, torsion, phase) и подходят для специфических видов материалов, где объекты имеют четко выраженную текстуру, отличную от подложки [3].

Разработанный квазипараллельный градиентно-волновой алгоритм выращивания областей (ГВВО) (рис. 1) будет сравниваться с алгоритмом волнового выращивания областей локальных максимумов с выбором пикселей в порядке убывания значений (ВОЛМА) [4], с алгоритмом маркерного водораздела с автоматической расстановкой маркеров (МВА) [5], с алгоритмом выращиванием областей (ВО) [6] по критериям, не требующим наличия эталонных (ground truth) изображений или оценки экспертов [1].

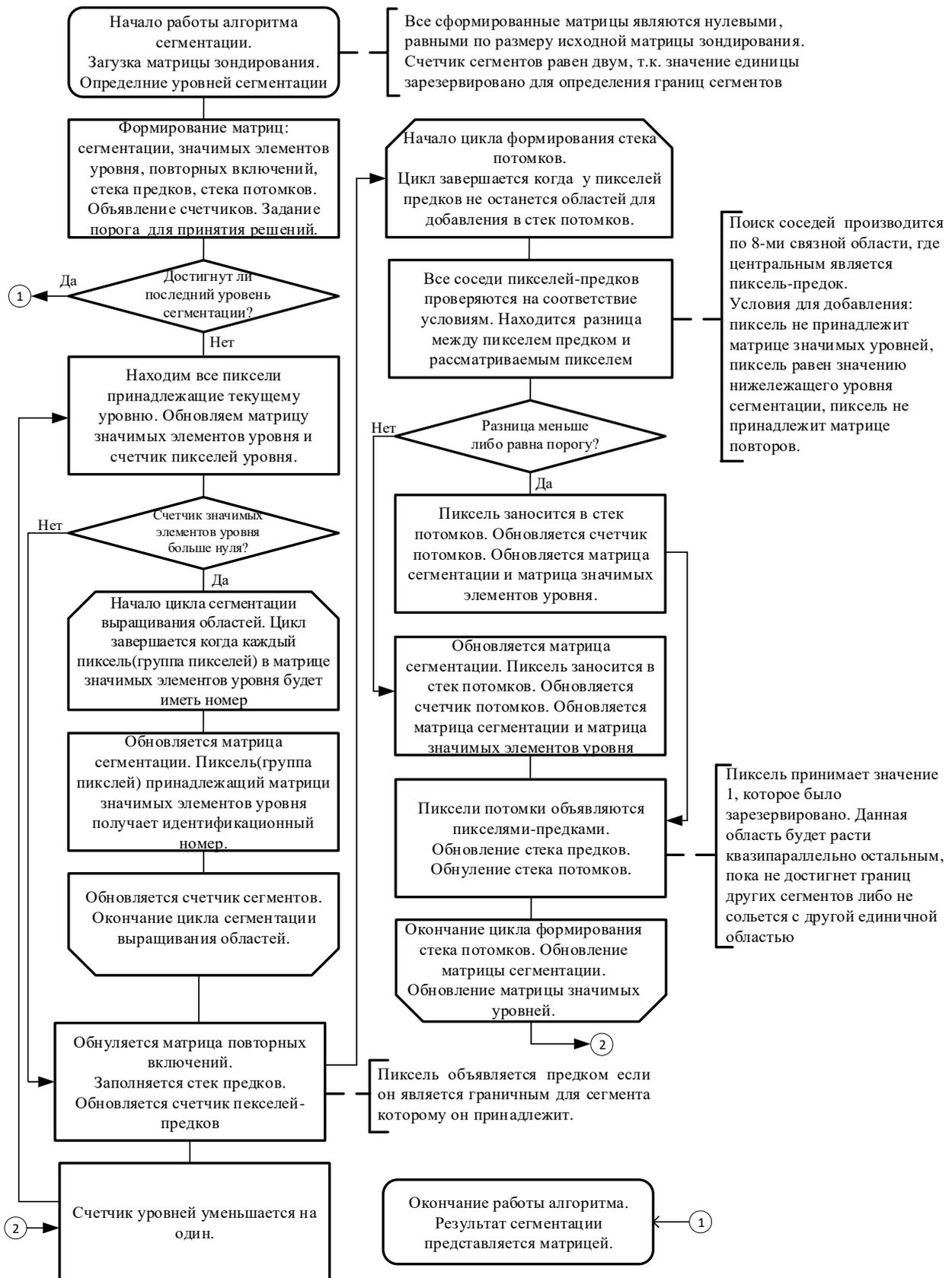


Рис. 1. Блок-схема градиентно-волнового алгоритма сегментации АСМ-изображений

Алгоритм градиентно-волновой сегментации АСМ-изображений

Разработанный алгоритм производит квазипараллельное выращивание областей вокруг начальных точек роста, в пределах одного уровня яркости. Для каждого пикселя в сегменте отслеживается скорость изменения градиента, при превышении которой пиксели помещаются в особую область, которая рассматривается как фон и разрастается параллельно остальным. Эта особенность алгоритма, увеличивает количество сегментов на 1. Работа алгоритма представлена на блок-схеме (рис. 1).

В результате работы алгоритма формируется матрица сегментов, внутри которых дисперсия признака минимальна. Разработанный алгоритм подходит для АСМ-изображений со структурой, имеющей отдельно стоящие объекты на подложке. Например, изображения микросхем, контактные окна микросхем, волокна целлюлозы, бактерии и т.д.

Критерии оценки качества сегментации

Критерии оценки качества сегментации можно разделить на две основные группы: субъективные и объективные [7, 8]. Субъективные методы используют групповую оценку экспертов [9]. Объективные методы не требуют наличия эксперта и в свою очередь делятся на системные и прямые. Системные методы способны оценить качество на основе конечных результатов, т.е. оценивают только матрицу сегментации на выходе системы [10]. Прямые методы работают как с алгоритмом, так и с результатом его работы. Прямые методы можно представить двумя группами: аналитические (которые рассматривают только алгоритм), и эмпирические (которые оценивают только результат работы). Эмпирические методы в свою очередь подразделяются на контролируемые (использующие для оценки эталон изображения) и автоматические или неконтролируемые (оценивают только качество разделения изображения на области по определенному признаку, не опираясь на эталоны) [7, 8].

Для оценки качества сегментации АСМ-изображений, которые представлены разнообразным количеством неорганических и органических материалов и структур, были выбраны критерии, не требующие наличия эталона и оценки эксперта: однородность признака внутри сегментов и контраст между соседними сегментами [1, 7, 11–14]. Так же было учтено количество полученных сегментов по сравнению с исходными тестовыми изображениями.

Как следует из [14] все алгоритмы определения контраста для специализированных изображений можно разделить на три категории: глобальный контраст, локальный пиксельный контраст и контраст в области интереса. Поскольку АСМ-изображения могут иметь артефакты сканирования, был выбран метод определения контраста для определенной области интереса.

Мера однородности (МО) основана на значении дисперсии признака сегмента [7] и вычисляется по формуле:

$$U_{Mz} = 1 - \sum_{S_i \in Mz} \frac{\omega_i \sigma_i^2}{N}, \quad (1)$$

где $\omega_i = \frac{A_i \sigma_i^2}{A_{Mz} \sigma_{\max}^2}$ – вес, определяющий вклад сегмента в АСМ-изображение, A_i – площадь

рассматриваемого сегмента, A_{Mz} – площадь всего региона интереса, σ_i^2 – дисперсия признака (здесь и далее под признаком понимается значение яркости или высоты для канала topography, АСМ-изображений) в рассматриваемом сегменте, которая определяется по формуле (2), σ_{\max}^2 – максимальное значение дисперсии признака, определяется по формуле (3), N – нормирующий множитель.

$$\sigma_i^2 = \sum_{i \in S} \frac{(f_i - \bar{f}_i)^2}{A_i}, \quad (2)$$

$$\sigma_{\max}^2 = \frac{1}{2}(f_{\max} - f_{\min})^2, \quad (3)$$

где f_{\max}, f_{\min} – максимальные и минимальные значения признака на всем регионе интереса (для тестовых изображений – вся матрица зондирования Mz), \bar{f}_i – среднее значение яркости в рассматриваемом сегменте. Мера контраста (МК) между соседними сегментами всего региона интереса определяется как:

$$C_{Mz} = \frac{\sum_{S_i \in Mz} \omega_i c_i}{\sum_{S_i \in Mz} \omega_i}, \quad (4)$$

где $c_i = \sum_{S_j} p_{ij} c_{ij}$, – контраст рассматриваемого сегмента, p_{ij} – коэффициент смежности

рассматриваемого сегмента; $c_{ij} = \frac{|\bar{f}_i - \bar{f}_j|}{\bar{f}_i + \bar{f}_j}$ – контраст двух соседних сегментов [7–12].

Оценка эффективности градиентно-волнового алгоритма сегментации

Выполнено сравнение предложенного градиентно-волнового алгоритма выращивания областей (ГВВО) с алгоритмами, реализованными в среде Matlab: маркерного водораздела с автоматической расстановкой маркеров (МВА) [5], выращиванием областей (ВО) [6], волнового выращивания областей локальных максимумов с выбором пикселей в порядке убывания значений (ВОЛМА) [4], которые обычно используются для анализа АСМ-изображений. Оценка работы (табл. 1) выбранных алгоритмов проводилась по количеству сегментов (КС), по контрасту между сегментами (МК), по однородности сегментов (МО). Для этого эксперимента были выбраны четыре, синтезированных в Gwyddion, изображения размером пикселей [15]: АСМ-1 (10 полусфер рис. 2, а), АСМ-2 (35 склеенных полусфер, рис. 2, б), АСМ-3 (8 полусфер с четкими границами, рис. 2, в), АСМ-4 (100 полусфер с наложением друг на друга рис. 2, г). И четыре изображения искусственно созданные в Matlab: АСМ-5 (100 разноуровневых полусфер, рис. 2, д), АСМ-6 (100 тороидов одинаковой высоты, рис. 2, е), АСМ-7 (100 разноуровневых тороидов, рис. 2, ж) и АСМ-8 (100 тороидов и 100 полусфер лежащих на разном уровне и имеющие общие границы, рис. 2, з). В табл. 1 приведены значения, полученные при сегментации указанных тестовых изображений. На рис. 3. представлены оценки для различных алгоритмов.

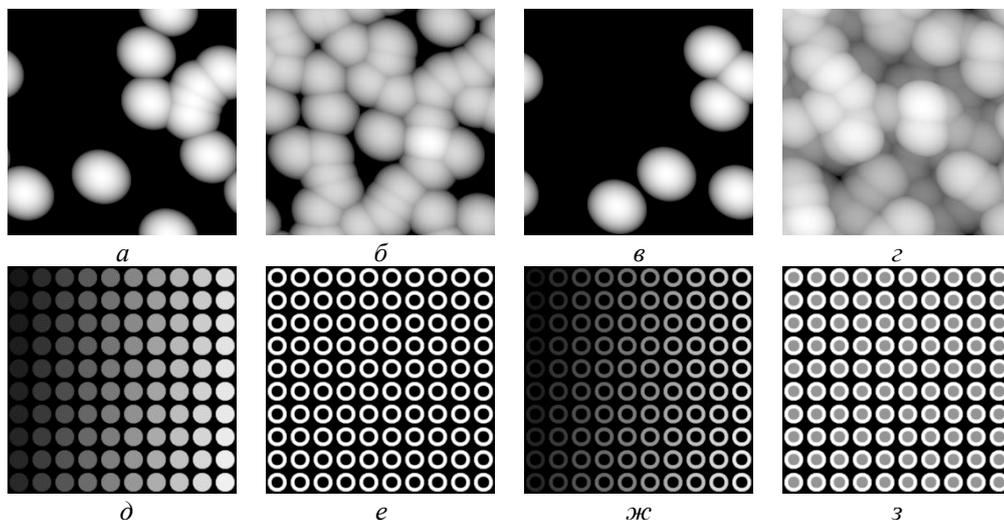


Рис. 2. Тестовые полутоновые изображения: а – АСМ-1; б – АСМ-2; в – АСМ-3; г – АСМ-4; д – АСМ-5; е – АСМ-6; ж – АСМ-7; з – АСМ-8

Таблица 1. Результат сравнения алгоритмов сегментации

№	Алгоритм сегментации											
	ВОЛМА			Градиент			Водораздел			RG		
	КС	МО	МК	КС	МО	МК	КС	МО	МК	КС	МО	МК
АСМ-1	14	0,7426	0,2530	15	0,9663	0,7187	8	0,8290	0,6775	19605	1	0,0000
АСМ-2	47	0,8657	0,0997	48	0,9684	0,2764	6	0,8564	0,0668	38388	1	0,0000
АСМ-3	8	0,7754	0,2431	9	0,9737	0,7976	6	0,9301	0,9752	13960	1	0,0000
АСМ-4	47	0,9520	0,0708	48	0,9833	0,1279	10	0,9762	0,1612	32220	1	0,0000
АСМ-5	100	0,7272	0,0383	101	0,9417	0,6566	97	0,9864	0,8650	60681	1	0,0000
АСМ-6	100	0,6244	0,0000	101	0,8163	0,7094	1	0,0000	0,0000	119401	1	0,0000
АСМ-7	100	0,7710	0,0383	101	0,9349	0,6255	1	0,0000	0,0000	101145	1	0,0000
АСМ-8	100	0,6518	0,0000	101	0,7816	0,5474	1	0,0000	0,0000	112201	1	0,0000

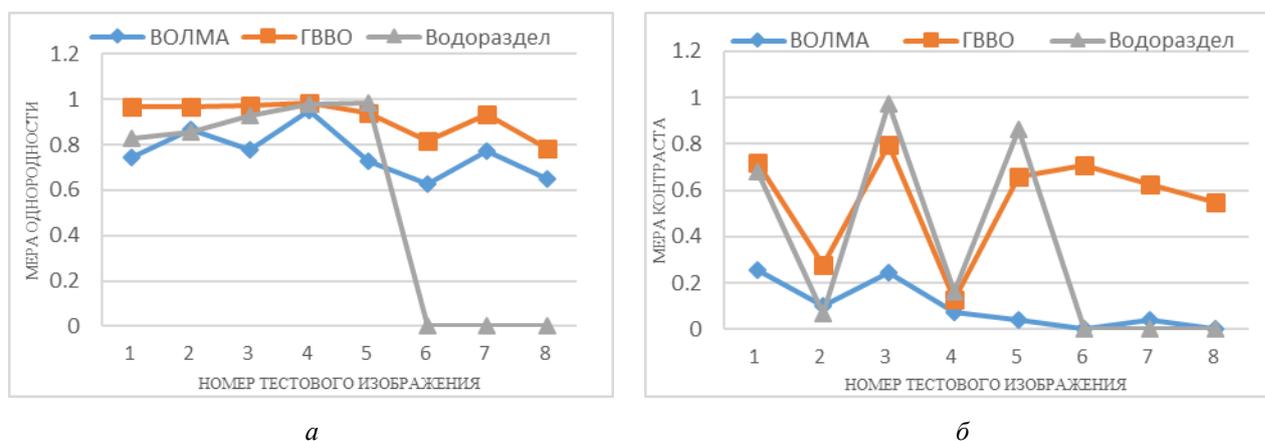


Рис. 3. Результат работы алгоритмов сегментации: *а* – мера однородности для трех алгоритмов; *б* – мера контраста для трех алгоритмов

Заключение

Как видно из полученных данных, разработанный градиентно-волновой алгоритм показывает стабильно высокую однородность яркости внутри сегментов, вне зависимости от расположения объектов друг относительно друга. Значения однородности для разработанного алгоритма изменяются в пределах от 0,7816 до 0,9762, для ВОЛМА от 0,6244 до 0,9762, для маркерного водораздела от 0 до 0,9864. Мера контраста в разрабатываемом алгоритме выше для изображений, на которых есть отдельно стоящие области и выделяемый фон. Для отдельно стоящий структур МК изменяется в пределах от 0,6255 до 0,7976, для ВОЛМА от 0 до 0,2530 и для водораздела от 0 до 0,9752. На тестовых изображениях алгоритм показывает устойчивый результат по контрасту между соседними областями. Так же разработанный алгоритм на ряду с алгоритмом ВОЛМА получают наиболее точное количество сегментов изображений по сравнению с другими алгоритмами.

GRADIENT-WAVE GROWING SEGMENTATION ALGORITHM OF AFM- IMAGES

V. V. RABTCEVICH, V. Yu. TSVIATKOU

Abstract. A quasi-parallel gradient-wave algorithm for growing regions around local maxima as a result of adjoining neighboring pixels to them, which do not exceed the gradient change the of threshold rate was proposed. A comparison with the basic segmentation algorithms applied to images of atomic force microscopy according to the number of segments, as the brightness into the segment and contrast between adjacent segments was made. It was shown that the developed algorithm was able to determine the segments boundaries according to the selected criteria.

Keywords: AFM, wave growth, gradient, uniformity measure, contrast measure.

Список литературы

1. Кольцов П.П., [и др.] // Компьютерная оптика. 2015. Т.39. № 4. С. 542–556.
2. Eaton P., West P. Atomic force Microscopy. Oxford University Press, 2010.
3. Marker-Controlled Watershed Segmentation [Electronic resource] URL: <https://www.mathworks.com/help/images/examples/marker-controlled-watershed-segmentation.html> (date of access: 02.10.2019).
4. Tiryaki V.M., [et. al.] // Cytometry. The journal of the International Society for Analytical Cytology. 2015. P. 1090 – 1100.
5. Рабцевич В.В., [и др.] // Вес. Нац. акад. навук Беларусі. Сер. фіз.-тех. навук. 2019. № 2. С. 215–231.
6. Альмияхи О.М., Цветков В.Ю., Конопелько В.К. // Докл. БГУИР. 2016. № 3 (97). С. 24–30.
7. Захаров А.В., [и др.] // Труды НИИСИ РАН. 2012. Т. 2. № 2. С.87–99.
8. Zhang H., Fritts J.E., Goldman S.A. // Computer Vision and Image Understanding. 2008. P. 260–280.
9. Gelasca E.D., [et. al.] // Proceedings of Conference on Computer Vision and Pattern Recognition Workshop. 2004. Vol. 4. P.576–579.
10. Shin M.C., Goldgof D., Bowyer K.// Computer vision and Image Understanding. 2001. Vol. 84 (1). P. 160–178.
11. Ouattara S., [et. al.] // Journal of Software Engineering and Applications. 2011. Vol. 4. P. 371–378
12. Sankur B., Sezgin M. // Pattern Recognition. 2001. Vol. 13 (1). P. 146–168.
13. Peters II R.A., [et. al.] // Automatic Target Recognizer System and Technology Conference. 1990. P. 1–17.
14. Remes V., Haindl M. // Kybernetika. 2018. Vol. 54. P. 978–990.
15. Gwyddion [Electronic resource]. URL: <http://gwyddion.net> (date of access: 02.04.2019).

УДК 621.391

ИНТЕЛЛЕКТУАЛЬНЫЙ КЛАССИФИКАТОР ОБНАРУЖЕНИЯ ВТОРЖЕНИЙ НА БАЗЕ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ

М.Г. МОЗДУРАНИ ШИРАЗ, В.А. ВИШНЯКОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 01 ноября 2019*

Аннотация. Предложена модель обнаружения вторжений на основе генетических алгоритмов. Найдена лучшая хромосома в цикле эволюции, которая применена к тестовому набору данных типовых атак KDD. Построен классификатор на базе модели, для которого получена хорошая степень распознавания атак и показана высокая скорость обнаружения вторжений, продемонстрировав результаты лучше, чем аналогичные. Также этот классификатор может обнаружить значительный процент атак R2L, который выше, чем у аналогов, разработанных ранее.

Ключевые слова: обнаружение вторжений, генетические алгоритмы, набор данных атак KDD, классификатор атак.

Введение

С развитием новых инфокоммуникационных технологий (ИКТ) и расширением сетевых компьютерных систем при значительном приросте интернет-трафика, конфиденциальные данные находятся под постоянной угрозой от различных атак. Вторжения являются распространенными угрозами для сетей, и с ростом количества злоумышленников, разработка эффективных инструментов безопасности затрудняется. Современные инструменты безопасности, такие как межсетевые экраны (файрволлы), системы обнаружения вторжений (IDS) и другие средства, недостаточны для обеспечения целостности безопасности. Возникает необходимость в обнаружении неизвестных вредоносных вторжений, т.е. вторжения становятся все более важной технологией, которая отслеживает сетевой трафик и распознает незаконное использование локальных компьютерных систем.

Разница между известными продуктами безопасности и системами обнаружения вторжений заключается в том, что последним требуется больше интеллекта. Они должны анализировать собранную информацию и выводить результаты в отсутствие сигнатур атак в базе. Основная проблема заключается в том, что эти системы не могут обнаружить новые атаки без известных шаблонов, а также производные шаблоны известных атак [1]. В последнее время IDS стали одним из важных предметов исследований в области информационной безопасности (ИБ), новые исследования сосредоточены на том, как сделать IDS интеллектуальными. Т. е. улучшить их таким образом, чтобы иметь возможность автоматически обнаруживать новые формы атак. Эти исследования сосредоточены на объединении методов искусственного интеллекта с механизмами обнаружения вторжений для повышения точности идентификаторов (т. е. производительности обнаружения), эффективности и удобства использования. Некоторые исследования используют методы интеллектуального анализа данных, другие – нейронные сети или генетические алгоритмы. Анализ систем обнаружения вторжений (СОВ) приведен в работе [2]. Ниже рассматривается построение интеллектуальной СОВ на базе генетических алгоритмов.

Описание наборов данных KDD99

Набор данных KDD99 типовых атак [2] очень важен для оценки идентификаторов обучения, он является эталоном для тестирования IDSs. Атаки этого набора подразделяются на четыре основные категории:

- DOS: отказ в обслуживании, например, SYN flood;
- R2L: несанкционированный доступ с удаленного компьютера, например, угадывание пароля;
- U2R: несанкционированный доступ к привилегиям локального пользователя (root), например, различные атаки переполнения буфера;
- Зондирование: наблюдение и другое зондирование, например, сканирование портов.

Набор данных KDD99 имеет два типа данных:

– набор данных обучения: он используется для обучения любого алгоритма идентификаторов обучения, поскольку он содержит тип атаки для каждого соединения. Этот набор данных имеет 24 типа обучающих атак и содержит 4898431 соединения.

– набор данных тестирования: используется для проверки эффективности идентификаторов в исследованиях, например, для проверки способности нового алгоритма обнаружения выявлять атаки в наборе данных обучения, кроме новых 19 атак, которые включены в набор данных тестирования. Набор данных тестирования содержит 311 029 соединений.

В данных KDD99 каждое соединение имеет 42 функции (включая метку класса), которые содержат информацию о сеансе. Эти функции используются для распознавания нормальных соединений от атакующих; их можно разделить на четыре категории.

1. Basic: содержит основные характеристики записи подключения.
2. Содержимое: создаются из полезной нагрузки пакетов трафика и содержат сведения, связанные с хостом, такие как количество сбоев входа в систему.
3. Трафик: содержит статистические данные, такие как количество подключений к одному и тому же хосту в двухсекундном временном окне.
4. Класс: указывает, является ли соединение нормальным или интрузивным; используется для обучения и оценки.

Алгоритм обучения классификатора

В задаче проектирования интеллектуальной системы обнаружения вторжений она была подвержена воздействию атак из наборов данных KDD99 [2]. Набор обучающих данных содержит около 5 миллионов записей подключения, каждая запись содержит 41 функцию. Процесс занимает много времени, чтобы обучить систему с учетом всех особенностей. Но если бы можно уменьшить количество функций, например, до 8 функций, сильно уменьшается количество вычислений.

Пространство набора данных можно разделить на:

- 2 основных класса: обычный трафик и атака трафик;
- более точно – 5 классов: нормальный, DoS, R2L, U2R, зонд.

Таким образом, пространство набора данных может быть описано следующими величинами:

X : случайная величина = {нормальный, DOS, U2R, R2L, зонд}

Y : функция подключения: (41 значение независимых случайных величин)

Пример: $Y(\text{Protocol_type}) = \{\text{ICMP, TCP, UDP}\}$

Теперь, чтобы извлечь значение функции соединения ($X.\text{protocol_type}$), можно вычислить «объем информации о X (нормальное соединение или DOS, R2L, U2R, зонд), содержащийся в Y (функция соединения X тип протокола)».

Оценка эффективности алгоритма

Подход к генетическому алгоритму (ГА) зависит от результатов функций селекции, имеющих параметры: длительность, обслуживание, байты источника (src_bytes), байты результата (dst_bytes) и количество [4]. Генетический классификатор называется *GACL* и может быть записан в виде:

$$GACL(conn_j) = \begin{cases} \text{if } \sum_{i=1}^n W_i f_i(conn_j) < \text{threshold} \text{ then } conn_j \text{ is } ATTACK \\ \text{else } conn_j \text{ is } NORMAL \end{cases},$$

где $conn_j = connection_j = (\text{duration}, \text{service}, \text{src_bytes}, \text{dst_bytes}, \text{count})$.

Неравенство было выбрано по причине того, что для предыдущих признаков большинство атакующих соединений в обучающем наборе KDD имеют низкие значения. Например, набор данных обучения KDD содержит 396743 записи атак, 396083 из которых имеют значение длительности, равное 0.

Шесть параметров управления доступом кодируются следующим образом:

– пять параметров представляют веса объектов (W_1, W_2, W_3, W_4, W_5);

– один параметр представляет собой пороговое значение.

Каждый параметр представляет собой двоичный ген длиной 11 бит: 10 бит для значения параметра, один бит для знака. Число 10 бит (значение параметра) было выбрано с точностью $2^{10} \approx 0,001$. Пороговое значение остается всегда положительным.

В качестве примера можно получить двоичное значение гена следующим образом:

$Gene1 = 11000000010 = 1 \& 1000000010$ (11-й бит = 1, так что знак отрицательный).

Десятичное значение гена равно:

$Gene1 = 1000000010 = -1 \cdot 2^9 + 0 \cdot 2^8 + 0 \cdot 2^7 + 0 \cdot 2^6 + 0 \cdot 2^5 + 0 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = -514$.

В табл. 1 показаны хромосомные гены.

Таблица 1. Кодирование параметров

Гены	Параметры генов	Границы генов	Значения представлений	
			Двоичное	Десятичное
<i>Gene1</i>	$W_{duration}$	[-1023,1023]	11000000010	-514
<i>Gene2</i>	$W_{service}$	[-1023,1023]	01000000010	514
<i>Gene3</i>	W_{src_bytes}	[-1023,1023]	00000000010	2
<i>Gene4</i>	W_{dst_bytes}	[-1023,1023]	11111111111	-1023
<i>Gene5</i>	W_{count}	[-1023,1023]	01111111111	1023
<i>Gene6</i>	<i>Threshold</i>	[0,1023]	00101011010	532

Хромосома предыдущих генов имеет вид, представленный в табл. 2.

