HAL open science

# Object Classification Using Sequences of Zernike Moments

Aneta Górniak, Ewa Skubalska-Rafajlowicz

## To cite this version:

Aneta Górniak, Ewa Skubalska-Rafajlowicz. Object Classification Using Sequences of Zernike Moments. 16th IFIP International Conference on Computer Information Systems and Industrial Management (CISIM), Jun 2017, Bialystok, Poland. pp.99-109, 10.1007/978-3-319-59105-6_9 . hal-01656212

**HAL Id: hal-01656212**

**https://inria.hal.science/hal-01656212v1**

Submitted on 5 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

# Object classification using sequences of Zernike moments

Aneta Górniak[1] (0000-0002-6353-5788) and Ewa Skubalska-Rafajłowicz[1] (0000-0002-2795-3835)

Wrocław University of Science and Technology, Wrocław, Poland
aneta.gorniak@pwr.edu.pl, ewa.skubalska-rafajlowicz@pwr.edu.pl

**Abstract.** In this paper we propose a method of object classification based on the sequences of Zernike moments. The method makes use of the pattern recognition properties of Zernike moments and expands it to the problem of classification. Since the distinctive features of the classified objects are carried over to the Zernike moments, the proposed method allows for a robust, rotation and translation invariant classification of complex objects in grayscale images. In this approach, each object class has defined a reference Zernike moment sequence that is used as the prototype of the class. The object's affiliation to the class is decided with the MSE criterion calculated for the object's Zernike moments sequence and the reference Zernike moments sequence of the class. The method is tested using grayscale images of handwritten digits and microscopic sections.

**Keywords:** Image processing ·Microscopic image ·Object matching ·Classification ·Zernike moments

## 1 Introduction

Image recognition is considered one of the main branches of image processing. It may pertain to such fields as pattern recognition and analysis or image description and finds application in many fields of study including medicine, astronomy, digital communication technologies, military industry and many more [12].

The core problem of image recognition lies within recognition of objects and characters regardless of their position, size and orientation in the image. Through the space of time many methods of image recognition were proposed. One of the approaches dedicated to solving this problem, and which is the focus of this paper, revolves around the application of geometrical moments and orthogonal polynomials [12]. Moments have a history of being used in image matching, recognition and classification [5], [9], [15]. They serve such purposes as feature detection, feature description or feature extraction [2]. One of the more explored variants of this approach involves the use of Zernike polynomials as the basis function in these moments [2], [7], [10], [11].

Zernike moments were originally introduced in the 1930s by physicist and Noble prize winner Fritz Zernike to describe optical aberrations. Because of

the orthogonal radial polynomials used as the basis of Zernike moments, the moments do not contain redundant information [11]. Zernike moments are known to be translation, rotation and scale invariant [3], [6], [7], [14]. It is possible to reconstruct an image from the set of Zernike moments [12]. The number of details in the reconstructed image and the resemblance to the original image depend on the level of the order used in the reconstruction process. For the low-order Zernike moments describe the general shape of the image, while the high-order Zernike moments cover more detailed aspects of the image [11].

Nowadays, Zernike moments are used in image shape feature extraction and description or content-based image retrieval [11], the region-based matching [2]. The more specific cases cover matching and recognition of characters and objects [7] or human faces [1], emblem detection and retrieval [4].

In this paper we propose a method of object classification based on the sequence of Zernike moments. Since it has been shown that Zernike moments can be successfully applied to pattern recognition, we chose to expand that property and apply it in the field of classification. The proposed method is based on the notion that the absolute value of Zernike moments of a given order will have an approximate value for similar-looking objects. Having the value of a Zernike moment at the span of a certain number of order levels allows for the construction of a Zernike moment sequence with a distinguished value pattern shared between objects of a similar shape.

The paper is organized as follows. Section 2 contains the necessary definitions of complex Zernike moments and Zernike polynomials. Section 3 shows Zernike moments as image descriptors and how they are applied in the construction of a Zernike moment sequence. Section 4 provides the description to the proposed method of classification using the sequences of Zernike moments. Section 5 describes the applied classification experiments and the data used in those experiments. Finally, in section 6, we present a brief conclusion to the paper.

