PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Neural eye-processing computer based on FPGA technologies

Kozhemyako, Andriy, Bezkrevnyi, Oleksandr

Andriy V. Kozhemyako, Oleksandr S. Bezkrevnyi, "Neural eye-processing computer based on FPGA technologies," Proc. SPIE 11137, Applications of Digital Image Processing XLII, 111371Z (6 September 2019); doi: 10.1117/12.2527879

Event: SPIE Optical Engineering + Applications, 2019, San Diego, California, United States

Neural eye-processing computer based on FPGA technologies

Andriy V. Kozhemyako[a], Oleksandr S. Bezkrevnyi*[a],
[a]Vinnytsia National Technical University, 95 Khmelnytsky highway, Vinnytsia, Ukraine 21021

ABSTRACT

One of the promising areas in the field of image processing and analysis is the hardware implementation of neural network for processing and analyzing of images based on FPGA technologies. The structural scheme of the multifunctional calculator was developed. It contains MxN cells in the form of matrix, a formation block of signs that has N control nodes, input of clock impulses, input of reset, output of common feature of zero and outputs of features of zero by columns and rows of matrix. The described structure is a classifier and it was written to the FPGA crystal[1].

The main task is to provide processing and analysis of images in real time. Hardware implementation allows to use it in many types of activities: biomedical engineering, industry, aerospace sphere, etc[2].

Keywords: image processing, FPGA, neural network, eye-processing, technical vision, parallel processing.

1. INTRODUCTION

An important task of image recognition is its qualitative description, that is, the creation of such an image model that would effectively represent the recognition object through its characteristic features and unambiguously match this representation with a set of image classes[2]. The quality of the image description directly affects both the recognition algorithm and its specific hardware implementation. Color, texture, brightness of the objects can be used as the characteristic classification features in the recognition tasks[1,2].

However, a special interest is the description of images as geometric outlined by contour models. The use of informative features, related specifically to the geometry of the image contour, is explained by their simplicity, high stability and productivity. The transition to processing the image contour reduces by several orders of magnitude the information size that comes in real time. In addition, the contours are invariant to the transformations of the image brightness. Outline signs stays one of the most informative characteristics of the depicted object when it is perceived by a person, which is a decisive factor in the construction of modern environment of image recognition[4].

The most perfect natural prototype of technical vision is the human eye. The action of the human eye is related to brain activity. One of the priority tasks is to create a theoretical justification for non-traditional methods of processing information that can work not on the usual principles, but approaches to the human brain processing form. The problem is to identify the images. The goal is to develop an optimal system of technical vision.

2. EYE-PROCESSOR COMPUTERS

In work [5] the definition of the eye-processor sounds like "a technical system, presented in the form of a visual environment of arbitrary figuration, identifies certain features of the environment, processes the selected features and makes decisions automatically or with the participation of the operator". Considering the process of intellectualizing the functionality of the eye-processor in the context of optoelectronic logic-time information computing environments, its definition has been changed. Now the eye-processor is seen as "information intellectual system that simulates the image of the world on the basis of visual information perception of arbitrary figuration, identifies certain properties and features of the environment, processes and makes corresponding decisions automatically or with the participation of the operator"[6]. The peculiarity of such an eye-processor is the ability of intelligent decision-making in processing, analysis and recognition of image[7-10].

alexvntu@gmail.com;

Applications of Digital Image Processing XLII, edited by Andrew G. Tescher, Touradj Ebrahimi, Proc. of SPIE
Vol. 11137, 111371Z · © 2019 SPIE · CCC code: 0277-786X/19/$21 · doi: 10.1117/12.2527879

The main functions of the eye-processor as an intellectual system are:

a) recognition of a certain object in a given class of images according to the standard or a priori features allocated during the analysis of images;

b) automatic tracking of the specified object parameters in the images in the conditions of its evolution, that is, changes in position, size, brightness, momentary features, etc.[3, 8-11].

In addition, the eye-processor performs a number of auxiliary functions that are related to image rendering, namely, zoom, shift, rotation, pre filtration, definition of the center and contours of the image, performing some logic operations[3,5,9-11].

Figure 1 shows the scheme of classification models for the creation of a figurative computer prototype based on an eye-processor at 12 levels, namely: problem statement - level 1, computing character - level 2, type of processing - level 3, type of transformation - level 4, method of processing – level 5, form of data submission - level 6, coding method - level 7, algorithms - level 8, basic information technologies - level 9, architectural solution - level 10, base structure - level 11, elemental basis - level 12.