Таблица 2. Хромосома предыдущих генов

<i>Gene6</i>	<i>Gene5</i>	<i>Gene4</i>	<i>Gene3</i>	<i>Gene2</i>	<i>Gene1</i>
00101011010	01111111111	11111111111	00000000010	01000000010	11000000010
532	1023	-1023	2	514	-514

Таким образом значение $GACL$ для этой хромосомы:

$$GACL(conn_j) = \begin{cases} \text{if } -514 f_{duration} + 514 f_{service} + 2 f_{src_byte} - 1023 f_{dst_byte} + 1023 f_{count} < 532 \\ \text{then } conn_j \text{ is } ATTACK \text{ else } conn_j \text{ is } NORMAL \end{cases}.$$

Перед определением структуры программы для работы классификатора приведем основные понятия из теории ГА [3]. Гены – это параметры равные сумме веса и порога. Хромосома – это решение – генетический классификатор $GACL$. Индивид – это то же самое, что хромосома с дополнительными переменными для хранения значения пригодности, показателей производительности и других параметров. Потомство – это ребенок, полученный в результате операции скрещивания (кроссовера). Популяция – совокупность индивидов.

Поколение – это то же самое, что и популяция с дополнительными переменными для хранения всей пригодности поколения, лучшего и худшего индивидуума и других.

Запуск: программа запуска операции эволюции начинается с генерации начальной случайной популяции, проходит через создание нескольких поколений и заканчивается созданием последнего поколения, которое содержит искомое решение, если оно существует. На рис. 1 представлена структура программы.

```

Type INDIVIDUAL // Структура
crom (1..gensNb*genLength) As Byte // Классификатор GACL 110101010000100110001100
decCode (1..gensNb) As Integer // десятичное значение (512 340 211 -20 -10 20)
fitVal As Double // значение функции пригодности
fitProb As Double // вероятность пригодности общей приспособленности популяции
reFit As Boolean //индикатор – нужно ли пересчитывать значение пригодности
TP As Long // number количество атакующих соединений, обнаруженных данной хромосомой
TN As Long // # of нормальных связей, которые были обнаружены этой хромосомой
FP As Long // # ложноположительные связи этой хромосомы
FN As Long // # ложноотрицательные связи этой хромосомы

```

Рис. 1. Структура программы

При запуске операции эволюции все результаты будут сохранены в базе данных в соответствии с предыдущей структурой данных, что позволит извлечь и построить много полезных статистических данных о найденных решениях. Предложенная функция пригодности совпадает со скоростью классификации

$$fitVal = fitnessValue = CR = \frac{\text{number of correctly detected connections}}{\text{size of dataset}} = \frac{TP + TN}{\text{size of dataset}}$$

Поскольку каждая хромосома в популяции является решением (классификатором), ее значение CR будет вычисляться в соответствии с обучающим набором данных KDD – 10 %, который содержит объекты соединений и метки (normal, attack).

Ниже приводится краткое представление о генетических операциях, используемых в предлагаемом обучении ГА: отбор, кроссовер и мутация [4].

Отбор – это действие выбора родителей, т.е. как выбрать индивидов в популяции, которые создадут потомство для следующего поколения, и сколько потомков создаст каждый. Цель отбора состоит в том, чтобы выделить более приспособленных особей в популяции в надежде, что их потомство, в свою очередь, будет иметь еще более высокую приспособленность. Существует множество стратегий выбора, таких как: «колесо рулетки», «масштабирование Сигмы», элитизм и др. [3].

Был принят метод выбора «колеса рулетки», потому что он прост в реализации и быстро сходится, но недостатком которого является то, что, если, например, индивидуальная пригодность составляет 90 %, существует низкий шанс для выбора других особей. Основная идея этого метода отбора заключается в том, что лучшие особи получают более высокие шансы, т.е. шансы пропорциональны значению пригодности. Отметим, что главной отличительной особенностью ГА является использование кроссовера (скрещивания).

Одноточечный кроссовер: это самая простая форма. Одна позиция пересечения выбирается случайным образом, и части двух родителей после позиции пересечения обмениваются, чтобы сформировать два потомства.

Двухточечный кроссовер: две позиции выбираются случайным образом, и сегменты между ними обмениваются.

N -точечное пересечение: выбирается n случайных точек пересечения, затем их разделяют вдоль этих точек и делают пересечение. Это обобщение одного точечного кроссовера

Использован двухточечный кроссовер для создания потомства, поскольку он лучше, чем одноточечный кроссовер и быстрее, чем n -точечный или равномерный кроссовер. Операция кроссовера будет выполнена в соответствии с вероятностью – P_c .

Предложенный ГА был выполнен с использованием следующих параметров:

Длина хромосомы – 66, количество генов – 6, размер популяции – 500, число поколений – 160, мутационная способность Pm – 0,02, вероятность кроссовера Pc – 0,7.

HighCR – это самая высокая скорость классификации и она представляет собой критерий остановки, может быть, например, в диапазоне от 0,975 до 1. Если это не достигнуто, то ГА остановится, когда будет достигнут номер предопределенного поколения. В процессе моделирования лучшая найденная хромосома имела вид:

001110100011111001010110001100000011101110010110100101100001100111.

Для нее значение классификатора:

$$GACL(conn_j) = \begin{cases} \text{if } 103f_{duration} + 843f_{service} + 953f_{src_byte} - 96f_{dst_byte} - 917f_{count} < 465 \\ \text{then } conn_j \text{ is } ATTACK \text{ else } conn_j \text{ is } NORMAL \end{cases}$$

Этот классификатор был применен для обнаружения атак в тестовом наборе данных KDD99, и были получены результаты, представленные в табл. 3.

Таблица 3. Показатели эффективности классификатора линейного генетического алгоритма

Класс атаки	Записи	Обнаружено	Соотношение	
			DR	NDR
NORMAL	60593	58330	96,27 %	–
DOS	229853	222142	96,65 %	16,34 %
PROBE	4166	3227	77,46 %	92,34 %
R2L	16347	4953	30,30 %	0,40 %
U2R	70	53	75,71 %	77,42 %
TP = DR			91,99 %	
NDR			14,89 %	
TN			96,27 %	
CR			92,82 %	
FP			3,73 %	
FN			8,01 %	

Сравнивая эти результаты с предыдущими идентификаторами, можно увидеть, что *GACL* имеет меньшую частоту обнаружения для нормальных, DOS и зондовых соединений, но имеет и другие преимущества:

- скорость выше, чем SF-5NN и SUS-5NN;
- лучшее обнаружение новых атак равно 14,89 %, в то время как в SF-5NN оно составляет 11,81 %;
- лучше частота выявления R2L, равная 30,30 %, а в SF-5NN – 7,99 %, причем атака R2L более опасна, чем DOS и зондовые атаки.

Заключение

1. Построен линейный классификатор на базе теории генетических алгоритмов, оценивающий суммы взвешенных признаков связи в сети. Если сумма была меньше пороговой, то соединение рассматривалось как атакуемое. Параметры, определяющие работу фильтра – сумма весов и порога. Для получения этих параметров был использован бинарный генетический алгоритм, в котором гены равны параметрам. Функция пригодности была определена как скорость классификации (*CR*), цикл эволюции занял 160 поколений, каждое поколение состояло из 500 хромосом.

2. Найденная лучшая хромосома *GACL* была применена к тестовому набору данных типовых атак KDD. Получена хорошая скорость классификации – 92,82 %, другим важным результатом была скорость обнаружения атаки, которая оказалась лучше, чем у известных решений (14,89 %). Также этот классификатор может обнаружить 30,30 % опасных атак R2L. Этот показатель лучше, чем имеют аналогичные подходы, разработанные ранее.

INTELLIGENT CLASSIFIER OF INTRUSION DETECTION ON BASES OF GENETIC ALGORITHMS

U.A. VISHNIAKOU, M.M. MOSDURANY SHIRAS

Abstract. Some detection model was proposed using genetic algorithms and KDD99 data sets. The best chromosome in evolution cycle was founded, which to test set of typical attacks was used. The classifier on the model base was designed, for which good result of attack detection was received and high speed of intrusion detection was shown. These results are better than analogous. So this classifier may detect some percent of R2L attacks high than analogue was designed.

Keywords: intrusion detection, genetic algorithms, attack set data KDD99, classifier.

Список литературы

1. Seagren S. Secure Your Network for Free. Syngress Publishing. 2007.
2. Lazarević A., Srivastava J., Kumar V. Data Mining For Intrusion Detection, Army High Performance Computing Research Center, Department of Computer Science, University of Minnesota, 2003.
3. Melanie M. An Introduction to Genetic Algorithms. Cambridge, Massachusetts London, England. 1999.
4. Вишняков В.А., М.Г. Моздурани Шираз. // Современные средства связи: материалы 24-й междунар. науч. конф. Минск. РИВШ. 2019. С.160.

УДК 621.391

ВОЗМОЖНОСТЬ НЕСАНКЦИОНИРОВАННОГО ДОСТУПА К ИНФОРМАЦИИ, ПЕРЕДАВАЕМОЙ ОТ АБОНЕНТА НА ДРУГОЙ АБОНЕНТСКОЙ РОЗЕТКЕ В ПАССИВНОЙ ОПТИЧЕСКОЙ СЕТИ, ПРИ ИСПОЛЬЗОВАНИИ ОТРАЖЕННОГО СИГНАЛА ОТ ОТВЕТВИТЕЛЯ

Н.Н. СЕРГЕЕВ, В.Н. УРЯДОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 05 ноября 2019*

Аннотация. Рассматривается возможность несанкционированного доступа к информации, передаваемой от абонента, при использовании отраженного сигнала от ответвителя на другой абонентской розетке. Проведен расчет уровня, полученного отраженного сигнала, а также расчет порога чувствительности приемника. Произведена оценка возможности атаки с другого абонентского окончания.

Ключевые слова: пассивная оптическая сеть, защита информации, несанкционированный доступ, сетевые атаки и угрозы, оптические предусилители, оптический модуль с pin-фотодетектором, отраженный сигнал, порог чувствительности приемника.

Введение

Недостатком PON-сети является незащищенность канала передачи от действий злоумышленников. В статье [1] показано, что для того, чтобы воспользоваться отраженным сигналом и снять информацию абонента с другой соседней оптической розетки, достаточно, чтобы величина отраженного сигнала была теоретически в рамках квантового предела детектируемости. Однако чувствительность в квантовом пределе недостижима, и обычно применяют приемный оптический модуль с оптическим предусилителем, обеспечивающим максимальную чувствительность.

Целью данной работы является исследование возможности применения для несанкционированного доступа в сети PON приемного оптического модуля с предусилителем.

Одним из основных модулей приема является оптический модуль с pin-фотодетектором и оптическим предусилителем, обеспечивающим чувствительность близкую к квантовому пределу.

Вид рассматриваемой угрозы

На канальном уровне проблема безопасности в PON практически решена, современные алгоритмы шифрования данных обеспечивают необходимый уровень защиты. На физическом уровне – неисправность лазера обратного канала или контроллера этого лазера может вывести из строя всю сеть. В принципе, такие случаи предусмотрены и отслеживаются системами управления, а включившийся по странному стечению обстоятельств на непрерывную передачу лазер может быть отключен обыкновенным watchdog'ом, контролирующим то, чтобы лазер не превышал отведенных ему временных интервалов. Куда серьезнее могут быть последствия действий злоумышленников: для нарушения (или даже прекращения) работы всего сегмента PON-сети достаточно засветить лазером в любое из абонентских окончаний или даже просто подключить к нему любое активное сетевое устройство. Достаточно просто реализуемо и снятие информации нисходящего потока, поскольку обычный приемник обеспечивает прием сигнала любого приемника, если использовать другой временной интервал. Более сложна в реализации снятия информации контролируемого абонента с другой абонентской розетки поскольку

используется отраженный сигнал от ответвителя. При использовании отраженного сигнала, необходимо учесть, что в пассивной оптической сети существуют возвратные потери в неоднородностях.

Расчет уровня полученного отраженного сигнала

Рассмотрим модель (рис. 1), состоящего из стационарного терминала OLT, оптической линии, разветвителя и нескольких абонентских терминалов ONT. Для снятия информации, передающейся от ONT-1 в восходящем (*upstream*) направлении, с ONT-2, будет использоваться отраженный сигнал от ответвителя, характеризуемого коэффициентом направленности $A_{отв}$. Для каждой оптической линии представим все потери в линии в виде суммы затуханий всех компонентов: оптического кабеля, разъемных соединениях, сварных соединениях, разветвителях:

$$A = A_{отв} + A_{аб}, \quad [дБ], \quad (1)$$

где A – суммарные потери; $A_{отв}$ – коэффициент направленности абонентской линии, $A_{аб}$ – потери в разъемном соединении. Коэффициент направленности стандартного PON-ответвителя составляет 48,5–51,8 [дБ]. Потери в абонентской линии порядка 1–2 [дБ]. Таким образом согласование приемника составляет 50,5–52 [дБ].

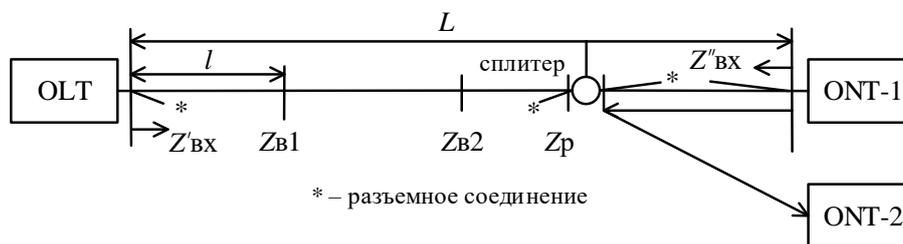


Рис. 1. Поток информации в пассивной оптической сети

Был получен уровень сигнала ONT-1, который придет на терминал ONT-2. Далее необходимо рассчитать, достаточен ли будет полученный уровень сигнала, для снятия информации на ONT-2, для этого найдем порог чувствительности приемника. Так, как одним из основных модулей приема на сегодняшний день является оптический модуль с pin-фотодетектором, то производить расчет следует именно для такого детектора [6].

Расчет порога чувствительности приемника с оптическим предусилителем

Предварительное усиление оптического сигнала перед фотодетектором повышает мощность входного сигнала, делая влияние шумов оптического приемника пренебрежимо малым по сравнению с шумом оптического предусилителя. Наиболее явным кандидатом на роль оптического предусилителя является EDFA, благодаря его широкополосности, существенному коэффициенту усиления и низкому коэффициенту шума.

Для расчета чувствительности приемного оптического модуля необходимо учесть все составляющие шумового тока, включая шум спонтанной эмиссии оптического предусилителя, который будет оказывать существенное влияние на характеристики приемника. Анализ показывает, чувствительность такого приемника определяется выражением:

$$\overline{P}_{in} = h\nu F_n \Delta f \left[Q^2 + Q \sqrt{\frac{\Delta\nu_{opt}}{\Delta f}} \right], \quad (2)$$

$$Q = \frac{RG(2P_{cp})}{\sqrt{\left(2(q\eta G)^2 F_n P_s \Delta f / h\nu\right)^2 + \left((q\eta G F_n)^2 \Delta\nu_{opt} \Delta f\right)^2 + (q\eta G F_n)^2 \Delta\nu_{opt} \Delta f}},$$

где G – коэффициент усиления; $R=(\eta \times q)/(h \times \nu)$; η – квантовая эффективность фотодетектора; h – постоянная Планка; q – заряд электрона; ν – частота оптического излучения; $h \times \nu$ – энергия фотона; $P_s = |E_s|^2$ – оптическая мощность сигнала до усиления; E_s – поле сигнала; F_n – шум-фактор оптического усилителя; $\Delta \nu_{opt}$ – полоса пропускания оптического фильтра; Δf – шумовая полоса приемника.

Шум-фактор современных волоконно-оптических предварительных усилителей на основе оптического волокна легированного ионами эрбия *EDFA (Erbium Doped Fiber Amplifier)* может достигать 3,5 дБ (например, *MAFA 5000 series EDFA* компании *Emcore*). Для получения предельной чувствительности примем полосу пропускания оптического фильтра $\Delta \nu_{opt}$ равную ширине спектра оптического сигнала. Существует обратная зависимость между полосой пропускания оптического фильтра и чувствительностью приемника. При увеличении ширины полосы пропускания оптического фильтра – будет увеличиваться шумовая составляющая, соответственно будет снижаться отношение сигнал/шум (ОСШ), что приведет к падению чувствительности. В то же время полосу оптического фильтра нельзя сделать уже ширины полосы спектра несущей с сигналом, т.к. это повлечет собой уменьшение средней мощности, падающей на фотодиод, а, следовательно, уменьшение ОСШ и чувствительности.

Для примера, вероятности ошибки $P_e = 10^{-6}$, скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм чувствительность приемника по средней мощности составляет $\overline{P_{in}} = -52,05$ [дБм].

Используя предложенные выражения, можно рассчитать кривые зависимости чувствительности приемных оптических модулей от скорости передачи и длины волны и использовать их при оценке защищенности линий ВОЛС от несанкционированного доступа.

Для тех же параметров, что и в предыдущих расчетах с учетом того, что полоса спектра сигнала одного канала $\Delta \nu$ для системы со скоростью B Мбит/с примерно равна $\Delta \nu = 2B$, построим зависимость чувствительности оптического приемника с *pin*-фотодиодом от скорости передачи при $k_1 = 0,5$ путем расчетов в *Matcad* на ЭВМ (рис. 2).

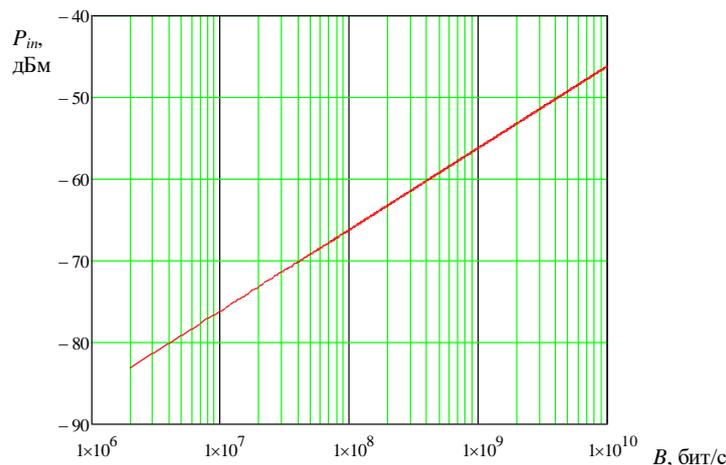


Рис. 2. Чувствительность приемника с предварительным усилителем, для $\lambda = 1550$ нм и $P_{\text{ош}} = 10^{-6}$

На рисунке показано, что оптический предусилитель может обеспечить чувствительность оптического приемника с *pin*-фотодетектором доступ к информации, передаваемой от соседнего абонента и ее обработку, что приводит к несанкционированному снятию информации.

Заключение

Показано, что в стандартной сети PON возможен скрытый доступ к передаваемой информации при атаке с любой соседней оптической розетки, выделяя тайм-слот прямого и отраженного сигнала в оптическом сплиттере. Для обратного канала, достаточно, чтобы величина этого сигнала была в рамках чувствительности приемного оптического модуля. Анализируя полученные результаты, можно предложить следующие методы защиты от несанкционированного доступа к информации, передаваемой на соседней абонентской розетке.

1. Увеличить затухание в абонентской линии;
2. Использовать оптические ответвители с высоким коэффициентом направленности;
3. Использовать разъемные соединения с низким коэффициентом отражения (с высокими возвратными потерями).

POSSIBILITY OF UNAUTHORIZED ACCESS TO INFORMATION TRANSFERRED FROM SUBSCRIBER TO OTHER SUBSCRIBER SOCKET IN A PASSIVE OPTICAL NETWORK, WHEN USING A REFLECTED SIGNAL FROM A SPLITTER

N.N. SERGEEV, V.N. URYADOV

Abstract. The possibility of unauthorized access to information transmitted from the subscriber when using the reflected signal from the coupler to another subscriber outlet was considered. The calculation of the received reflected signal level and the calculation of the sensitivity threshold of the receiver were carried out. The possibility of an attack from another subscriber's ending was evaluated.

Keywords: passive optical network, information protection, unauthorized access, network attacks and threats, optical preamplifiers, optical module with pin photodetector, reflected signal, receiver sensitivity threshold.

Список литературы

1. Птицын Г.А. Живучесть динамических сетей телекоммуникаций. Учебное пособие. М.: МТУСИ. 2008.
2. Dovid Gutierrez, Jinwoo Cho, Leonid G. Kozovsky // Optical Fiber Communication and the National Fiber Optic Engineers Conference. Optical Society of America. 2007. P. 87–89.
3. Булавкин И.А. // Технологии и средства связи. 2006. С. 104–108.
4. Рекомендация МСЭ-Т G.983.1. Широкополосные оптические сети доступа на базе пассивных оптических сетей.
5. Урядов В.Н., Алишев Я.В. Перспективные информационные технологии в волоконно-оптических сетях телекоммуникаций. Минск, 2003.
6. Урядов В.Н., Сергеев Н.Н., Шишпоренок С.С. // Докл. БГУИР. 2018. № 4 (114). С. 5–10.

УДК 621.391

РАСПРЕДЕЛИТЕЛЬНАЯ СЕТЬ ИНТЕРНЕТ-ВЕЩАНИЯ ПО ТЕХНОЛОГИИ ОТТ

Э.Б. ЛИПКОВИЧ, А.А. ПАВЛОВСКИЙ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 30 октября 2019*

Аннотация. Выполнен анализ перспективной распределительной сети интернет-вещания по технологии Over the Top (ОТТ), рассмотрена структура построения и определены требования к ее составу. В соответствии с этой технологией на стороне передачи создается множество версий видеопотоков с различным разрешением и с адаптивно изменяемыми параметрами под пропускную способность сети и технические возможности стационарных и мобильных оконечных устройств (IP-приставки, смартфоны, планшеты, ноутбуки). Данная технология не имеет административных ограничений на размеры зон обслуживания, не содержит схем привязки абонентов к конкретным сетям, позволяет расширить их архитектуру, обеспечить конкурентоспособность и окупаемость.

Ключевые слова: многоформатный контент, адаптивная передача данных, сегментация пакетов, транскодирование потока данных.

Введение

Нынешнее многообразие способов передачи информации по каналам связи при множестве типов пользовательских приемных устройств с повышенной функциональностью заставляют провайдеров видеослужб искать перспективные решения, позволяющие расширить абонентскую сеть, обеспечив ей конкурентоспособность и окупаемость.

Развитие базовых принципов IPTV-вещания привело к созданию технологии Over the Top (ОТТ), которая основывается на адаптации передаваемого медиаконтента под пропускную способность телекоммуникационных каналов, а также под используемый формат видеоразрешения и тип разверток оконечных устройств [1]. Технология обеспечивает доставку затребованного пользователями видеоматериала по открытой интернет-сети в понятном формате для его отображения на стационарных (Smart-TV, IP-приставки) и мобильных (смартфоны, планшеты, ноутбуки и др.) устройствах.

Распределительная сеть интернет-вещания по технологии ОТТ

В соответствии с технологией ОТТ на стороне передачи создается множество видеопотоков одинакового содержания с различным разрешением и с динамически изменяемыми параметрами под пропускную способность сети и технические возможности приемных устройств. Данная технология не имеет административных ограничений на размеры зон обслуживания, не содержит схем привязки абонентов к конкретной IPTV-сети, но вместе с тем она создает ощутимую нагрузку на магистрали интернет-операторов, с которыми заключается договор на проведение трансляций.

Трансляция независимых потоков в адрес пользователей чаще всего осуществляется по протоколу HLS (HTTP Live Streaming), поддерживающему принцип «каждому устройству – требуемый вид услуг», с учетом совместимости характеристик медиаконтента с характеристиками многоэкранных типов оконечных устройств. Протокол HTTP (Hyper-Text Transfer Protocol) работает поверх протокола TCP (Transmission Control Protocol) и служит для передачи файлов различных веб-сайтов. Запрос на требуемый контент осуществляется

по технологии, принятой при организации запросов веб-страниц с использованием команды GET протокола HTTP.

Протокол TCP обеспечивает высокую надежность доставки персональных пакетов к месту назначения, поскольку осуществляется контроль за последовательностью их поступления, в случае необходимости организуется повторная передача потерянных пакетов. Для реализации этого протокола требуется прямая связь между вещательным сервером ОТТ-сети и абонентским устройством. Однако с применением этого протокола исключается потоковый (многоадресный) режим доставки, сокращается скорость передачи данных из-за необходимости подтверждений и повторной пересылки потерянных пакетов, увеличивается нагрузка на вещательный сервер и сказывается на зависимости величины пропускной способности сети от удаленности абонентского устройства от информационного сервера. Переход к протоколу UDP, который не требует подтверждений при передаче и допускает потоковый режим доставки пакетов, предполагает наличие специального клиентского приложения на оконечных устройствах, входящих в сеть ОТТ. Поскольку многие существующие устройства приема не оснащены требуемым ПО, то рассматриваются различные сценарии решения этой проблемы. В частности, предлагается вводить избыточное кодирование, новые виды протоколов, использовать видеоархивы, кэширование, большое число видеосерверов на стороне пользователей и многие другие решения [1, 2].

Учитывая большой объем данных, содержащийся в запрашиваемых видеофайлах, их передача в адрес пользователей ведется сегментами (чанками) с динамически изменяемым размером посылок. На приемной стороне принятые сегменты данных подвергаются процедуре буферизации, которая исключает пробелы в передаче и обеспечивает непрерывность воспроизведения. Скорость передачи посылок зависит от пропускной способности соединения с сетью Интернет. Одной из задач мультисервисного IP-вещания является отладка открытой сети для достижения стабильной работы и качественной доставки многоформатного контента на различные абонентские устройства.

В отличие от закрытых для несанкционированных пользователей IPTV-сетей с полным контролем трафика, скорости передачи и полосы пропускания канала, контроль в ОТТ-сетях достаточно ограничен и состоит в оценке потерь пакетов, времени буферизации данных и сборе статистики. Причины пониженного уровня контроля состоят в наличии джиттера, скачков в пропускной способности, задержек и сниженной ответной реакции сервера на запросы при большом числе обращений.

Структура построения интерактивной сети ОТТ-вещания (рис. 1) на мультиэкранные устройства пользователей во многом близка к архитектуре сетей IPTV. Определяющими компонентами этой сети являются: подсистема приема и обработки информации; средства многоформатного кодирования и транскодирования; система условного доступа; пакетизатор (упаковщик); вещательные видеосерверы обработки и хранения различных версий контента; подсистема Middleware (MW); сеть Интернет; пользовательский сегмент.

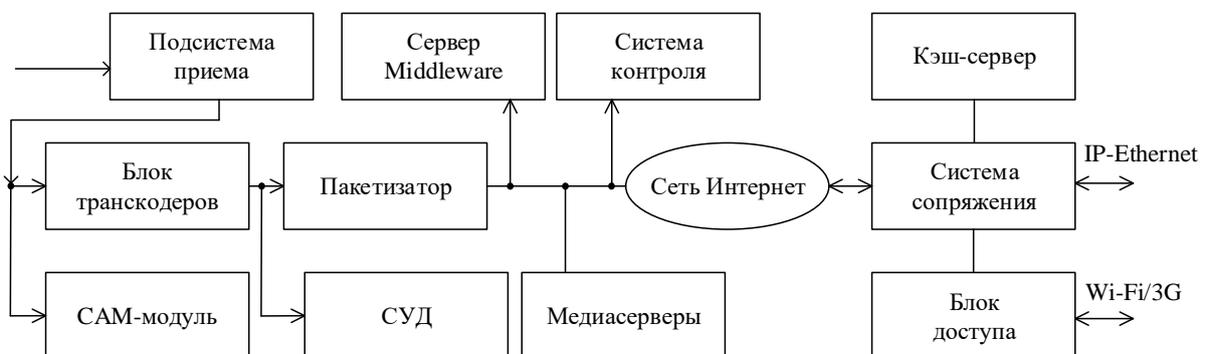


Рис. 1. Структурная схема системы ОТТ-вещания

Подсистема приема и обработки информации включает многофункциональные модули стандартов DVB, потоков ASI, SDI, IP-Ethernet, а также устройства захвата и обработки сигналов с SD- и HD-видеокамер для последующего кодирования, транскодирования и формирования однопрограммных потоков SPTS. Транскодер системы позволяет изменять формат компрессии потока данных и создавать необходимый набор уровней с различным разрешением, пропускной

способностью, частотой и размером кадров. С его помощью поток данных с HD-разрешением может трансформироваться в группу низкопрофильных форматов, например, с разрешением 320×240 и скоростью данных 0,3 Мбит/с.