## 2 Zernike Moments

The basis of a complex Zernike moment is a set of Zernike complete orthogonal polynomials defined over the interior of the unit disc in the polar coordinate space, i.e., $x^2 + y^2 = 1$ [12]. Let us denote the set of Zernike polynomials as $V_{nm}(x, y)$ and defined with

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta). \tag{1}$$

The Zernike polynomial is split into the real part $R_{nm}$ (the radial polynomial) and the complex part $\exp(jm\theta)$. In this equation $n$ is a positive integer and $m$ is a positive (and negative) integer subjected to constraints

$$m \in \{0, \pm 1, \ldots, \pm |n| \quad | \quad n - |m| \quad even\}, \tag{2}$$

$\rho$ is the length of vector from origin to $(x, y)$ pixel, $\theta$ is an angle between vector $\rho$ and x-axis in a counter-clockwise direction. The radial polynomial $R_{nm}$ is defined as

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \rho^{n-2s}, \tag{3}$$

where $R_{nm}(\rho) = R_{n(-m)}(\rho)$.

Let $f(x, y)$ be the continuous image intensity function. The two-dimensional complex Zernike moment of order $n$ and repetition $m$ is defined as

$$A_{nm} = \frac{n+1}{\pi} \int_x \int_y f(x, y)[V_{nm}(x, y)]^* dxdy \tag{4}$$

where $[V_{nm}(x, y)]^*$ is the complex conjugate of Zernike polynomial $V_{nm}(x, y)$ that follows $[V_{nm}(x, y)]^* = V_{n(-m)}(x, y)$.

For the computer digital image, let us denote the intensity of the image pixel as $I(x, y)$, so the equation (4) can be represented as

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y I(x, y)[V_{nm}(x, y)]^* \tag{5}$$

Having all moments $A_{nm}$ of the image function $f(x, y)$ up to the given order of $n_{max}$ it is possible to reconstruct a discrete function $\hat{f}(x, y)$ with matching moments $A_{nm}$ [7]

$$\hat{f}(x, y) \approx \sum_{n=0}^{n_{max}} \sum_m A_{nm} V_{nm}(\rho, \theta) \tag{6}$$

where $m$ is the subject to the same constraints as in equation (2).

## 3 Image Description with the Use of Zernike Moments

Each image can be represented with a sequence of Zernike moments $A_{nm}$

$$\{A_{0,0}, A_{1,1}, A_{2,0}, A_{2,2}, \ldots, A_{n_{max},m}\}, \tag{7}$$

where $n = 0, \ldots n_{max}$ and $m$ is subject to usual constraints (2).

Each moment carries a different piece of information pertaining to the image. A fair number of Zernike moments allows for a detailed image characteristic and the focus on some moments allows for the characteristic of a singular image feature. It is important to note that the number of Zernike moments affects the quality of image reconstruction and shows the way the number of moments may influence the general image characteristics [11].

We used the respective moments $A_{nm}$ for $n = 0, \ldots, n_{max}$ to construct the image characteristic. Since moments $A_{nm}$ are dependent on the values of $n$ and $m$, thus they cannot be unequivocally ordered in a linear manner. Therefore to simplify the image description the moments were grouped by ascending order $n$ following

$$A_n = \sum_m A_{nm}. \tag{8}$$

where $m$ is (2). Since our goal is to have a descriptor that is rotation and translation invariant we used the absolute values of subsequent moments $A_n$, thus the final sequence is of the form

$$\{|A_1|, |A_2|, \ldots, |A_{n_{max}}|\}. \tag{9}$$

In the final image description we skipped the value of $A_0$.

## 4 Applying Zernike Moments to Image Classification

We assume that the images belonging to the same group of objects will share similar features (equation 9). The notion allows for the construction of a referential Zernike moment sequence for a class of images and may allow for image classification based on this sequence (prototype of the class).

