A Homogenous Parameter Set for Image Recognition Based on Area

Stepan Belan[1] & Sergey Yuzhakov[1]

[1] State Economic and Technologies Transport University, Kiev, Ukraine

Correspondence: Stepan Belan, State Economic and Technologies Transport University, Kiev, Lukashevicha str., 19, 03049, Ukraine. Tel: 380-445-915-127. E-mail: bstepan@ukr.net

Abstract

This paper considers a method of image recognition, which is produced by the transformation of the object into a set of functions representing the area of intersection of the figure and its copy at its shift in one direction. The object of this research is image shapes of individual objects in the visual picture. The efficiency of different methods of optical image recognition has been conducted depending on the scale and rotation angle.

Keywords: optical image recognition, function of the area of intersection, parallel shift, image noise, image area

1. Introduction

The development of computer vision systems is one of the main directions of creation of artificial intelligence. A computer vision system must enable to recognize the image, to locate the object and to determine the parameters of its motion.

There are now a lot of recognition methods that describe the process of recognizing an object and assign it to a given class (Rafael, Richard, & Steven, 2004). However, they focus on a specific array of user-defined features specified in advance. To determine the geometric image uses a lot of features that provide high quality of recognition and form the vector of recognition. However, the processing speed and complexity of processing of the array of features gives us the task of reducing their number. One of the features involved into the vector formation, is the area of the image shape. Area of the figure itself contains information only on the object size but not about its form. Knowing the area of the figure and the unit area can give an approximate idea of the size of the figures, but the image recognition in this case is impossible (Rafael, Richard, & Steven, 2004; Chen, Rau, & Wang, 1995). To reduce the number of features of image recognition and limit their set only one - the area, in this paper the problem of calculating the dynamics of change of the area during a parallel shift of copy of a recognizable image (Belan & Yuzhakov, 2002, 2008).

The proposed method of image recognition is based on the definition of the area of intersection of the original figure and its copy, which is shifted in parallel direction. The resulting function of the area of intersection (FAI) depending on the shift will be the same for a group of objects. Using the definition of the FAI of the original figure in the other direction, we can get another group of objects, for which the new FAI is the same. The image belongs to the intersection of the two resulting of groups of objects. In addition, this method will allow work with the image as a whole rather than as a set of pixels. It can be used in transactions with not bitmap images.

The determination of the area and shift can be easily implemented in hardware. Particularly effective is the use of cellular automata (S. Belan & N. Belan, 2012; Bandini, Bonomi, & Vizzari, 2012; Ioannidis, Andreadis, & Sirakoulis, 2012). With their help it is easy to implement parallel processing of data arrays.

This paper contains the following sections:

- formation of the image characteristic features based on the determination of the FAI;
- mathematical model of obtaining of FAI;
- image recognition based on FAI analysis and integral indicators;
- investigation of effect of changes in the scale and orientation of figures on the recognition process images with FAI;
- conclusion.

Descriptions of the proposed method by other developers for authors are unknown. Use of parallel shift in the

recognition process was not described in the scientific literature. Parallel shift is used by other developers only at the stage of normalization of the image. Calculations described in this paper were performed by the authors.

2. Formation of the Image Characteristic Features Based on the Determination of the FAI

FAI(x) is determined as the area of intersection of the image (Figure 1a) and its copy, that is shifted in a parallel direction φ (Figure 1b). The shift is performed from the position of the complete coincidence of the image and its copy to the moment of zero intersection. The maximal distance of the image shift will be called maximal shift (X_{max}). The obtained FAI (Figure 2a) will be equal for the whole class of images (X), that differ only by their mutual position of components, situated in the shift direction.

To distinguish the images that belong to one class X, one more FAI for direction $\varphi_1 \neq \varphi$ ($\varphi_1 \in \overline{\varphi \pm \pi}$) should be determined. Thus we obtain a Y image class with equal FAI(y) (Figure 2c). The image that is recognized belongs to the intersection of classes X and Y. The shifts are performed in orthogonal directions to facilitate the technical implementation of method.

The basic parameters of the FAI at image recognition are: initial image area (S_0), maximal shift in φ direction (X_{max}), maximal shift in φ_1 direction (Y_{max}).
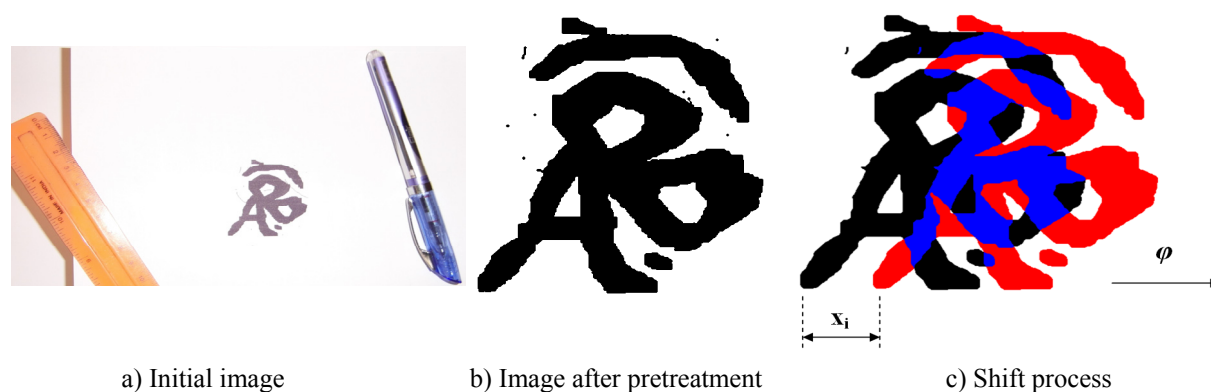