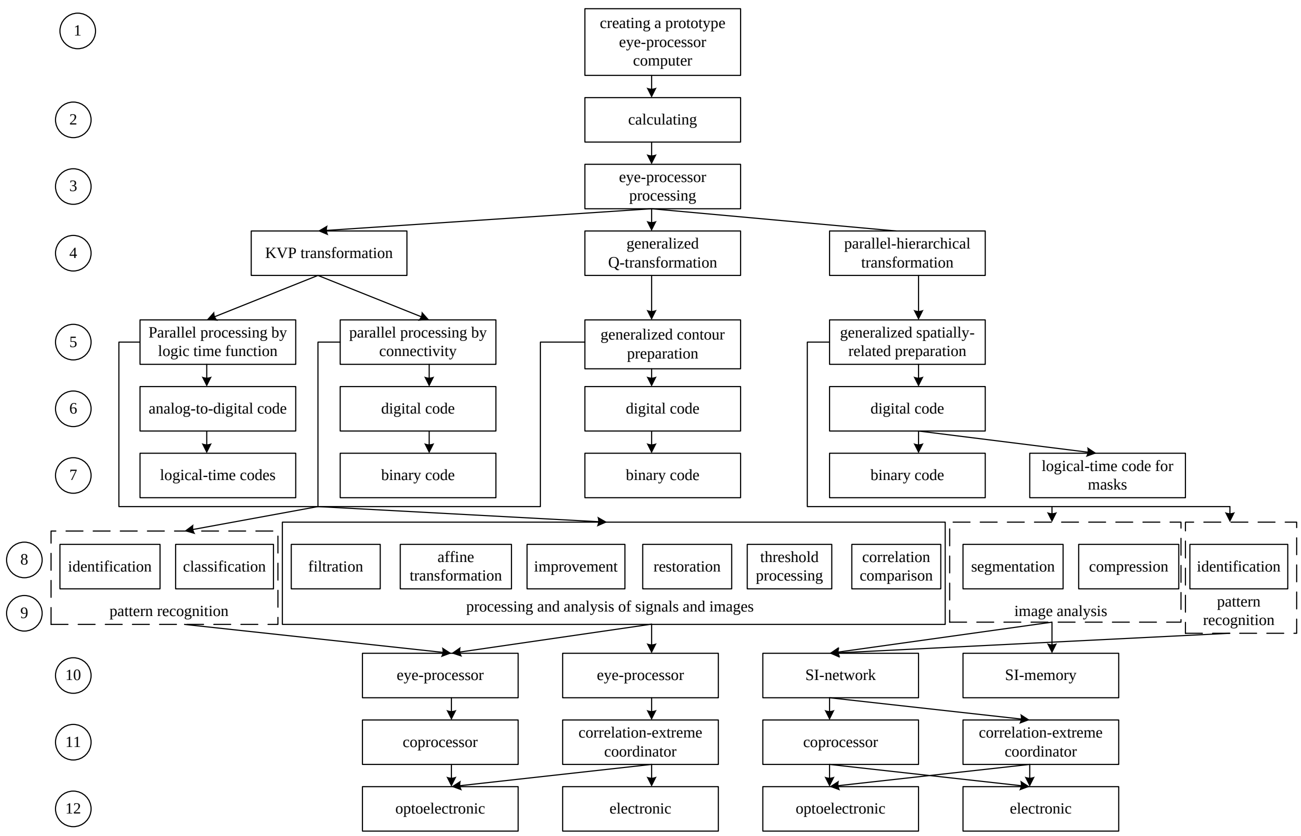

Figure 1. Classification of prototype figurative computer models.

Eye-processor is selected as a type of processing in the given classification, due to these publications [3, 5, 8-11] its basic principles are:

a) parallel processing of all elements of the input data array, based on KVP transformation;

b) the iterative nature of image processing;

c) the selection of a set of features in the pre-processing of images, combining pre-processing with the primary analysis of the input data array;

d) basic operations - extracting the general component of all elements of the data array and hierarchical (parallel) summing (accumulation) of the current general components for each iteration;

e) formation of an evolution database for the recognition of input images (objects);

f) hierarchical organization of the process of analysis and image recognition.

The most adequate eye-computer processing is realized by a device whose structure is shown in Fig. 2[4,12] and contains the following nodes: functional integrated l_{ij} synthesizers, information analyzer 2, hierarchical structure 3 as a block of synthesizers-generators of features, block of formation knowledge base and the choice of reference 4, comparison scheme 5. The device can be used in systems of technical vision for recognition of visual images in works for modern systems of search, observation, guidance, biomedical diagnostics and technological control. The technical result of this approach is to simplify the procedure of recognizing images with the possibility of organizing an evolving knowledge base, increasing the speed and authenticity of recognition[13,14].

