

Application of Zernike moments in computer vision problems for infrared images

Dmitriy Otkupman^{1*}, Sergey Bezdidko¹ and Victoria Ostashenkova¹

¹Moscow State University of Geodesy and Cartography, Moscow, Russia

Abstract: The efficiency of using Zernike moments when working with digital images obtained in the infrared region of the spectrum is considered to improve the accuracy and speed of an autonomous thermal imaging system. The theoretical justification of the choice of Zernike moments for solving computer (machine) vision problems and the choice of a suitable threshold binarization method is given. In order to verify the adequacy and expediency of using the chosen method, practical studies were conducted on the use of Zernike methods for distorting various thermal images in shades of gray.

Keywords: Zernike polynomials, pattern recognition, thermal imaging systems, binarization.

1 Introduction

One of the main tasks in computer (machine) vision is feature matching. In this regard, a large number of algorithms have been proposed using local feature descriptors that are used to recreate or recognize patterns. Despite the rapid development of autonomous recognition systems, in which artificial neural networks (ANN) are often used, approaches based on the use of local descriptors remain relevant today [1-3]. Their main positive qualities are ease of use, speed and resistance to various types of noise.

There are a large number of varieties of such algorithms, such as: scale-invariant feature transform (SIFT) [4]; accelerated up robust features (SURF) [5]; local binary patterns (LBP) [6], etc. This article examines the Zernike moments (ZM) as one of many local descriptors in image recognition problems. However, among the above descriptors, they represent an easily embedded mathematical apparatus.

ZM is a descriptor used to characterize the object in the image. Initially, geometric moments were used in image processing tasks. Their main problem was information redundancy. The use of ZM, built on a set of orthogonal radial Zernike polynomials, helped to solve this problem. The main advantage of the ZM is the use of orthogonal functions, since there is no redundancy of information between the moments, which makes them more reliable and discriminant.

The article [7] considers the introduction of ZM into the adaptive local feature descriptor and provides results showing easy embeddability of moments. This suggests that the ZM can be used, among other things, as an additional mathematical apparatus for one or another algorithm. However, one of the main tasks of this article is the possibility of using the ZM in problems of object recognition in digital images captured in the infrared region of the spectrum.

2 Mathematical Description

In 1934, F. Zernike introduced a system of polynomials that are orthogonal inside the unit circle and have invariance properties, described in detail in [8]. Based on a set of orthogonal radial Zernike polynomials, Teague [9] described the ZM as a two-dimensional function.

Two-dimensional ZM (Z_{nm}) of order n with repetition m are defined in polar coordinates (r, θ) inside the unit circle as

$$Z_{nm} = \frac{n+1}{\pi} \int_0^1 \int_0^{2\pi} [R_{nm}(r) \cdot e^{jm\theta} \cdot f(r, \theta) \cdot r] dr d\theta, \quad (1)$$

where n is the order of the moment; m is the harmonic order (repetition), while $0 \leq |m| \leq n$, $n - |m|$ is even; $R_{nm}(r)$ – is the n -th order Zernike radial polynomial, given by the formula:

* Corresponding author: odvk@ya.ru

$$R_{nm}(r) = \sum_{k=0}^{(n-|m|)/2} \left[(-1)^k \frac{(n-k)!}{k![(n-2k+|m|)/2]![(n-2k-|m|)/2]!} r^{n-2k} \right] \quad (2)$$

If we take into account that $R_{nm}(r) \cdot e^{-jm\theta}$ is a complex conjugation of Zernike polynomials defined on the unit circle in polar coordinates, and represented as $R_{nm}(r) \exp(-jm\theta) = V_{n,m}^*(x,y)$, then we can easily express ZM of order n with a repetition of m of this conjugation for a discrete digital image as [7]:

$$A_{n,m} = \frac{n+1}{\pi} \sum_{x^2+y^2 \leq 1} f(x,y) V_{n,m}^*(\rho,\theta), \quad (3)$$

in this case, $V_{n,m}^*(\rho,\theta) = V_{n,m}^*(x,y)$, a $f(x,y)$ is a two – dimensional image function.