a) Initial image   b) Image after pretreatment   c) Shift process

Figure 1. Image shift for FAI formation

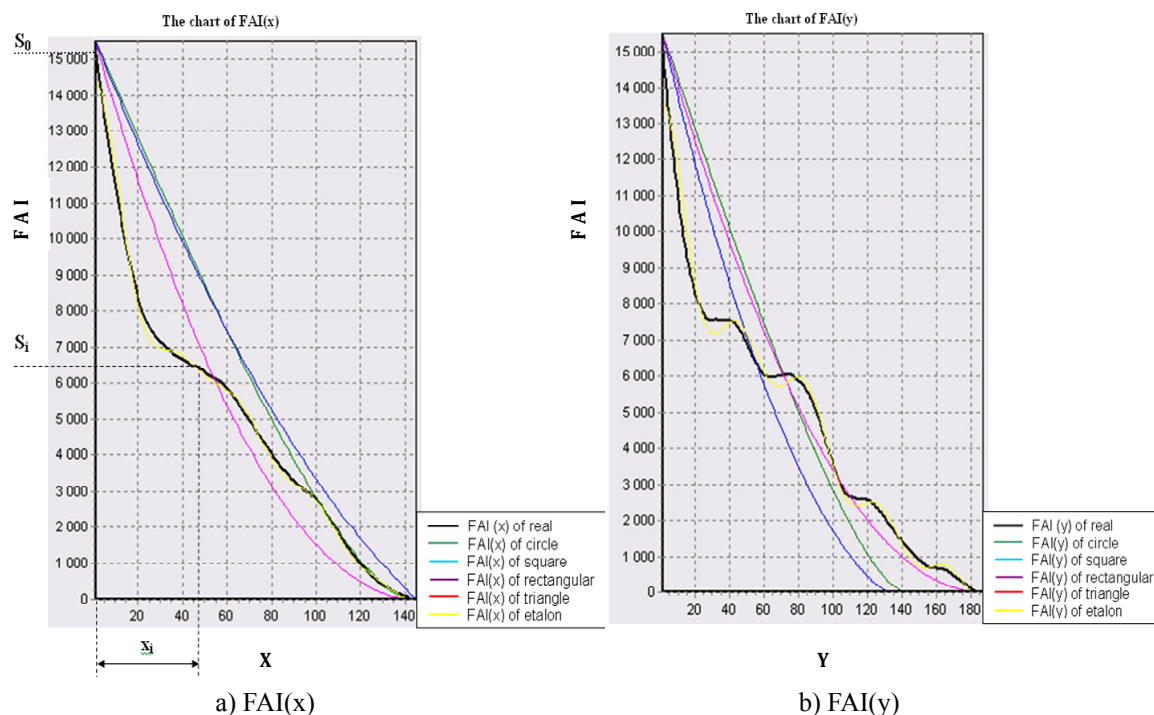a) FAI(x)   b) FAI(y)

Figure 2. Functions of the area of intersection

On the chart of function shown on Figure 2 the areas of intersection of simple images are built by substitution into analytically calculated formulas of the basic parameters of the real FAI. Etalon FAI was taken from the etalon database. It was obtained during the learning process of image recognition systems for the image shown on the Figure 1b.

3. Mathematical Model of Obtaining of FAI

Input image (Figure 3a) is limited above by the function $f_2(x)$ and below by the function $f_1(x)$. Image in Figure 3b is aligned at the bottom edge of the input image. It is bounded below by the x-axis, and the top – by the function $f_3(x)$. Since the image data differ only by mutual position is located in the direction of displacement (in this case the vertical), the $FAI(y)_a = FAI(y)_b$. $FAI(y)$ of the image when vertical shift of y_1 shown on Figure 3c.