Транскодеры реализуются на сочетании аппаратных и программных решений. Они используют специализированные процессоры обработки цифровых видео- и аудиопотоков и являются многофункциональными. Например, транскодер INCA 4430 компании WISI (Германия) способен из 18 HD- или 32 SD-программ сформировать 72 варианта видеопотоков для интернет-вещания. Для обеспечения безопасной доставки видеоконтента на оконечные устройства, предусматривающие различные требования к уровню защиты контента и сервиса от информационных атак, разработано значительное число сетей условного доступа (СУД). Выбор конкретной системы при построении OTT-сети зависит от таких ее показателей, как производительность, масштабируемость, простота инсталляции и обслуживания. В большинстве крупных OTT-проектах используется СУД от Verimatrix, которая отвечает требованиям безопасности передачи контента и по сравнению с другими системами защиты информации лучше адаптирована для интерактивных сетей с разными видами бескарточных абонентских устройств на базе iOS и Android.

Пакетизатор на программном уровне осуществляет сегментацию видеопотоков на отдельные фрагменты и подготавливает их для OTT-стриминга в протоколах передачи HLS, MPEG-DASH или Smooth Streaming [1]. Сформированные сегменты данных с объемом каждого в несколько секунд передачи поступают на вещательные серверы для трансляции по сети Интернет.

Процесс доставки информации от ближайшего к пользователю медиасервера осуществляется под контролем и управлением подсистемы Middleware. Она осуществляет распределение сетевой нагрузки между компонентами системы, авторизацию абонентских устройств, обработку запросов, взаимодействие с элементами учета и тарификации услуг, а также сбор статистической информации. Кроме того, часть функций, относящихся к СУД, возлагается на подсистему MW, которая поддерживает работу с различными моделями IP-приставок. Контроль качества системы состоит в мониторинге нагрузки на серверы и в контроле передаваемых данных по каналам.

Среди известных компаний, специализирующихся на реализации оборудования и построении OTT-сетей, следует отметить Envivo, Harmonic, Wisi, Ericsson, Broadbury Lab.

Заключение

Представлен анализ функционирования обобщенной структуры построения распределительной сети интернет-вещания по технологии OTT. Определены достоинства данной архитектуры и указаны необходимые технические решения по обеспечению ее конкурентоспособности и окупаемости. Указаны необходимые требования к характеристикам базовых устройств, входящих в систему, для обеспечения адаптивной передачи многоформатного контента в адрес стационарных и мобильных пользователей.

OTT TECHNOLOGY DISTRIBUTION NETWORK

E.B. LIPKOVICH, A.A. PAULOUSKI

Abstract. The analysis of perspective distribution network of Internet broadcasting on Over the Top technology (OTT) was carried out, the structure of construction is considered and requirements to its structure were defined. In accordance with this technology, many versions of video streams with different resolutions and adaptively variable parameters were created on the transmission side for the network bandwidth and technical capabilities of stationary and mobile terminal devices (IP-set-top boxes, smartphones, tablets, laptops). This technology has no administrative restrictions on the size of service areas, does not contain schemes of binding of subscribers to specific networks, and allows to expand their architecture to ensure competitiveness and payback.

Keywords: multi-format the content-adaptive transmission of data, segmentation of packets, transcoding the data stream.

Список литературы

1. Сэнджой П. Распределение цифрового видео по широкополосным, телевизионным, мобильным и конвергентным сетям. Тенденции, проблемы и решения. Техносфера, 2012.
2. Jeremy A. Internet Television Is An Open Platform [Electronic resource]. – URL: http://www.masternewmedia.org/news/2005/05/17/internet_television_is_an_open.html. (date of access: 20.10.2019)

УДК 621.391

АЛГОРИТМ АНАЛИЗА ИЗОБРАЖЕНИЙ В ВИДЕОАНАЛИТИКЕ

С.И. НАЛИВАЙКО, В.А. ВИШНЯКОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 31 октября 2019*

Аннотация. Рассмотрен алгоритм анализа изображений в видеопотоке. Показано, что недостатком быстрого механизма обнаружения движущихся объектов на видеоизображениях, является необходимость предварительной обработки кадров с целью устранения шумов и межкадровых деформаций фона, возникающих, например, при перемещении камеры.

Ключевые слова: видеоаналитика, алгоритм видеоанализа, обработка кадров.

Введение

На сегодняшний день, функциональные возможности систем видеонаблюдения получают все больше средств для автоматического анализа видеoinформации. Качественная современная система видеонаблюдения должна не только производить запись и выводить изображение на экран, но и осуществлять ряд аналитических функций. Одним из приоритетных направлений цифрового видеонаблюдения является видеоаналитика.

Видеоаналитика – это интеллектуальный анализ потока видеоданных от камеры (последовательно поступающих видеоизображений) и автоматическое выявление различного рода данных и детектирование заранее запрограммированных ситуаций. Аналитика ведется как в режиме реального времени, так и при работе с архивом. Результатами работы видеоаналитики являются события, которые могут быть переданы оператору системы видеонаблюдения в виде сообщения или записаны в архив для последующего поиска по ним и составления отчетов [1].

Имеется ряд основных задач, которые успешно решает видеоаналитика. Самыми распространенными являются следующие задачи:

- распознавание номеров (автомобильных, на денежных купюрах, документах);
- обнаружение опасных ситуаций (скопления людей, оставленные предметы, возгорания и задымления и т.п.);
- распознавание человеческих лиц и поиск их в базах данных;
- распознавание с целью подсчета людей и транспорта.

Применение видеоаналитики предоставляет возможность в автоматическом режиме, без участия человека, в процессе видеонаблюдения решать сложные задачи, которые обычно под силу только человеческому зрению.

У видеоаналитики пока наблюдаются некоторые ограничения, но появляется все больше высокотехнологичных и гибких систем, которые позволяют решать различные задачи. Они направлены на разные сегменты рынка и могут как обеспечивать безопасность, так и быстро извлекать необходимую информацию из видеоархивов, в результате чего повышается эффективность работы предприятий. С улучшением аналитических алгоритмов и увеличением вычислительных мощностей процессоров, используемых в серверах и камерах видеонаблюдения, ресурсы таких инструментов значительно расширятся, а видеоаналитика станет более доступной и удобной в использовании [2].

Алгоритм анализа изображения

Современные системы видеоаналитики работают с видеопотоком в режиме реального времени или из файлов архива. И в том, и в другом случае это последовательность кадров, которые передаются с некоторой частотой в секунду, которая называется fps (frames per second). Каждый кадр является статичным изображением, который несет какие-то данные. В зависимости от формата (количеством бит на пиксель и их интерпретацией: какие биты за какую составляющую цвета отвечают) и разрешения каждый кадр имеет конкретный размер. Каждый кадр, обусловленный параметрами системы, может быть цветным или черно-белым, сжатым или нет. Так же из-за надобности хранения сообщений о нарушениях и сопоставление кадров между собой системам видео аналитики требуется параметр, который был бы единственным ключом в этих обстоятельствах. Таким ключом является timestamp – уникальная временная метка каждого кадра. Исходя из вышеперечисленного, видеопоток обладает следующими основными параметрами:

- частота кадров в секунду;
- формат;
- разрешение;
- timestamp.

Алгоритм видеоанализа складывается из нескольких стадий. На выходе любого из них данные о происходящем в кадре расширяются все большими подробностями. В свой черед могут быть обратные связи между этапами, чтобы более тонко реагировать на изменения в кадре. Отдельно взятая реализация данных этапов может быть различна. Рассмотрим главные методы и алгоритмы распознавания событий в видеопотоке. Вычисление межкадровых отличий является весьма популярным методом первоначального распознавания движения, после выполнения которого, с уверенностью можно утверждать, имеет ли место в потоке кадров движение. При этом, видеокadres предварительно должны быть обработаны перед вычислением отличий между ними. Схема представлена на рис. 1.



Рис. 1. Схема работы анализа видеопотока

В цветном изображении, который имеет формат RGB24, любой отдельно взятый пиксель обладает размером в 3 байта. Сохранить информативность, при этом снизив размер можно путем перевода картинку из цветного в формат серого изображения Y8, где для каждого пикселя используется 1 байт информации. Следовательно, размер потока кадров можно сократить в три раза.

В свою очередь можно уменьшить число кадров в секунду в видеопотоке, что так же положительно отразится на объеме обрабатываемой информации и скорости видеоанализа. Актуальные системы видеоаналитики могут работать с видео, обладающим частотой 12 кадров в секунду [3].

Для определения движения, идентификации объектов, создания трека движения необходимо создать фон, посредством которого можно определить инварианты и меняющиеся пиксели в кадрах. Получается это с помощью поккадрового сравнения с шаблоном. Наиболее распространенным примером является создание фоновой модели, построенной на пустой сцене. В следствии того, что в системе сопоставление кадров с изображением, обладающим форматом Y8, то происходит попиксельное сравнение с фоном фреймом в градациях серого. Для минерализации «шумов» и увеличения скорости действия процесса получается бинаризация пикселей – все пиксели, которые больше 0 изменяют свое значение на 255, что приводит к появлению черного-белого изображения. На следующем этапе появляется отделенный фон и объект с четкими границами, но в некоторых случаях, так же возможно «присоединение» к объекту его тени, или других подвижных деталей, не относящихся к анализируемому объекту. Следовательно, на данном шаге нужно классифицировать объекты, чтобы не принимать во внимание природные факторы, влияния на камеру или шумы изображения: блики, образование пыли на объективе камеры, артефакты, которые могут иметь место при сжатии кадров и т.д.

Работа по выделению объектов детектором движения, построенным на алгоритме вычисления попиксельных отличий между кадрами, начинается с анализа маски движения. Настоящим объектам в ней обычно соответствуют пиксели, которые образуют связные группы, поэтому разумно выделить объект с точки зрения детектора движения как группу взаимосвязанных пикселей в маске движения. Обозначение объектов происходит при последовательном обходе пикселей маски движения. Когда в маске движения обнаружена единица, запускается процесс поиска пикселей, сопряженных с данным, которым также соответствуют единицы в маске движения. Найденные пиксели помечаются, чтобы не допустить повторного их обнаружения. Термин «трассирование» подразумевает отслеживание перемещающегося объекта и значений его параметров в течение всего времени его пребывания в кадре. При правильном трассировании объекта доступна возможность просмотра траектории движения объекта с любым уровнем детализации (с точностью до кадра, до 3, 5 и т.д. кадров). Данная возможность обеспечивается сохранением положения центральной точки объекта на каждом из кадров в специальном массиве, который введен для этой цели при определении параметров для каждого из движущихся объектов. Алгоритмы, трассирующие движущиеся объекты, не должны приравнивать один объект к другому, терять объект, если он остановился на небольшой промежуток времени или скрылся за каким-либо препятствием (например, человек, скрывшийся за деревом), а потом появился вновь [4].

В архитектуре могут содержаться обратные связи, повышающие работу предшествующих этапов. Вероятно, основное решение – применять информацию об объектах в сцене при формировании фона, например, так можно отличить отставленные предметы и не делать его частью фона. Или с другой стороны: если при создании фона на сцене был человек то, когда он уйдет, на его месте появится объект-«призрак». Понимая, что в этом месте начинается траектория движения объекта, можно быстро убрать «призрака» в фон [5].

В итоге мы получаем метаданные по объекту. Каждый из них характеризуется размерами, плотностью, скоростью, траекторией, направлением движения и другими параметрами.

Эти метаданные и используются при аналитике сцены. Можно определить нахождение объекта в определенных зонах или движение в неправильном направлении. Можно подсчитать количество объектов в заданной зоне, оставленные предметы и множество других событий.

Заключение

Задача отслеживания движущихся объектов на видеоизображениях является актуальной. Известные методы, предполагающие распознавание объектов и оценку параметров их движения, требуют значительных вычислительных ресурсов. Недостатком быстрого механизма обнаружения движущихся объектов на видеоизображениях, основанного на расчете межкадровой разности, является необходимость предварительной обработки кадров с целью устранения шумов и межкадровых деформаций фона, возникающих, например, при перемещении камеры.

IMAGE ANALYSIS ALGORITHM IN VIDEO ANALYTICS

S.I. NALIVAİKA, V.A. VISHNYAKOU

Annotation. An algorithm for analyzing images in a video stream was considered. It was shown that a drawback of the fast mechanism for detecting moving objects in video images was the need for preliminary processing of frames in order to eliminate noise and interframe deformations of the background that arise, for example, when moving the camera.

Keywords: video analytics, video analysis algorithm, frame processing.

Список литературы

1. Возможности современной видеоаналитики. [Электронный ресурс]. URL: <https://www.videomax-server.ru/support/articles/vozmozhiiosti-sovrememioy-videoanalitiki/> (дата доступа: 30.10.2019).
2. Видеоаналитика на практике. [Электронный ресурс]. URL: <https://www.osp.ru/lan/2015/03/13045265> (дата доступа: 30.10.2019).
3. Как устроена видеоаналитика. [Электронный ресурс]. URL: <https://habrahabr.ru/post/271207/> (дата доступа: 30.10.2019).
4. Гулящева Л.Н. // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем. 2014. № 1 (9). С. 107–111.
5. Торстен Анштедт, Иво Келлер, Харальд Лутц. Видеоаналитика: Мифы и реальность: Security Focus, 2012.

УДК 621.391

МЕТОД РЕВЕРСИВНОГО КОДИРОВАНИЯ НУЛЬ-ДЕРЕВЬЕВ

В.В. НОВИЦКИЙ, В.Ю. ЦВЕТКОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 18 ноября 2019*

Аннотация. Предложен метод эффективного кодирования цифровых изображений на основе реверсивного кодирования нуль-деревьев битовых плоскостей. Он обладает низкой вычислительной сложностью и высокой степенью параллелизма. Численное моделирование на реальных аэрокосмических снимках показало приблизительно равный SBHP коэффициент сжатия.

Ключевые слова: реверсивное кодирование нуль-деревьев, аэрокосмические снимки.

Введение

Современные космические аппараты дистанционного зондирования Земли (КА ДЗЗ), такие как Worldwiev-3, GeoEye-2, KOMPSAT-5, ASRANO, SPOT-7, Pleiades-1B имеют на борту съемочную аппаратуру сканирующего типа, формирующую цифровые кадры сверхвысокого разрешения [1]. Скорость формирования этих цифровых кадров значительно превышает пропускную способность радиолинии передачи данных, что вынуждает использовать сжатие в режиме реального времени в бортовой аппаратуре приведенных выше КА.

В качестве бортовых систем сжатия чаще всего используют кодеки JPEG2000 и ICER [2, 3], которые базируются на вычислительно сложном методе контекстного битового моделирования с последующим арифметическим кодированием, так называемом MQ кодере [4–6]. Данный кодер имеет низкую пропускную способность из-за необходимости тройного прохода по каждой битовой плоскости при контекстном моделировании. Этот факт приводит к необходимости параллельной работы большого количества таких кодеров для сжатия потока кадров сверхвысокого разрешения, что в свою очередь приводит к существенному росту ресурсоемкости подобного решения.

В тоже время известен метод эффективного кодирования SBHP [7], который осуществляет рекурсивное кодирование нуль-деревьев битовых плоскостей изображения и имеет в четыре раза меньшую вычислительную сложность по сравнению с MQ кодером. Данный метод входит в верификационную модель стандарта JPEG2000 и позволяет сжимать изображения с небольшим проигрышем в коэффициенте сжатия по сравнению с базовым методом стандарта.

Целью данной работы является разработка нового метода эффективного кодирования на основе SBHP, хорошо подходящего для реализации на базе программируемой логики (ПЛИС) и не уступающего по качеству сжатия прототипу.

Математическая модель реверсивного кодирования нуль-деревьев

Предлагаемый метод эффективного кодирования заключается в реверсивном кодировании иерархических квадродеревьев битовых плоскостей кодовых блоков, на которые равномерно делится изображение после его пространственно-частотного преобразования (рекомендуется вейвлет-преобразование). Метод назван FBQT – Fast Block Quad Tree coding.

Сущность метода заключается в построении L -уровневого иерархического квадродерева для битовой плоскости кодового блока по формуле:

$$m_{HQT} \left(\sum_{m=0}^{l-1} \frac{N_{CB}}{4^m} + p \right) = \prod_{k=0}^3 m_{HQT} \left(\sum_{m=0}^{l-1} \frac{N_{CB}}{4^m} - \frac{N_{CB}}{4^{l-1}} + 4p + k \right), \quad (1)$$

где $l = \overline{0, L-1}$, $p = 0, \frac{N_{CB}}{4^l} - 1$, L – количество уровней иерархического квадродерева, N_{CB} – размер битовой плоскости после Z -перестановки пикселей кодового блока. При этом нулевым уровнем иерархического квадродерева является сама битовая плоскость, а каждый последующий уровень меньше предыдущего в четыре раза.

После построения иерархического квадродерева оценивается компактность его уровней по следующим формулам:

$$C_L = \sum_{l=1}^L \left(4 \sum_{p=0}^{\frac{N_{CB}-1}{4}} m_{HQT} \left(\sum_{m=0}^{l-1} \frac{N_{CB}}{4^m} + p \right) \right), \quad (2)$$

$$C_1 = \frac{N_{CB}}{4} + 4 \sum_{p=0}^{\frac{N_{CB}-1}{4}} m_{HQT} (N_{CB} + p), \quad (3)$$

и выбирается минимальная.

На основе данной оценки компактности строится карта значимости, состоящая из значимых четырехбитных кластеров, которая наиболее компактно описывает текущую битовую плоскость. Значимость кластера оценивается по следующей формуле:

$$S_{CL}(V_p, p) = \begin{cases} 1, & \text{если } \sum_{i=0}^3 V_p(p+i) > 0 \\ 0, & \text{иначе} \end{cases}. \quad (4)$$

Описание алгоритма для метода FBQT

Предлагаемый метод эффективного кодирования FBQT состоит из следующих шагов.

1. Преобразование исходного двумерного массива $C = \|c(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$ элементов (вейвлет-коэффициентов, пикселей изображения) в одномерный массив $M_Z = \|m_Z(n)\|_{(n=\overline{0, N_{CB}-1})}$

с помощью рекурсивной Z -перестановки [8], где $N_{CB} = Y \times X$. Рекурсивная Z -перестановка сохраняет пространственную корреляцию элементов, существующую в двумерном массиве, после переноса их в одномерный массив и позволяет упростить процесс кодирования битовых плоскостей. В процессе перестановки осуществляется анализ значений элементов для определения номера MSB старшей битовой плоскости и знаков $S_K = \|s_K(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$ значимых элементов с помощью выражений

$$MSB = \left\lfloor \log_2 \left(\max(|c(y, x)|) \right) \right\rfloor, \quad S_K = \begin{cases} 1, & \text{если } c(y, x) \leq 0, \\ 0, & \text{иначе} \end{cases} \quad \text{при } y = \overline{0, Y-1}, \quad x = \overline{0, X-1}, \quad \text{где } \lfloor \rfloor -$$

операция округления до целого с недостатком. Значимыми считаются все ненулевые элементы.

2. Инициализация массивов и указателей: $b = 0$ – номер обрабатываемой битовой плоскости; $M_{SP} = \|m_{SP}(n)\|_{(n=\overline{0, N_{CB}-1})}$ – массив значимых бит ($m_{SP}(n) = 0$ при $n = \overline{0, N_{CB}-1}$);

$P_{SM} = 0$ – указатель на обрабатываемый бит карты значимости $SM = \|sm(n)\|_{(n=\overline{0, N_{CB}-1})}$;

$M_{HQT} = \left\| m_{HQT}(n) \right\|_{(n=0, N_{CB}-1)}$ – массив нулевого уровня квадродерева

$m_{HQT}(n) = \begin{cases} 0, & \text{если } m_Z(n) \wedge 2^b = 0, \\ 1, & \text{иначе} \end{cases}$ – заполнение нулевого уровня $m_{HQT}(n)$ битами текущей

битовой плоскости; $L = \log_4(N_{CB})$ – количество уровней квадродерева.

3. Заполнение нулевого уровня $m_{HQT}(n)$ битами текущей битовой плоскости согласно

выражению $m_{HQT}(n) = \begin{cases} 0, & \text{если } m_Z(n) \wedge 2^b = 0, \\ 1, & \text{иначе} \end{cases}$, где b – номер обрабатываемой битовой

плоскости.

4. Формирование бинарного L -уровневого квадродерева. Каждый уровень квадродерева, начиная с первого, заполняется на основе предыдущего в результате оценки значимости его 4-элементных кластеров с помощью формулы (1).

5. Оценка компактности карты значимости. Поскольку SM формируется из иерархического квадродерева путем включения только значимых кластеров, то по построенному дереву можно определить ее оптимальную длину: используя формулу (3) для одноуровневого квадродерева и формулу (2) для L -уровневого квадродерева. Если выполняется условие $C_L \leq C_1 \leq N_{CB}$, то осуществляется переход к шагу 6, если выполняется условие $C_1 \leq N_{CB}$, то осуществляется переход к шагу 7, иначе – переход к шагу 8.

6. Формирование карты значимости на основе L -уровневого квадродерева (в случае, если оценка компактности показала, что исходная битовая плоскость содержит много незначимых кластеров). Если верхний уровень квадродерева равен 0, то флагу F присваивается значение 0 и осуществляется переход к шагу 8, иначе флагу F присваивается значение 1, и осуществляется перенос в SM всех значимых кластеров M_{HQT} и переход к шагу 9.

7. Формирование карты значимости на основе 1-уровневого квадродерева (в случае, если оценка компактности показала, что исходная битовая плоскость содержит немного незначимых кластеров и целесообразнее формировать карту значимости на основе 1-уровневого квадродерева). Флагу F присваивается значение 2, осуществляется перенос в SM уровня 1 M_{HQT} и всех значимых кластеров уровня 0 M_{HQT} , переход к шагу 9.

8. Передача битовой плоскости без формирования ее карты значимости (в случае, если оценка компактности показала, что исходная битовая плоскость является шумом и кодирование нецелесообразно). Флагу F присваивается значение 3, осуществляется перенос в SM уровня 0 M_{HQT} и переход к шагу 9.

9. Окончание формирования карты значимости для текущей битовой плоскости. Проверка условия $b = MSB$. Если условие выполняется, то осуществляется переход к шагу 10, иначе – $b = b + 1$ и переход к шагу 3.

10. Окончание эффективного кодирования FBQT исходного двумерного массива C элементов. В конец выходного потока, компактно описывающего текущий кодовый блок, вставляется код знаков значимых пикселей S_K . На рис. 1 представлена структура выходного потока кодера FBQT.

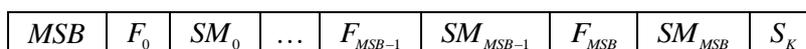


Рис. 1. Структура выходного потока кодера FBQT

Структура блочно-поточного кодера, реализующего метод FBQT

На рис. 2 представлена функциональная схема блочно-поточного кодера, реализующего метод FBQT. Особенность данной схемы заключается в множественном распараллеливании вычислительных операций при построении иерархического квадродерева и оценки

компактности его уровней, а также их конвейеризации при Z-перестановке кодового блока, кодировании текущей битовой плоскости и формировании выходного байтового потока.

В реверсивном кодере одновременно формируются все уровни иерархического квадродерева и оценивается их компактность при чтении контроллером текущей битовой плоскости по кластерам. Таким образом, быстродействие реверсивного кодера зависит от размера кластера (чем больше размер, тем выше быстродействие и ресурсоемкость). Размер кластера 8 бит является оптимальным и позволяет достигать пропускной способности кодера до 2000 Мбит/с при сжатии не квантованного изображения.

