

# DEVELOPMENT OF METHODS FOR CRYSTALLOGRAMMS IMAGES CLASSIFICATION BASED ON TECHNIQUE OF DETECTION INFORMATIVE AREAS IN THE SPECTRAL SPACE

N. Kravtsova<sup>1</sup>, R. Paringer<sup>1,2</sup>, A. Kupriyanov<sup>1,2</sup>

<sup>1</sup>Samara National Research University", Samara, Russia

<sup>2</sup>Image Processing Systems Institute - Branch of the Federal Scientific Research Centre  
"Crystallography and Photonics" of Russian Academy of Sciences", Samara, Russia

**Abstract.** We propose a new approach to classifying diagnostic crystallographic images. The classification procedure uses a three-nearest neighbor algorithm based on the Euclidean distance. The image segmentation is conducted in a spatial domain, with energy values in each segment serving as features. Based on the value of the separability criterion used in discriminant analysis, most informative features and their respective segments are selected. With the classification using only informative segments, the classification error is shown to be reduced by 2% when compared with the use of the entire image.

**Keywords:** diagnostic crystallogram, spatial spectrum, discriminant analysis, k-NN classification.

**Citation:** Kravtsova N, Paringer R, Kupriyanov A. Development of methods for crystallogramms images classification based on technique of detection informative areas in the spectral space. CEUR Workshop Proceedings, 2016; 1638: 357-363. DOI: 10.18287/1613-0073-2016-1638-357-363

## 1 Introduction

Biological fluids may serve as indicators of abnormal metabolic processes associated with organ pathologies. The fluid composition is representative of metabolism changes associated with various pathologies. In the case of pathology, there occur multiple changes in the molecular composition of tissue and biological fluids. One way to identify the correlation between the constituent elements in a system involves the phase transition of fluid from one state to another. In the laboratory diagnostics, the phase state of a biological fluid can be most readily changed by crystallization. Changes in the properties of the crystals thus obtained are the result of changing physical-chemical properties of the original biological fluid. Medical crystallograms represent structures formed as a results of salt crystallization when drying out a biological fluid. These structures are characterized by key parameters, such as a predominant direction of crystallization lines and their density at each point in the image. The lach-

rymal fluid turned out to be the most readily available and fairly informative object for crystallographic analysis. The crystallographic method of lachrymal fluid analysis has been recommended as a test for diagnosing inflammatory, tumor, and dystrophic diseases of the organ of vision [1].

In recent years, computer-aided methods of medical image processing have become one of key research tools, enhancing the effectiveness of early diagnostics of various diseases. With an automated analysis, abnormalities in the medical crystallogram structure can be estimated not just qualitatively but quantitatively as well. In clinical practice, the crystallograms are analyzed using their photographs. It is not always possible to identify visually key signs of pathology. This has prompted the use of methods for digital crystallographic image processing. Advantages of the computerized image analysis include its objectivity and feasibility to conduct a quantitative image analysis. In the course of image analysis the following problem arises: the information contained in a crystallographic image is structurally redundant. It has also been known that if the initial crystallographic image is characterized by parallel lines of a definite direction, its Fourier transform is also dominated by the same-direction lines. This property can be put to use when analyzing medical crystallograms [2-5].

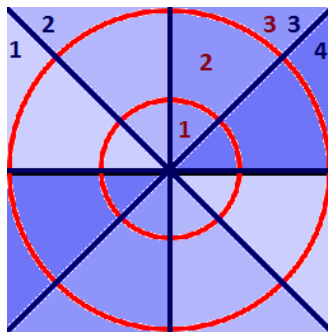
## 2 Description of features used

If the image function and its Fourier transform  $F(u, v)$  are considered in a spatial domain, then the magnitude  $|F(u, v)|^2$  defines an energy spectrum of the image. The energy spectrum of the image can be directly analyzed as a whole or partially.

In this work, we analyzed features derived by calculating the total energy of a selected domain of the spectrum image. The spectrum image in the domain of interest was segmented using a formula:

$$C_{r_1 r_2 \theta_1 \theta_2} = \sum_{r=r_1}^{r_2} \sum_{\theta=\theta_1}^{\theta_2} |F(r, \theta)|^2,$$

where  $r = \sqrt{u^2 + v^2}$ ,  $\theta_1$  и  $\theta_2$  – are the bounding angles of the sector (Fig. 1).



**Fig. 1.** Schematic diagram of the segmentation technique

Since the spectral image symmetrical relative to the center, then you-division signs only half of the image will be used to eliminate the signs of recurrence.

### 3 Implementation of classification and analysis features

In this work, with a view of accumulating statistics we utilized crystallographic images of a lachrymal fluid. The images were divided in two classes based on a visual analysis. The training set was composed of 100 samples of 256×256 pixels, containing 50 images of each class. The images under testing had similar parameters, with the samples containing 50 images in each class. Based on this data, image classification was done. As a criterion of classification quality, we used the classification error, which defines the percentage of faulty decisions and is calculated by the formula:

$$\varepsilon = \frac{m}{n} \cdot 100\% ,$$

where  $m$  is the number of faulty classifications and  $n$  is the general number of the images under testing.