$$f_3(x) = f_2(x) - f_1(x) \tag{1}$$

In the Formula (1) of the function $f_3(x)$ provided the basic meaning of the method of image recognition based on technology of parallel shift. The function $f_3(x)$ is obtained as a superposition of the functions $f_2(x)$ and $f_1(x)$, which form the contour of the object. This property can be used for research of form of the input image.
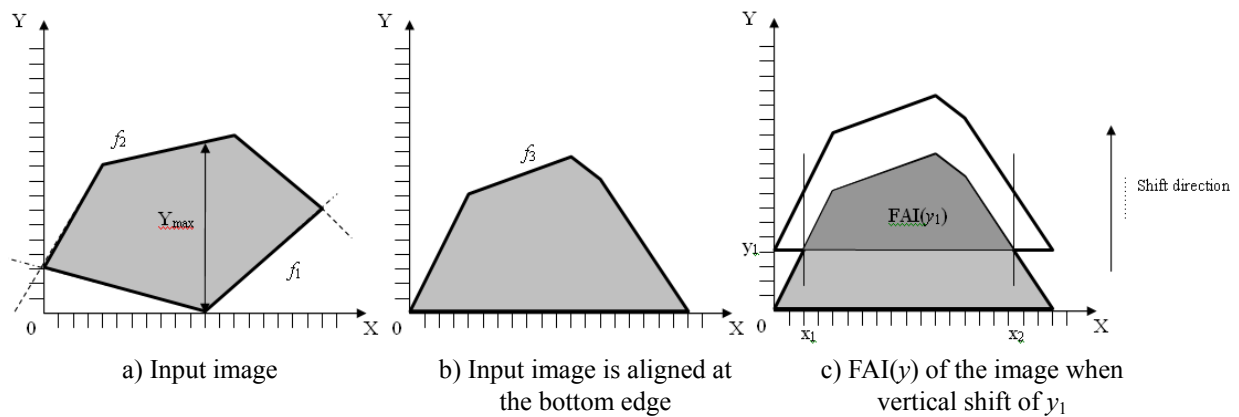

a) Input image   b) Input image is aligned at the bottom edge   c) FAI(y) of the image when vertical shift of y_1

Figure 3. Mathematical model of obtaining of FAI(y)

$$FAI(y) = \int_{x_1}^{x_2} f_3(x)dx - y \cdot (x_2 - x_1) \tag{2}$$

The values of x_1 and x_2 are determined from the system (3)

$$\begin{cases} y = f_3(x_1) \\ y = f_3(x_2) \end{cases} \tag{3}$$

where y – is the value of the shift.

Obtaining FAI(x) is similar (Figure 4). The shift in this case proceeds in the other direction (in this case along the x-axis).
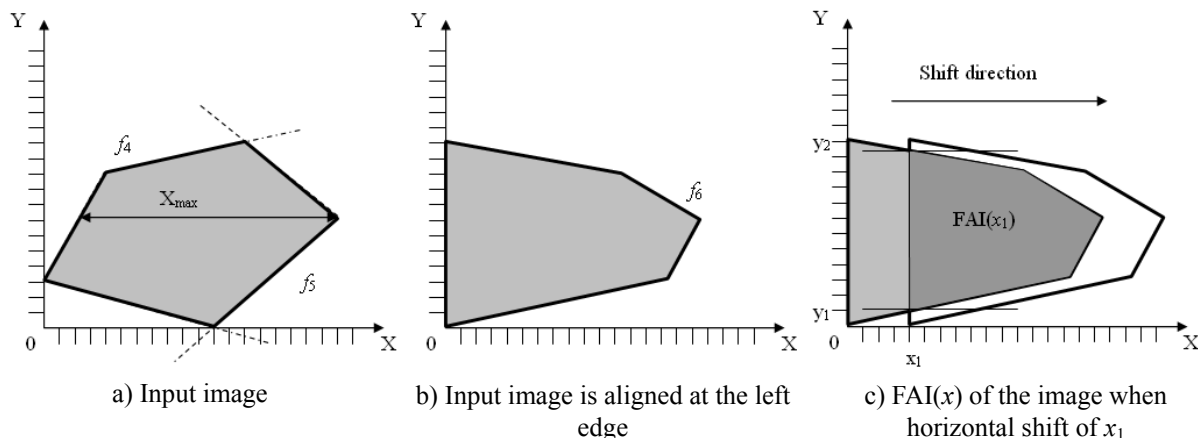

a) Input image   b) Input image is aligned at the left edge   c) FAI(x) of the image when horizontal shift of x_1

Figure 4. Mathematical model of obtaining of FAI(x)

$$f_6(y) = f_5(y) - f_4(y), \tag{4}$$

where $f_5(y)$ and $f_4(y)$ – functions that limit the input image on the right and on the left.

$$FAI(x) = \int_{y_1}^{y_2} f_6(y)dy - x \cdot (y_2 - y_1) \tag{5}$$

The values of y_1 and y_2 are determined from the system (6)

$$\begin{cases} x = f_6(y_1) \\ x = f_6(y_2) \end{cases} \tag{6}$$

where x – is the value of the shift.

As can be seen from Figures 3 and 4, the functions $f_2(x)$ and $f_1(x)$, on which is based the calculation of FAI, include the same point of the input image, that the functions $f_5(y)$ and $f_4(y)$. This feature increases the redundancy of the recognition method.

In determining the FAI of figures, which are not convex or have voids within the image, you need to take into account such changes of shape.

FAI and integral indicators are constructed using a function that displays a form shapes. They themselves become more complex due to the increased complexity of the forms. So they can be used for image recognition.

It is possible to analytically calculate FAI of some images on the basis of a mathematical model. Images, FAI which can be analytically determined, are called simple images. Simple images include circle, triangle, square and rectangle.