3 Practical Description and Image Reconstruction

From the mathematical description given in equations (1)-(2), it follows that the Zernike polynomials are orthogonal along a disk with radius r , which is given in polar coordinates. This property makes them applicable in computer vision problems and the description of the shape.

Figure 1 (a) shows some selected shape of the object, in this case, the binary image of the tank, for which it is necessary to calculate the ZM. The object or feature is placed in a disk with a radius of r , describing the selected shape. In order to cover the entire area of the figure, the radius r must be set correctly, but in practice this restriction is usually relaxed to ensure a consistent description of the image.

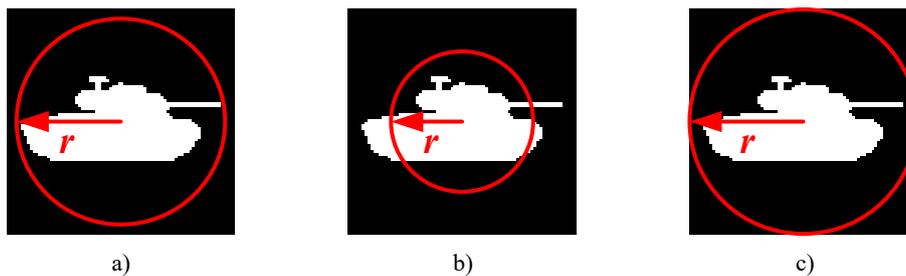


Fig. 1. The selected object of interest placed in a disk with radius r : a) - an example of choosing the right radius; b) - an example of choosing a smaller radius; c) - an example of choosing a larger radius.

To calculate the ZM, they are usually set by two parameters: the radius of the disk (r) and the order (degree) of the polynomial (n). It is worth noting that the radius is the area in which the polynomials are defined.

First, the input image is placed in a disk with radius r , where the center of the image coincides with the center of the disk. For the choice of the radius r , a certain tolerance is allowed without negative effects on the result. However, a large deviation of the value when determining the radius can introduce undesirable effects.

Figure 1 (b)-(c) shows examples of incorrect selection of the disk radius. If the pixels of an object or feature fall outside the disk, they are ignored and not calculated. If the radius was chosen more than the object of interest or feature of interest to us, then other features and objects may also accidentally fall into the disk area when calculating the ZM.

After selecting the radius of the disk describing the object or feature, the ZM is calculated to the specified degree and used as a feature vector. The size of the returned feature vector directly depends on the degree of the polynomial. The larger the degree, the larger the feature vector.

ZM can describe the shape of an object or feature with a function. With the help of the inverse transformation, the image can be restored by the moments of a certain order, which is based on the theorem on the uniqueness of the expansion [10]. The results of the reconstructed thermal image are shown in Figure 2. It can be seen from the figure that with a larger order of the polynomial, the image is restored more clearly. However, due to the fact that the use of the ZM requires a shape and a clear outline of the object of interest, such an initial image is not suitable for further research.



Fig. 2. Restoration of a thermal image of imperfect quality using ZM: a) - the original image; b) - the restored image at $n = 40$

Thus, for a more accurate description of the object of interest and the highlighting of the necessary features, the described shape can be represented as a binary image, or as the boundary of the object (that is the "contour" of the shape). Figure 3 shows the recovery of a binary image using ZM. In comparison with Figure 2, the outline of the object in this case is described more clearly.



Fig. 3. Restoration of a binary image using ZM: a) - the original image; b) - the reconstructed image at $n = 40$

It is worth noting that when there is an overabundance of polynomials, the so-called "collapse" effect occurs. It consists in the fact that the outlines of the object of interest or feature disappear in the reconstructed image. This is another additional factor that shows the importance of choosing the order of the polynomial.