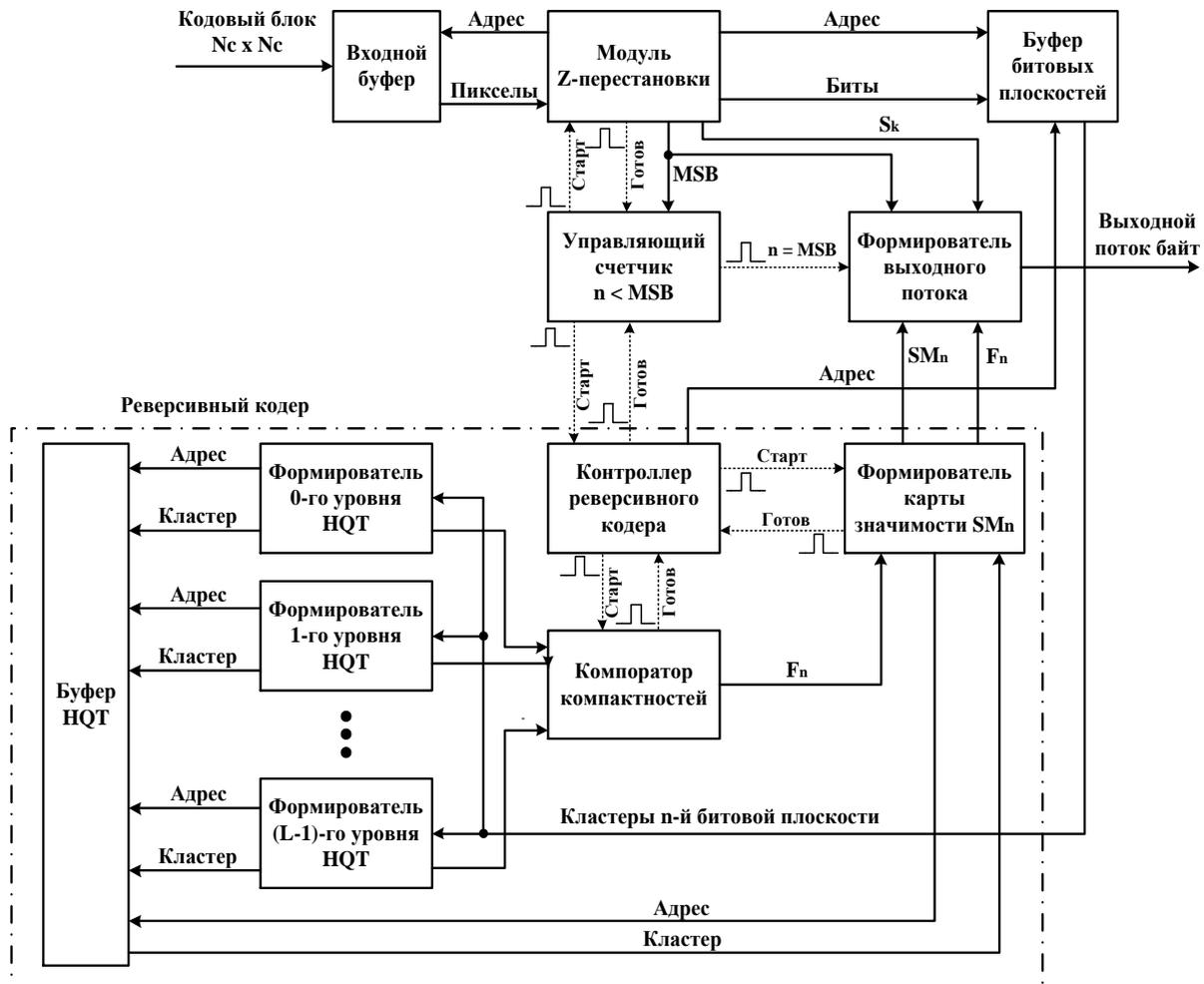
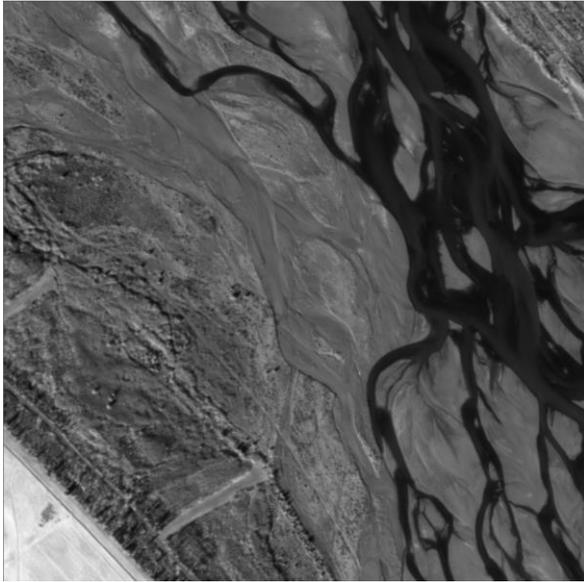


Рис. 2. Функциональная схема кодера FBQT

Результаты численного моделирования и их обсуждение

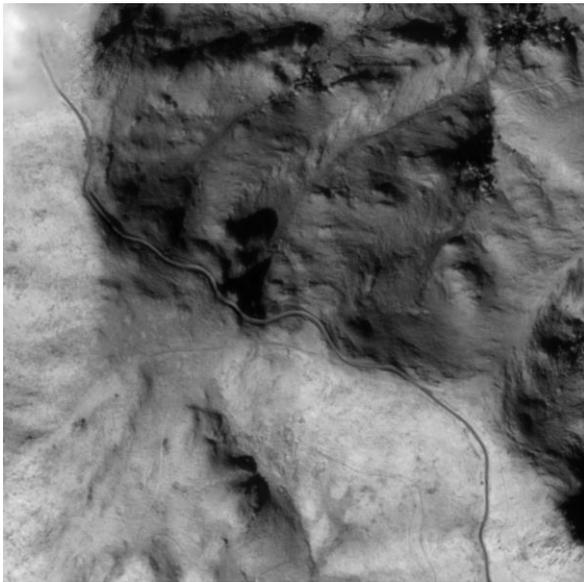
С учетом того факта, что предлагаемый метод FBQT разрабатывался с целью реализации в бортовой аппаратуре спутника ДЗЗ для сжатия потока цифровых кадров сверхвысокого разрешения в режиме реального времени целесообразно для оценки качества сжатия использовать спутниковые снимки сверхвысокого разрешения. Для подобной оценки был выбран показательный набор кадров в панхроматическом спектральном канале объемом 60 кадров размера 768×768 и разрядности 12 бит. На рис. 3 отображены 4 кадра данного маршрута, а на рис. 4 – график зависимости среднего контраста соседних пикселей от номера кадра, который показывает разнообразие статистических свойств и классов изображений в данном маршруте.



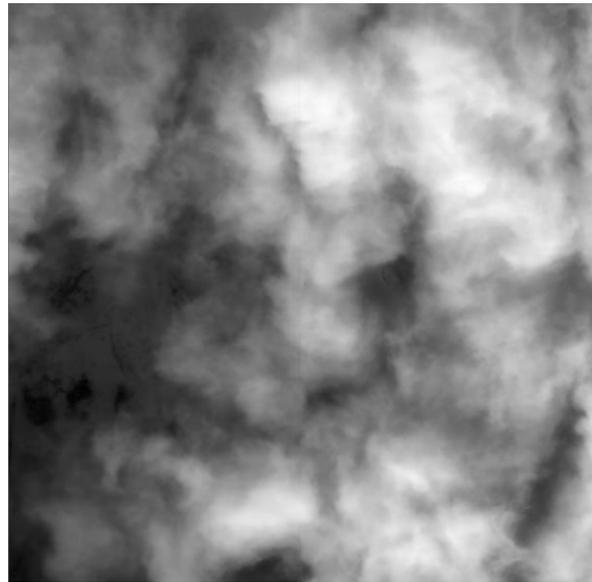
a



б



в



г

Рис. 3. Выборочные кадры панхроматического канала показательного маршрута: *a* – 7-й кадр, *б* – 23-й кадр, *в* – 38-й кадр, *г* – 56-й кадр

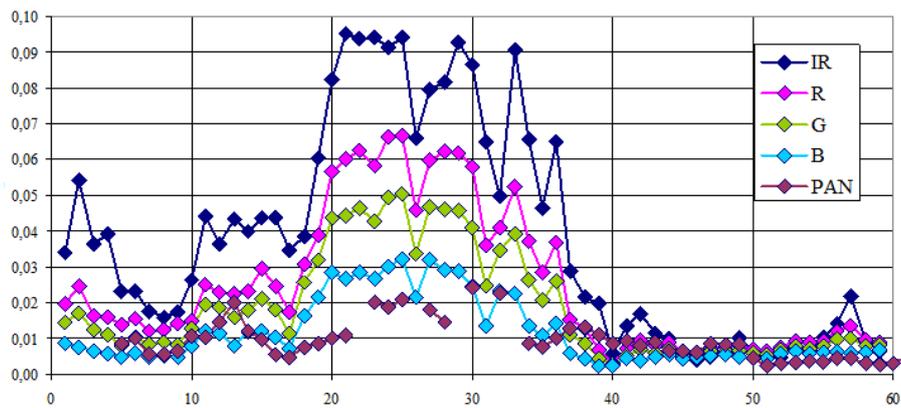


Рис. 4. Зависимость среднего контраста CT от номера кадра

В табл. 1 даны минимальное, максимальное и среднее значение коэффициента сжатия для всех кадров маршрута при их целочисленном вейвлет-преобразовании и эффективном кодировании методами MQ, SBHP и FBQT.

Таблица 1. Значения коэффициента сжатия для всех кадров маршрута

Метод эффективного кодирования	Минимальный CR для 60 кадров	Максимальный CR для 60 кадров	Средний CR для 60 кадров
MQ	1,708	2,497	2,212
SBHP	1,657	2,412	2,155
FBQT	1,665	2,424	2,163

Как видно из табл. 1 предложенный метод FBQT выигрывает у прототипа (без дополнительного энтропийного кодирования) порядка 1 % в коэффициенте сжатия и проигрывает значительно более вычислительно сложному методу MQ порядка 5 %.

Заключение

Предложен новый метод эффективного кодирования FBQT, который заключается в реверсивном кодировании иерархических квадродеревьев битовых плоскостей кодовых блоков. Данный метод предназначен для реализации на программируемой логике и имеет значительные упрощения в кодировании по сравнению с прототипом, что позволило реализовать множественное распараллеливание вычислительных операций. Также предложена структура блочно-поточного кодера, реализующего метод FBQT, которая подтверждает возможность работы устройства сжатия в режиме реального времени.

Численное моделирование показало, что предлагаемый метод практически идентичен по качеству сжатия с прототипом и уступает значительно более вычислительно сложному методу MQ стандарта JPEG2000.

METHOD OF REVERSIVE ZERO-TREE CODING

V.U. NAVITSKI, V.Yu. TSVIATKOU

Abstract. An effective coding method of digital images based on reversive zero-tree coding of bit planes was proposed. It is a low complexity and highly paralleled method. It was shown that the method has a little bit higher compression ratio in comparison with SBHP at aerospace images.

Keywords: reversive zero-tree coding, aerospace images.

Список литературы

1. Bormin H. // Sat. Data Compression. Springer New York DordrechtHeidelberg. London. 2011. P. 33–45.
2. Taubman D.S., Marcellin M.W. JPEG2000: Fundamentals, Standards and Practice. Kluwer Academic Publishers, Boston, 2002.
3. Kielyand A., Klimesh M. The ICER Progres. Wavelet Im. Compressor. IPN Progress Report 42-155, 2003.
4. Ghodhbani R., [et. al.] // Journal of Real-Time Processing. 2019. Vol. 16. P. 1595–1606.
5. Horrigue. L., [et. al.]. An efficient, high speed architecture for JPEG2000 MQ-coder. IEEE IPAS, 2014.
6. Modrzyk D., Staworko M. A high-performance architecture of JPEG2000 encoder. IEEE ESPP, 2011.
7. Chrysafis C., [et. al.] // IEEE ICASSP, 2000. Vol. 4. P. 2035–2038.
8. Morton G.M. A computer oriented geodetic data base and a new technique in file sequencing. Technical report, IBM Ltd., Ottawa, Ontario, Canada, 1966.

УДК 621.391

**ПРЯМОЕ ФОТОДЕТЕКТИРОВАНИЕ И КОГЕРЕНТНЫЙ ПРИЕМ
С РАЗЛИЧНЫМИ ВИДАМИ МОДУЛЯЦИИ
В ОТКРЫТЫХ СИСТЕМАХ ОПТИЧЕСКОЙ СВЯЗИ
СО СТАЦИОНАРНЫМИ КОСМИЧЕСКИМИ ОБЪЕКТАМИ**

С.А. ЛУКАШЕВИЧ, В.Н. УРЯДОВ, Т.В. ПОЛУЯН

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

Поступила в редакцию 01 ноября 2019 года

Аннотация. Рассмотрена эффективность прямого фотодетектирования и методы когерентного приема оптических сигналов инфракрасного диапазона длин волн с различными видами модуляции в открытых системах связи со стационарными космическими объектами.

Ключевые слова: гомодинный метод приема, гетеродинный метод приема, метод прямого фотодетектирования с предварительным усилителем, амплитудная манипуляция, инфракрасное излучение.

Введение

При разработке оптических космических систем возникает необходимость выбора метода приема и вида модуляции оптической несущей. Целью данной работы является сравнение использования различных видов модуляции и методов приема сигналов в открытых системах оптической связи со стационарными космическими объектами (спутниками).

Эффективность прямого фотодетектирования при связи со стационарными космическими объектами и различными видами модуляции

Квантовый предел чувствительности при амплитудной (ООК) модуляции

Квантовый предел детектирования определяется шумами, связанными с сигналами. Чувствительность приемника по средней мощности будет равна [1]:

$$\overline{P}_{in} = \overline{N}_p h\nu B = \frac{\overline{N}_p hcB}{\lambda},$$

Для вероятности ошибки $P_e = 10^{-6}$, скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм чувствительность приемника по средней мощности составляет $\overline{P}_{in} = 1,237 \cdot 10^{-9}$, что соответствует $-59,07$ дБм, чувствительность по пиковой мощности $P_{in} = -56,06$ дБм. Для длины волны $\lambda = 1060$ нм: $\overline{P}_{in} = -57,42$ дБм, $P_{in} = -54,41$ дБм.

Чувствительность приемного оптического модуля (ПРОМ) с p-i-n фотодетектором

Чувствительность оптического приемника по средней мощности рассчитаем, используя следующее выражение:

$$\overline{P}_{in} = \frac{A_\lambda}{\eta} \sqrt{4kT(2\pi C_\Sigma)B^2 \left[\frac{I_{n2}}{2K} + \frac{I_{n3}(2\pi C_\Sigma)BF_{\text{ПТ}}}{q_m} \right]}.$$

Для вероятности ошибки $P_e = 10^{-6}$, скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм чувствительность приемника по средней мощности составляет $\overline{P}_{in} = -36,7$ дБм. Чувствительность приемника по пиковой мощности составляет $P_{in} = -33,7$ дБм. Расчет производился при следующих параметрах: $\eta = 0,85$; $A_\lambda = 3,8$ Вт/А; $C_\Sigma = 0,5$ пФ, $In_2 = 0,55$; $In_3 = 0,085$; $q_m = 35 \cdot 10^{-3}$ см $^{-1}$; $F_{\text{ПТ}} = 1,5$; $T = 300$ К; $K = 20$. Для длины волны $\lambda = 1060$ нм чувствительность приемника составила $\overline{P}_{in} = -35$ дБм, $P_{in} = -32$ дБм.

Чувствительность ПРОМ с ЛФД

Для случая $P_0 = 0$ чувствительность ЛФД рассчитывается следующим образом [2]:

$$\overline{P}_{in} = \frac{A_\lambda}{\eta} \left(\frac{\sqrt{\overline{i}_{pin}^2}}{M} + qQ_{\text{ош}} F_A In_2 B \right).$$

Для применения при заданных длины волны и скорости передачи наиболее подходит ЛФД на основе InGaAs. Типовое значение коэффициента ионизации для данного материала составляет $k = 0,45$ [3]. Для вероятности ошибки $P_e = 10^{-6}$, скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм оптимальный коэффициент лавинного умножения $M_{\text{ошт}} = 14,2$, чувствительность приемника по средней мощности составляет $\overline{P}_{in} = -44,88$ дБм, а чувствительность по пиковой мощности $P_{in} = -41,88$ дБм. Расчет производился при следующих параметрах: $\eta = 0,85$; $A_\lambda = 3,8$ Вт/А; $C_\Sigma = 0,5$ пФ; $In_2 = 0,55$; $In_3 = 0,085$; $q_m = 35 \cdot 10^{-3}$ см $^{-1}$; $F_{\text{ПТ}} = 1,5$; $T = 300$ К; $K = 20$. Для длины волны $\lambda = 1060$ нм чувствительность приемника составила $\overline{P}_{in} = -43,23$ дБм, $P_{in} = -40,23$ дБм.

Фотодетектор ПРОМ с оптическим предусилителем

Чувствительность приемника по средней мощности при использовании предварительного усилителя определяется выражением [1]:

$$\overline{P}_{in} = hvF_n \Delta f \left[Q^2 + Q \sqrt{\frac{\Delta v_{\text{opt}}}{\Delta f}} \right].$$

Для вероятности ошибки $P_e = 10^{-6}$, скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм чувствительность приемника по средней мощности составляет $\overline{P}_{in} = -52,05$ дБм, а по пиковой мощности $P_{in} = -49,05$ дБм. Для длины волны $\lambda = 1060$ нм чувствительность приемника составила $\overline{P}_{in} = -50,4$ дБм, $P_{in} = -47,4$ дБм.

Таким образом прием методом прямого фотодетектирования модуляции ООК наиболее эффективен при использовании оптического предусилителя. Чувствительность при его использовании для связи с геостационарными спутниками, составляет минус 51 дБм для длины волны 1550 нм и скорости передачи 1,25 Гбит/с.

Прием сигналов с DPSK-модуляцией

Чувствительность приемника определяется следующим выражением:

$$\overline{P}_{in} = \overline{N}_p hvB = \frac{\overline{N}_p hcB}{\lambda}$$

Для вероятности ошибки $P_e = 10^{-6}$, скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм чувствительность приемника составляет $\overline{P}_{in} = P_{in} = -56,06$ дБм. Для длины волны $\lambda = 1060$ нм: $\overline{P}_{in} = P_{in} = -54,41$ дБм.

Вероятность ошибки балансного приемника сигналов с DPSK определяется выражением [1]:

$$P_e = \frac{1}{2} \exp\left(\frac{-\eta N_p}{n_{sp}}\right) \left(1 + \frac{\eta N_p}{4n_{sp}}\right),$$

где n_{sp} – коэффициент спонтанной эмиссии предварительного усилителя, $n_{sp} \approx F_n / 2$.

Для получения $P_e < 10^{-6}$ необходимое число фотонов на бит составляет $N_p \geq 20$ ($\eta = 0,85$; $F_n = 3,5$ дБ). Чувствительность приемника для скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм составляет $\overline{P_{in}} = P_{in} = -54,94$ дБм. Для длины волны $\lambda = 1060$ нм: $\overline{P_{in}} = P_{in} = -53,29$ дБм.

Произведенные расчеты показывают, что при использовании DPSK модуляции приемник с предварительным усилителем обладает лучшей чувствительностью приблизительно на 3 дБ по сравнению с модуляцией интенсивности, что согласуется с результатами работ [4–6].

Прием сигналов с QPSK-модуляцией

Для получения $P_e < 10^{-6}$ необходимое число фотонов на бит составляет $N_p \geq 23$ ($\eta = 0,85$). Чувствительность приемника определяется следующим выражением:

$$\overline{P_{in}} = \overline{N_p} h\nu B = \frac{\overline{N_p} hcB}{\lambda}.$$

Для вероятности ошибки $P_e < 10^{-6}$, скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм чувствительность приемника составляет $\overline{P_{in}} = P_{in} = -54,33$ дБм. Для длины волны $\lambda = 1060$ нм: $\overline{P_{in}} = P_{in} = -52,68$ дБм.

Вероятность ошибки данного приемника при использовании оптического предварительного усилителя определяется выражением [1]:

$$P_e = Q_1(a,b) - \frac{1}{2} I_0(ab) \exp\left[-\frac{a^2+b^2}{2}\right] + \frac{b^2-a^2}{8ab} I_1(ab) \exp\left[-\frac{a^2+b^2}{2}\right],$$

$$a = \sqrt{(2-\sqrt{2}) \frac{\eta N_p}{n_{sp}}}, \quad b = \sqrt{(2+\sqrt{2}) \frac{\eta N_p}{n_{sp}}},$$

где Q_1 – Q-функция Маркума, I_0, I_1 – модифицированные функции Бесселя нулевого и первого порядка соответственно.

Для получения $P_e < 10^{-6}$ необходимое число фотонов на бит составляет $N_p \geq 23$ ($\eta = 0,85$; $F_n = 3,5$ дБ). Чувствительность приемника для скорости передачи $B = 1,25$ Гбит/с и длины волны $\lambda = 1550$ нм составляет $\overline{P_{in}} = P_{in} = -53,8$ дБм. Для длины волны $\lambda = 1060$ нм: $\overline{P_{in}} = P_{in} = -52,15$ дБм.

Когерентный прием в открытых системах связи со стационарными космическими объектами

Гомодинный приемник с балансным включением и ООК модуляцией

Параметр Q для гомодинного приема при применении ООК модуляции будет равен:

$$Q_{ООК,CR} = \sqrt{(2\eta N_{avg})}.$$

Выразим коэффициент ошибок:

$$BER_{НОМ,ООК} = \frac{1}{2} \operatorname{erfc}(\sqrt{\eta N_{avg}}).$$

При этом среднее число фотонов $\overline{N_{avg}}$ – можно выразить через мощность на входе фотоприемника как:

$$\overline{N}_{avg} = \frac{10^{10} 10^{-3} \lambda}{hcB}$$

где \overline{P}_{in} – средняя уровень мощности на входе фотоприемника. Следует учитывать, что при расчете вероятности ошибки принято использовать среднее число фотонов/бит, что соответствует средней мощности (чувствительности) на входе фотоприемника.

Результаты расчетов: для вероятности ошибки 10^{-6} и $\eta = 0,85$ – значение средней чувствительности при длине волны:

– 1550 нм составит: $\overline{P}_{in} \approx -56,71$ дБм; $P_{in} \approx -53,7$ дБм; $\overline{N}_{avg} \approx 13,3$; $N \approx 26,6$;

– 1060 нм составит: $\overline{P}_{in} \approx -55,06$ дБм; $P_{in} \approx -52,05$ дБм.

Гомодинный прием с балансным включением для модуляции BPSK

В случае модуляции BPSK и гомодинном приеме вероятность ошибки рассчитывается как [1]:

$$BER_{НОМ, BPSK} = \frac{1}{2} \operatorname{erfc}(\sqrt{\eta 2 \overline{N}_{avg}})$$

Результаты расчетов: для вероятности ошибки 10^{-6} и $\eta = 0,85$ – значение средней чувствительности при длине волны:

– 1550 нм составит: $\overline{P}_{in} = P_{in} \approx -59,72$ дБм; $\overline{N}_{avg} = N \approx 6,6$;

– 1060 нм составит: $\overline{P}_{in} = P_{in} \approx -58,08$ дБм.

Гомодинный прием с балансным включением для модуляции QPSK

Отношение сигнал/шум в данном случае уменьшается вдвое по сравнению с BPSK модуляцией. В данном случае вероятность ошибочного приема рассчитывается как [1]:

$$BER_{НОМ, QPSK} = \frac{1}{2} \operatorname{erfc}(\sqrt{\eta \overline{N}_{avg}})$$

Результаты расчетов: для вероятности ошибки 10^{-6} и $\eta = 0,85$ – значение средней чувствительности при длине волны:

– 1550 нм составит: $\overline{P}_{in} = P_{in} \approx -56,71$ дБм; $\overline{N}_{avg} = N \approx 13,3$;

– 1060 нм составит: $\overline{P}_{in} = P_{in} \approx -55,07$ дБм.

Синхронный гетеродинный прием для модуляции BPSK

Ввиду биений сигналов локального генератора и входного сигнала формируется сигнал на промежуточной частоте, состоящий из положительной (действительной) и отрицательной (мнимой) частей. Шумы, при этом, вкладываются в действительную и удваиваются. Чувствительность падает на 3 дБ по сравнению с гомодинным методом приема и BPSK модуляцией. Вероятность ошибки рассчитывается как [1]:

$$BER_{НЕТ, BPSK} = \frac{1}{2} \operatorname{erfc}(\sqrt{\eta \overline{N}_{avg}})$$

Результаты расчетов: для вероятности ошибки 10^{-6} и $\eta = 0,85$ – значение средней чувствительности при длине волны:

– 1550 нм составит: $\overline{P}_{in} = P_{in} \approx -56,71$ дБм; $\overline{N}_{avg} = N \approx 13,3$;

– 1060 нм составит: $\overline{P}_{in} = P_{in} \approx -55,07$ дБм.

Асинхронный гетеродинный прием для модуляции DBPSK

Детектирование, в данном случае, происходит в фазовом детекторе, содержащим линию задержки, после вычитания сигналов пришедших с фотодиодов. Вероятность ошибочного приема рассчитывается как [1]:

$$BER_{НЕТ, DBPSK} = \frac{1}{2} \exp(-\eta \overline{N}_{avg}).$$

Результат дает ухудшение чувствительности на 0,5 дБ по сравнению с синхронным гетеродинным приемом.

Результаты расчетов: для вероятности ошибки 10^{-6} и $\eta=0,85$ – значение средней чувствительности при длине волны:

- 1550 нм составит: $\overline{P_{in}} = P_{in} \approx -56,1$ дБм; $\overline{N_{avg}} = N \approx 15,4$;
- 1060 нм составит: $\overline{P_{in}} = P_{in} \approx -54,42$ дБм.

Синхронный гетеродинный прием для модуляции QPSK

Для данного метода приема чувствительность будет равна чувствительностям QPSK при гомодинном приеме и BPSK при гетеродинном приеме. Вероятность ошибочного приема рассчитывается как [1]:

$$BER_{\text{HET.QPSK}} = \frac{1}{2} \operatorname{erfc}(\sqrt{\eta N_{avg}}).$$

Результаты расчетов: для вероятности ошибки 10^{-6} и $\eta=0,85$ – значение средней чувствительности при длине волны:

- 1550 нм составит: $\overline{P_{in}} = P_{in} \approx -56,71$ дБм; $\overline{N_{avg}} = N \approx 13,3$;
- 1060 нм составит: $\overline{P_{in}} = P_{in} \approx -55,07$ дБм.

Асинхронный гетеродинный прием с балансной схемой включения для модуляции FSK (с двумя фильтрами)

Вероятность ошибочного приема рассчитывается как [1]:

$$BER_{\text{HET.FSK}} = \frac{1}{2} \exp\left(-\frac{\eta N_{avg}}{2}\right).$$

Результаты расчетов: для вероятности ошибки 10^{-6} и $\eta=0,85$ – значение средней чувствительности при длине волны:

- 1550 нм составит: $\overline{P_{in}} \approx -53,05$ дБм; $P_{in} \approx -50,04$ дБм; $\overline{N_{avg}} \approx 30,9$; $N \approx 61,8$;
- 1060 нм составит: $\overline{P_{in}} \approx -51,4$ дБм; $P_{in} \approx -48,39$ дБм.

Графики зависимостей коэффициента ошибок от чувствительности, для гомодинных и гетеродинных приемников для различных типов модуляции приведены на рис. 1 и 2.

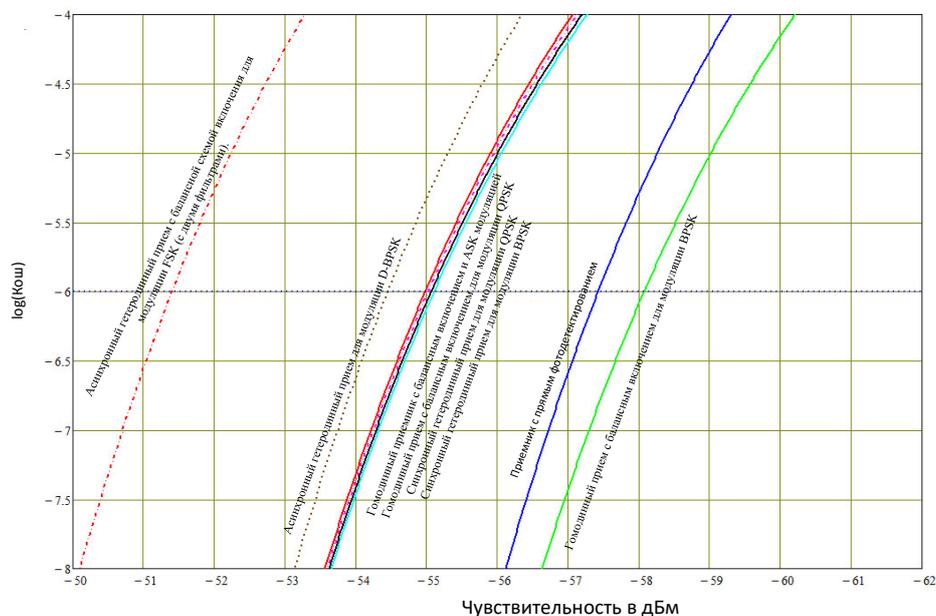


Рис. 1. Зависимость коэффициента ошибок от чувствительности приемника для длины волны 1060 нм

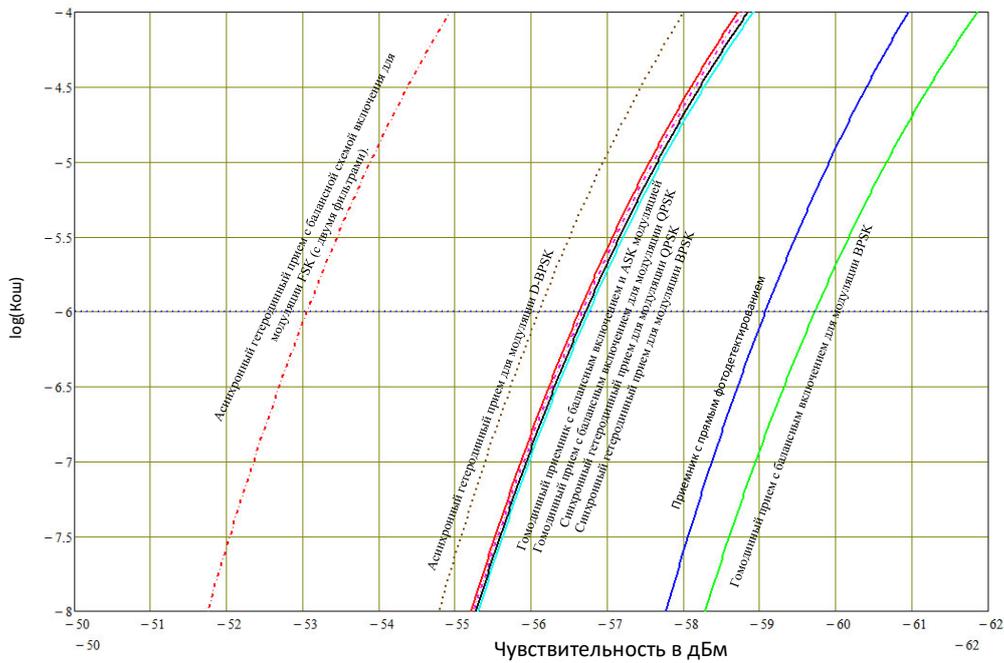


Рис. 2. Зависимость коэффициента ошибок от чувствительности приемника для длины волны 1550 нм

Заключение

Сравнивая (рис. 1 и 2) чувствительности когерентных приемников и приемника прямого фотодетектирования, следует отметить, что последний в идеальном случае в среднем, требует 8 фотонов/бит для достижения коэффициента ошибок $< 10^{-6}$. Это значение лишь немногим больше по сравнению с гомодинамным приемником с использованием BPSK модуляции и существенно ниже сравнительно с гетеродинамными приемниками. Тем не менее, такая чувствительность приемника прямого фотодетектирования на практике не достижима, вследствие наличия теплового шума, темнового тока и многих других факторов. Вследствие этого реальная чувствительность приемника прямого фотодетектирования с предусилителем составляет минус 52,05 дБм, а по пиковой мощности минус 49,05 дБм. В случае использования когерентного приема можно получить чувствительности порядка минус 59,7 дБм при коэффициенте ошибок равном 10^{-6} .