FAI for a circle.

$$FAI(x) = 2 \cdot \frac{S_0}{\pi} \cdot (arctg(\sqrt{\frac{4 \cdot S_0}{\pi \cdot x^2} - 1})) - x \cdot \sqrt{\frac{S_0}{\pi} - \frac{x^2}{4}} \tag{7}$$

where x – shift in φ direction.

$$FAI(y) = 2 \cdot \frac{S_0}{\pi} \cdot (arctg(\sqrt{\frac{4 \cdot S_0}{\pi \cdot y^2} - 1})) - y \cdot \sqrt{\frac{S_0}{\pi} - \frac{y^2}{4}} \tag{8}$$

where y – shift in φ_1 direction.

If $x = 0$ ($y = 0$), then $FAI(x) = S_0$ ($FAI(y) = S_0$).

For a rectangle with sides sized a and b and the tilt angle α to the φ direction:

$$FAI(x) = (b - x \cdot \sin(\alpha))(a - x \cdot \cos(\alpha)) = S_0 - Bx + \frac{x^2}{2}\sin(2\alpha) \tag{9}$$

$$FAI(y) = (b - y \cdot \cos(\alpha))(a - y \cdot \sin(\alpha)) = S_0 - Ay + \frac{y^2}{2}\sin(2\alpha) \tag{10}$$

where $A = \frac{S_0}{Y_{max}} + \frac{Y_{max}}{2} \cdot \sin(2\alpha)$, $B = \frac{S_0}{X_{max}} + \frac{X_{max}}{2} \cdot \sin(2\alpha)$.

$$FAI(x) = (X_{max} - x)(\frac{S_0}{X_{max}} - \frac{x}{2}\sin(2\alpha)) \tag{11}$$

$$FAI(y) = (Y_{max} - y)(\frac{S_0}{Y_{max}} - \frac{y}{2}\sin(2\alpha)) \tag{12}$$

For a rectangle, depending on ratio the size of side (a and b) and the tilt angle α to the φ direction, there are several variants determination of $sin(2\alpha)$:

$$\sin 2\alpha = \frac{2 \cdot S_0}{X_{max} \cdot Y_{max}} \quad \text{or} \quad \sin 2\alpha = \frac{2 \cdot X_{max} \cdot Y_{max}}{X_{max}^2 + Y_{max}^2} \quad \text{or} \quad \sin 2\alpha = \frac{2 \cdot \sqrt{S_0 \cdot (X_{max} \cdot Y_{max} - S_0)}}{X_{max} \cdot Y_{max}} \tag{13}$$

For a square where $X_{max} = Y_{max}$:

$$FAI(x) = (X_{max} - x) \cdot \frac{X_{max} \cdot S_0 - x \cdot \sqrt{S_0 \cdot (X_{max}^2 - S_0)}}{X_{max}^2} \tag{14}$$

$$FAI(y) = (X_{max} - y) \cdot \frac{X_{max} \cdot S_0 - y \cdot \sqrt{S_0 \cdot (X_{max}^2 - S_0)}}{X_{max}^2} \tag{15}$$

For a triangle:

$$FAI(x) = \frac{S_0 \cdot (X_{max} - x)^2}{X_{max}^2} \tag{16}$$

$$FAI(y) = \frac{S_0 \cdot (Y_{max} - y)^2}{Y_{max}^2} \tag{17}$$

Proceeding from mathematical model the concurrence FAI is possible only for a triangle and square, which sides are located under a angle of 45 degrees to a horizontal. However in this case for the square $X_{max}=Y_{max}$ and $X_{max}*Y_{max}=2S_0$.

To determine the classes of images X and Y, at the intersection of which is located recognizable image, in the case of simple images, the following operations must be performed.

    1)   Obtain an FAI for a real image in two directions and determine their basic parameters (S_0, X_{max}, Y_{max}).

    2)   The substitution of these basic parameters in analytical formulas of FAI of simple images.

    3)   Compare the FAI of real image for two directions with calculated formulas of the area of intersection of simple images.

    4)   In the case that FAI with given accuracy *dS* in both directions for some simple shapes, the input image is defined as this figure.

To generate FAI for simple images, patterns of analytically calculated FAI must be stored in the memory of system of image recognition.

If the process of recognition of simple figures is successful, use of base parameters (S_0, X_{max}, Y_{max}) will allow to determine the quantitative characteristics (radius of a circle, sizes of the parties and tilt angle to the horizontal direction for a square and rectangular).

4. Image Recognition Based on FAI Analysis and Integral Indicators

For the most images it is impossible to calculate FAI analytically. Therefore, each image is represented of etalon FAI for each direction of the shift. The etalon of image is stored as an etalon surface, which consists of the functions of the area of the intersection of the shape in the $\overline{0,\pi}$ shift direction. This surface has a central symmetry, so it is unnecessary to store FAI for $\overline{0,2\pi}$ directions. The etalon surfaces are formed during learning of recognition system. FAI can be stored in different types. For example, in a sequence of numbers or as a harmonic of the Fourier transform. In any case, the final stage of recognition will be the comparison etalon of functions of the area of intersection with FAI of real object.