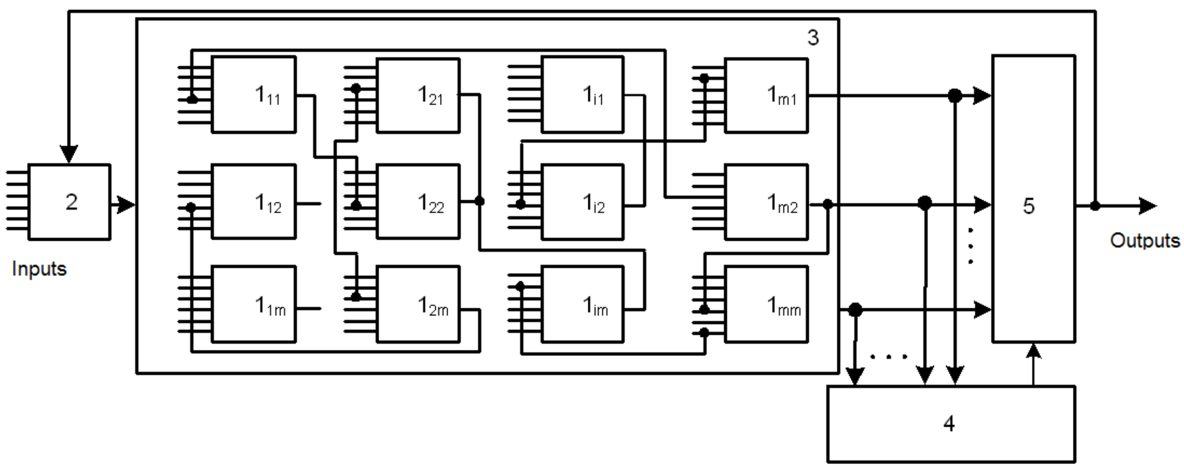

Figure 2. Structure of the eye-computer processing device.

3. FPGA CLASSIFICATION FOR THE IMPLEMENTATION OF AN EYE-COMPUTER PROCESSING DEVICE

When an elemental base of signal processing systems are selected, the following selection criteria are usually followed: speed; logical capacity sufficient for implementation of the algorithm; circuit design and structural parameters of the FPGA, reliability, operating temperature range, resistance to ionizing radiation; the cost of ownership of development tools, which includes both the cost of software, the availability and cost of hardware debugging; the cost of equipment for programming FPGA or configuration ROM and cost of chips; availability of methodical and technical support.

The products of the company are reviewed Altera, Table 1 shows the main parameters of FPGA

After considering and comparing the characteristics of FPGA data of leading manufacturers, it was decided to select the Arria family of company Altera, namely the EP2AGX260FF35I5 Arria II GX family, and to implement the image processing neural network based on this FPGA.

As one of the main criteria for selecting a FPGA was the agreement with the camera interface and compliance with certain requirements, namely: the number of registers, the amount of memory and the cost of the circuit, Arria II GX 2008 circuit is suitable for designing and implementation of the project, due to compliance, despite the relatively long year of issue.

Table 1. The main characteristics of the FPGA company Altera.

	Year of manufacture	Process technology, nm	Held in memory, KB	Number of registers	Number of logical elements	Number of outputs
Arria II GZ	2010	40	11115	179200	224	878
Arria II GX	2009	40	8550	205200	244	692
Statrix V	2010	28	13312	356000	236	664
Statrix IV	2008	40	14283	182400	228	584
Cyclone IV	2009	60	6480	92400	150	508
Cyclone III	2007	65	3888	85600	119	347

The main parameters of the FPGA of the firm Xilinx are considered, and are presented below in Table 2[14].

Table 2. The main characteristics of the FPGA company Xilinx.

	Year of manufacture	Process technology, nm	Held in memory, KB	Number of registers	Number of logical elements	Number of outputs
Spartan 7	2009	28	1100	102400	240	400
Artix 7	2010	28	2888	215360	210	500
Kintex 7	2010	28	6788	356160	325	500
Virtex 7	2011	28	13275	305400	350	600

A camera that works through the Camera Link interface shows high performance. Since the FPGA has the ability to directly connect to the Camera Link interface, the best choice will be implemented on CMOS technologies - the camera Q-12A65 from the series QUARTZ, the company Adimec.

Table 3. The main characteristics of the camera's company Adimec, QUARTZ series.