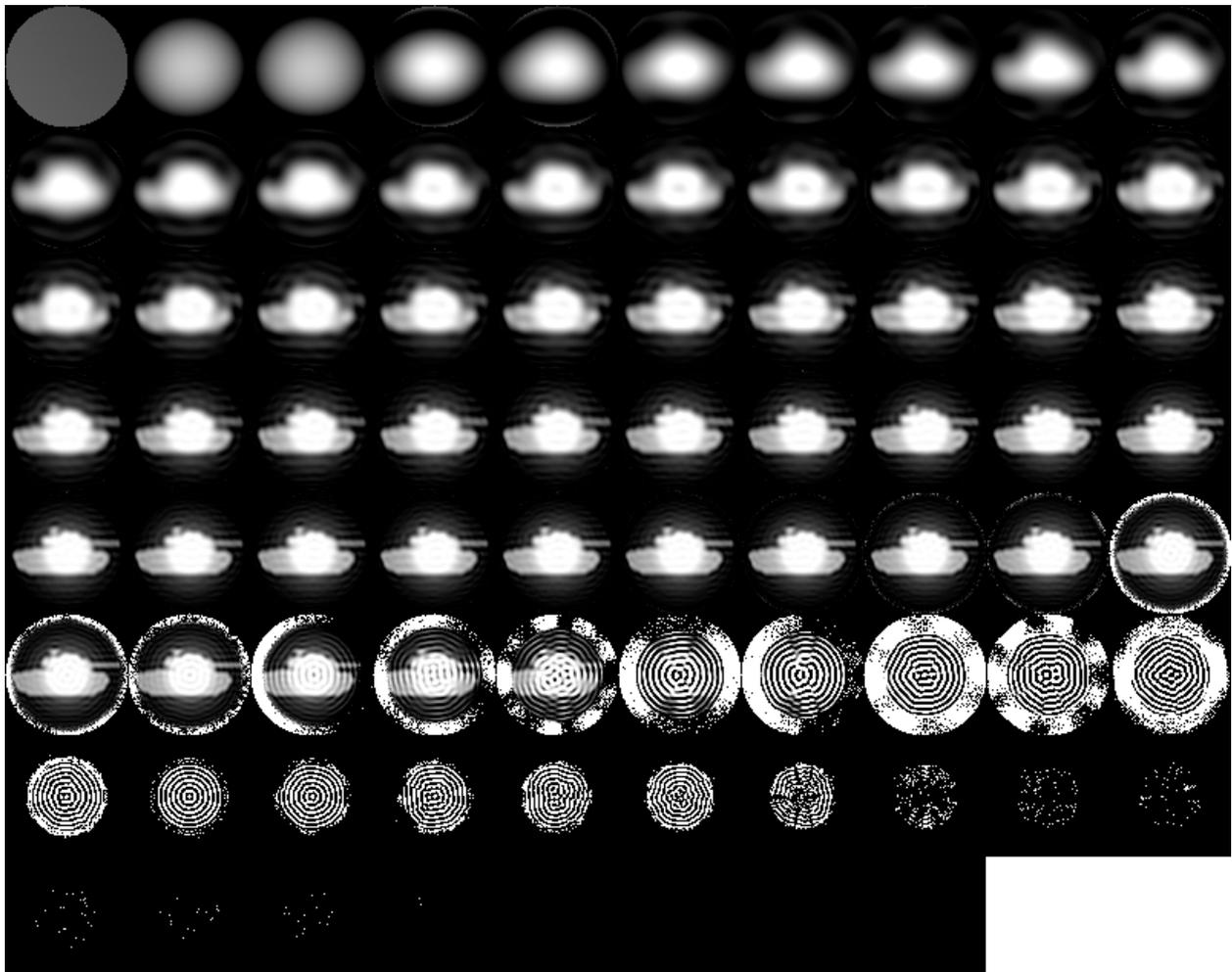


Fig. 4. An example of the "collapse" effect at $n = 0 \dots 88$

4 Results of Practical Research

4.1 Binarization

Before the practical application of the ZM, it is necessary to choose a method of digital filtering of a thermal image to highlight the necessary features or an object of interest. Due to the lower susceptibility to noise, the threshold filtering (binarization) method was chosen. Figure 5 shows the types of binary images. These images were modeled using a special library of computer vision algorithms, image processing and general-purpose numerical algorithms with open source – OpenCV. It is worth noting that the choice of a specific binarization method will depend on other conditions and the quality of the images obtained. Obviously, the greatest effect can be achieved with the help of adaptive algorithms based, for example, on the use of a local or global image histogram.