This approach is time consuming, depending on the process of calculating the FAI and further comparing the numerical sets. To reduce the number of comparison operations in the array of etalons introduced integral indicator k_x for each direction. It has the ability to not change with scale and based on the basic parameters of the function of the area of intersection (S_0, X_{max}, Y_{max}).

$$k_x = \frac{\int_0^{X_{max}} FAI(x)dx}{S_0 \cdot X_{max}} = const \tag{18}$$

where *x* and X_{max} – single (basic) and maximal shift in φ direction (Figure 5). Integral indicator – a dimensionless quantity.
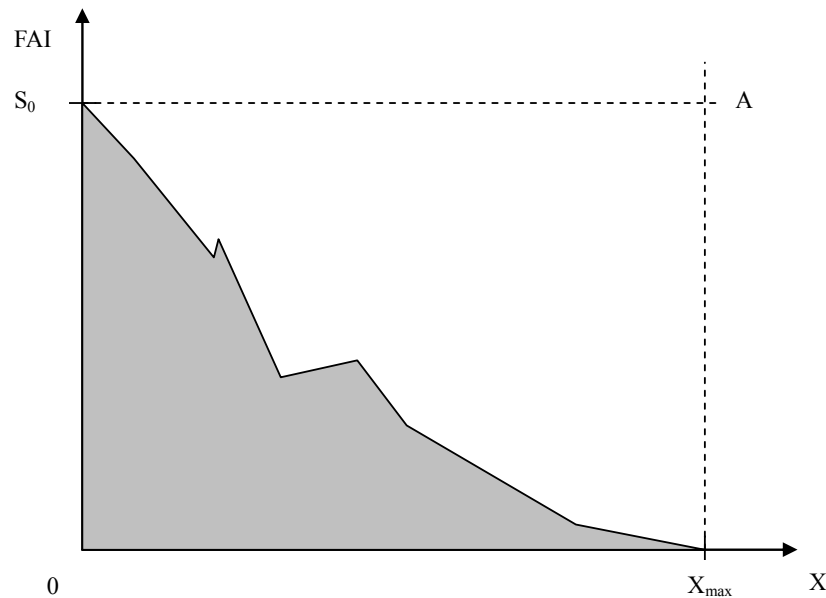
Figure 5. Example of image representation by FAI

k_x value determines the ratio of the filled area to the rectangle 0 S_0 A X_{max}. This ratio remains constant at any scaling graphic of FAI. In determination of k_y, x is substituted by y, and X_{max} by Y_{max}.

Database of etalons must also include a set of integral indicators for the corresponding directions of the shift.

The use of integral indicators can significantly reduce the number FAI of etalons, which are selected for phase of comparison with FAI of input image.

Integral indicator for simple figures can be calculated analytically using formulas (1) ÷ (18) (Table 1).

Table 1. Integral indicator for simple figures

The name of figures	Integral indicators for φ direction shift
Circle	$\dfrac{4}{3 \cdot \pi}$
Triangle	$\dfrac{1}{3}$
Square	$\dfrac{1}{2} - \dfrac{1}{6}\sqrt{\dfrac{X_{max}^{2}}{S_0} - 1}$
Rectangle	$\dfrac{1}{2} - \dfrac{1}{12} \cdot \dfrac{X_{max}^{2}}{S_0} \cdot \sin 2\alpha$

The algorithm of image recognition with the help of integral indicators consists in the following.

(1) With the help of two directional shift (φ and $\varphi_1 = \varphi \pm \dfrac{\pi}{2}$) real FAIs of the input image are obtained and their basic parameters (S_0, X_{max}, Y_{max}) and integral indicators (k_x and k_y) are determined.

(2) k_x with the preset accuracy value dk is compared with integral indicators for all etalons and all directions, in case of coincidence class X of images is obtained. The directions φ for X class, for which this coincidence occurred are saved.

(3) k_y with the preset accuracy value *dk* is compared with integral indicators for all etalons for X class for

$\varphi_1 = \varphi \pm \dfrac{\pi}{2}$ directions, in case of coincidence class Y of images is obtained. The directions φ_l for Y class of images, for which this coincidence occurred are saved. Etalons for which occurred a coincidence form a class Y ($Y \in X$).

(4) The real FAI(x) with the preset accuracy value *dS* is compared with etalon FAI(x) for the class Y of images for φ direction, in case of coincidence class X1 of images($X1 \in Y$) is obtained. In the process of comparing we must execute the rescaling of functions of etalon to bring them to the basic parameters of the real FAI.

(5) The real FAI(y) with the preset accuracy value *dS* is compared with etalon FAI(y) for the class X1 of images, in case of coincidence class Y1 of images is obtained. The processed image belongs to the Y1 class ($Y1 \in X1$). In the process of comparing we must execute the rescaling of functions of etalon to bring them to the basic parameters of the real FAI. The recognition of the input image is possible with an accurate choice of *dk*, *dS* and in the presence of appropriate of etalons.