	Q-2A340	Q-4A150	Q-4A180	Q-12A-65
Max resolution	2048x1088	2048x2048	2048x2048	4096x3072
Max sustained speed at full resolution	340 fps (2000x1088 CL 10 tap)	156 fps (2048x2048 CL 8 tap)	180 fps (2000x2048 CL 10 tap)	65 fps (4080x3072 CL 10 tap)
Max burst speed at full resolution	340 fps (2000x1088 CL 10 tap)	180 fps (2000x2048 CL 10 tap)	180 fps (2000x2048 CL 10 tap)	75 fps (4080x3072 CL 10 tap)

4. PARALLEL AND SEQUENTIAL PROCESSING

Figure 3 shows a structural scheme of a calculator that contains MxN cells in the matrix form, a shaping block of features that has N control nodes, input of clock impulses, reset input, output of the common feature of zero and outputs of the zero features by columns and rows of matrix[15].

Figure 4 shows the functional scheme of three cells (first, second, and n-th) of the i-th row of the calculator in the form of a homogeneous structure with parallel data recording [15]. Each cell $1_{i1},\ldots,1_{in}$ of the i-th row of a homogeneous structure (Fig. 4) contains a counter, two multiplexers, an element NO, element AND, a configuration block. The configuration block for each cell 1_{ij}, except for the first and the last one, consists of element AND, NO, OR and element EQUALITY. The first cell 1_{i1} in each i-th row of a homogeneous structure contains a counter, a multiplexer, elements NO, AND, and the configuration block which contains of two elements AND and the element EQUALITY. The last cell 1_{i1} in each i-th row of a homogeneous structure contains a counter, multiplexer, element AND, and configuration block contains of element AND and the element EQUALITY[15].

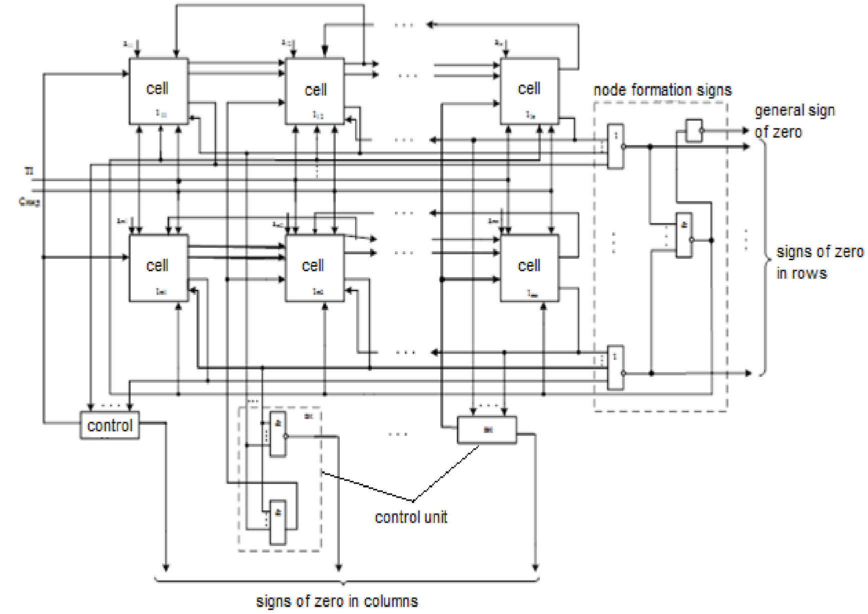

Figure 3. The block diagram of the calculator.

The function of parallel recording is realized in the counters of all cells of the calculator for the input data in the form of weighted input signals as elements of the matrix. Performing the function of simultaneous removal of a minimal element in the columns of the calculator activates the inputs of the reciprocal number in the counters of all cells, and when performing the transposition function in each line, zero elements move to the right with the simultaneous exchange of data in the neighboring elements[14,15].