The referential sequence (class prototype - $cp$) constructed for the class $k$ of images was obtained as a mean value of all Zernike moment sequences for a given class $k$

$$cp_k = \{\frac{1}{N}\sum_{i=1}^{N}|A_1^i|, \frac{1}{N}\sum_{i=1}^{N}|A_2^i|, \ldots, \frac{1}{N}\sum_{i=1}^{N}|A_{n_{max}}^i|\} \tag{10}$$

where $k$ is the index of the class, $N$ is the number of images used in building of the reference and $A_n^i$ is the value of the Zernike moment of order $n$ of the image $i$ in the series where $i = 1, \ldots, N$ and $n = 1, \ldots, n_{max}$.

In order to use a reference Zernike moment sequence effectively, there is a need to assign a certain margin of error for any classification attempt. Therefore for every point in the reference we calculated an acceptable deviation value from the point for the image being classified

$$d_k = \{d_1^k, d_2^k, \ldots, d_{n_{max}}^k\} \tag{11}$$

The deviation for the point of reference $n$ in the class prototype $k$ is as follows

$$d_k(n) = 3 \cdot \sigma = 3 \cdot \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_n^i - cp_k(n))^2}, \quad i = 1, \ldots, N, \tag{12}$$

where $cp_k(n)$ is the value of $A_n$ from the class prototype $cp_k$ in (10). It is the assumed maximal possible deviation for the reference point $n = 1, \ldots, n_{max}$ in the complete image set of sample size $N$.

The image falls into a class when the absolute difference between the absolute value of its Zernike moments and the value of the reference point falls into the ascribed deviation margin

$$|cp_k(n) - A_n| \leq d_k(n). \tag{13}$$

The need for a different deviation value for every point stems from the Zernike moment's property where the low-order moments respond to a general image feature (like its general shape or size) and go into more detail with the high-order moments (small distortions in the object's general shape). Therefore it is prudent to leave a wider deviation margin for the high-order reference points, where we expect greater differences between images. If the end goal is to classify images only on the basis of the general shape it may be beneficial to cut off the high-order moments completely from calculations. The purpose of the inclusion of this condition is to accept into the class only sequences that follow its Zernike moment pattern. Otherwise, the criterion of the match remains vague and does not discourage the mismatch on the basis of the minor deviations from the reference.

As mentioned earlier we assume that images of the same object will share a similar Zernike moments sequence, therefore the difference between the reference sequence and the image sequence should fall within the permitted deviation margin. As a criterion of classification we used the standard mean square error (MSE) in the form

$$\min_k Err(k) = MSE(k) = \frac{1}{n_{max}} \sum_{n=1}^{n_{max}} (cp_k(n) - A_n)^2. \tag{14}$$

For an object to belong to the class its Zernike moment sequence has to fulfil the $MSE$ criterion from (14) and the deviation condition from (13), otherwise the image remains unclassified.

## 5 Experiments and Discussion

(a) A (b) A (c) B (d) B (e) C (f) C (g) D (h) D (i) E (j) E

Fig. 1: Sample images from classes A to E containing animal shapes. The MPEG-7 Core Experiment CE-Shape-1 Test Set

The experiment data sets consist of images in grayscale. Image pixels are encoded in the range of 0 to 1, where the value of 0 corresponds to black and 1 to white. The data sets vary based on the subject they represent and the size of the images. The first experiment is performed on the series of microscopic images divided into five classes of objects. The size of the image

in this group is 256 × 256 pixels. There are 41 images in this data set (I - 5, II - 10, III - 9 , IV - 9, V - 8). The second experiment is run on the set of digital scans of handwritten digits from number 1 to 9 and the image size of 28 × 28 pixels. There is a total of 900 images in this data set divided into subsets of 100 images for every digit. The data set comes from the MNIST dataset of handwritten digits (http://yann.lecun.com/exdb/mnist/). In the third experiment we used the MPEG-7 Core Experiment CE-Shape-1 Test Set from http://www.dabi.temple.edu/ shape/MPEG7/dataset.html that consists of binary images of object and animal shapes. We chose 5 groups of different animal shapes labelled from A to E and of 20 images each (100 images in the total). The standard image size in this group is 256 × 256 pixels. The fourth experiment was performed on the data set of various object images shown in different stages of 3D rotation. We used 8 complete groups of object sets numbered from o1 to o8, each set containing 72 images of the object. The final data set consisted of 572 object images of the size 128 × 128 pixels. This data comes from the Columbia University Image Library (COIL-20) (http://www1.cs.columbia.edu/CAVE/software/softlib/coil-20.php).