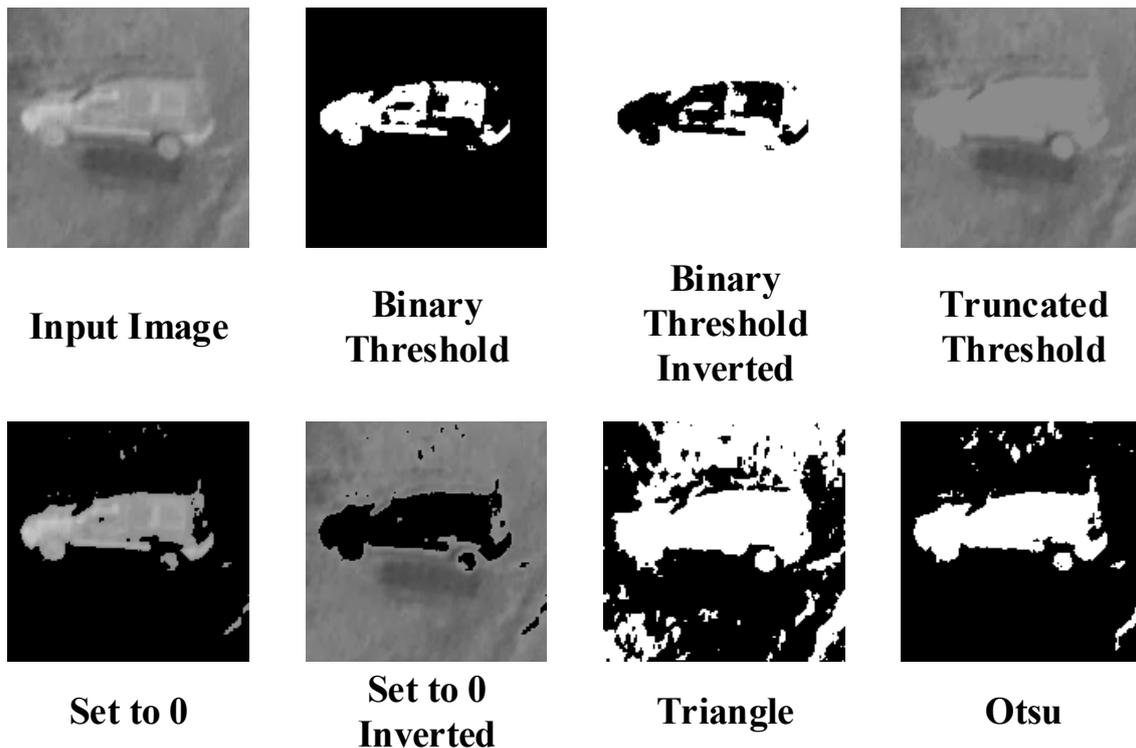


Fig. 5. Examples of binary images at a given threshold $\gamma = 140$.

In the course of comparing the visual representation of a binary thermal image, the method of Analysis was chosen for further research. In particular, it is worth noting that such a threshold filtering method as a multi-level threshold transformation is not suitable for solving such a problem.

4.2 Thermal Image Recognition Using ZM

To implement the ZM when recognizing the object of interest in the image, code written in the Python programming language using special libraries and modules was used. Thus, the value of the disk radius r was calculated using the program and set automatically, with the order of the polynomial $n = 8$. Figure 6 shows the results of autonomous recognition of the object of interest using the ZM. The image was taken using a microbolometric module on vanadium oxide with a resolution of 640×480 and a step of 17 microns, as well as a thermal imaging lens with passive athermalization, a focal length of 100 mm and a relative aperture of 1:4.