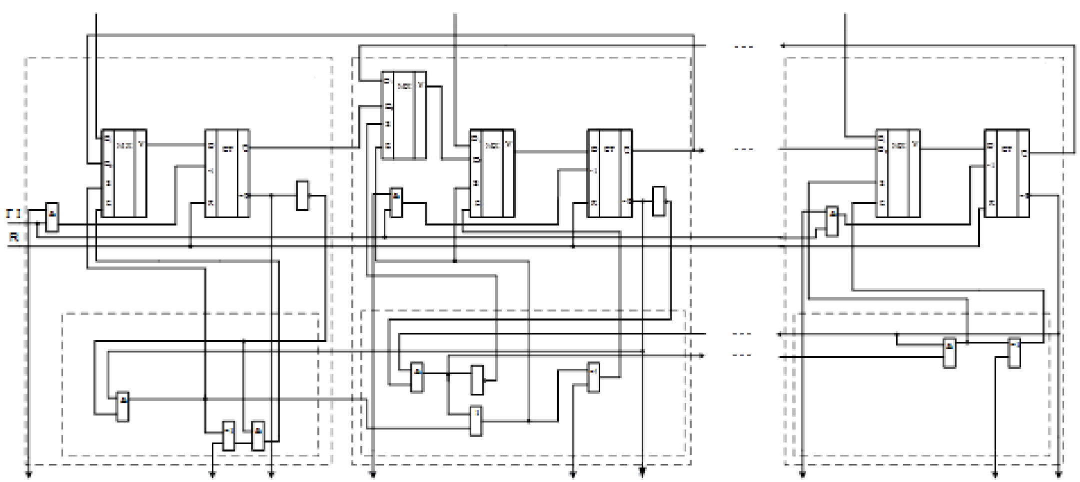

Figure 4. Functional scheme of a row of cells of a homogeneous structure with a parallel data recording.

Another variant is a homogeneous structure with sequential data recording (Fig. 5).

The cell 1_{ij} in i-th row of such a homogeneous structure (Figure 5), except the first and the last one, contains a counter, a multiplexer, an element NO, element AND, a configuration block. The configuration block for each cell 1_{ij}, except for the first and the last one, consists of elements AND, OR, NO. The last cell 1_{in} in each i-th row of a homogeneous structure contains a counter, a multiplexer, an element AND, and the configuration block contains element AND and the element EQUALITY. The first cell 1_{i1} in each i-th row of a homogeneous structure contains a counter, elements NO, AND, a switch, and the configuration block contains an element AND.

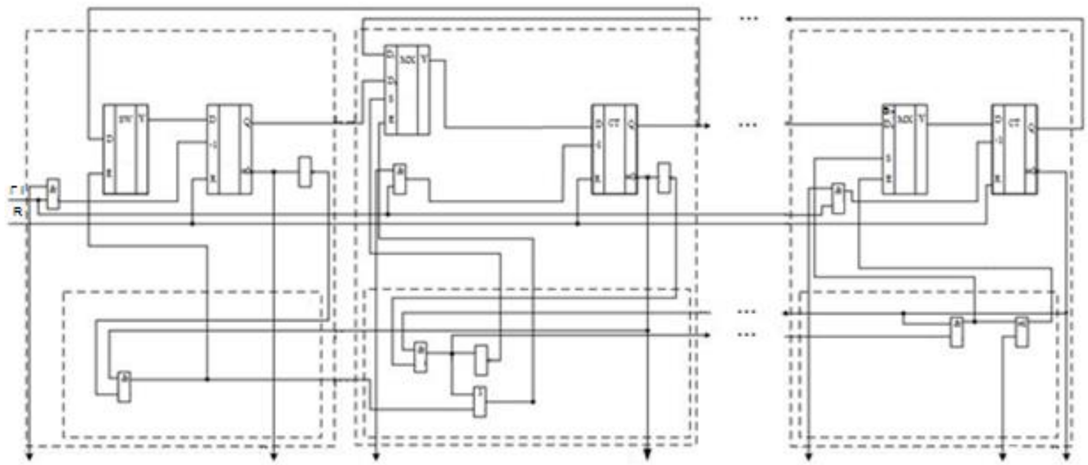

Figure 5. Functional scheme of a row of cells of a homogeneous structure with sequential data recording.

The analysis of the schemes in Figures 4 and 5 shows that in the second variant (Fig. 5) the hardware complexity of the cells decreases in comparison with the first variant (Fig. 4), but the time of writing the matrix in such a homogeneous structure has a dependence O(*n*).

5. IMPLEMENTATION OF A NEUROCOMPUTER CELL FOR IMAGE PROCESSING ON A FPGA

Fig. 6 shows a simulated computational scheme, the structural diagram of which is shown in Fig. 3 in the previous section.