Длина волны также влияет на чувствительность фотоприемника, следует отметить, что при переходе от 1060 к 1550 нм, чувствительность увеличивается приблизительно на 1,7 дБ, что соответствует уменьшению мощности необходимой для приема сигнала с заданным коэффициентом ошибки в 1,5 раза (предполагалось большее КПД лазера для длины волны 1060 нм по сравнению с 1550 нм, причем увеличение выходной мощности было именно в 1,5 раза). С точки зрения длины волны, стоит обращать внимание именно на те компоненты, которые являются наиболее распространенными на рынке (1550 нм) и, тем самым, снизить стоимость конечного продукта.

С точки зрения выбора метода приема и типа модуляции, если не принимать во внимание сложность конструкции приемника и множество проблем при его проектировании, стоит выделить гомодинамный прием с использованием BPSK модуляции, т.к. именно такая комбинация приемника и типа модуляции дает наибольший выигрыш чувствительности для стационарных объектов (геостационарных спутников).

**DIRECT PHOTODETECTION AND COHERENT RECEPTION
WITH DIFFERENT MODULATION TECHNIQUES
IN OPEN OPTICAL COMMUNICATION SYSTEMS
WITH STATIONARY SPACE OBJECTS**

S.A. LUKASHEVICH, V.N. URJADOV, T.V. POLUYAN

Abstract. The efficiency of direct photodetection and methods of coherent reception of optical signals in infrared wavelength range with different modulation techniques in open communication systems with stationary space objects were considered.

Keywords: homodyne reception method, heterodyne reception method, direct photodetection method with pre-amplifier, amplitude manipulation, infrared radiation.

Список литературы

1. Govind P. Fiber-Optic Communications Systems, Third Edition. Agrawal, 2002.
2. Тсанг У. Техника оптической связи. Фотоприемники. М., 1988.
3. А Фриман Р.Л. Волоконно-оптические системы связи. М., 2003.
4. Xu C., Liu X., Wei X. // IEEE J. Sel. Top. Quantum Electron. 2004. Vol. 10, № 2, P. 281–293.
5. Ho K.-P., Gnauck A. // Optical Fiber Communication Conference 2003, paper ThE1.
6. Gnauck A.H, [et. al.] // IEEE Photonics Technol. Lett. 16. 2004. P. 909–911.

УДК 621.391

ДЕТЕКТИРОВАНИЕ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ ПРИ ПОМОЩИ ФРЕЙМВОРКА TENSORFLOW

А.Л. КАЛОША, А.В. САСКЕВИЧ, М.В. СТЕРЖАНОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 01 ноября 2019*

Аннотация. Разработана системы для локализации и классификации объектов на изображениях.

Ключевые слова: TensorFlow, CUDA, нейронная сеть, машинное обучение.

Введение

Человеческий мозг хранит множество загадок. С появлением современной электроники, начались попытки аппаратно-программного воспроизведения его работы. Бурное развитие и применение аппарата искусственных нейронных сетей пришло на последние полвека. На текущий момент задача обнаружения и классификации объектов на изображениях достаточно актуальна. Повсеместно ведется разработка программного обеспечения для автоматического управления транспортными средствами, валидации сканируемых документов и т.д. Другими словами, разрабатываются алгоритмы и программы компьютерного зрения [1].

Основная часть

Правильно обученная нейронная сеть позволяет с высокой точностью найти на изображении объекты и определить их классы.

Нейронная сеть – это громадный распределенный параллельный процессор, состоящий из элементарных единиц обработки информации, накапливающих экспериментальные знания и предоставляющих их для последующей обработки [2].

Для решения данной задачи использовалась сверточная нейронная сеть, которая относится к нейронным сетям глубокого обучения.

Под термином глубокое обучение часто подразумевают такие модели, как рекуррентные нейронные сети, сети с долгой краткосрочной памятью, рекуррентные тензорные нейронные сети, сверточные нейронные сети и генеративно-сопоставительные сети, получившие большую популярность в последние годы [3].

Часто под глубокими нейронными сетями подразумевают нейронные сети с несколькими внутренними слоями, термин «глубокое обучение» в действительности мало чем отличается от искусственных нейронных сетей. Однако, разные архитектуры реализуют уникальные возможности в слоях, что дает им возможность моделировать очень сложные данные [3].

Сверточные нейронные сети (Convolutional Neural Networks, CNN), например, объединяют многослойные перцептроны со сверточным слоем, который итеративно строит карту для извлечения важных признаков, а также имеют этап свертки, уменьшающий размерность признаков, но сохраняющий наиболее информативные их компоненты. Сверточные сети очень эффективны для моделирования изображений и решения задач классификации и обобщения [3].

Для обучения нейронной сети была выбрана библиотека TensorFlow как один из лучших инструментов машинного обучения. TensorFlow – это библиотека программного обеспечения с открытым исходным кодом для численного расчета с использованием графиков потока данных [4].

Для ускорения процесса обучения используется вычислительная мощность видеокарты, а именно технология CUDA. CUDA – это архитектура параллельных вычислений от NVIDIA, позволяющая существенно увеличить вычислительную производительность благодаря использованию GPU (графических процессоров) [5].

В данной работе используется Tensorflow Object Detection API, который упрощает процесс обучения нейронной сети. Например, предоставляет несколько недообученных моделей для обнаружения объектов с разной производительностью и точностью, основанных на датасете COCO (Common Objects in Context). В данной работе используется модель SSD-MobileNet, основанная на алгоритме SSD (Single Shot Detector). Использование данных моделей позволяет сократить время обучения нейронной сети в несколько раз.

Чтобы обучить нейронную сеть новым классам, которые не содержатся в базовой модели, нужно использовать инструмент LabelImg, который упрощает разметку искомым объектам на изображениях. Файл разметки элемента обучающей выборки содержит такие данные как: имя файла изображения, его размер, количество цветов, имя класса, координаты верхнего левого и нижнего правого прямоугольника, содержащего объект (рис. 1).

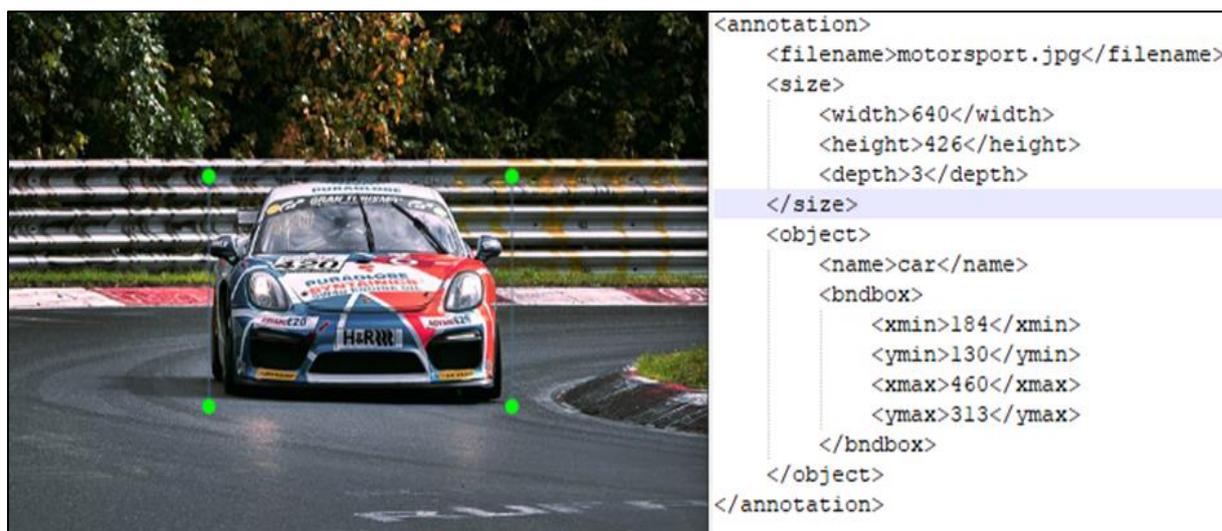


Рис. 1. Пример эталонной разметки изображений обучающей и тестовой выборки

Чтение данных является узким местом обучения. Частично эту проблему решает новый формат данных TFRecord, который предназначен для хранения последовательности двоичных записей. Данный формат позволяет легко комбинировать несколько наборов данных и интегрируется с функциями импорта и предварительной обработки данных.

Для обучения нейронной сети должны использоваться как позитивные, так и негативные примеры областей на изображении. Как правило, негативных примеров требуется в несколько раз больше. При обучении нейронной сети с использованием Tensorflow API можно не указывать негативные примеры. В таком случае все области на изображении, не являющимися позитивными, по умолчанию считаются негативными. Данная возможность позволяет сократить время на подготовку данных для обучения.

Реализация метода SSD осуществлена на языке программирования Python 3 с использованием библиотек глубокого обучения TensorFlow и Keras, а также технологии NVIDIA CUDA.

Обучение и тестирование модели SSD было произведено на компьютере со следующими характеристиками: процессор Intel Core i5-8600; оперативная память 16,0 ГБ; видеокарта Nvidia GeForce GTX 2060, частота графического процессора 1365 МГц, 6144 Мб видеопамати GDDR6, частота видеопамати 3500 МГц, число универсальных процессоров 1920.

Результат работы приложения представлен на рис. 2.

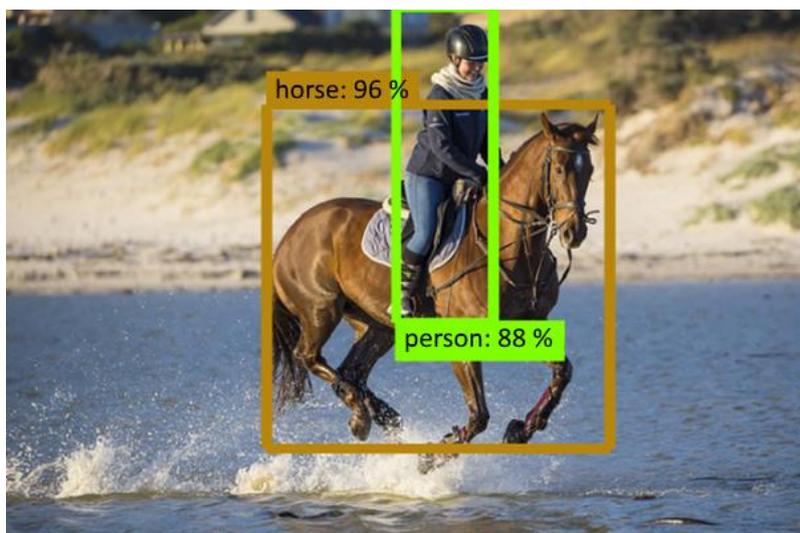


Рис. 2. Результат работы приложения

Заключение

Обучение сети производилось на более чем 3000 изображений, что занимает от 3 до 6 часов, в зависимости от глубины обучения и точности результата. В результате система способна локализовать и классифицировать объекты на изображении. Средняя точность классификации при оптимальных условиях составляет 80 %. Для увеличения точности классификации можно использовать другой тип модели или увеличить выборку для обучения.

OBJECTS DETECTION AND RECOGNITION ON IMAGES USING TENSORFLOW

A.L. KALOSHA, A.V. SASKEVICH, M.V. STERJANOV

Abstract. Software for object detection and recognition on images using TensorFlow was developed.

Keywords: TensorFlow, CUDA, neural network, machine learning.

Список литературы

1. Ситкин И.В. // Научное сообщество студентов XXI столетия. Технические науки: сб. ст. по мат. LXIV междунар. студ. науч. – практ. конф. № 4(63). [Электронный ресурс] URL: [https://sibac.info/archive/technic/4\(63\).pdf](https://sibac.info/archive/technic/4(63).pdf). (дата доступа: 30.10.2019).
2. Хайкин С. Нейронные сети: полный курс, 2-е издание. М.: Издательский дом «Вильямс», 2006.
3. Бенгфорт Б., Билбро Р., Охеда Т. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка. СПб.: Питер, 2019.
4. Library for numerical computation using data flow graphs [Электронный ресурс] URL: <https://www.tensorflow.org>. (дата доступа: 30.10.2019).
5. Параллельные вычисления CUDA [Электронный ресурс] URL: <http://www.nvidia.ru/object/cuda-parallel-computingru.html>. (дата доступа: 30.10.2019).

УДК 621.391

СЖАТИЕ ИЗОБРАЖЕНИЙ БЕЗ ПОТЕРЬ НА ОСНОВЕ КОМБИНИРОВАННОГО ЭФФЕКТИВНОГО КОДИРОВАНИЯ БИТОВЫХ ПЛОСКОСТЕЙ

Б.Д.С. САДИК, В.Ю. ЦВЕТКОВ, М.Н. БОБОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 02 ноября 2019*

Аннотация. Разработана структура и исследована эффективность комбинированного кодека сжатия изображений без потерь в пространственной области, использующего алгоритмы арифметического и RLE-кодирования для компактного представления различных битовых плоскостей изображения.

Ключевые слова: комбинированное эффективное кодирование, сжатие изображений без потерь.

Введение

На изображениях различных типов наибольший коэффициент сжатия показывают разные кодеки. Это обусловлено тем, что распределения значений в битовых плоскостях изображений отличаются. Можно также предположить, что различные битовые плоскости одного изображения могут лучше сжиматься с помощью разных кодеков. При этом для изображений различных типов (спутниковых, медицинских, тепловизионных, гиперспектральных и т.д.) наибольшие коэффициенты сжатия возможны при различных комбинациях кодеков битовых плоскостей. Такой подход, предполагающий использование для сжатия изображений кодера с комбинированной структурой, сочетающей несколько различных кодеров, систему их подключения к различным битовым плоскостям и систему управления этими подключениями, в настоящее время практически не исследован. В современных кодеках сжатия изображений [1–6] используется один алгоритм обработки для всех битовых плоскостей. Косвенные подтверждения эффективности комбинированного кодирования можно обнаружить в структурах кодеков EZW, SPIHT, SPECK [7–9], обеспечивающих отдельное кодирование битовых плоскостей и предусматривающих возможность использования младших битовых плоскостей с равновероятным повтором нулей и единиц без кодирования, а также в структуре кодека JPEG 2000 [6], предусматривающего возможность использования двух алгоритмов для сжатия изображений без и с потерями. Однако, в рассмотренных кодеках не предусмотрено применение разных алгоритмов кодирования к различным битовым плоскостям. Эффективность отдельного кодирования битовых плоскостей анализируется в работах [10, 11], но только для кодера RLE [12]. Результаты исследования эффективности использования комбинации арифметического и RLE-кодирования приведены в работе [13], но только для нескольких типов изображений. Недостаточное исследование комбинированного подхода к кодированию изображений связано с ростом вычислительной сложности: временной (из-за дополнительных затрат на выбор алгоритма кодирования) и пространственной (из-за дополнительных затрат памяти для реализации нескольких алгоритмов кодирования). Однако, с развитием элементной базы рост вычислительной сложности кодирования становится менее критичным в сравнении с увеличением коэффициента сжатия, особенно для приложений, предусматривающих передачу

изображений в реальном времени по каналам с ограниченной пропускной способностью. Кроме того, возможно использование комбинации кодеков со сложной и простой структурой, что ведет к относительно небольшому росту вычислительной сложности. Наибольший интерес в этом плане представляют арифметический [14, 15] и RLE [12] кодеки. Арифметический кодек входит в состав ядра кодера JPEG 2000 и позволяет достичь высоких коэффициентов сжатия изображений. RLE-кодек имеет относительно низкую вычислительную сложность, благодаря чему нашел широкое применение в составе различных кодеков сжатия изображений и архиваторов.

Цель работы: разработка структуры и исследование эффективности комбинированного кодера сжатия изображений различных типов без потерь в пространственной области на основе алгоритмов арифметического и RLE-кодирования.

Раздельное эффективное кодирование битовых плоскостей изображений

Для разработки структуры комбинированного кодера сжатия изображений без потерь в пространственной области проведено исследование эффективности раздельного кодирования битовых плоскостей изображений с помощью арифметического и RLE-кодеров. Сущность раздельного эффективного кодирования состоит в использовании для сжатия старших и младших битовых плоскостей изображения или их комбинаций независимых кодеров одного типа или одного кодера, подключаемого к плоскостям поочередно (рис. 1).

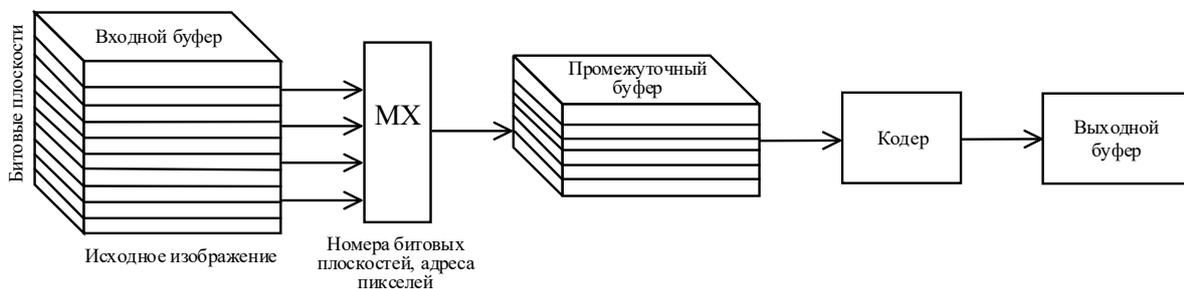


Рис. 1. Раздельное эффективное кодирование битовых плоскостей изображения

Битовые плоскости $B(r)$ формируются из одинаковых разрядов r пикселей $i(R, y, x)$ R -разрядного изображения $I(R) = \|i(R, y, x)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$ и представляют собой матрицу $B(r) = \|b(r, y, x)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$, состоящую из нулей и единиц ($b(r, y, x) = \{0, 1\}$), размер $Y \times X$ которой совпадает с размером $Y \times X$ исходного изображения $I(R)$. Значения $i(R, y, x)$ и $b(r, y, x)$ связаны выражением

$$i(R, y, x) = \sum_{r=0}^{R-1} 2^r b(r, y, x)$$

при $y = \overline{0, Y-1}$, $x = \overline{0, X-1}$.

Комбинация из нескольких r_C битовых плоскостей ($0 < r_C < R$) с r_L по r_H ($r_H > r_L$, $r_C = r_H - r_L + 1$) представляет матрицу $I_C(r_L, r_H) = \|i_C(r_L, r_H, y, x)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$, значения $i_C(r_L, r_H, y, x)$ которой имеют r_C разрядов. В этом случае

$$i_C(r_L, r_H, y, x) = \sum_{r=r_L}^{r_H} 2^{r-r_L} b(r, y, x) \text{ при } y = \overline{0, Y-1}, x = \overline{0, X-1}.$$

Для частного случая $r_H = r_L$ справедливо равенство $i_C(r_L, r_H, y, x) = b(r_L, y, x)$.

В табл. 1–3 приведены средние коэффициенты сжатия битовых плоскостей 8-разрядных ($R = 8$) спутниковых, портретных, медицинских полутоновых и ландшафтных тепловизионных изображений, а также разностей спектральных каналов 16-разрядных ($R = 16$) гиперспектральных изображений (ГСИ), полученные с использованием RLE и арифметического кодеров (AC). Усреднение значений коэффициентов сжатия произведено по 8-ми тестовым изображениям каждого типа.

Кодер RLE применен отдельно для каждой битовой плоскости ($f_{RLE}(B(r))$ при $r = \overline{0, R-1}$), где f_{RLE} – функция RLE-кодирования. На основе получаемого при этом объема $\langle f_{RLE}(B(r)) \rangle$ кода (в битах) вычисляются частичные коэффициенты сжатия с помощью выражения

$$CR_{RLE}(r) = YX / \langle f_{RLE}(B(r)) \rangle \text{ при } r = \overline{0, R-1},$$

где $\langle \rangle$ – оператор вычисления объема кода.

Кодирование тестовых изображений целиком (без разделения на битовые плоскости) с помощью кодера RLE не приводит к сжатию ($\langle f_{RLE}(I(R)) \rangle \geq RYX$).

Таблица 1. Средние коэффициенты сжатия битовых плоскостей изображений для кодера RLE

Битовые плоскости	Частичные и полные коэффициенты сжатия битовых плоскостей изображений				
	Полутоновые спутниковые	Полутоновые портретные	Полутоновые медицинские	Ландшафтные тепловизионные	Разности спектральных каналов
7 (15)	1,74466	1,56655	10,47579	1,17699	18318,4105 (1,4104)
6 (14)	1,12429	0,86344	3,50286	0,5182	18318,8455(50909,6075)
5 (13)	0,63831	0,6439	2,12782	0,34866	1157,6183(50909,6075)
4 (12)	0,42969	0,43127	1,10178	0,36688	1,1011 (50909,6075)
3 (11)	0,39203	0,40328	0,74446	0,33853	0,2947 (50909,6075)
2 (10)	0,40096	0,48656	0,50431	0,33297	0,2553 (50909,6075)
1 (9)	0,39986	0,50185	0,38566	0,36361	0,2450(45954,3425)
0 (8)	0,40006	0,44505	0,33411	0,38925	0,2494(36585,6001)
Все	0,51390	0,52599	0,73092	0,40438	0,81049

Таблица 2. Средние коэффициенты сжатия битовых плоскостей изображений для арифметического кодера

Битовые плоскости (старшие/младшие)	Частичные (полные) коэффициенты сжатия старших/младших битовых плоскостей изображений			
	Полутоновые спутниковые	Полутоновые портретные	Полутоновые медицинские	Ландшафтные тепловизионные
7/6–0	25,7676/1,0804 (1,2272)	13,0135/1,2382 (1,3941)	13,0135/1,6706 (1,8720)	27,4541/1,1168 (1,2689)
7–6/5–0	3,5595/1,0246 (1,2150)	2,2076/1,1635 (1,2982)	6,39613/1,5153 (1,8473)	2,8637/1,0581 (1,2550)
7–5/4–0	3,0243/1,0015 (1,3113)	1,9414/1,1058 (1,2886)	3,9399/1,3923 (1,7935)	2,2112/1,0244 (1,2816)
7–4/3–0	2,2277/0,9912 (1,3543)	1,7919/1,0650 (1,3209)	3,2933/1,2982 (1,8250)	1,9459/1,0045 (1,3239)
7–3/2–0	1,8680/0,9870 (1,3905)	1,6408/1,0244 (1,3295)	2,8619/1,2243 (1,8878)	1,7454/0,9925 (1,3578)
7–2/1–0	1,6625/0,9806 (1,4111)	1,5466/1,0208 (1,3577)	2,5539/1,1666 (1,9586)	1,5976/0,9827 (1,3810)
7–1/0	1,5246/0,9621 (1,4186)	1,4963/0,9358 (1,3821)	2,3217/1,0958 (2,0332)	1,4915/0,9649 (1,3958)

Арифметический кодер применен отдельно для старших ($f_{AC}(i_C(r_{HL}, r_{HH}, y, x))$) и младших ($f_{AC}(i_C(r_{LL}, r_{LH}, y, x))$) битовых плоскостей ($r_{HH} > r_{HL}, r_{LH} = r_{HL} - 1, r_{LH} > r_{LL}, r_{LL} \geq 0$), где f_{AC} – функция арифметического кодирования. Частичные коэффициенты сжатия вычисляются с помощью выражений

$$CR_{AC}(r_{HL}, r_{HH}) = (r_{HH} - r_{HL} + 1)YX / \langle f_{AC}(I_C(r_{HL}, r_{HH})) \rangle,$$

$$CR_{AC}(r_{LL}, r_{LH}) = (r_{LH} - r_{LL} + 1)YX / \langle f_{AC}(I_C(r_{LL}, r_{LH})) \rangle.$$

Таблица 3. Средние коэффициенты сжатия битовых плоскостей разностей каналов ГСИ для арифметического кодера

Битовые плоскости (старшие/младшие)	Частичные (полные) коэффициенты сжатия старших/младших битовых плоскостей изображений
15/14-0	0,5706/2,4806 (2,0414)
14/13-0	15,6852/2,3152 (2,6180)
14-13/12-0	31,3705/2,1498 (2,6180)
14-12/11-0	47,0557/1,9844 (2,6180)
14-11/10-0	62,7409/1,8191 (2,6180)
14-10/9-0	78,4262/1,6537 (2,6180)
14-9/8-0	94,1114/1,4883 (2,6180)
14-8/7-0	109,7933/1,3235 (2,6191)
14-7/6-0	119,2902/1,1660 (2,6355)
14-6/5-0	92,6252/1,0025 (2,6136)
14-5/4-0	94,0904/0,8537 (2,6491)
14-4/3-0	93,4257/0,7415 (2,7321)
14-3/2-0	22,6167/0,6888 (2,9558)
14-2/1-0	6,5078/0,6497 (3,0008)
14-1/0	3,9643/0,6073 (3,0477)

Битовая плоскость 15 разностей спектральных каналов ГСИ содержит коды знаков значений разностей: 0 – плюс, 1 – минус и кодируется отдельно ($f_{AC}(B(15))$). В таблицах приведены также полные коэффициенты сжатия $CR_{RLE}(R)$ и $CR_{AC}(R)$ всех битовых плоскостей изображений, учитывающие объемы кодов отдельных битовых плоскостей и их сочетаний, вычисляемые с помощью выражений

$$CR_{RLE}(R) = \frac{RYX}{\sum_{r=0}^{R-1} \langle f_{RLE}(B(r)) \rangle},$$

$$CR_{AC}(R) = \frac{RYX}{\langle f_{AC}(I_C(r_{HL}, r_{HH})) \rangle + \langle f_{AC}(I_C(r_{LL}, r_{LH})) \rangle}.$$

Из табл. 1 следует, что кодер RLE позволяет сжимать две старшие ($r=7, 6$) битовые плоскости медицинских и одну старшую ($r=7$) битовую плоскость портретных 8-разрядных полутоновых изображений, а также 9–11 старших битовых плоскостей, начиная с 14-й, разностей спектральных каналов ГСИ. Для остальных битовых плоскостей его использование не приводит к сжатию. Причем, даже в случае сжатия RLE-кодер уступает в эффективности арифметическому кодеру на всех типах изображений ($CR_{RLE}(r) > 1$, $CR_{RLE}(r) < CR_{AC}(r_{HL}, r_{HH})$, $CR_{RLE}(r) < CR_{AC}(r_{LL}, r_{LH})$), кроме разностей каналов ГСИ, что связано с относительно высокой вероятностью нулей в их старших битовых плоскостях.