(6) If a void class (X = Ø, Y = Ø, X1 = Ø or Y1 = Ø) is obtained at any image processing stage, it means that there is either no etalon for such image or the preset accuracy parameters have been wrongly chosen. In this case it is necessary to adjust the preset accuracy values or to scan the input image in $\overline{0, \pi}$ direction, thus obtaining a new etalons surface and array of integral parameters.

Class names (X, Y, X1, Y1) are chosen arbitrarily. X, Y – the result of comparison of integral indicators. X1, Y1 – the result of comparison of FAI.

The process described in parts 2 and 3 of the algorithm is called fast image recognition stage and the process described in 4 and 5 – detailed stage. On a fast recognition stage is a comparison of numerical values (integral indicators), and on the detailed recognition stage is a comparison of functions, which is much longer.

For simple images, the comparison set forth in positions 2-5 of this algorithm is conducted not with etalon but with the analytically determined components on the basis of basic parameters (S_0, X_{max}, Y_{max}).

It should be noted that in pos.1 and 3 we have chosen the $\varphi_1 = \varphi \pm \dfrac{\pi}{2}$ direction because it facilitates the hardware implementation of searching for FAI of input image. Theoretically, any direction can be chosen $\varphi_1 \neq \varphi$ (where $\varphi_1 \in \overline{\varphi \pm \pi}$).

5. Investigation of Effect of Changes in the Scale and Orientation of Figures on the Recognition Process Images with FAI

Due to rounding in certain mathematical calculations, loss of information during data array transformations and the possible presence of noise inclusions in receptive field, permissible errors (PE) should be used at each stage of the image recognition algorithm based on functions of the area of intersection.

PE *dk* – is used in comparisons of integral indicators of FAI at the fast image recognition stage.
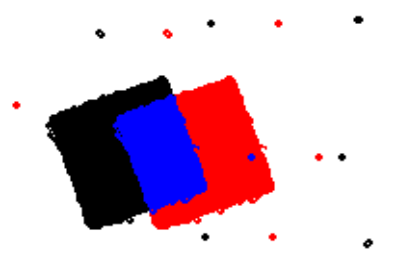
PE *dS* – is used in FAI comparison at the detailed stage of image recognition.

PE *dN* – is used in the formula of determination of maximal shift values.

Image noise (Figure 6) leads to horizontal FAI line, that doesn't allow determination of zero intersection of image and its copy. FAI must be limited if its horizontal components occur. To achieve it, additional adjustments are introduced, that are calculated as noise threshold (*dN*). *dN* parameter determined a rate of components of the input image area (S_0), that are attributed to noise.
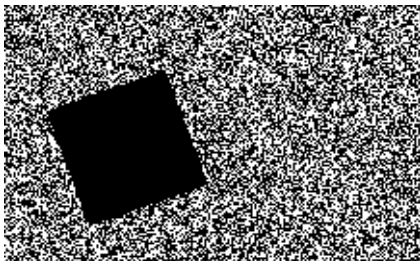
a) Initial image before noise reduction (PSNR=9.12 dB)   b) Some noises are not removed by pretreatment

Figure 6. The absence of zero-crossing due to the presence of noise

Taking into consideration the threshold processing, the parameters formula for maximal of shifts will be as follows:

$$X_{max} = x\text{-}1, \text{ if } FAI(x\text{-}1) = FAI(x) \text{ and } FAI(x) \le dN^*S_0. \tag{19}$$

$$Y_{max} = y\text{-}1, \text{ if } FAI(y\text{-}1) = FAI(y) \text{ and } FAI(y) \le dN^*S_0. \tag{20}$$

Values obtained with the help of these formulae (19, 20) completely match the variant of maximal shift which obtained with the help of the zero intersection in case of absence of noise.

The simulation of the recognition of objects of arbitrary shape was performed. The efficiency of recognition of simple and etalon images in different sizes and spatial orientation was investigated.

As input data used to programmatically generated and scanned images of simple figures, as well as other images (a set of numbers and letters of the Latin and Cyrillic alphabet printed of Arial font). It should be noted that this system is not specialized for the recognition of printed fonts. This type of etalons chosen for ease of reference creation images of different shapes.