We studied a technique for selecting informative segments in a spectrum image. The informativeness of segments was evaluated using separability criteria of discriminant analysis.

The separability criterion for a sample composed of  $n$  elements grouped into  $g$  classes and containing  $p$  features is given by

$$J = tr((\mathbf{B} + \mathbf{W})^{-1}\mathbf{B}) ,$$

where  $\mathbf{B}$  is the intergroup scattering matrix, whose elements are found from:

$$b_{ij} = \sum_{k=1}^g n_k (\bar{x}_{ik} - \bar{x}_i)(\bar{x}_{jk} - \bar{x}_j), \quad i, j = \overline{1, p} ,$$

$\mathbf{W}$  is the intragroup scattering matrix, whose elements are derived from:

$$w_{ij} = \sum_{k=1}^g \sum_{m=1}^{n_k} (x_{ikm} - \bar{x}_{ik})(x_{jkm} - \bar{x}_{jk}), \quad i, j = \overline{1, p} ,$$

$x_{ikm}$  is the value of the  $i$ -th feature for the  $m$ -th element in class  $k$ ,

$\bar{x}_{ik} = 1/n_k \sum_{m=1}^{n_k} x_{ikm}$  is the average value of the  $i$ -th feature in class  $k$ ,

$\bar{x}_i = (1/n) \sum_{k=1}^g n_k \bar{x}_{ik}$  is the average value of the  $i$ -th feature over all classes,

$n_k$  is the number of elements in class class  $k$ .

For instance, for a set of features whose parameters include 3 rings and 4 sectors (Fig. 1) the following sets were found to be most informative: (sector 1, ring 2), (sector 3, ring 2). Figures 2–4 illustrate images from each class, their spectra, and most informative spectrum fragments.

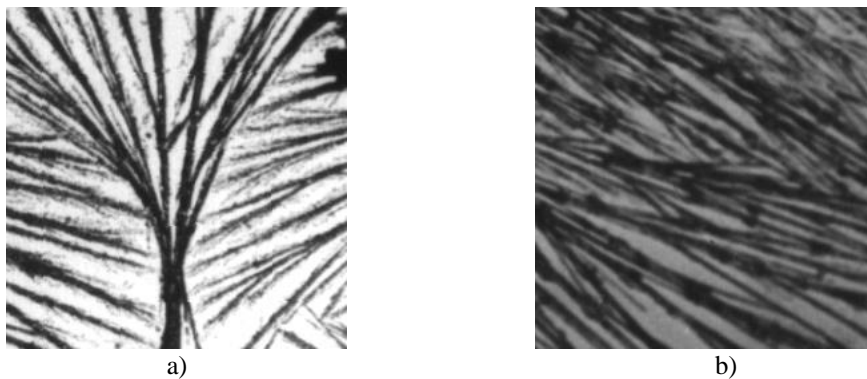
In this work, we presented the results of crystallographic image classification using local features of the spatial spectrum. The classification was done using a three-nearest neighbor technique based on the Euclidean distance. A minimal error of 6%

was obtained for feature sets derived by breaking down the image into 4 sectors and 4-8 rings (highlighted in Table 1) [6-8]. Then, discriminant analysis was conducted [9-11], with the individual separability criterion calculated for each feature. Subsequent classification involved the search of features using the following procedure:

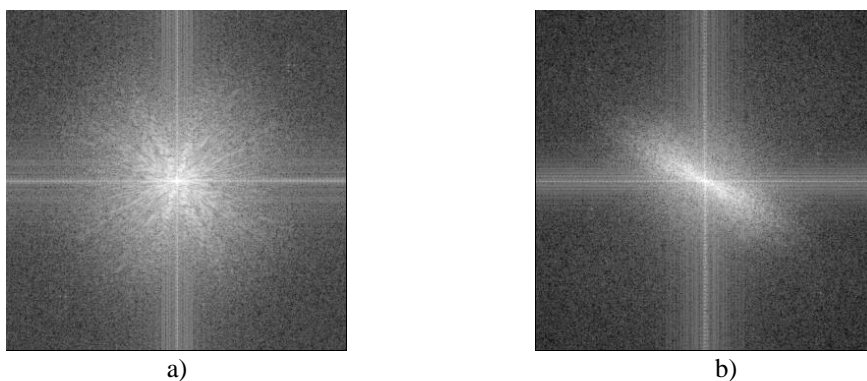
- A single feature with the maximum criterion was selected, with the classification done based on that feature;
- The feature set was extended by adding one more feature with the second maximum separability criterion, with the classification done based on two features;
- The feature set was extended again by adding the third maximum separability criterion, with the procedure being reiterated until all features had been added.

Figure 5 illustrates in which way the number of features affects the value of error for the image broken down into 4 sectors and 8 rings.

In our study, the result characterized by the least error was taken as a classification result following the feature selection (Table 2). The least classification error is highlighted in Table 2.