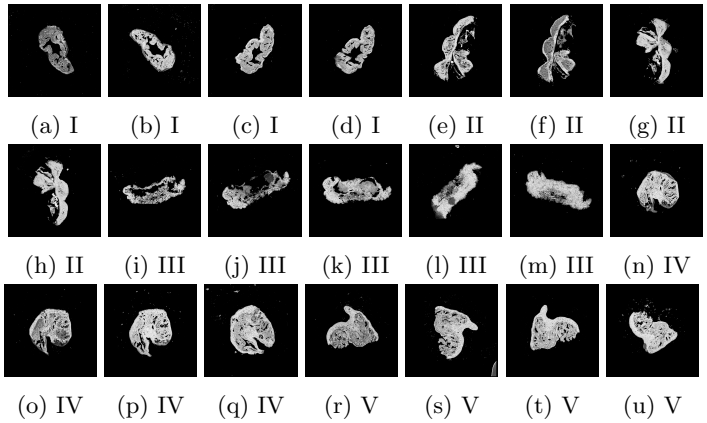



Fig. 2: Sample images from classes I to V containing microscopic images.

The goal of this experiment is to determine when the Zernike moment descriptor can be used to classify object images and when to refrain from using it. The experiment consists of three stages:

1. Calculation of Zernike moment sequence for every image in the data set.
2. Construction of the class prototype using Zernike moment sequence.
3. Testing of image classification on the data set using the reference sequence.

In the beginning, we calculated the Zernike moment sequence for every image in the data set according to (eq. 9). Next, we divided every image set into a learning and a testing set. The learning set is used to create the prototype of

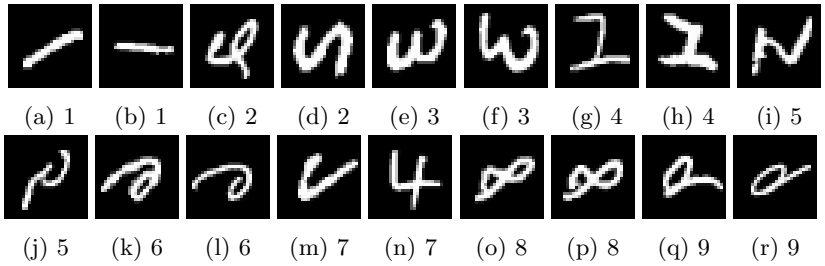



Fig. 3: Sample images from classes 1 to 9 containing handwritten digits images. MNIST dataset of handwritten digits.

the class (eq. 10). The remaining images are put into the testing set. Depending on the sample size, we used either half (the digit and the 3D object data sets) or all the images from the class (the microscopic image and the animal shapes data sets) to obtain the class prototype Zernike sequence. When constructing the sequence it is advisable to use a diversified learning set so that the deviation margin of the reference sequence includes a wide array of class objects. The purpose of dividing the data set into two subsets was to have a reference sequence that is based on actual images from the class and an unbiased testing set to test the classification method.

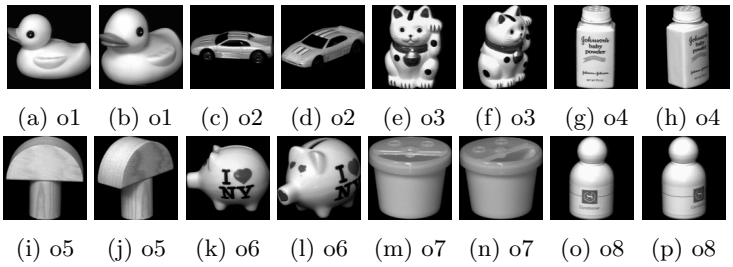
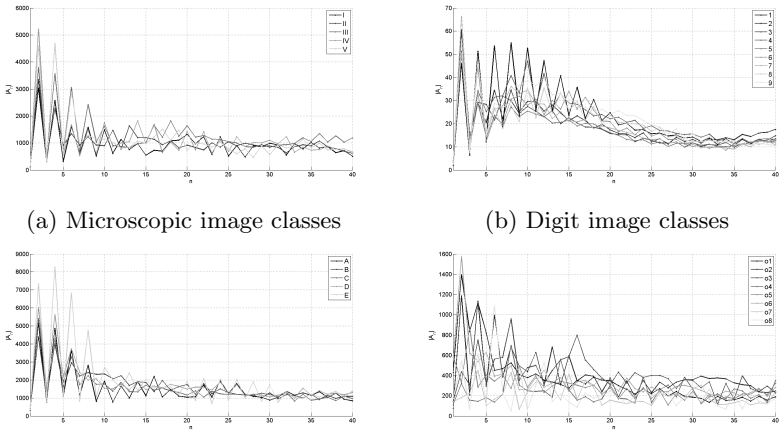