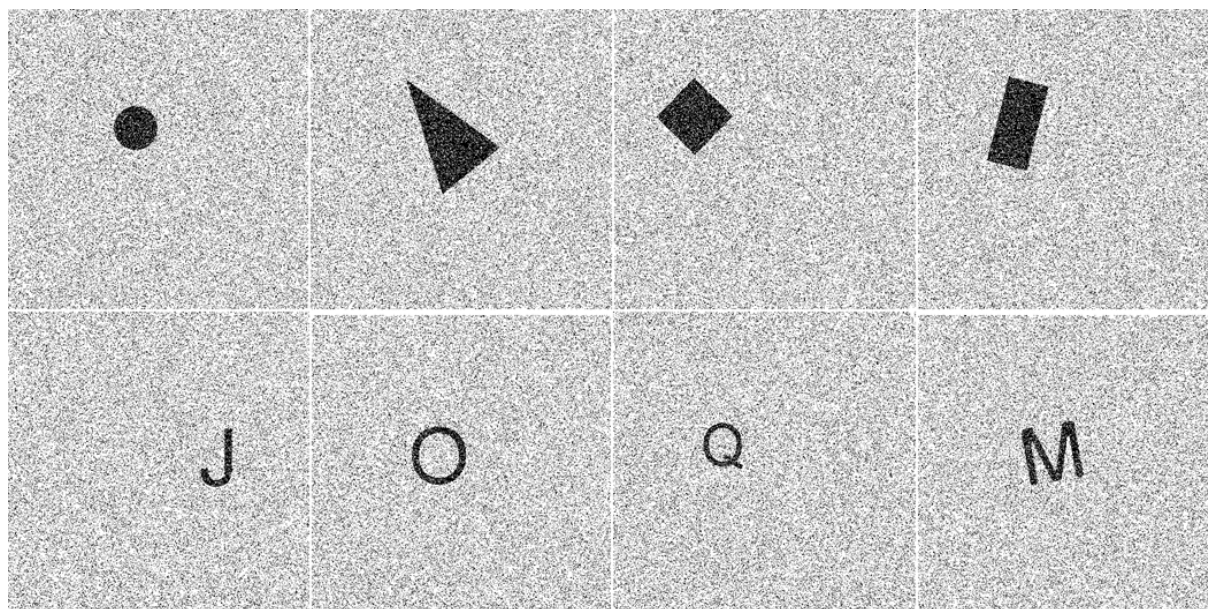The examples of the images recognized by system at PSNR=20 dB are submitted in Figure 7.

Figure 7. The examples of the images recognized by the system at PSNR = 20 dB

Prevent the recognition of the small size of the image (removed in the pretreatment like noise), high noise interference, high contouring (complicated scaling). The examples of images are not recognized by the system shown in Figure 8.

Figure 8. The examples of images are not recognized by the system

Efficiency ratio of the system (*eR*) calculated by the next formula:

$$eR = \frac{cR - cnR}{cR} \tag{21}$$

where *cR* - the number of attempts for recognition of images, *cnR* - number of not recognized image.

Correlation between efficiency ratio of recognition of simple images and their size (in pixels) at *dk* = 0.04, *dS* = 0.1, *dN* = 0.01 is shown on the Figure 9, and the correlation of the efficiency ratio for recognition of etalon images and their size at *dk* = 0.01, *dS* = 0.1, *dN* = 0.01 – on the Figure 10.
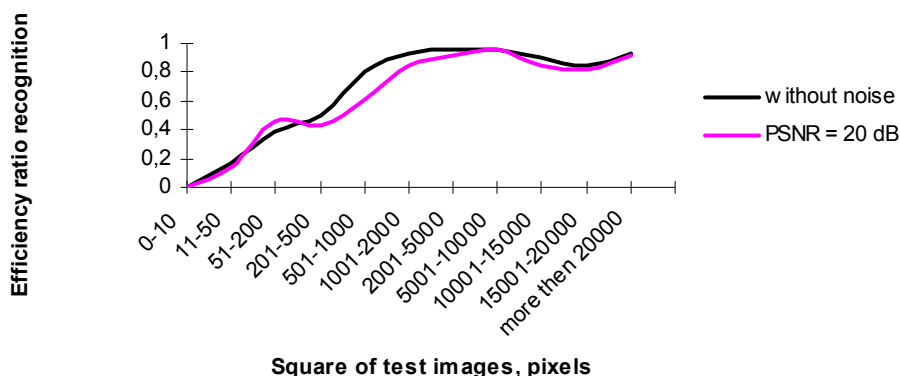
Figure 9. Correlation between efficiency ratio of recognition of simple images and their size at *dk*=0.04, *dS*=0.1, *dN* = 0.01

Local peaks and valleys in the graph 9 caused by the features of generation of simple figures in the modeling program. The rather high efficiency of recognition (eR > 0.8) was obtained for images an area of more than 2000 of pixels. Small-size distorted in the process of removing noise.

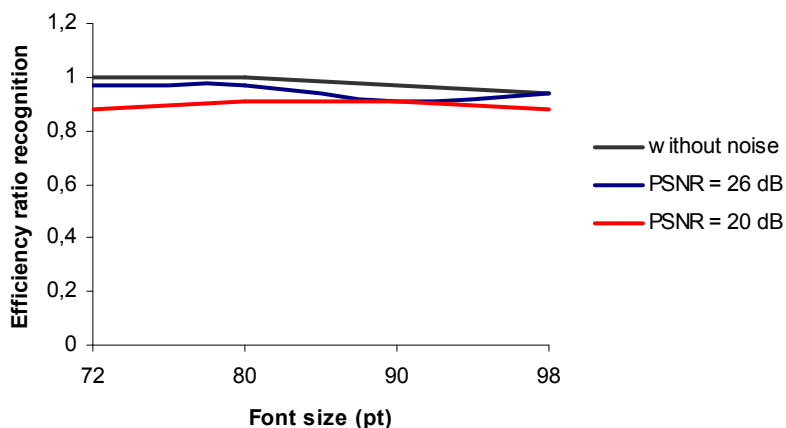Figure 10. Correlation of the efficiency ratio of recognition of etalon images and their size at *dk*=0.01, *dS*=0.1, *dN* = 0.01

The area selected for this study fonts 900-4500 pixels. This corresponds to a range of effective recognition of simple images. If you change the font size detection efficiency ratio of recognition does not change. This confirms the invariance of image recognition using FAI respect to changes size.