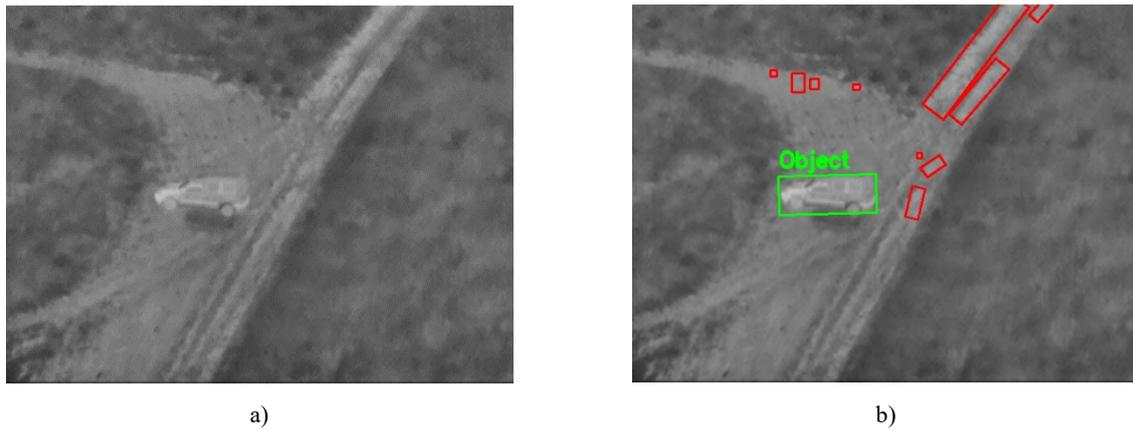
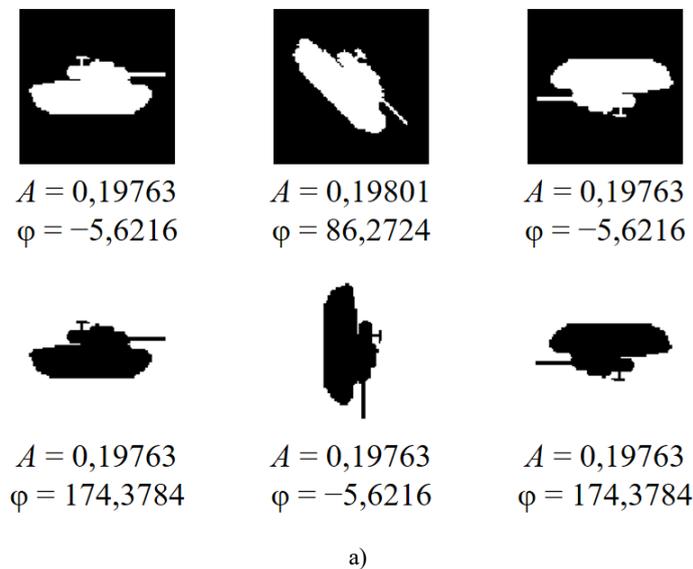


Fig. 6. Autonomous recognition on a thermal image using the ZM: a) - the original image of the scene, b) - recognition of the object of interest in the image of the scene.

4.3 Investigation of Rotation Invariance

Despite the positive result in the autonomous recognition of the object of interest on a thermal image using ZM, their main advantages remain invariance and resistance to various types of distortion and noise. This is of particular interest for moving carriers subject to all kinds of turbulence and rapidly changing foreshortening. For further research, the captured image of Figure 6 (a) was subjected to digital distortions to test the theory of the use of ZM.

In [11], the authors provide a mathematical proof of the invariance of the ZM to rotations. In this paper, this statement was studied practically, its results are presented in Figure 7. Note that the invariance of ZM to rotations is revealed from the equal values of the image amplitude A and the phase φ , the values of which are directly proportional to the rotation. At the same time, in Figure 7 (b), the object of interest is highlighted with a green rectangle, while the red rectangles are similar in some features, but below the specified threshold value.