Fig. 4: Sample images from classes o1 to o8 containing 3D object images. Columbia University Image Library (COIL-20).

The prototype Zernike moment sequence for a class was constructed as presented in equations (10) and (12). We constructed 5 prototype Zernike moment sequences for microscopic images classes (denoted from I to V), 9 prototype sequences for digit classes (denoted from 1 to 9 respectively), 5 prototype sequences for animal shape classes (denoted from A to E) and 8 prototype sequences for object images (denoted from o1 to o8). The obtained prototype sequences are presented in Figure 5. Due to the numerical calculation constraint we calculated the Zernike moments up to $n_{max} = 40$. It was noted during the testing that

the longer Zernike moment sequences generated more accurate classification. Therefore, it is advisable to use the sequence of the highest order $n$ possible.




(a) Microscopic image classes

(b) Digit image classes

(c) Animal image classes

(d) 3D object image classes

Fig. 5: Presentation of the Zernike moment sequences for all class prototypes within the four image groups.

We tested the proposed method of classification on the sample and the testing set combined. Due to the restriction caused by the differences in the size of the images we ran four testing experiments: one for each image data set.

We ascertained the accuracy of the classification by the percentage of the images classified correctly as belonging to their class of origin:

$$p = \frac{\text{\# of images classified correctly}}{\text{initial \# of the images in the class}}. \tag{15}$$

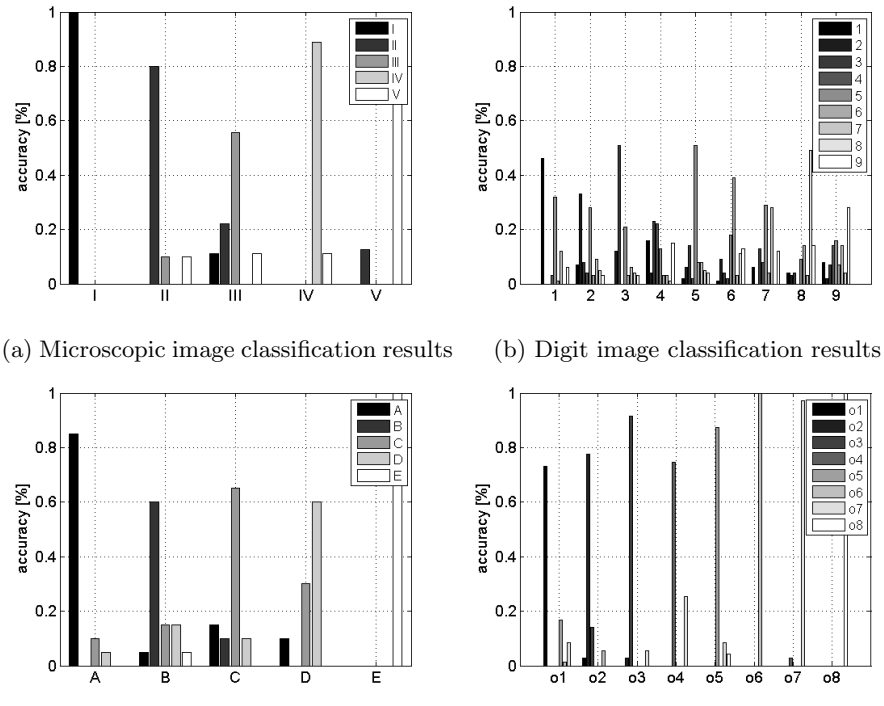
Table 1: Results of classification for all classes.