Использование арифметического кодера для младшей битовой плоскости 8-разрядных полутоновых изображений ($\langle f_{AC}(I_C(0,0)) \rangle > YX \Rightarrow CR_{AC}(0,0) < 1$) и 2-х младших битовых плоскостей 16-разрядных разностей каналов ГСИ ($\langle f_{AC}(I_C(0,1)) \rangle > 2YX \Rightarrow CR_{AC}(0,1) < 1$), как правило, не эффективно. В остальных случаях арифметический кодер обеспечивает сжатие отдельных битовых плоскостей и их комбинаций. Причем, при кодировании старших битовых плоскостей разностей каналов ГСИ арифметический кодер показывает худшие результаты в сравнении с кодером RLE.

При кодировании тестовых изображений целиком (без деления на битовые плоскости) с помощью арифметического ($f_{AC}(I(R))$) и RLE ($f_{RLE}(I(R))$) кодеров получены следующие

средние коэффициенты сжатия: спутниковые изображения – $CR_{AC}=1,22$ и $CR_{RLE}=0,95$ раз; портретные изображения – $CR_{AC}=1,35$ и $CR_{RLE}=1,04$ раз; медицинские изображения – $CR_{AC}=1,86$ и $CR_{RLE}=1,62$ раз; ландшафтные тепловизионные изображения – $CR_{AC}=2,13$ и $CR_{RLE}=1,69$ раз; разности каналов ГСИ – $CR_{AC}=2,29$ и $CR_{RLE}=1,02$ раз. Сопоставление этих результатов с данными, приведенными в табл. 1–3, показывает, что раздельное кодирование битовых плоскостей ($f_{RLE}(B(r))$ при $r=0, \overline{R-1}$ и $f_{AC}(I_C(r_{HL}, r_{HH})), f_{AC}(I_C(r_{LL}, r_{LH}))$ при $r_{HH}=R-1, r_{HH} > r_{HL}, r_{LH}=r_{HL}-1, r_{LH} > r_{LL}, r_{LL}=0$) позволяет повысить коэффициент сжатия изображений по сравнению с их непосредственным кодированием ($f_{RLE}(I(R))$ и $f_{AC}(I(R))$).

Комбинированное эффективное кодирование битовых плоскостей изображений

Для повышения коэффициента сжатия изображений без потерь предлагается использовать комбинированное эффективное кодирование. Его сущность состоит в использовании для сжатия старших битовых плоскостей изображения или их комбинаций нескольких кодеров различных типов ($f_{RLE}(B(r))$ и $f_{AC}(I_C(r_{HL}, r_{HH}))$), лучше учитывающих распределение их значений, и непосредственном включении в результат кодирования младших битовых плоскостей ($f_{NC}(B(r))$ при $r \geq 0$ или $f_{NC}(I_C(0, r_{LH}))$ при $r_{LH} \geq 0$), кодирование которых не эффективно ($\langle f_{RLE}(B(r)) \rangle \geq YX$ при $r \geq 0$ или $\langle f_{AC}(I_C(0, r_{LH})) \rangle \geq (r_{LH} + 1)YX$ при $r_{LH} \geq 0$) (рис. 2), где f_{NC} – функция непосредственного переноса младших битовых плоскостей без кодирования.

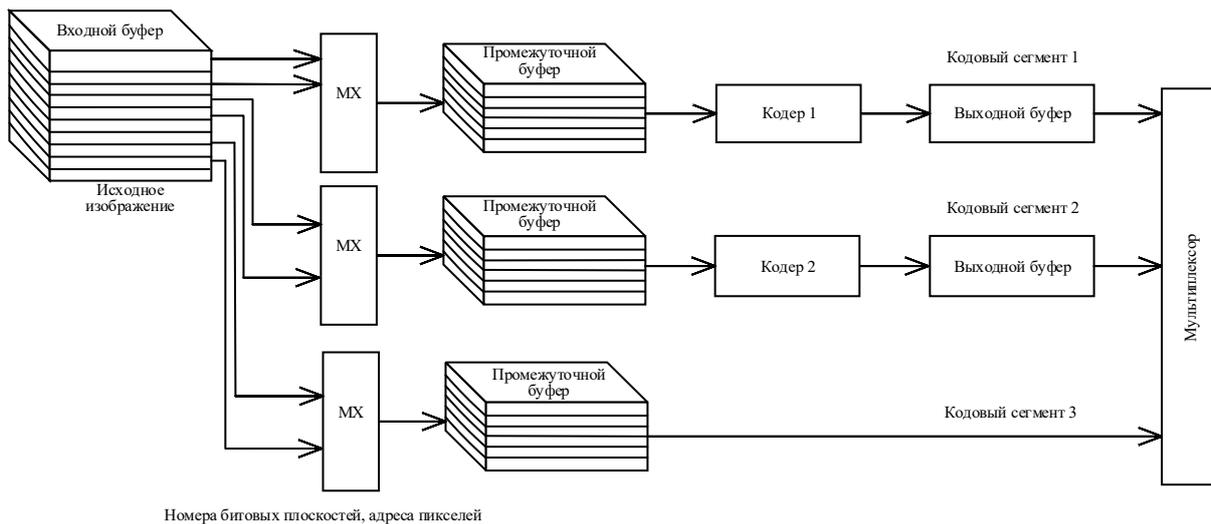


Рис. 2. Схема комбинированного эффективного кодирования битовых плоскостей изображения

Из табл. 1–3 следует, что комбинированное кодирование эффективно для сжатия разностей каналов ГСИ. В табл. 4 приведены наиболее эффективные комбинации арифметического и RLE кодеров для битовых плоскостей четырех разностей соседних каналов тестовых ГСИ, взятых из базы Aviris [16], и коэффициенты сжатия CR_C , соответствующие этим комбинациям.

Коэффициент сжатия CR_C при комбинированном эффективном кодировании вычисляется с помощью выражения

$$CR_C = \frac{RYX}{\sum_{m=1}^M \langle f_{RLE}(B(r(m))) \rangle + \sum_{n=1}^N \langle f_{AC}(I_C(r_L(n), r_H(n))) \rangle + \sum_{p=0}^{P-1} \langle f_{NC}(B(p)) \rangle}$$

при $R = M + N + P$, $r_H(n) > r_L(n)$, $M \geq 0$, $N \geq 0$, $P \geq 0$,

где M – число битовых плоскостей, кодируемых с помощью кодера RLE; N – число битовых плоскостей, кодируемых с помощью арифметического кодера; P – число битовых плоскостей, непосредственно переносимых в результирующий код.

Таблица 4. Наиболее эффективные комбинации арифметического и RLE кодеров и коэффициенты сжатия

Битовые плоскости	Комбинации арифметического и RLE кодеров для разностей каналов ГСИ							
	Разность 1		Разность 2		Разность 3		Разность 4	
15 7	RLE	RLE	AC	RLE	AC	AC	RLE	AC
14 6	RLE	RLE	RLE	RLE	RLE		RLE	
13 5	RLE	RLE	RLE	RLE	RLE		RLE	
12 4	RLE	AC	RLE	AC	RLE		RLE	
11 3	RLE		RLE		RLE	RLE	RLE	RLE
10 2	RLE	NC	RLE	NC	RLE	NC	RLE	NC
9 1	RLE	NC	RLE		RLE	NC	RLE	NC
8 0	RLE	NC	RLE	NC	RLE	NC		NC
CR_C	3,72917		3,16748		3,15484		2,80657	
CR_{AC}	2,53492		2,29310		2,16147		2,18485	
CR_{RLE}	1,02866		0,97902		1,04260		1,03313	
$\overline{CR_C}$	2,46175		3,16773		3,07913		1,93187	

Основываясь на табл. 4, для повышения коэффициента сжатия предлагается следующее правило комбинированного кодирования разностей соседних каналов ГСИ, определяющее формирование результирующего кода C_C : необходимо использовать арифметический кодер для старшей знаковой плоскости и битовых плоскостей 8–2, кодер RLE – для битовых плоскостей 14–9, перенос без кодирования – для битовых плоскостей 1–0, что определяется выражением

$$C_C \leftarrow f_{AC}(B(15)) \oplus f_{RLE}(B(14)) \oplus \dots \oplus f_{RLE}(B(9)) \oplus f_{AC}(I_C(2,8)) \oplus f_{NC}(I_C(0,1)),$$

обеспечит объем кода $\langle C_C \rangle$, равный

$$\langle C_C \rangle = \langle f_{AC}(B(15)) \rangle + \sum_{r=9}^{14} \langle f_{RLE}(B(r)) \rangle + \langle f_{AC}(I_C(2,8)) \rangle + \langle f_{NC}(I_C(0,1)) \rangle$$

и коэффициент сжатия $\overline{CR_C}$, равный

$$\overline{CR_C} = \frac{16YX}{\langle C_C \rangle},$$

где \oplus – оператор конкатенации, обеспечивающий формирование результирующего кода из фрагментов, соответствующих битовым плоскостям и их комбинациям.

Из табл. 4 следует, что комбинированное кодирование позволяет повысить коэффициент сжатия разностей соседних каналов ГСИ в среднем в $CR_C/CR_{AC} = 1,4$ и $CR_C/CR_{RLE} = 3,1$ раза по сравнению с арифметическим и RLE-кодированием соответственно.

Предложенное правило позволяет синтезировать структуру кодера ГСИ (рис. 3). Она состоит из комбинированного кодера разностей каналов ГСИ и кодера опорного канала ГСИ. Комбинация арифметического и RLE кодеров в части кодера разностей каналов ГСИ отличается от эффективных комбинаций, рассмотренных в табл. 4, что приводит к снижению коэффициента сжатия, однако не требует использования в кодеках дополнительных коммутационных

и управляющего блоков. Выигрыш в коэффициенте сжатия разностей соседних каналов ГСИ составляет в среднем в $\overline{CR}_C/\overline{CR}_{AC} = 1,2$ и $\overline{CR}_C/\overline{CR}_{RLE} = 2,6$ раза по сравнению с арифметическим и RLE-кодированием соответственно. Для кодирования опорного канала ГСИ используется арифметический кодер, на который подаются 12 старших битовых плоскостей опорного канала, а 4 младших битовых плоскости непосредственно переносятся в результирующий код. Такая структура следует из анализа эффективности отдельного кодирования битовых плоскостей каналов ГСИ, который показывает, что, как правило, от 2 до 6 младших битовых плоскостей не сжимаются с помощью арифметического кодера [17]. Наиболее вычислительно сложными элементами кодера ГСИ являются блоки арифметического кодирования (AC), работающие на частотах $YXC_{AC}f$ и $YXC_{AC}f/(N-1)$ в части кодеров разностей спектральных каналов и опорного канала ГСИ соответственно, где N – число спектральных каналов; $C_{AC} > 1$ – коэффициент учитывающих вычислительную (временную) сложность арифметического кодера в сравнении с RLE (работает на частоте YXf); f – тактовая частота $f = Nf_i$; f_i – частота формирования ГСИ. Экспериментально установлено, что $C_{AC} \approx 11$.

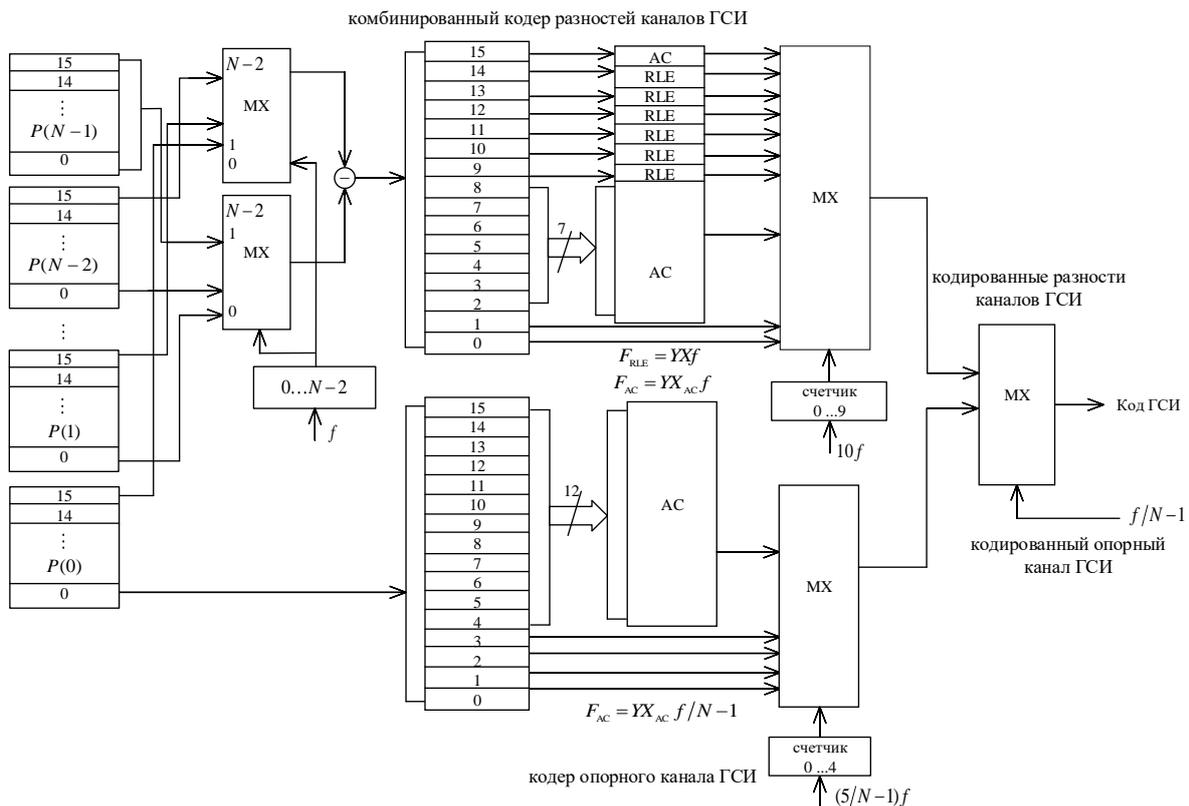


Рис. 3. Структура кодера гиперспектральных изображений

Структура кодера ГСИ включает также мультиплексоры спектральных каналов, работающие на частоте f по входам управления и на частоте $16YXf$ по входам данных; мультиплексор кодов битовых плоскостей разностей спектральных каналов, работающий на частоте $10f$; мультиплексор кодов битовых плоскостей опорного канала, работающий на частоте $5f/(N-1)$; мультиплексор кодов опорного и разностей каналов ГСИ, работающий на частоте $f/(N-1)$. При последовательной реализации вычислительная сложность комбинированного кодера составляет примерно $2YXC_{AC}f + 6YXf$ при большом числе спектральных каналов, что примерно в 2,5 и 28 раз больше по сравнению с арифметическим и RLE-кодированием соответственно.

Заклучение

Разработана структура комбинированного кодера сжатия битовых плоскостей изображений без потерь в пространственной области на основе арифметического и RLE-кодеров. Установлена эффективность комбинированного кодирования для сжатия разностей каналов гиперспектральных изображений. Выбор оптимальных комбинаций битовых плоскостей для арифметического и RLE кодирования позволяет повысить коэффициент сжатия разностей соседних каналов ГСИ в среднем в 1,4 и 3,1 раза по сравнению с арифметическим и RLE-кодированием соответственно. Разработано правило комбинированного кодирования разностей соседних каналов гиперспектральных изображений, использующее арифметическое кодирование для старшей знаковой плоскости и битовых плоскостей 8–2, кодирование RLE – для битовых плоскостей 14–9, перенос без кодирования – для битовых плоскостей 1–0, что позволило повысить коэффициент сжатия разностей соседних каналов гиперспектральных изображений в среднем в 1,2 и 2,6 раза по сравнению с арифметическим и RLE-кодированием соответственно при увеличении вычислительной сложности в 2,5 и 28 раз. На основе данного правила разработана структура кодера гиперспектральных изображений.

LOSSLESS IMAGE COMPRESSION BASED ON THE COMBINED EFFICIENT CODING OF BIT PLANES

B.J.S. SADIQ, V.Yu. TSVIATKOV, M.N. BOBOV

Abstract. A structure of the efficiency of the combined image compression codec without loss in the spatial domain using arithmetic and RLE-coding algorithms for compact representation of various bit planes of the image was developed and researched.

Keywords: combined efficient coding, lossless image compression.

Список литературы

1. Ватолин Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М: Диалог-МИФИ, 2003.
2. Сэлмон Д. Сжатие данных, изображений и звука. М: Техносфера. 2004.
3. Тропченко А.Ю., Тропченко А.А. Методы сжатия изображений, аудиосигналов и видео: учеб. пособие. СПб: СПбГУ ИТМО. 2009.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М: Техносфера. 2006.
5. Ziv J., Lempel A. // IEEE Transactions on Information Theory. IT-23. 1977. P. 337–343.
6. Taubman D.S., Marcellin M.W. JPEG2000: image compression fundamentals, standards, and practice. Springer-Verlag. 2002.
7. Shapiro J.M. // IEEE Trans. Signal Processing. 1993. № 41. P. 3445–3462.
8. Said A., Pearlman W.A. // IEEE Trans. on Circuits and Systems for Video Technology. 1996. Vol. 6. P. 243–250.
9. Islam A., Pearlman W.A. // ISO/IEC/JTC1/SC29, WG1. 1998. № 873. P. 312–326.
10. Аль-Бахдили Х.К., [и др.] // Докл. БГУИР. 2016. № 2 (96). С. 63–69.
11. Al-Bahadily H.K., [et. al.] // International Journal of Advanced Computer Science and Applications. 2016. Vol. 7. P. 250–255.
12. Golomb S.W. // IEEE Transactions on Information Theory. 1966. P 399–401.
13. Abdmouleh M.K., [et. al.] // Journal of Software Engineering and Applications. 2012. № 5. P. 41–44.
14. Abramson N. Information Theory and Coding. McGraw-Hill, New York. 1963.
15. Rubin F. // IEEE Trans. Inf. Theory IT-25, 6. 1979. P. 672–675.
16. Airborne Visible/Infrared Imaging Spectrometer [Electronic resource]. URL: <https://aviris.jpl.nasa.gov/data/index.html>. (date of access: 02.11.2019).
17. Садик Б.Дж., Бобов М.Н. // Телекоммуникации: сети и технологии, алгебраическое кодирование и безопасность данных: материалы междунар. научно-технич. семинара, ч. 2. Минск. БГУИР, 2016. С. 61–67.

УДК 621.391

ОЦЕНКА ЭФФЕКТИВНОСТИ КАНАЛОВ РАСПРЕДЕЛИТЕЛЬНОЙ СИСТЕМЫ СТАНДАРТА DVB-C2

Э.Б. ЛИПКОВИЧ, А.А. ПАВЛОВСКИЙ, А.А. СЕРЧЕНЯ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 30 октября 2019*

Аннотация. Рассмотрены принципы построения и функционирования передающего тракта цифровой распределительной сети стандарта DVB-C2. Представлены расчетные соотношения, по которым определены значения информационной скорости и спектральной эффективности каналов DVB-C2 с различными видами и порядком модуляции. Приведены математические модели расчета помехоустойчивости стандарта DVB-C, использующий многопозиционные виды модуляции и блочное кодирование по алгоритму Рида – Соломона (РС). На основании расчетных моделей выполнен сравнительный анализ эффективности каналов DVB-C2 и DVB-C и отмечены существенные преимущества систем второго поколения. Дана оценка показателей эффективности реальных каналов DVB-C2 по сравнению с теоретически достигаемыми по Шеннону.

Ключевые слова: помехоустойчивость, многопозиционные виды модуляции, кодирование, энергетическая эффективность, вероятность ошибки.

Введение

Возросшие требования к пропускной способности и энергетической эффективности кабельных сетей цифрового телевизионного вещания и мультисервисных услуг привели к разработке нового стандарта DVB-C2 EN 302769 [1]. Стандарт DVB-C2 и рекомендация к нему TS 102 991 [2] утверждены ETSI в 2011 г. Основные положения нового стандарта базируются на спецификациях транспортных стандартов DVB-S2 [3] и DVB-T2 [4], зарегистрированных в ETSI в 2006 и 2008 гг. соответственно.

К числу отличительных особенностей стандарта DVB-C2 по сравнению с DVB-C следует отнести:

- возможность обработки нескольких входных потоков, переносящих разные виды информационных услуг (концепция PLP – Physical Layer Pipe);
- двухступенчатое кодирование принятых данных по схеме BCH+LDPC;
- адаптивный выбор требуемых параметров передачи;
- использование высокоформатных видов модуляции QAM-1024 и QAM-4096 для повышения спектральной эффективности сети;
- применение устойчивого к интерференции многочастотного режима передачи с защитными интервалами между OFDM-символами;
- внесение в передаваемый поток несущих пилот-сигналов, необходимых для коррекции искажений в распределительной сети и быстрого вхождения приемника в режим синхронизма.

Принципы построения и функционирования передающего тракта системы

На рис. 1 приведена структурная схема передающего тракта системы DVB-C2. На входы тракта могут параллельно поступать независимые потоки данных TS/GS, содержащие транспортные пакеты MPEG-2 TS, IP-пакеты по протоколу GSE (Generic Stream Encapsulation) и непакетированные данные.

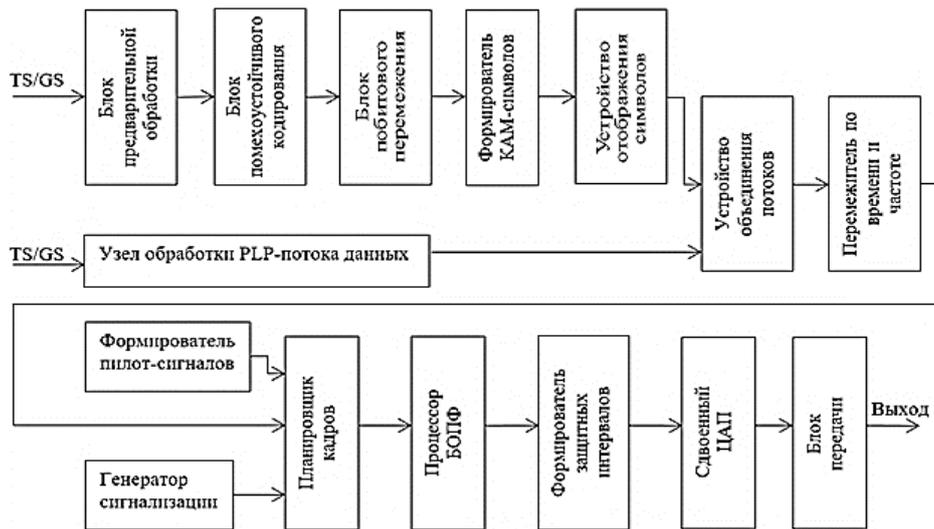


Рис. 1. Структурная схема передающего тракта системы DVB-C2

Каждый отдельный входной информационный поток проходит блок предварительной обработки данных (БПОД), или модуль адаптации режима, структурная схема которого приведена на рис 2. По своей структуре и решаемым задачам этот блок подобен аналогичным модулям обработки PLP-потоков стандартов DVB-S2 и DVB-T2.

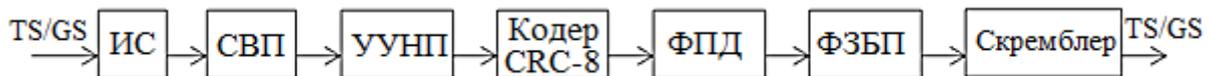


Рис. 2. Блок предварительной обработки PLP-потока

В состав БПОД входят: интерфейс сопряжения (ИС); синхронизатор входного потока (СВП) для синхронизации устройств приема по тактовой частоте $F_r = 8\Delta f_k / 7$; устройство удаления нулевых пакетов (УУНП) из транспортного потока; кодер CRC-8 для обнаружения на приеме ошибок в пакете; формирователь полей данных (ФПД) требуемых размеров; формирователь заголовка базовой полосы (ФЗБП) и скремблер. Заголовок имеет длину 80 бит, содержит сведения о характеристиках потока и располагается перед полем данных.

По выделенному на приеме идентификатору PLP ID все пакеты, не принадлежащие к затребованной услуге, не обрабатываются. Скремблер выполнен на сдвиговом регистре с обратной связью, описывается полиномом $G(x) = 1 + x^{14} + x^{15}$ и упрощает восстановление тактовой частоты на приеме при длинных сериях нулей и единиц на стороне передачи. Исходное состояние скремблера устанавливается через интервалы поля данных инициализирующей последовательностью 100101010000000. С выхода БПОД скремблированный поток поступает на внешний кодер с кодом Боуза-Чоудхури-Хоквингема (BCH – Bose Chaudhuri Hocquenghem) блока помехоустойчивого кодирования. Кодер BCH является циклическим, двоичным и работает с длинными кодами. Длина информационной части кода $K_{BCH} = (n_L \cdot R_L - 16t)$. Она зависит от относительной скорости внутреннего LDPC-кодера с низкой плотностью проверок на четность $R_L = k_L / n_L$ и является частью (рис. 3) общей длины кодового слова $n_L = 64\ 800$ бит (стандартный формат), или $n_L = 16\ 200$ бит (сокращенный формат), где t – исправляющая способность кода BCH, $16t$ – проверочная часть кода BCH.

Относительная скорость кода BCH близка к единице и рассчитывается по формуле

$$R_{BCH} = K_{BCH} / N_{BCH} = 1 - (16t / R_L \cdot n_L). \quad (1)$$

Для кодера LDPC значения R_L установлены равными: 2/3, 3/4, 4/5, 5/6, 9/10 при $n_L = 64\ 800$ бит и 1/2, 2/3, 3/4, 4/5, 5/6, 8/9 при $n_L = 16\ 200$ бит. Задача процесса кодирования состоит в определении состава кодового слова и его проверочной части ($n_L - k_L$).

Результирующая кодовая скорость R_k определяется по формуле

$$R_k = R_L - (16t / n_L). \quad (2)$$

Принятый в DVB-C2 двухступенчатый способ кодирования с разной структурой построения отдельных устройств по схеме BCH + LDPC обеспечивает лучшую помехоустойчивость приема по сравнению с одноступенчатым декодированием по алгоритму Рида – Соломона, принятому в DVB-C.