Results of altering of efficiency ratio of recognition of etalon images with etalons images rotated by any angle at *dk* = 0.01, *dS* = 0.1, *dN* = 0.01 are shown on the Figure 11.
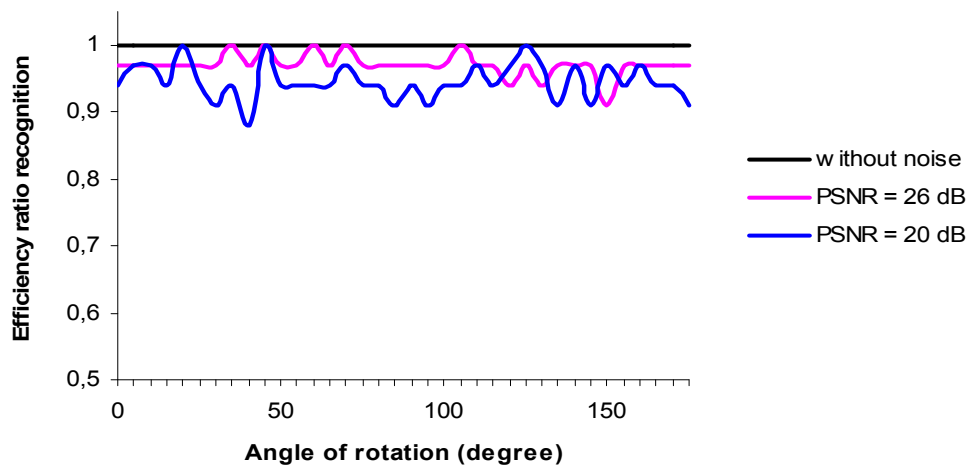
Figure 11. Values of the efficiency ratio of recognition of etalon images depending on the rotation angle at $dk = 0.01$, $dS = 0.1$, $dN = 0.01$

Obtained a result of high efficiency ratio of recognition (eR > 0.9) at different angles of rotation of the image. This proves the invariance of the method with respect to the rotation of images.

Large number of etalons in the etalons database can lead to the decrease in the efficiency ratio of recognition. Unable to recognize mirror images due to the central symmetry of the etalons surfaces. Solution to these problems would be to increase the number of directions the shift for recognition. This requires create a specialized camera where creation FAI will be implemented in hardware. This will get rid of the presence in the database information about the FAI. Comparison can be made exclusively to the database of integral indicators. This will get rid of the processes of scaling, which lead to errors at the stage of comparison.

6. Conclusion

The results of studies of the image recognition using FAI showed high efficiency of recognition, which is determined by more than 80% of correct results when the peak signal to noise ratio of 20 dB. These studies confirm that the proposed method for recognition is invariant under affine transformations such as resizing and rotating.

The time needed to process one image is reduced due to the use of one a homogeneous parameter for the etalon code.

Use of parameter of the area allows perform of simple of geometric operations with images. This will enable you to create the image recognition system based on analysis of FAI. In the future it is necessary to conduct investigations of image edge detection, determining the spatial orientation and movement parameters of objects with help of a parallel shift.

References

Bandini, S., Bonomi, A., & Vizzari, G. (2012). An Analysis of Different Types and Effects of Asynchronicity in Cellular Automata update Schemes. *Natural Computing, 11*(2), 277-287. http://dx.doi.org/10.1007/s11047-012-9310-4

Belan, S., & Belan, N. (2012). Use of Cellular Automata to Create an Artificial System of Image Classification and Recognition. *Springer-Verlag Berlin Heidelberg, ACRI2012, LNCS 7495*, 483-493.

Belan, S., & Yuzhakov, S. (2002). Recognition of plane figures based on analysis of the area of intersection of similar objects. *Herald VPI, 3*, 74-78.

Belan, S., & Yuzhakov, S. (2008). Method of image recognition on the basis of the process of parallel shift for systems of object recognition at railway crossings. *Collection of scientific works of the State Economic and Technologies Transport University, 13*, 216-226.

Chen, C. H., Rau, L. F., & Wang, P. S. P. (1995). *Handbook of pattern recognition and computer vision.* Singapore-New Jersey-London-Hong Kong: World Scientific Publishing Co. Pte. Ltd.

Ioannidis, K., Andreadis, I., & Sirakoulis, G. Ch. (2012). An Edge Preserving Image Resizing Method Based on Cellular Automata. *Springer-Verlag Berlin Heidelberg, ACRI2012, LNCS, 749*, 375-384.

Rafael, C. G., Richard, E. W., & Steven, L. E. (2004). Digital Image Processing using MATLAB.