| **class** | I | II | III | IV | V |
|---|---|---|---|---|---|
| **p(%)** | 1.00 | 0.80 | 0.44 | 0.89 | 0.88 |

| **class** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| **p(%)** | 0.42 | 0.26 | 0.46 | 0.11 | 0.44 | 0.35 | 0.26 | 0.44 | 0.20 |

| **class** | A | B | C | D | E |
|---|---|---|---|---|---|
| **p(%)** | 0.85 | 0.60 | 0.50 | 0.50 | 1.00 |

| **class** | o1 | o2 | o3 | o4 | o5 | o6 | o7 | o8 |
|---|---|---|---|---|---|---|---|---|
| **p(%)** | 0.69 | 0.75 | 0.79 | 0.59 | 0.73 | 1.00 | 0.79 | 1.00 |




(a) Microscopic image classification results

(b) Digit image classification results

(c) Animal shapes classification results

(d) Object classification results

Fig. 6: Presentation of the classification results. The graph shows how the images from the class were distributed by the proposed method of classification.

Results of the classification are shown in Table 1. The first noticeable finding is the discrepancy of results for the handwritten digit data set from other data sets. The accuracy of that classification in every case is below 0.5. The results for the macroscopic image classification (bar sample III) falls into the range of $0.8 - 1.0$ accuracy and for the 3D object image classification it is the $0.7 - 1.0$ range (bar sample o4). The animal shape image classification is the most mixed group as some of the results fall within the $0.85 - 1.00$ accuracy range (samples A and E) and some into the $0.5 - 0.6$ accuracy range (samples B-D). Studying the graphic presentation of the reference Zernike moment sequence of the classes in Figure 5 there is a noticeable correlation between the Zernike moment values and the classification results. When the characteristics share too many similar features there is difficulty with differentiation between classes as is noticeable in the case of the handwritten digit data set and some of the classes in the animal shape data set. The distribution of misclassified objects is shown in Figure 6.

We can differentiate classes using their Zernike moment sequence. In cases of microscopic image series and the 3D object image series the differences between classes of objects is easily noticeable. Therefore, the classification results tend to have a higher level of accuracy. The classes of handwritten digits aren't that distinct and some of the numbers' reference Zernike moment sequences share a similar shape. The type of the object used has an influence on the shape of the characteristic based on the Zernike moment sequence. The handwritten digit image series contains very little distinctive information as most of the analysed objects are composed only of basic curves and lines. Since the module of a Zernike moment is rotation invariant there is little chance to differentiate a six from a nine. That lack of distinctive information is a cause of a common misclassification of images in this data set. A similar case is present in the animal shape classification. The images in this data set, while containing a very diverse representation of shapes within the class, carry only black and white pixel values, which translates into less diverse information to be processed. Both of those properties seem to heavily influence the results of classification and show that the Zernike moments have difficulty with differentiating based only on binary information. The opposite results were shown in the microscopic image data set containing objects that vary in shape, volume and are subject to rotation. However, this data set's images are encoded in grayscale. The diversified pixel values and the information they carry result in higher accuracy of classification, when using Zernike moments. The results of 3D object classification seem to support that notion as most of the images in this data set were also classified correctly. The last two data sets contain objects that are complex, vary in size and are subjected to rotation, both two- and three-dimensional. Despite that the proposed classification method was able to ensue accurate results.

While the $MSE$ criterion allows for accurate classification on its own, it does not ensure that an object from an unknown class will not be classified as one of the available classes. This is the reason why we apply a second condition that demands from the image sequence to mimic the Zernike moment sequence of the class prototype with some deviation. While the $MSE$ criterion will show the

similarity on the level of Zernike moment values, the second condition will keep this similarity within acceptable parameters. The image sequence that is close to the reference sequence in values, but does not follow the pattern of these values will not be classified into the class.