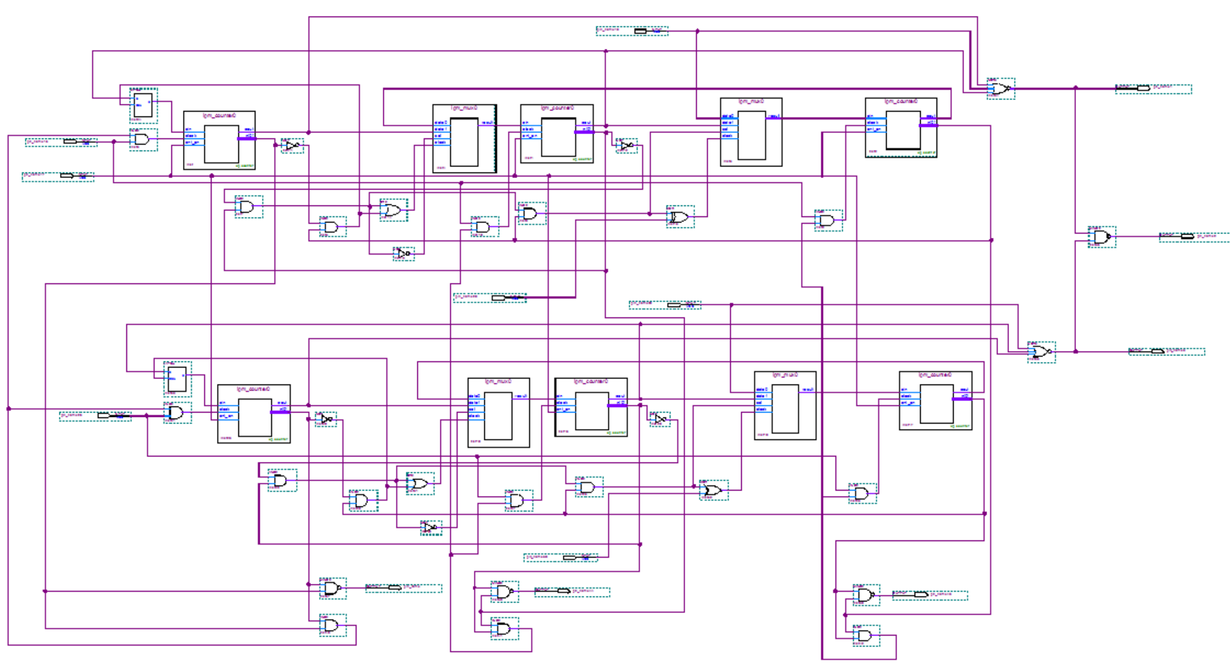

Figure 6. Scheme of the calculator.

6. MATHEMATICAL METHODS OF IMAGE PROCESSING FOR IMPLEMENTATION ON FPGA

Roberts, Prewitt and Sobel methods are often used to find the boundaries of an object and separate it from the main background[16,17].

Each of the methods is based on the spatial filtering process. Spatial filtering is based on order move the mask of the filter from the first pixel of the image to the last one. Each pixel with coordinates (x, y) is calculated as the sum of multiplication of the pixel values under the filter mask to their corresponding filter coefficients. If sharp drops of brightness are observed in the image, then the derivatives of the first and second order are used. In order to determine the first derivative of a one-dimensional function f (x), it is necessary to calculate the difference in the values of the neighboring elements of the image:

$$\frac{df}{dx} = f(x+1) - f(x) \tag{1}$$

A partial derivative is used to preserve the same allocations for the variables f (x, y) if it is necessary to work on two axes of coordinates with partial derivatives. The second derivative is calculated here as the difference between adjacent values of the first derivative:

$$\frac{d^2 f}{dx^2} = f(x+1) + f(x-1) - 2f(x) \tag{2}$$

The calculation of the first derivative of image is based on discrete approximations of a two-dimensional gradient. The image gradient f (x, y) in the pixel (x, y) is a vector:

$$\Delta f = \frac{G_x}{G_y} = \frac{{df}/{dx}}{{df}/{dy}} \tag{3}$$

The trajectory of the vector gradient coincides with the trajectory of the maximum rate of change of the function f in the pixel (x, y). The module of this vector plays a major role in determining the contours in the image. It is denoted by the symbol Δ f and equated with:

$$\Delta f = |\Delta f| = \sqrt{G_x^2 + G_y^2} \tag{4}$$

In all methods of defining the boundaries of an object, a matrix of pixels is used size 3x3 $\begin{pmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{pmatrix}$

The easiest way to determine the first partial derivatives at the Z_5 point is to use Roberts's cross gradient operator:

$$G_x = (z_9 - z_5);\, G_y = (z_8 - z_6); \tag{5}$$

The values of Gx and Gy can be obtained by converting an entire image using one of Roberts's masks:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

The Prewitt operator, as well as the Roberts operator, works with a floating mask with 3x3 pixels dimension. Unlike the Roberts operator in the Prewitt operator, the matrices are given by other formulas:

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3); \quad G_y = (z_3 + z_6 + z_9) - (z_7 + z_4 + z_1);$$

To apply these formulas for selecting the edges in the image, the Prewitt operator[16,18] is used. This operator is given by the following masks dimension 3x3:

$$\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \tag{6}$$

Sobel operator also uses a dimension matrix 3x3. The main difference from the Prewitt operator is the application of the weighting factor 2 for the center pixels:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3); \quad G_y = (z_3 + 2z_6 + z_9) - (z_7 + 2z_4 + z_1); \tag{7}$$

Masks that are used in the Sobel method with a dimension of 3x3:

$$\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}.$$

Component values of the gradient Gx and Gy are calculated by using matrices. To determine the gradient value, these components need to be operated simultaneously:

$$\Delta f \approx |G_x| + |G_y| \qquad (8)$$

One of the methods of image markup is the ABC mask. It allows to label image objects in one pass. This is done to determine the contour of the image, the area, geometric characteristics of the object. Passage through the image is left-to-right and top-to-down. It is believed that there are no objects outside the image, so if B or C enters there, this requires additional scan verification[19].