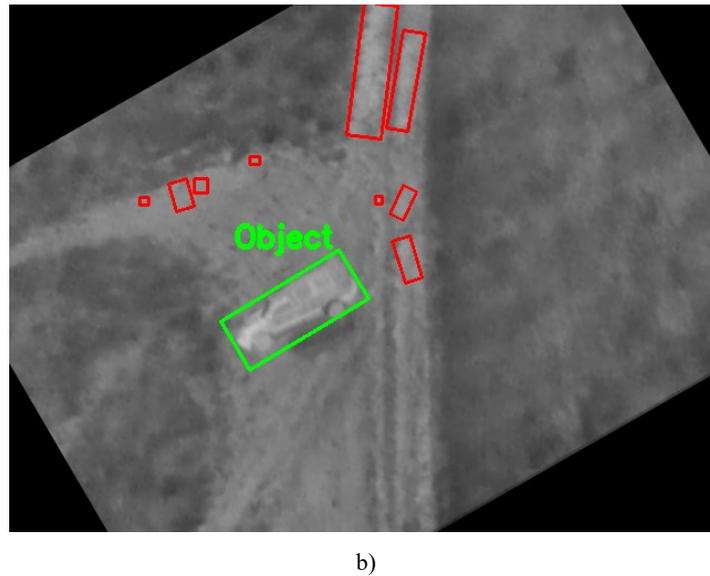


Fig. 7. Results of the study of the use of ZM when the image is rotated: a) - comparison of the values of the amplitude A and phase φ of the images, at $m = 2, n = 4$; b) - the result of recognition of the object of interest in the thermal image, when the image of the scene is rotated by 30° .

4.4 Investigation of Scale Invariance and Noise Resistance

Additional studies of the use of ZM on a thermal image when recognizing an object of interest were carried out in order to verify the invariance of moments to scale and resistance to noise. The thermal image size of 640×480 was enlarged to the format of 1000×750 and reduced to 200×150 . However, despite the change in the size of the object of interest, the program was able to accurately recognize the selected object. The results are shown in Figure 8.

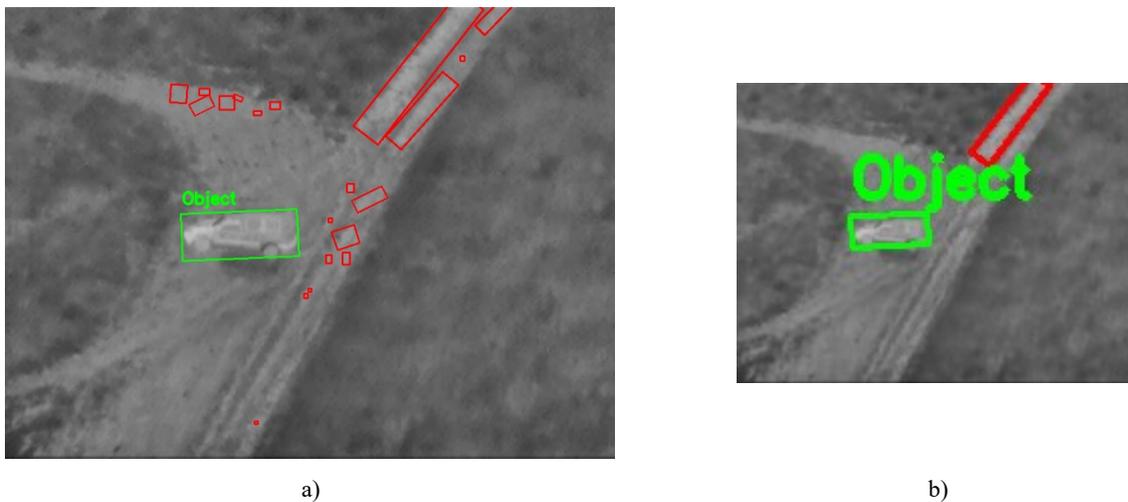


Fig. 8. Results of recognition of the object of interest on a thermal image to check the invariance of the ZM to size changes: a) - an enlarged image of the 1000×750 format; b) - a reduced image of the 200×150 format.

Then the median blur and impulse noise of the "salt/pepper" type were digitally superimposed on the original thermal image. The result of recognizing an image distorted by noise is shown in Figure 9.