С выхода кодера с кодом с низкой плотностью проверок на четность (LDPC – Low Density Parity Checking) данные поступают на блок побитового перемежения, в котором по определенному правилу в зависимости от значений R_L и $Q_L = 180(1 - R_L)$ осуществляется перемежение проверочных бит. После этой операции следует повторное перемежение на базе блочного перемежителя со сдвигом начала стартовых позиций в столбцах. Запись данных в блочный перемежитель производится по столбцам, считывание – по строкам. Число столбцов и строк зависит от кратности модуляции m . При $m = 4, 6, 8, 10$ и $=64$ 800 бит их число составляет $2m$ и $n_L / 2m$ соответственно. Например, при использовании QAM-16 ($m = 4$) число столбцов равно 8, а число строк в каждом столбце – 8100. Для QAM-4096 ($m = 12$) число столбцов и строк равно 12 и 5400 соответственно. Стартовая позиция в столбцах сдвигается на определенную величину, принятую в стандарте. Например, для QAM-16 с восемью столбцами величина сдвига в столбцах равна 0; 0; 2; 4; 5; 7; 7. В формирователе QAM -символов поток данных демультиплексируется на кортежи. Их число составляет n_L/m , а число бит в картеже – m . Имеющееся сочетание бит в каждом кортеже с учетом кода Грея отображается на сигнальном созвездии модулятора определенным вектором сигнала. Положение этого вектора представляется в устройстве отображения символов комплексным числом $\dot{Z}q$. Для усреднения выходной мощности передатчика при использовании в системе разных видов модуляции значения $\dot{Z}q$ нормируются на величину $\sqrt{E_{cp}} = \sqrt{2(M-1)}/3$, т. е. на $\sqrt{2}, \sqrt{10}, \sqrt{42}, \sqrt{70}, \sqrt{170}, \sqrt{682}, \sqrt{2138}$, для QAM с M , равными 4, 16, 64, 256, 1024 и 4096 соответственно.

Поскольку PLP-потoki могут переносить разные объемы данных и ввиду того, что кабельная распределительная сеть является закрытой средой без ограничений на выбор полосы каналов, то в стандарте DVB-C2 рекомендуется объединять потоки в устройстве объединения потоков (УОП) в слайсы (блоки). Ширина полосы каждого блока определяется объемом данных и числом несущих, выделенных ему для передачи. При базовом формате передачи в режиме работы 4k суммарное число несущих $N_\Sigma = 3408$, частотный разнос между ними $\Delta f_p = 1/T_{ин} = 2,232$ кГц и ширина полосы блока $\Delta f_c = N_\Sigma \cdot \Delta f_p = 7,61$ МГц. Здесь $T_{ин}$ – длительность информационной части OFDM-символа, равная 448 мкс для $\Delta f_k = 8$ МГц. Блоки могут объединяться в общую группу без защитных полос между ними (рис. 3), что повышает скорость передачи данных и эквивалентную полосу $\Delta f_{эк}$, занимаемую сигналом в полосе канала Δf_k .



Рис. 3. Структура кадра системы DVB-C2

Для сохранения единства в ширине полосы приемников разных стандартов в DVB-C2 рекомендован сегментированный способ передачи и приема, подобно принятому в японской системе ISDB-T. Суть сегментации состоит в выделении для конкретного вида услуг и объема переносимых данных определенного числа несущих. Объединенные блоки данных PLP-поток подвергается временному и частотному перемежению. При временном перемежении задействованы два блока буферной памяти: в первый вносятся данные, а из второго они считываются. Далее происходит переключение режимов. При частотном перемежении осуществляется изменение порядка следования QAM-символов в цифровой последовательности, что исключает их передачу на близлежащих несущих OFDM-символов. В этом отношении частотное перемежение можно трактовать как способ борьбы с узкополосными помехами, поскольку перемежение QAM-символов эквивалентно частотному перемежению несущих.

Следующим устройством в системе DVB-C2 является планировщик кадров. Планирование кадров начинается с формирования преамбулы (заголовка), которая занимает полосу частот 7,61 МГц и состоит из одного или нескольких OFDM-символов длительностью $T_{\text{и}}$. Каждая шестая несущая преамбулы предназначена для передачи пилотных данных, поступающих от генератора сигнализации L1. В частности, в преамбулу вносятся сведения о начальных и конечных значениях частот отдельных блоков (слайсов), значение оптимальной частоты настройки приемника на требуемый блок, данные, описывающие состав преамбулы и др. Благодаря этой информации достигается точная настройка приемника на требуемый сегмент общей полосы, переносящий заданный вид услуг.

Преамбула и 448 OFDM-символов с $T_{\text{и}} = 448$ мкс образуют кадр общей длительностью около 200 мс. Для обеспечения устойчивого приема сигналов при меняющихся параметрах в среде передачи в каждый кадр вносятся несущие пилот-сигналов, отличающиеся от информационных несколько большей амплитудой. Делятся пилотные несущие на три группы: с постоянными позициями в OFDM-символе, рассеянные по кадру и граничные, указывающие на края каждого символа. Общее число пилотных несущих $N_{\text{пс}}$ в OFDM-символе может варьироваться и составлять некоторую часть от общего числа информационных несущих $N_{\text{и}}$. Наибольшее применение при построении системы данного стандарта получили два значения отношений $N_{\text{пс}} / N_{\text{и}}$, равные 1/48 и 1/96.

Величина спектральных потерь от наличия в полосе канала пилотных несущих составляет

$$N_{\Sigma} / N_{\text{и}} = 1 + (N_{\text{пс}} / N_{\text{и}}). \quad (3)$$

Согласно (3) для $N_{\Sigma} = 3800$ и отношений 1/48 и 1/96 число пилотных несущих равно 70 и 36, а число информационных несущих – 3338 и 3372 соответственно.

В процессорном блоке быстрого обратного преобразования Фурье (БОПФ) осуществляется процедура модуляции несущих, состоящая в матричном перемножении дискретных значений несущих N_{Σ} и комплексных значений QAM-символов. В базовом режиме работы $4k$ размерность процессорного блока составляет 4096×4096 . С увеличением общей полосы, превышающей Δf_K , используются режимы работы $8k$, $16k$, $32k$, для которых процессорная емкость в 2, 4 и 8 раз больше, чем в режиме $4k$.

Для борьбы в кабельных сетях с интерференцией и эхо-сигналами, возникающими в виде задержанных копий, стандартом предусматривается размещение между OFDM-символами защитных интервалов (ЗИ) длительностью T_3 . Длительность T_3 обычно составляет небольшую часть от информационной длительности символа $T_{\text{и}}$ и превышает типовую задержку эхо-сигналов в каналах. В стандарте DVB-C2 отношение $T_3 / T_{\text{и}}$ имеет два значения 1/64 и 1/128. Эти значения относительно принятых в DVB-T/T2 небольшие и оказывают гораздо меньшее влияние на снижение информационной скорости данных, чем в системах НЦТВ [4]. На выходе

блока БОПФ присутствуют двухканальный ЦАП, фильтры нижних частот в I - и Q -каналах и блок передачи, оснащенный переносчиками частоты радиосигнала на частоту передачи.

Математические модели расчета эффективности передающего тракта системы

Информационная скорость данных в тракте системы с учетом потерь от наличия ЗИ, пилотных несущих, заголовка и преамбулы составляет

$$B_0 = \frac{\Delta f_{\text{эк}} \cdot m \cdot R_k}{(1 + N_{\text{пс}}/N_{\text{и}}) \cdot (1 + T_3/T_{\text{и}}) \cdot (1 + 80/n_k) \cdot (1 + \delta/448)}, \text{ бит/с}, \quad (4)$$

где $\Delta f_{\text{эк}}$ – эквивалентная ширина полосы, занимаемая информационным сигналом в канале, Гц; δ – число OFDM-символов в преамбуле.

В табл. 1 приведены результаты расчета B_0 для используемых в DVB-C2 видов модуляции и кодовых скоростей R_L при условии, что $\Delta f_{\text{эк}} = 7,61$ МГц, $T_3/T_{\text{и}} = 1/128$, $N_{\text{пс}}/N_{\text{и}} = 1/96$, $n_L = 64800$ и $\delta = 2$.

Таблица 1. Значения информационной скорости передачи данных, бит/с

R_L	QAM-16	QAM-64	QAM-256	QAM-1024	QAM-4096
2/3	20,22	30,32	40,44	50,55	60,66
3/4	22,74	34,11	45,48	56,85	68,22
4/5	23,69	35,53	47,38	59,22	71,07
5/6	24,7	37,04	49,39	61,74	74,09
9/10	26,69	40,03	53,38	80,72	80,07

При оценке эффективности характеристик системы DVB-C2 важно исследовать спектральные, энергетические и информационные показатели.

Спектральная эффективность каналов γ_c и удельная скорость данных γ_0 рассчитываются по формулам

$$\gamma_c = B_0/\Delta f_k = m \cdot R_k/b_p \cdot L_{\text{и}}, \text{ бит/с} \cdot \text{Гц}; \quad (5)$$

$$\gamma_0 = B_0/B_c = m \cdot R_k/L_{\text{и}}, \text{ бит/симв}; \quad (6)$$

$$L_{\text{и}} = (1 + T_3/T_{\text{и}}) \cdot (1 + 80/n_L) \cdot (1 + \nu/448), \quad (7)$$

где B_c – символьная скорость, симв/с; $L_{\text{и}}$ – потери в скорости передачи; b_p – коэффициент, определяющий потери в использовании полосы канала, равный

$$b_p = \Delta f_k \cdot (1 + N_{\text{пс}}/N_{\text{и}})/\Delta f_{\text{эк}}. \quad (8)$$

Для типовых параметров канала ($\Delta f_k = 8$ МГц, $\Delta f_{\text{эк}} = 7,61$ МГц, $N_{\text{пс}}/N_{\text{и}} = 1/48$ и $1/96$) значения коэффициентов b_p равны 1,073 и 1,062 соответственно. Эти значения меньше принятого в системе DVB-C $b_p = 1,15$ и, следовательно, в DVB-C2 лучше используется полоса канала.

В табл. 2 приведены значения γ_c (бит/с·Гц) и отношения несущая/шум ρ_c (дБ) для каналов формата DVB-C2.

Таблица 2. Значения параметров γ_c , γ_0 и ρ_c для стандарта DVB-C2

R_k	QAM-16		QAM-64		QAM-256		QAM-1024		QAM-4096	
	γ_c	ρ_c								
2/3	2,53	8,9	3,79	13,5	5,05	17,8	6,32	–	7,58	–
3/4	2,84	10,0	4,26	15,1	5,68	20,0	7,11	24,8	8,53	–
4/5	2,96	10,7	4,44	16,1	5,92	21,3	7,40	–	8,88	–
5/6	3,09	11,4	4,63	16,8	6,18	22,0	7,72	27,2	9,26	32,4
9/10	3,34	12,8	5,01	18,5	6,68	24,0	8,34	29,5	10,01	35,0

Чтобы выявить преимущества в помехоустойчивости и спектральной эффективности каналов DVB-C2 по сравнению с DVB-C, выполним соответствующие расчеты для каналов DVB-C, использующих на стороне приема декодирование по алгоритму Рида – Соломона с жестким принятием решения. Величина отношения, несущая/шум для заданной вероятности ошибки на выходе декодера РС взаимосвязана с величиной ОСШ h_k и рассчитывается по формулам

$$\rho_k = h_k + 10 \lg \gamma_0, \text{ дБ.} \quad (9)$$

$$h_{kPC} = 10 \lg [2,3(B_i - \lg \sqrt{2,3B_i / \mu_{PC}}) / \mu_{PC}], \text{ дБ.} \quad (10)$$

$$B_i = -\lg P_b - \lg(\sqrt{\pi q_i / C_i}) + \lg \sqrt{\mu_{PC} / q_i}, \quad (11)$$

$$\mu_{PC} = R_{PC}(t+1)\beta_{PC}q_i \quad (12)$$

$$\beta_{PC} = [1 - (L_{PC} / (3,5L_{PC}\sqrt{P_b} - \lg P_b))] / Q_{PC}, \quad (13)$$

$$L_{PC} = \lg[R_{PC}(t+1)\sqrt{d_M q_i}], \quad Q_{PC} = 1 + \lg[(t+1)n / 2t(-\lg P_b)]. \quad (14)$$

где μ_{PC} – показатель эффективности процесса декодирования; β_{PC} – показатель взаимосвязи между параметрами кода и вероятностью ошибки на бит P_b ; $t = (n-k)/2$ – количество исправляемых символов декодером РС в кодовом слове; $d_M = (2t+1)$ – минимальное кодовое расстояние по Хэммингу; q_i – квадрат коэффициента помехоустойчивости; $R_{PC} = k/n = 188/204$ – кодовая скорость двоичного кода РС, принятая в системе DVB-C; $\gamma_0 = m \cdot R_k$ – удельная скорость; бит/симв.

Спектральная эффективность каналов DVB-C составляет

$$\gamma_c = \gamma_0 / b_p = m \cdot R_k / 1,15, \text{ бит/с} \cdot \text{Гц.} \quad (15)$$

Результаты расчета значений ρ_k , h_k и γ_c для $R_{PC} = k/n = 188/204$ и пяти видов модуляции в предположении вероятности ошибки на выходе декодера РС $P_b = 10^{-11}$ представлены в табл. 3.

Таблица 3. Результаты расчетов значений ρ_k , h_k и γ_c для каналов стандарта DVB-C, дБ

Параметр	Вид модуляции				
	QAM-16	QAM-32	QAM-64	QAM-128	QAM-256
ρ_k , дБ	17,76	21,02	23,85	26,94	29,78
h_k , дБ	12,1	14,39	16,43	18,85	21,11
γ_c , бит/с·Гц	3,205	4,0	4,808	5,61	6,41

В табл. 4 приведены результаты сравнения энергетической эффективности каналов стандартов DVB-C2 и DVB-C при одинаковом формате модуляции и близких для них значений спектральных эффективностей γ_c .

Таблица 4. Результаты сравнения ОНШ стандартов DVB-C2 и DVB-C

R_k	Вид модуляции		
	QAM-16	QAM-64	QAM-256
5/6	6,3	7,05	7,78
9/10	4,96	5,35	5,78

Данные табл. 4 отражают заметный выигрыш в ОНШ при использовании составных схем кодирования/декодирования, принятых в DVB-C2, по сравнению с кодированием/декодированием по алгоритму РС с жестким решением. Сравнение данных табл. 2 и 3 показывает, что при QAM-1024 для DVB-C2 требуемое значение ОНШ

($\rho_k = 29,5$ дБ, $R_L = 9/10$) близко к ОИШ при QAM-256 ($\rho_k = 29,78$ дБ) для каналов DVB-C. Однако при равенстве значений ρ_k новый стандарт позволяет при QAM-1024 получить выигрыш в спектральной эффективности 48,6 %.

Таким образом, выполненные исследования и полученные результаты позволяют количественно оценить преимущества DVB-C2 над DVB-C по базовым показателям, а также установить, насколько энергетическая эффективность каналов DVB-C2 проигрывает пределу по Шеннону.

Проигрыш в ОИШ для реального канала по сравнению с потенциально возможным значениям $\rho_{ш}$ по Шеннону рассчитывается по формуле

$$\delta\rho = \rho_k - \rho_{ш} = \rho_k - 10\lg(10^{0,3\gamma} - 1), \text{ дБ.} \quad (16)$$

Подставляя в (16) данные из табл. 2, получим, что энергетический проигрыш $\delta\rho$ в случае $\gamma = \gamma_0 = \gamma_c b_p$ находится в диапазоне значений от 2 до 3 дБ. Этот результат показывает, что при условии обеспечения квазибезошибочного приема сигналов в системе DVB-C2 ее характеристики еще не позволяют вплотную приблизиться к границе Шеннона.

Заключение

Изложены механизмы функционирования передающего тракта распределительной системы стандарта DVB-C2. Разработаны модели расчета спектральной эффективности и информационной скорости передачи данных по каналам системы. Приведены расчетные соотношения для оценки помехоустойчивости систем стандарта DVB-C, использующие многопозиционные виды модуляции и блочное кодирование по алгоритму Рида – Соломона. Дана сравнительная оценка показателей эффективности каналов стандартов DVB-C2 и DVB-C. Определены отклонения энергетической эффективности реальных каналов DVB-C2 от пороговой эффективности по Шеннону.

EVALUATION OF EFFICIENCY OF DVB-C2 DISTRIBUTION SYSTEM CHANNELS

E.B. LIPKOVICH, A.A. PAULOUSKI, A.A. SERCHENYA

Abstract. The principles of construction and functioning of the transmission path of the digital distribution network of DVB-C2 standard are considered. The calculated relations are presented, according to them the values of information speed and spectral efficiency of DVB-C2 channels with different types and modulation order are determined. Mathematical models of noise immunity calculation of DVB-C standard using multi – position modulation and block coding by Reed-Solomon algorithm (RS) are presented. On the basis of the calculated models, a comparative analysis of the efficiency of DVB-C2 and DVB-C channels was performed and significant advantages of the second generation systems were noted. The estimation of efficiency indicators of real DVB-C2 channels in comparison with theoretically reached on Shannon was given.

Keywords: noise immunity, multi-position modulation types, coding, energy efficiency, error probability.

Список литературы

1. ETSI EN 302 769. Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for a second generation digital transmission system for cable systems (DVB-C2) [Electronic resource]. URL: <http://www.etsi.org/deliver>. (date of access: 21.10.2019).
2. ETSI TS 102 991. Digital Video Broadcasting (DBV); Implementation Guidelines for a second generation digital cable transmission system (DVB-C2). [Electronic resource]. URL: <http://www.etsi.org/deliver>. (date of access: 21.10.2019).
3. ETSI EN 302 307. Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2). [Electronic resource]. URL: <http://www.etsi.org/deliver>. (date of access: 21.10.2019).
4. ETSI TS 302 775. Digital Video Broadcasting (DBV); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2). [Electronic resource]. URL: <http://www.etsi.org/deliver>. (date of access:21.10.2019).
5. ETSI EN 300 429. Framing structure, channel coding and modulation for cable systems [Electronic resource]. URL: <http://www.etsi.org/deliver>. (date of access:21.10.2019).

УДК 621.391

РАЗРАБОТКА ВЕБ-ПРИЛОЖЕНИЯ ДЛЯ ОБРАБОТКИ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Н.И. ЛИСТОПАД, О.Н. ВИНИЧУК

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

Поступила в редакцию 01 ноября 2019

Аннотация. Данная работа посвящена анализу существующих веб-приложений, выявлению их достоинств и недостатков, а также реализации веб-приложения для обработки цифровых изображений

Ключевые слова: Веб-приложения, векторное и растровое представления, многостраничные и одностраничные приложения.

Введение

На сегодняшний день достаточно сложно подобрать графический редактор для обработки цифровых изображений ввиду огромного количества как десктоп версий, так и веб-приложений, каждый из которых обладает рядом функциональных возможностей и, так или иначе, дополняет друг друга.

При создании html-страниц цифровые изображения подвергаются коррекции. Зачастую, материал для верстки html-страниц представлен не в psd-макете (исходный материал для верстки сайта), а в качестве скомпонованных по разделам файлов, которые и требуют последующей обработки.

Все компьютерные изображения, которые размещаются в качестве графического представления на html-странице, как известно, можно разделить на два формата – векторные и растровые, – различающиеся, прежде всего, уровнем абстракции, примененной к изображению [1]. При этом если векторная графика пытается имитировать восприятие изображений человеком, то растровый формат хранит графику в том виде, в каком она легче всего обрабатывается компьютером. Соответственно, векторная графика в большинстве своем создается пользователем с нуля используя векторные редакторы. Основным поставщиком растровых изображений – цифровое изображение, то есть в существенной своей части автоматический процесс с легко оцифровываемыми результатами.

Анализ методов и моделей обработки изображений

В последнее время список сервисов и ресурсов в Интернете достаточно расширился. Интернет превратился из однообразных статических html-страниц в мощный инструмент интерактивности и общения с конечными пользователями. В связи с этим веб-приложения в настоящее время приобрели небывалую популярность, ввиду предоставления различных преимуществ, которые отсутствуют в обычных десктоп-приложениях. Исходя из этого, была сформулирована гипотеза: веб-приложения можно выделять как отдельное звено из всей цепочки приложений. К сожалению, десктоп-приложения развиваются не так быстро и стремительно, как веб-технологии. Большое множество компаний переходят с обычных приложений на веб-приложения, именно потому, что видят в них будущее. На основании

вышеописанной характеристики десктоп-приложений и веб-приложений были выделены основные преимущества работы с веб-приложениями [2]:

1. Установка веб-приложений дешевле и намного проще. Для работы с веб-приложениями, как правило, не требуется установка дополнительного ПО, пользователю для полноценной работы достаточно стандартного браузера, который идет в пакете установленных программ любой операционной системы (далее ОС), а вот при работе с десктоп-приложениями обязательно требуется установка данного приложения и, возможно, для бесперебойной работы, дополнительного программного обеспечения (далее ПО, как правило, это дополнительные драйверы или расширения).

2. Обновление веб-приложений дешевле и намного проще. Для того чтобы совершить обновление веб-приложения, его необходимо обновить только на сервере и для того, чтобы работать с новой версией пользователю достаточно перезагрузить страницу. Обновление ПО десктоп-версии очень похоже на его установку, поэтому преимущества, которые были упомянуты выше имеют, место и в данной ситуации.

3. Веб-приложения более универсальны и практичны для конечного пользователя. Ввиду того, что веб-приложение хранится на сервере, доступ к нему можно получить с любого устройства, где подключен Интернет и установлен браузер, то есть можно сделать вывод, что доступ к веб-приложению возможен из любой ОС и с любого устройства. К сожалению, для реализации таких возможностей десктоп-приложений существует только один выход: разработка отдельных приложений под ОС и отдельных приложений для различных устройств, однако, в случае, если приложения будут иметь общий доступ к БД, которая должна храниться на сервере, необходимо будет наличие сети Интернет;

4. Веб-приложения облегчают организацию хранения данных. Если есть необходимость обращаться к одним и тем же данным из разных мест, то намного проще организовать их хранение в одном месте, как вариант, облачное хранилище данных (модель онлайн-хранилища, в котором данные хранятся на многочисленных распределенных в сети серверах, предоставляемых в пользование клиентам).

Чтобы обеспечить стабильную работу сложных веб-приложений, желательно использовать технологии, которые дадут наилучшую производительность и скорость [2]. На сегодняшний день существует два основных способа разработки веб-приложений:

- SPA (Single Page Applications) – одностраничные приложения;
- MPA (Multi Page Applications) – многостраничные приложения.

MPA – это традиционное многостраничное веб-приложение. При таком подходе каждый раз, когда приложение запрашивает данные либо отправляет их на сервер, оно вынуждено получать новую страницу в полном объеме, а затем отображать ее в браузере. При использовании такого способа разработки веб-приложений тратится время и снижается производительность на отображение одних и тех же элементов, что влияет на скорость и производительность в целом. Многие веб-разработчики для того, чтобы хоть как-то повысить скорость и уменьшить нагрузку на сервер используют библиотеку jQuery.

Для относительно простых веб-приложений такой способ вполне подходит, но ввиду того, что веб-приложение для обработки изображений большого объема обладает рядом функциональных возможностей, которые будут нагружать веб-приложение, то речь может идти о том, как ускорить сам процесс обработки данных. Все это ведет к тому, что размер трафика у MPA заметно увеличивается, и при многостраничном подходе неминуемо появляются проблемы с производительностью.

Для решения проблемы с производительностью был выбран способ SPA для создания веб-приложений. Архитектура у SPA устроена таким образом, что при переходе на новую страницу, обновляется только часть контента. Таким образом, нет необходимости повторно загружать одни и те же элементы. Это очень удобно не только для пользователей, но и для разработчиков.

Для разработки SPA используется один из самых популярных языков программирования – JavaScript, а также библиотека jQuery.

При создании html-страниц с использованием способа SPA можно выделить ряд достоинств по отношению к MPA:

- более быстрая загрузка веб-страниц;
- улучшенное восприятие пользовательского интерфейса, поскольку загрузка данных с сервера происходит в фоновом режиме;
- нет необходимости разрабатывать код на сервере для визуализации страницы;
- разделение на Front-end (клиентская сторона пользовательского интерфейса к программно-аппаратной части сервиса.) и Back-end (программно-аппаратная часть сервиса) – разработку;
- упрощенная разработка под мобильные приложения, а именно, возможность использования одного и того же серверного кода как для веб-приложений, так и для мобильных приложений.

Проведенный анализ показал, что SPA – это своеобразная эволюция способа MPA и технологии обмена данными в фоновом режиме AJAX. При таком подходе только структура веб-страницы строится на сервере, все остальное генерируется средствами JavaScript. SPA запрашивает html-разметку и данные отдельно и визуализирует результаты непосредственно в браузере.

Описанный выше подход представлен на рис. 1.

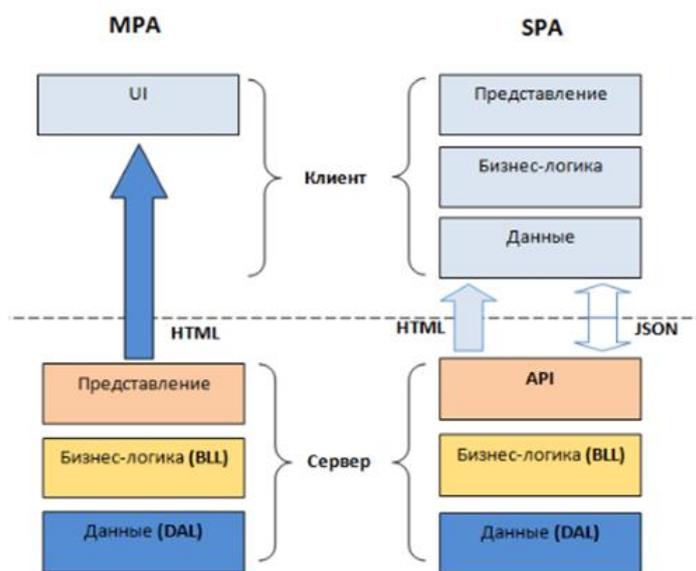


Рис. 1. MPA и SPA: способы загрузки и отображения данных

Заключение

Данная работа посвящена анализу существующих веб-приложений, выявлению их достоинств и недостатков, а также реализации веб-приложения для обработки цифровых изображений. Использование вышеописанных моделей и алгоритмов позволяет сократить объем памяти, необходимой для хранения изображений, времени обработки изображений, а также обеспечить качественное отображение изображений на веб-страницах.

Таким образом, предлагаемая одностраничная основа SPA может быть эффективно применена при создании веб-приложения для обработки изображений большого объема.

DEVELOPMENT OF A WEB APPLICATION FOR PROCESSING DIGITAL IMAGES

N.I. LISTOPAD, O.N. VINICHUK

Abstract. This work was devoted to the analysis of existing web applications, identifying their advantages and disadvantages, as well as the implementation of a web application for processing digital images.

Keywords: Web applications, vector and raster representations, multi-page (MPA) and single-page (SPA) applications.

Список литературы

1. Флэнаган Д. JavaScript. Подробное руководство. СПб: Символ-Плюс, 2008.
2. Прохоренок Н.А. HTML, JavaScript, PHP и MySQL. Джентльменский набор Web-мастера. СПб.: БХВ-Петербург, 2010.