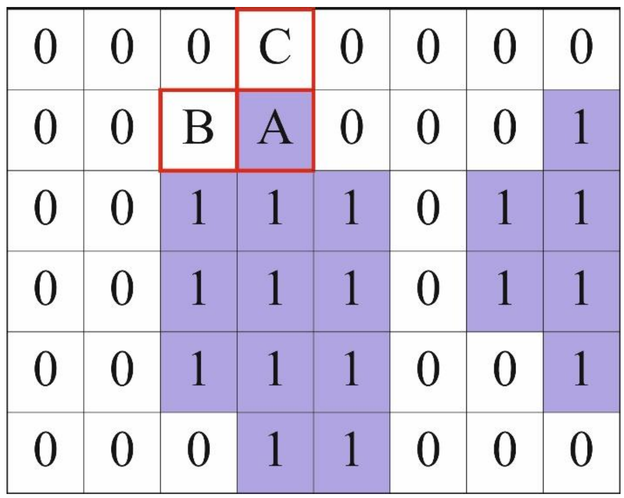

Figure 7. ABC mask.

5 possible positions for the ABC mask (Figure 8) [16-19]:

1. When all three components of the mask are not marked - the pixel is skipped

2. When only element A is marked, a new object (new number) is created.

3. When element B is marked - the current pixel A is marked with a label located in B.

4. When the C element is marked, the current pixel A is marked with a label located in C.

5. When the elements B and C are marked, the labels (object numbers) B and C are related - that is equivalent and the pixel A may be denoted either as B or as C. In the event that B is not equal to C then it is renumbered all processed pixels are labeled as C in the B mark.

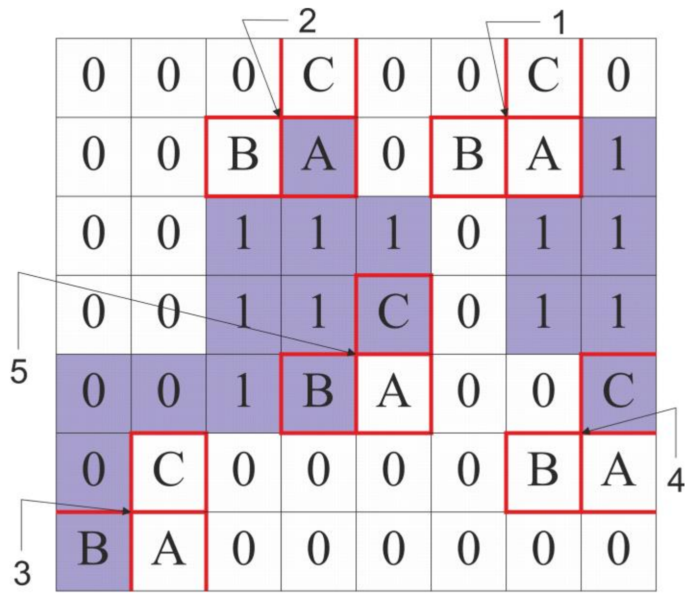

Figure 8. 5 possible positions of the ABC mask.

After mask passes throughout the image, all unconnected objects will be marked (Fig. 9).

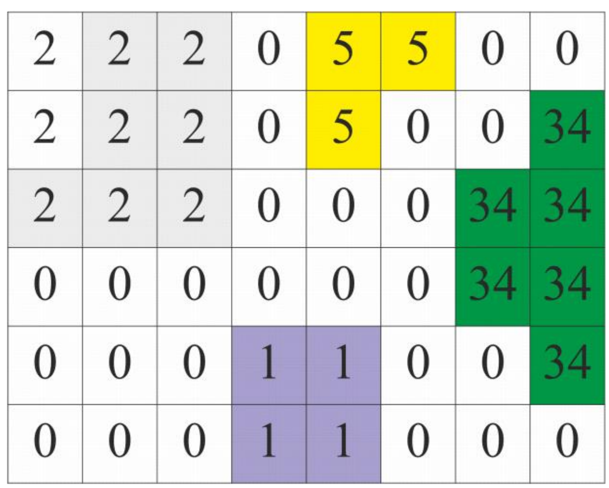

Figure 9. The result of the passage of the ABC mask by the image.