## 6 Conclusion

In this paper we presented a classification method based on the sequences of Zernike moments. As the results show, the proposed method can be applied to classification of images that contain a substantial amount of information to process like images in grayscale. The approach makes use of the distinctive shape and volume of the object that get translated into calculated Zernike moments. The sequence constructed with the Zernike moments carrying such information, allows for a unique image description that is shared between similar-looking objects. The other advantage of using Zernike moments is their scale, translation and rotation invariance property that allows for omitting some of the preprocessing stages of image processing.

The proposed method of classification is shown to be a useful tool that is simple in application and allows for a robust, scale, rotation and translation invariant classification of complex objects in grayscale images.

*Acknowledgements.* The authors express their special thanks to Dr. Agnieszka Malińska and to professor Maciej Zabel from the University School of Medicine, Poznan, Poland for the microscopy section images. This research was supported by 0401/0114/16 grant at the Faculty of Electronics, Wrocław University of Science and Technology.

## References

1. Alirezaee, S., Ahmadi, M., Aghaeinia, H., Faez, K.: A Weighted Pseudo-Zernike Feature Extractor for Face Recognition. IEEE International Conference on Systems, Man and Cybernetics, 3, 2128-2132 (2005). doi: 10.1109/ICSMC.2005.1571463
2. Chen, Z., Sun, S.K., A Zernike Moment Phase-Based Descriptor for Local Image Representation and Matching. IEEE Transactions on Image Processing, 19(1), 205-219 (2010). doi: 10.1109/TIP.2009.2032890
3. Chong, C.W., Raveendran, P., Mukundan, R., Translation invariants of Zernike moments. Pattern Recognition, 36(8), 1765-1773 (2003). doi: 10.1016/S0031-3203(02)00353-9
4. Cura, E., Tepper, M., Mejail, M., Content-Based Emblem Retrieval using Zernike Moments. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, LNCS, 6419, 79-86 (2010). doi: 10.1007/978-3-642-16687-7_15
5. Hu, M.K., Visual pattern recognition by moment invariants. IRE Transactions on Information Theory, 8(2), 179-187 (1962). doi: 10.1109/TIT.1962.1057692
6. Kamila, N.K., Mahapatra, S., Nanda, S. Invariance image analysis using modified Zernike moments. Pattern Recognition Letters, 26(6), 747-753 (2005). doi: 10.1016/j.patrec.2004.09.026

7. Khotanzad, A., Hong, Y.H., Invariant Image Recognition by Zernike Moments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(5), 489-497 (1990). doi: 10.1109/34.55109
8. Kim, W.Y., Kim, Y.S., A region-based shape descriptor using Zernike moments. Signal Processing: Image Communication, 16, 95-102 (2000). doi: 10.1016/S0923-5965(00)00019-9
9. Liao, S.X., Pawlak, M., On Image Analysis by Moments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(3), 254-266 (1996). doi: 10.1109/34.485554
10. Liao, S.X., Pawlak, M., On the Accuracy of Zernike Moments for Image Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(12), 1358-1364 (1998). doi: 10.1109/34.735809
11. Liu, M., He, Y., Ye, B., Image Zernike moments shape feature evaluation based on image reconstruction. Geo-Spatial Information Science, 10(3),191-195 (2007). doi: 10.1007/s11806-007-0060-x
12. Pawlak, M., Image Analysis by Moments: Reconstruction and Computational Aspects. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław (2006).
13. Seo, J.S., Yoo, C.D., Image watermarking based on invariant regions of scale-space representation. IEEE Transaction on Signal Processing, 54(4), 1537-1549 (2006). doi: 10.1109/TSP.2006.870581
14. Suk, T., Flusser, J., Zitova, B., Moments and Moment Invariants in Pattern Recognition. John Wiley and Sons Ltd, Chichester (2009).
15. Teague, M.R., Image analysis via the general theory of moments. Optical Society of America, 70(8), 920-930 (1979). doi: 10.1364/JOSA.70.000920
16. Yoon, Y., Lee, L.K., Oh, S.Y., Semi-Rotation Invariant Feature Descriptors using Zernike Moments for MLP classifier. 2016 International Joint Conference on Neural Networks (IJCNN), 3990-3994 (2016). doi: 10.1109/IJCNN.2016.7727718