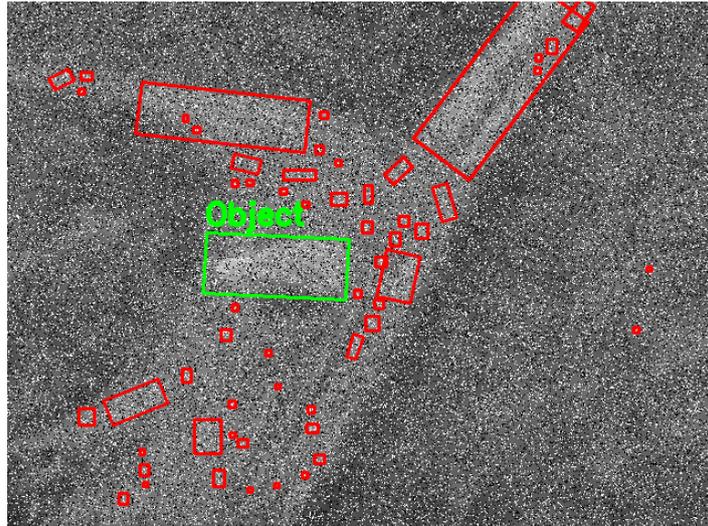


Fig. 9. The result of recognizing the object of interest in a thermal distorted image to check the stability of the ZM to noise.

4.5 Investigation of the Stability of ZM Under Heavy Lubrication

The following experimental study was conducted to test the possibility of recognizing the object of interest when using ZM on a thermal image with a significant blur. Figure 10 shows a detailed diagram of the study. First, a more or less high-quality thermal image was taken and the selected object of interest (car) was used as a feature. After that, ZM was used on a heavily blurred image to recognize the object of interest, while the program was able to easily recognize this object, despite the highly distorted image quality.