There are many unconnected objects (this depends on the pre-processing of the image after binaryization. Therefore, it is needed to set certain initial conditions. Before further processing, it is not superfluous to use the median filter, this will avoid unnecessary checks in the contour allocation.

7. CONCLUSION

It is possible to get a basic model of a shaped optoelectronic computer based on components such as the camera Q-12A65 from series QUARTZ the Adimec company and the FPGA Arria II GX family of EP2AGX260FF35I5. A cell of a homogeneous structure is written on the crystal of the FPGA, which implements the method of pre-processing of the image, namely the segmentation, the algorithm of which is described in the previous section. Simulated a cell of a homogeneous structure. A calculator scheme is also modeled, which plays a key role in image processing described in this article.

REFERENCES

[1] K. Verhagen, R. Dane, F. Grun and others, [Pattern Recognition: State and Prospects], Radio and communication, Moscow, (1985).

[2] Amin A. “Recognition of Hand – Printed Latin Characters Based on Generalized Hough Transform And Decision Tree Learning Techniques”. International Journal of Pattern Recognition and Artificial Intelligence, 14(3), 269 – 287 (2000).

[3] S.V. Ablameiko, D.M. Lagunovsky, [Image processing: technology, methods, application], Inst. Tech. Cybernetics of the National Academy of Sciences of Belarus, Minsk, (1999).

[4] Komashinsky V.I., Smirnov D.A. [Neural networks and their application in control and communication systems], Hotline – Telecom, Moscow, (2003).

[5] Kozhemyako, V.P. “A look at the nature of artificial intelligence”. Bulletin of the VPI, 1, 26-30, (1997).

[6] Kozhemyako V.P., Timchenko L.I., Yarovoy A.A., [Parallel hierarchical networks as a structural and functional basis for the construction of special models of a figurative computer], Universum, Vinnytsia, (2005).

[7] Kozhemyako V.P., Yarovoy A.A., “Scientific concept of a figurative video - computer of an eye-processor type in the context of modern methodology of artificial intelligence”, Optoelectronic information and energy technologies, 2, 84-89, (2001).

[8] Ivakhnenko A.G. “Structures of a figurative computer based on the principle of self-organization of models with an enhanced generalization property”, USiM, 4, 22-25, (2003).

[9] Kozhemyako V.P. [Optoelectronic time-based information and computing environments], Metsniereba, Tbilisi (1984).

[10] Kozhemyako V.P., Natroshvili O.G. “Problems and features of building a new generation of intellectual structures using optoelectronic information and energy technologies”, Proc. "Photonics ODS-2005", 10-12, 2005.

[11] Kozhem'yako V.P., Lishchynska L.B., Ponura O.G. “Features of the formation of system of attributes in the recognition of images in systems of technical vision”, Bulletin of the VPI, 1, 47-56, (1998).

[12] Kozhemyako V.P., Olshevskaya O.V., Kozhemyako A.V., Novytsky R. M. “Nanotechnological principles for the implementation of arithmetic-logical operations for optoelectronic computers of the logical-time type”, Optoelectronic information and energy technologies, 5-6(1-2), 74-87, (2003).

[13] Kozhemyako, A.V., Bezkrevnyi, O.S., Vasilkiva O.S. and Prokophiev O.V., "Geoinformation system for the classification of objects on geographic maps". Opto-electronic information and energy technologies, 33(1), 33-39 (2017).

[14] Martynyuk, T.B., Kozhemyako, A.V., Krupelynitsky, L.V., Perebeinis, O.M. and Bezkrevnyi O.S., "Implementation models of the matrix calculator for the biomedical data classifier". Information Technology and Computer Engineering 36(2), 43-51 (2016)

[15] T.B. Martynyuk, G.D. Doroshenkov, O.M. Hutsol, “Schematic solutions of base blocks for image classifier”, Bulletin of the Vinnitsa Polytechnic Institute, 3, 132-141, (2012).

[16] T.B. Martynyuk, A.V. Kozhemyako, A.V. Melnik, “Features of parallel-cut processing of matrix elements for the classification of objects”, Optoelectronic information and energy technologies, 26(2), 28-33, (2013).

[17] R. Gonzalez, R. Woods. [Digital Image Processing], Technosphere, Moscow, (2004).

[18] Lee C.C. “Fuzzy logic in control systems: fuzzy logic controller - part I, II”, IEEE Trans, on Systems, Man and Cybernetics, 20(2), 404-435, (1990).

[19] Nobuyuki Otsu, «A threshold selection method from gray-level histograms», IEEE Trans. Sys., Man., Cyber., 9, 62–66, (1979).