УДК 621.391

ВОЗМОЖНОСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И МАШИННОГО ОБУЧЕНИЯ В ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЯХ

Д.А. БЕЛЯНКОВ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 31 октября 2019*

Аннотация. Телекоммуникационная отрасль всегда сталкивалась со сложными системами, стохастическими контекстами, комбинаторными проблемами и жестким прогнозированием пользователей. В статье представлены некоторые основные концепции машинного обучения и способ их применения в контексте телекоммуникационных сетей, в частности в оптических и беспроводных сетях.

Ключевые слова: Оптические сети, беспроводные сети, эталонный формат машинного обучения.

Введение

К 2021 году глобальный IP-трафик увеличится в три раза по сравнению со значениями 2016 года согласно Индексу визуальных сетевых технологий 2017 года от Cisco [1]. Прогнозируется, что 63 % этого трафика будет происходить с устройств, не относящихся к ПК, что показывает наибольший прирост трафика с планшетов, смартфонов и межмашинных устройств. Таким образом, к 2021 году трафик от беспроводных устройств будет намного выше, чем у проводных. Также значительно увеличится количество устройств, подключенных к Интернету через беспроводную сеть, не только потому, что будет увеличиваться количество смартфонов, планшетов и носимых устройств, но и из-за новых парадигм, таких как «Межмашинные коммуникации» (Machine-to-Machine – M2M) и «Интернет вещей» (Internet of Things – IoT), где все виды электронных устройств будут иметь возможности беспроводной связи. В отчете прогнозируется, что количество подключенных устройств превзойдет человеческое население в три раза. Также важным является тот факт, что большая часть этого приращения трафика будет происходить из-за увеличения видео высокого разрешения (HD), которое будет генерировать увеличение необходимых скоростей передачи данных. В этом контексте разрабатываются сети 5G, чтобы справиться с таким ростом спроса на трафик и скорость передачи данных. Чтобы удовлетворить все эти требования, необходимы новые технологии на всех уровнях сети, от беспроводного доступа к проводной базовой сети и от физических плоскостей и плоскостей данных до плоскостей управления и контроля. Например, срезы сети были недавно введены и тщательно изучаются как средство преодоления неоднородности обслуживания и требований сетей 5G. На этом же уровне такие методы, как виртуализация сетевых функций (Network Function Virtualization – NFV) [2] и программно-определяемые сети (Software Defined Networks – SDN) [3], являются ключевыми инструментами для развертывания 5G. С другой стороны, оптические сети в ядре и на краю сети предлагаются для удовлетворения потребностей с низкой задержкой и высокой скоростью передачи данных. Однако все эти новые технологии, развернутые на разных уровнях и в разных частях сети, должны взаимодействовать друг с другом, чтобы удовлетворить требования пользователей. Более того, требования пользователей будут сильно различаться – от видео приложений с низкой задержкой и высокой скоростью передачи данных до устройств IoT с очень низкими требованиями к производительности. В этом сценарии предполагается, что сети должны будут добавить интеллектуальность и автономность, чтобы адаптироваться к этой огромной

гетерогенности и соответствовать требованиям [4]. Сети с такими характеристиками называются самоорганизующимися сетями (Self-Organizing Networks – SON). В частности, SON – это сеть, которая автономно адаптируется к контексту, чтобы оптимизировать свои цели производительности и может извлекать уроки из среды, а также из своих предыдущих действий. В целом SON делятся на три категории: самоконфигурация, самооптимизация и самовосстановление [4]. Другим фундаментальным аспектом, является тот факт, что доходы операторов сокращаются, поскольку ключевыми игроками являются компании OTT (Over The Top), которые получают выгоду от трафика, генерируемого контентом и приложениями, в частности, WhatsApp, Facebook, YouTube и др.), без инвестиций в сетевую инфраструктуру. В этом сценарии ключевое значение имеет максимальное использование при одновременном реагировании на требования пользователей, и операции на основе искусственного интеллекта имеют решающее значение для достижения этой цели. Проектирование и реализация SON выполняются с помощью двух основных методов: оптимизации и машинного обучения (Machine learning – ML) [4].

Методы оптимизации традиционно использовались в телекоммуникационных сетях и показали отличные результаты в самых разных областях. Однако, в целом, чтобы применить некоторый оптимизационный подход, необходим важный предварительный шаг моделирования и упрощения реальности, который может потребовать значительных объемов работы и привести во многих случаях к неэффективным решениям. Таким образом, после недавнего ажиотажа в области искусственного интеллекта были предложены методы машинного обучения, позволяющие решать самые сложные проблемы, связанные с моделированием реальности. В этой статье выделены некоторые преимущества применения ML.

Применение машинного обучения в телекоммуникациях

Искусственный интеллект (Artificial intelligence – II) – это наука о создании интеллектуальных машин, способных автономно принимать решения на основе воспринимаемой ими среды. Машинное обучение (ML) – ветвь искусственного интеллекта. Как правило, она рассматривается как универсальный инструмент, используемый для задач классификации, определения подходящей категории для нового набора наблюдений и регрессионных задач, оценки взаимосвязи среди выборок данных. Фактически, это разнообразное поле, состоящее из различных компонентов и требующее: программную экосистему, включающую мониторинг и преобразование данных, выбор и оптимизацию моделей, оценку производительности, визуализацию и интеграцию моделей. Очевидно, что ML относится к вычислительному представлению явления, нацеленного на выполнение задачи, дающей определенную производительность и основанной на заданной среде. Подходы ML могут быть классифицированы на основе целей обучения, где они могут быть направлены на идентификацию шаблона для классификации, предсказания, обучение действию или методы индуктивного обучения. Алгоритмы ML могут быть далее классифицированы по трем различным семействам обучения, то есть контролируемому обучению, неконтролируемому обучению и обучению с подкреплением. Полуконтролируемое обучение, или гибридное обучение, иногда считается четвертой ветвью, заимствуя функции из контролируемых и неконтролируемых категорий.

Контролируемое обучение (Supervised learning – SL) [5] использует известные выходные функции, называемые метками, чтобы получить вычислительную связь между входными и выходными данными. Алгоритм итеративно создает модель ML, обновляя ее весовые коэффициенты на основе сопоставления набора входных данных с соответствующими им выходными характеристиками. Кроме того, SL может быть разделено на задачи классификации и регрессии в зависимости от того, используются дискретные или непрерывные выходные характеристики.

Неконтролируемое обучение (Unsupervised learning – USL). В то время как SL обеспечивает чистый подход к построению модели ML, на практике помеченные данные не всегда доступны или доступны в избытке. USL направлено на построение представления заданного набора данных без механизмов обратной связи, управляемых какими-либо метками.

USL может быть далее классифицировано как кластеризация данных в сходные группы или выявление правил ассоциации, идентифицирующих взаимосвязи между функциями.

Обучение с подкреплением (Reinforcement learning – RL) относится к механизмам ML без явной фазы обучения. RL стремится создать и обновить модель ML на основе взаимодействия агента с его собственной средой. Ключевым отличием в отношении SL метода является то, что помеченные функции ввода–вывода не предоставляются, но взаимосвязь скорее изучается посредством применения начальной модели для тестирования данных. Эти методы должны позволять выводить из данных, полученных различными типами мониторов (например, качество сигнала, выборки трафика и т.д.) полезные характеристики, измерение которых затруднено. Некоторые предполагаемые приложения в сети могут включать в себя прогнозирование сбоев, обнаружение вторжений, безопасность, маршрутизацию, низкопрофильное проектирование, реконфигурацию пропускной способности с учетом трафика и многие другие.

Варианты использования машинного обучения

Инициализация в оптических сетях.

Эластичная оптическая сеть (Elastic optical networking – EON) использует динамическое и адаптивное обеспечение световых путей для эффективного и гибкого использования сетевых ресурсов. Однако такая гибкость и динамичность требуют схем автоматического обеспечения с надежной оценкой качества передачи (Quality of Transmission – QoT), основанных на реальных условиях сетевых соединений.

Когнитивная многодоменная архитектура эластичной оптической сети, схематическая структура которой показана на рис. 1, работает в соответствии с циклом наблюдения–анализа–действия. Плоскость данных каждого домена оптической сети содержит переключатели с изменяемой полосой пропускания (bandwidth-variable transponders – BVT) и переключатели с полосой пропускания с переменной длиной волны (bandwidth-variable wavelength selectiveswitches – BV-WSSs). Элементы контроля производительности (Performance Monitoring – PM) развернуты в определенных местах для сбора состояния сети в реальном времени (наблюдение). Менеджеры доменов (Domain Managers – DM) применяют парадигму программно-определяемых сетей (software-defined networking – SDN) для управления своими доменами оптической сети [6]. В частности, контроллеры SDN, находящиеся в плоскости контроля и управления доменом, могут собирать данные и осуществлять предоставление услуг, взаимодействуя с агентами программно-определяемых сетей, которые размещаются совместно с оптическими устройствами. На основании технических показателей и полученных PM, DM могут выполнять когнитивное предоставление услуг на уровне домена.

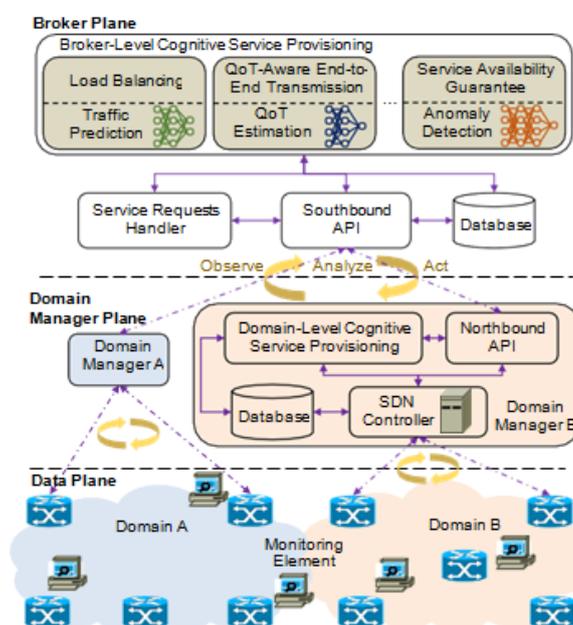


Рис. 1. Схема многодоменной оптической сети

На рис. 2 представлен принцип действия предложенной когнитивной многодоменной оптической сети. Каждый DM постоянно проводит РМ как для внутренних, так и для чужих длин волн. Затем выполняется процесс обучения на уровне домена (этап а). Полученные результаты могут быть использованы либо для внутриобластной инициализации (этап b), либо отправлены в посредническую плоскость для изучения характеристик междоменных сквозных световых трактов (этапы с – d). На этапе инициализации, после приема запроса световода (этап 1), DM сначала определяет, является ли это передачей внутри или между доменами. Если это внутренний запрос, DM немедленно пытается рассчитать решение маршрутизации, формата модуляции и назначения спектра (Routing, Modulation format, and Spectrum Assignment – RMSA), которое удовлетворяет требованию качества передачи. В частности, DM может вычислить несколько путей кандидата, используя RMSA. Затем DM вызывает модель оценки качества передачи на уровне домена для оценки качества передачи каждого возможного решения пути, предоставленного RMSA. DM выберет решение RMSA, которое удовлетворяет требованию качества передачи и обеспечивает качество передачи существующих соединений. С другой стороны, если запрос направлен на междоменную связь, запрос перенаправляется на уровень посредника (этап 2). Для расчета междоменного светового пути плоскость посредника запрашивает внутридоменную информацию от соответствующих DM (этап 3). Каждый вовлеченный DM, в свою очередь, отвечает несколькими внутренними потенциальными световыми путями вместе с их оцененным качеством передачи (этапы 4 и 5). Плоскость посредника агрегирует полученные данные, строит многодоменную виртуальную топологию и вычисляет междоменную сквозную RMSA с гарантией качества, опираясь на модель оценки качества передачи на уровне брокера (шаг 6, с подходом иерархического обучения). Наконец, каждый связанный DM устанавливает сегмент тракта в своем домене для обеспечения услуги междоменной инициализации (этап 8).

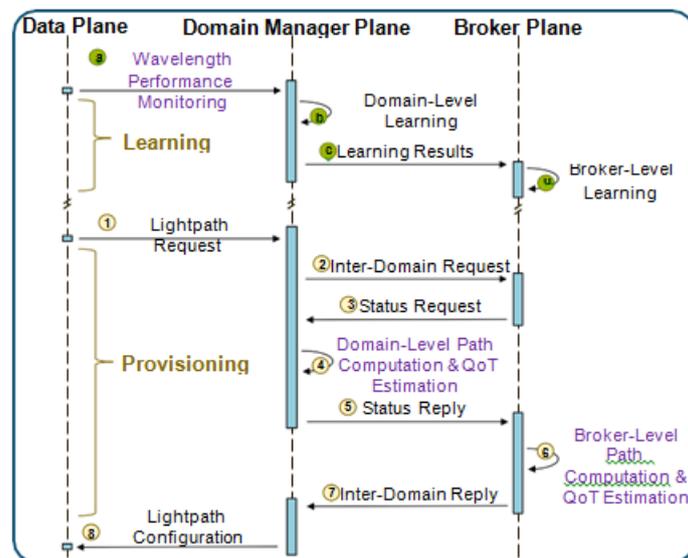


Рис. 2. Принцип действия многодоменной оптической сети с циклом обучения «Наблюдение–анализ–действие» и плоскостью брокера, взаимодействующей с несколькими управляющими доменами

Устойчивость в оптических сетях.

Локализация отказов – очень полезный метод, поскольку он помогает значительно сократить время восстановления после отказа. Когда в оптическом слое происходит серьезный сбой, измененный трафик должен быть немедленно повторно сохранен с использованием доступных в настоящее время ресурсов [7, 8]. Тем не менее, некоторые серьезные сбои начинаются как мягкие сбои, и они могут быть обнаружены как начальная деградация. Следовательно, желательно предвидеть и устранять серьезные сбои, чтобы планировать правильные действия по изменению маршрута. Несмотря на то, что эволюция некоторых мягких отказов может занять длительное время, они могут влиять на качество световых путей

в нерабочем состоянии. Для выявления мягких отказов, затрагивающих отдельные световые пути, такие как дрейф лазера, смещение фильтра или сжатый свет требуются современные встроенные методы мониторинга для анализа и оценки качества отдельных оптических световых траекторий. В связи с этим, хотя оптико-спектральные анализаторы (Optical Spectrum Analyzers – OSA) можно было использовать для анализа спектра оптических сигналов, до недавнего времени использование OSA в сети было очень ограниченным из-за высокой стоимости точных OSA. Тем не менее, технологии OSA совершенствуются, и в настоящее время новое поколение экономически эффективных анализаторов с разрешением ниже 1 ГГц недоступно для интеграции в оптические узлы нового поколения. Кроме того, OSA и другие системы мониторинга требуют сложных алгоритмов машинного обучения, способных выявлять и локализовать сбои. Эти алгоритмы могут быть развернуты в сетевом контроллере, а также в агентах узлов, близких к точкам мониторинга, чтобы уменьшить объем данных мониторинга, которые должны быть переданы на панель управления. После обнаружения причину сбоя необходимо локализовать с помощью методов пассивного мониторинга, чтобы облегчить перенаправление светового пути. На рис. 3 показано использование OSA для локализации мягких отказов, когда работает световой путь. OSA получают весь спектр C – диапазона, а затем извлекаются данные для части спектра, выделенной исследуемому световому тракту. Пассивный мониторинг OSA осуществляется во входном и каждом промежуточном узле (но не в выходном). Два модуля, работающие в агенте узла, отвечают за анализ спектра: 1) модуль «Извлечение объектов» (Feature Extraction – FeX) – представляет набор соответствующих точек в спектре сигнала, которые используются для вычисления значимых характеристик сигнала; 2) модуль проверки спектра сигнала (Signal Spectrum Verification – SSV), который нацелен на анализ извлеченных признаков для обнаружения неправильной конфигурации, то есть проблем смещения центральной частоты и фильтрации. Алгоритм ML, работающий в сетевом контроллере, отвечает за управление модулями в узлах и получение диагностики, а также соответствующих сигналов от них, чтобы локализовать отказ и оценить его величину. Стоит отметить, что алгоритм ML должен уметь различать фактические сбои и нормальные эффекты, которые могут привести к аналогичным свидетельствам: специфически непроницаемым эффектам фильтрации из-за каскадирования фильтра нормального сигнала. Алгоритм ML использует преимущества модуля сравнения спектра сигналов (Signal Spectrum Comparison – SSC), который генерирует диагностику одного сигнала, фокусирующегося конкретно на проблемах фильтрации. Два полезных алгоритма ML для обнаружения отказов, идентификации и локализации – это классификация и регрессия. При классификации цель состоит в том, чтобы классифицировать неизвестные принятые данные, например, оптический сигнал, и решить, принадлежит ли сигнал нормальному классу, классу FS или классу FT. В регрессии цель состоит в том, чтобы предсказать поведение; например, регрессия может использоваться для оценки величины отказа. Несмотря на то, что несколько алгоритмов ML подходят для одного и того же вывода, выбор лучшего является решением, зависящим от проблемы, где их производительность должна быть изучена для конкретного случая.

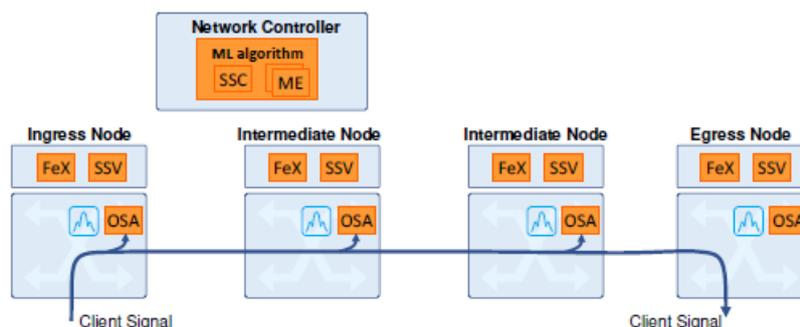


Рис. 3. Пассивный мониторинг OSA для локализации сбоев в работе

Разрезание беспроводных сетей.

В беспроводных сетях 5G требуется учитывать растущую неоднородность предлагаемых услуг, требований к производительности и технологий связи. В последнее время концепции

разделения сети появились в качестве метода для комплексного управления этими сетями. Разрезание сети состоит в разделении сетевой инфраструктуры на изолированные сегменты сети с собственным набором ресурсов и, возможно, с конкретными требованиями к производительности. Основными заинтересованными сторонами в этой парадигме являются: арендатор, который запрашивает срез и задающий его параметры производительности, и поставщик услуг, который контролирует инфраструктуру и передает ресурсы арендатору. Нарезка классифицируется следующим образом: срезы качества обслуживания (Quality-of-Service Slicing – QoSS), где срезы должны обеспечивать требуемое качество обслуживания, независимо от необходимых ресурсов, и срезы с разделением инфраструктуры (Infrastructure-Sharing Slicing – ISS), где каждому срезу выделяется набор ресурсов для его собственного использования, и он изолируется от остальной части сети. В варианте QoSs поставщик срезов должен удовлетворять запрошенным целям производительности для каждого потока в срезе, в то время как в варианте ISS арендатор требует, чтобы набор ресурсов был зарезервирован для среза. Реализация этих вариантов нарезки в беспроводных сетях представляет несколько проблем и существенно отличается от проводных сетей. В области проводной связи, принято управлять полосой пропускания канала в качестве выделяемого источника. Однако в беспроводных сетях полоса пропускания беспроводной линии связи не фиксирована, поскольку она зависит от характеристик канала. Пропускная способность беспроводной линии связи является переменной и зависит от нескольких факторов, таких как расстояние между узлами, физические препятствия, которые должен пройти сигнал через другие передачи, использующие тот же канал и помехи. Следовательно, в беспроводных сетях ресурсы состоят из радиоспектра (разделенного по времени, частоте или пространству) и времени передачи (эфирное время). В этом сценарии любой механизм, который предлагает управлять и распределять ресурсы, чтобы реализовать нарезку в беспроводных устройствах потребует знания характеристик канала, вероятно, с постоянным мониторингом и с оценкой того, как канал будет вести себя в будущем. Поскольку обстоятельства, влияющие на производительность канала, разнообразны и широки, в общем случае трудно правильно моделировать беспроводной канал. Как следствие, методы машинного обучения являются очень хорошими кандидатами для использования в реализации механизмов предоставления ресурсов для разрезания беспроводных сетей.

Заключение

В этой статье представлены основы машинного обучения и набор вариантов его использования. Рассмотрены две основные технологии физического уровня – оптическая и беспроводная. В случае оптических сетей представлены две наиболее сложные проблемы: обеспечение световых путей надежными требованиями QoT. Для беспроводных сетей исследован другой, но связанный, тип оптимизации: способ оптимизации, при котором сеть может быть «нарезана» или совместно использоваться для сложных аналитических целей. Во всех этих случаях предложено использовать методы машинного обучения, чтобы раскрыть модели воспроизведения, которые были бы слишком сложными, чтобы их можно было понять или сформулировать аналитически. Именно в этом и заключается преимущество машинного обучения. Периодическое человеческое взаимодействие с телекоммуникационными сетями и услугами всех видов содержит скрытые информационные камни, которые при правильной интерпретации позволяют иметь дело с физическими взаимодействиями, которые, на первый взгляд, недоступны. Кроме того, технологии машинного обучения могут быть использованы для выявления взаимодействия этих процессов с использованием пользователей. Промышленность в настоящее время применяет их в широком спектре приложений. Таким образом, машинное обучение позволяет вычислительным системам получать знания из опыта. Эта возможность определяется моделями и правилами, автоматически извлекаемыми из больших объемов данных, а также из постоянного наблюдения за окружающей средой, составления выводов и автоматизации процесса. Тем не менее, следует помнить, что технологии машинного обучения – это ничто иное, как набор алгоритмов, которые можно применить для конкретной функции или модели.

POSSIBILITIES OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN TELECOMMUNICATION NETWORKS

D.A. BELIANKOU

Abstract. The telecommunications industry has always faced complex systems, stochastic contexts, combinatorial problems and strict user predictions. Some basic machine learning concepts and the way they are applied in the context of telecommunications networks, particularly in optical and wireless networks, were presented.

Keywords: Optical networks, wireless networks, reference format of machine learning.

Список литературы

1. Cisco Visual Networking Index: Forecast and Methodology, 2016–2021. White Paper. Cisco Systems. 2017.
2. Mijumbi R., [et. al.] // Network Function Virtualization: State-of-the-Art and Research Challenges. 2016. P. 236–262.
3. F. Hu, Q. Hao, K. Bao. // IEEE Communications Surveys Tutorials. 2014. P. 2181–2206.
4. Aliu O.G., [et. al.] // A Survey of Self Organisation in Future Cellular Networks. 2013. P. 336–361.
5. Kotsiantis S. B., Zaharakis I., Pintelas P. // Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering. 2007. P. 3–24.
6. Castro A., [et. al.] // Journal of Lightwave Technology. 2016.
7. Kozat U.C., [et. al.] // IEEE/ACM Transactions on Networking. 2016. P. 2899–2912.
8. Ogino N., Yokota H. // Journal of Lightwave Technology. 2014. P. 467–482.

СВЕДЕНИЯ ОБ АВТОРАХ

1. Аксенов Вячеслав Анатольевич – старший преподаватель кафедры инфокоммуникационных технологий БГУИР
2. Алексеенко Александр Эдурдович – магистрант кафедры инфокоммуникационных технологий БГУИР
3. Белянков Дмитрий Андреевич – магистрант кафедры инфокоммуникационных технологий БГУИР
4. Бобов Михаил Никитич – д.т.н., профессор кафедры инфокоммуникационных технологий БГУИР
5. Виничук Ольга Николаевна – аспирантка кафедры информационных радиотехнологий БГУИР
6. Вишняков Владимир Анатольевич – д.т.н., профессор кафедры инфокоммуникационных технологий БГУИР
7. Воронов Александр Анатольевич – к.т.н., доцент, старший научный сотрудник лаборатории идентификации систем ОИПИ НАН Беларуси
8. Ганченко Валентин Валентинович – к.т.н., старший научный сотрудник лаборатории идентификации систем ОИПИ НАН Беларуси
9. Гондаг Саз Мостафа М. – магистр т.н., стажер кафедры коммуникационных технологий БГУИР
10. Доан Тхе Хоанг – аспирант кафедры инфокоммуникационных технологий БГУИР
11. Дудкин Александр Арсентьевич – д.т.н., профессор, заведующий лабораторией идентификации систем ОИПИ НАН Беларуси

12. Жэнь Сюньхуань – аспирант кафедры инфокоммуникационных технологий БГУИР
13. Инютин Александр Владимирович – научный сотрудник лаборатории идентификации систем ОИПИ НАН Беларуси
14. Калоша Андрей Леонидович – магистрант кафедры информатики БГУИР
15. Конопелько Валерий Константинович – д.т.н., профессор кафедры инфокоммуникационных технологий БГУИР
16. Курилович Андрей Владимирович – старший преподаватель кафедры инфокоммуникационных технологий БГУИР
17. Липкович Эдуард Борисович – доцент кафедры инфокоммуникационных технологий БГУИР
18. Листопад Николай Измаилович – заведующий кафедрой информационных радиотехнологий БГУИР, д.т.н., профессор
19. Лукашевич Сергей Александрович – старший преподаватель кафедры инфокоммуникационных технологий БГУИР
20. Ма Цзюнь – аспирант кафедры инфокоммуникационных технологий БГУИР
21. Марушко Евгений Евгеньевич – научный сотрудник лаборатории идентификации систем ОИПИ НАН Беларуси
22. Мохаммад Г. Моздурани Шираз – стажер кафедры инфокоммуникационных технологий БГУИР
23. Наливайко Сергей Иванович – магистрант кафедры инфокоммуникационных технологий БГУИР

24. Нгуен Ань Туан – аспирант кафедры инфокоммуникационных технологий БГУИР
25. Новицкий Виталий Владимирович – аспирант кафедры инфокоммуникационных технологий БГУИР
26. Павловский Александр Александрович – магистрант кафедры инфокоммуникационных технологий БГУИР
27. Полуян Татьяна Владимировна – магистрант кафедры инфокоммуникационных технологий БГУИР
28. Рабцевич Виолетта Викторовна – ассистент кафедры инфокоммуникационных технологий БГУИР
29. Садик Б.Д.С. – стажер кафедры инфокоммуникационных технологий БГУИР
30. Саломатин Сергей Борисович – к.т.н., доцент кафедры инфокоммуникационных технологий БГУИР
31. Саскевич Алексей Владимирович – аспирант кафедры информатики БГУИР
32. Сергеев Николай Николаевич – аспирант кафедры инфокоммуникационных технологий БГУИР
33. Серченя Анна Андреевна – магистрант кафедры инфокоммуникационных технологий БГУИР
34. Смоляк Сергей Владимирович – руководитель группы оптимизации радио и опорной сети УП «А1»
35. Стержанов Максим Валерьевич – к.т.н., доцент кафедры информатики БГУИР

36. Урядов Владимир Николаевич – д.т.н., доцент кафедры инфокоммуникационных технологий БГУИР
37. Урядов Владимир Николаевич – к.т.н., доцент, доцент кафедры инфокоммуникационных технологий БГУИР
38. Филиппов Николай Сергеевич – магистрант кафедры информационных радиотехнологий БГУИР
39. Хоменок Михаил Юлианович – доцент кафедры инфокоммуникационных технологий БГУИР, к.т.н.
40. Цветков Виктор Юрьевич – д.т.н., заведующий кафедрой инфокоммуникационных технологий БГУИР
41. Шайа Бахаа Хикмат – аспирант кафедры инфокоммуникационных технологий БГУИР