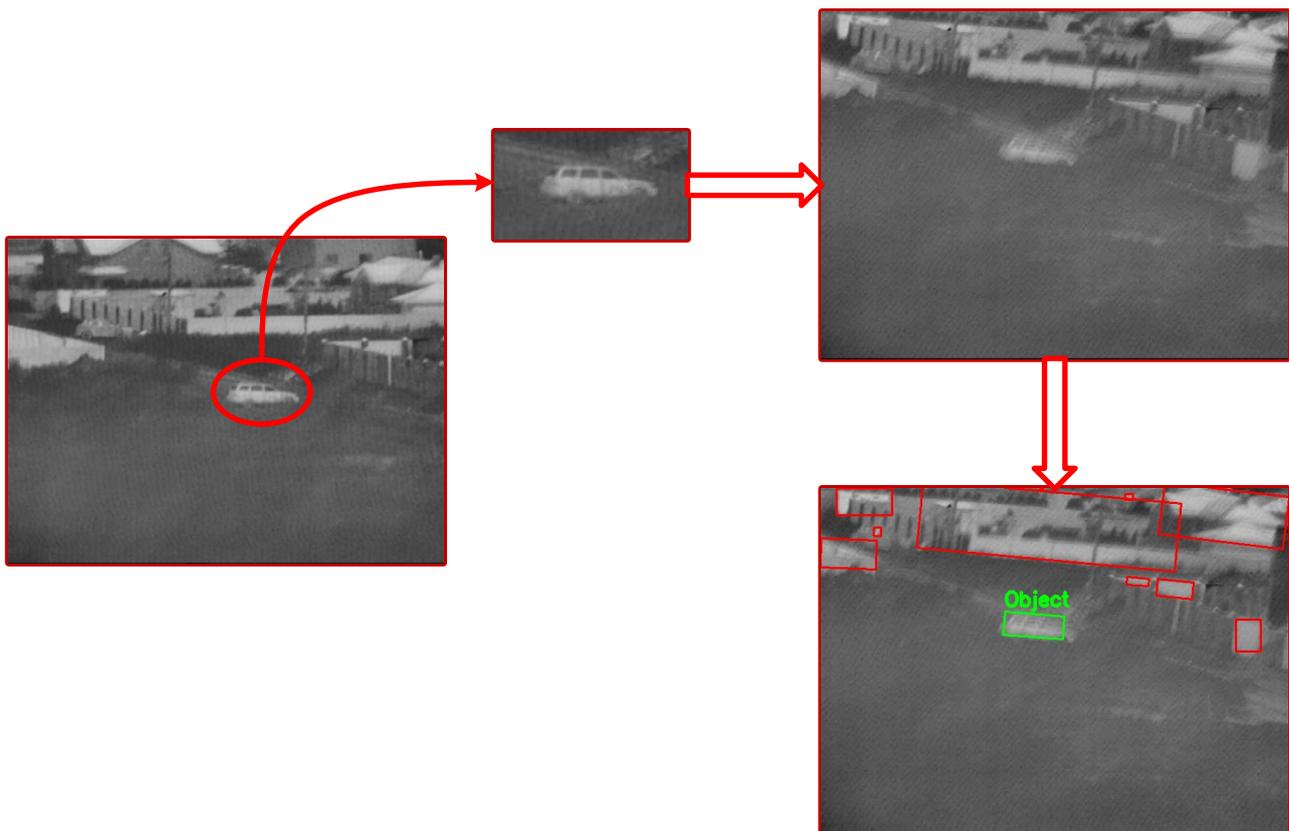


Fig. 10. The scheme of recognition of the object of interest in the thermal image with strong blur.

5 Conclusion

Based on the conducted research, the following conclusions can be drawn:

- 1) Due to the properties of invariance to rotation and scale, as well as relative resistance to noise (interference) and blur, the ZM can be used when working with distorted images.

2) In [7], studies were conducted, which provides a justification for the speed of the MC in contrast to other modern methods of pattern recognition. However, based on the research conducted in our article, it should be borne in mind that the input parameters significantly depend on the application conditions, that is, on a specific task.

3) Due to the relatively simple mathematical apparatus, ZM is successfully implemented into other feature descriptors to increase accuracy and speed, so in the future they can be used together with convolutional neural networks.

4) Due to the fact that classical thermal images do not have color information, but carry information only about the temperature of objects, which is represented by grayscale or pseudo-colors, it is more expedient to apply binarization to such images, moreover, without significant loss of information, which is an integral part of the ZM calculations.

References

1. C. Leng, H. Zhang, B. Li, G. Cai, Z. Pei, L. *Local Feature Descriptor for Image Matching: A Survey*, IEEE Acc., **7**, P. 6424-6434 (2018)
2. Y. Song, D. Zhu, J. Li, Y. Tian, M. Li, *Learning Local Feature Descriptor with Motion Attribute for Vision-based Localization*, in International Conference on Intelligent Robots and Systems, IEEE/RSJ, 4-8 Nov. 2019, Macau, China (2019)
3. E. Karami, S. Prasad, M. Shehata *Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images*, in Proceedings of the 2015 Newfoundland Electrical and Computer Engineering Conference, CCECE, 3-6 May 2015, St. Johns, Canada (2015)
4. D. G. Lowe, Int. J. Comput. Vis., **60**, 2 (2004)
5. S. Leutenegger, M. Chli, R. Siegwart, in Proc. IEEE Int. Conf. Comput. Vis., ICCV, 6-13 Nov. 2011, Barcelona, Spain (2011)
6. C. Singh, E. Walia, K. P. Kaur, Pattern Recognit., **76**, P. 50-68 (2018).
7. B. Zhou, X.-M. Duan, W. Wei, D.-J. Ye, R. Damaševicius, IEEE Acc., **7**, P. 183971 – 183984 (2019)
8. M. Born, E. Wolf Principles of optics (Camb. Univ. Pr., UK, 1999)
9. M.R. Teague, Opt. Soc. Am., **70**, P. 920-930 (1980)
10. M. Liu, Y. He, B. Ye, Geo-spat. Infor. Sc., **10**, 3 (2007)
11. S.-J. Ryu, M.-J. Lee, H.-K. L. *Detection of Copy-Rotate-Move Forgery Using Zernike Moments*, in Information Hiding - 12th International Conference, IH 2010, 28-30 June 2010, Calgary, AB, Canada (2010)