**Fig. 2.** Image examples for each class: (a) class 1 and (b) class 2



**Fig. 3.** Examples of spectrum images for each class: (a) class 1 and (b) class 2

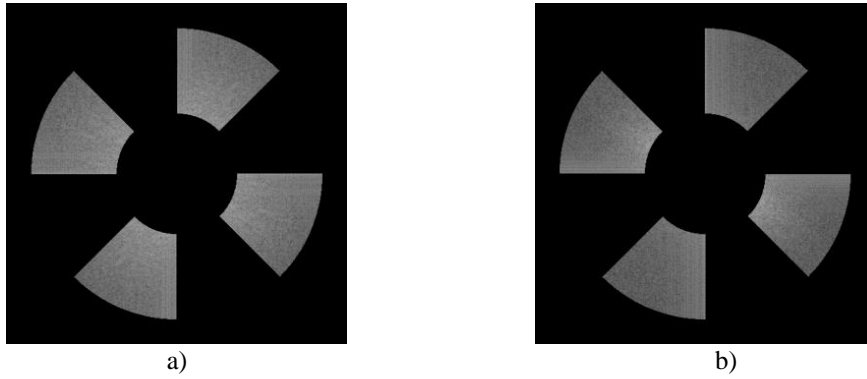


Fig. 4. Images of informative segments for each class: (a) class 1 and (b) class 2

Table 1. Classification error prior to feature selection, %

		Number of rings							
		1	2	3	4	5	6	7	8
Number of sectors	1	10	12	11	11	11	11	12	12
	2	9	7	8	8	8	8	7	8
	3	9	7	7	7	7	7	7	7
	4	9	7	7	6	6	6	6	6
	5	7	7	7	7	7	7	7	7
	6	8	7	7	7	7	7	7	7
	7	8	7	7	7	7	7	7	7
	8	7	7	7	7	7	7	7	7

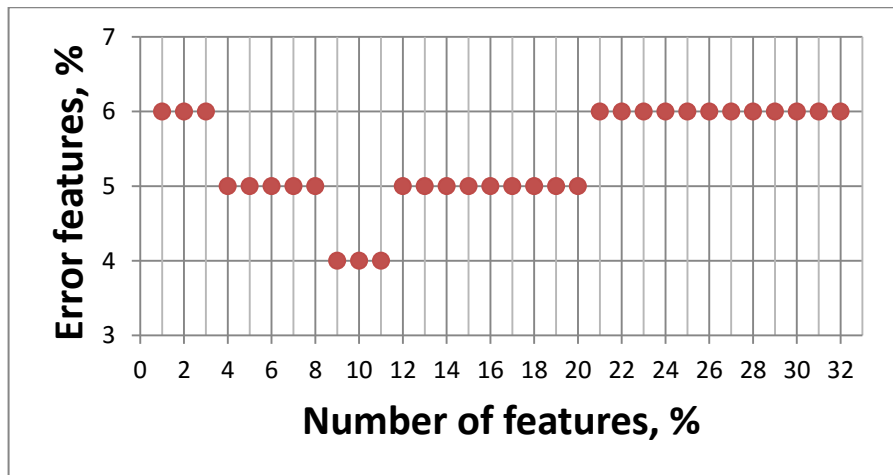


Fig. 5. Error vs. the number of features

**Table 2.** Classification error following feature selection, %

		Number of rings							
		1	2	3	4	5	6	7	8
Number of sectors	1	9	10	9	9	9	9	10	10
	2	8	8	7	7	7	7	7	7
	3	8	7	6	6	7	7	7	7
	4	7	7	6	6	5	5	4	4
	5	7	7	7	6	6	6	7	7
	6	7	7	7	7	7	7	7	7
	7	7	7	7	7	7	7	7	7
	8	7	7	7	7	7	7	7	7

## 4 Conclusion

We have proposed a technique for selecting informative segments in the spectrum image. The study was conducted using medical crystallogram images. With the entire spectrum image broken down into segments used for classification, the breakdown into 4 sectors and 5-8 rings has been found to be optimal, with the classification error being equal to 6%. The use of the separability criterion borrowed from discriminant analysis for selecting the informative features has made it possible to reduce the classification error. For the image broken down into 4 segments and 5-6 rings, following the selection and use of informative features, the error was reduced from 6% to 5%, while the breakdown composed of 4 sectors and 7-8 rings allowed the error to be reduced from 6% to 4%.

The analysis of the results suggests that post-selection informative feature sets enable a more accurate classification of medical crystallograms of a lachrymal fluid.

## Acknowledgements

This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 14-07-97040, # 15-29-03823, # 15-29-07077, # 16-57-48006); by the ONIT RAS program # 6 "Bioinformatics, modern information technologies and mathematical methods in medicine" 2016.

## References

1. Ilyasova NYu, Kupriyanov AV, Khramov AG. Information technologies of image analysis in medical diagnostics. Moscow: "Radio i svyaz" Publisher, 2012; 424 p. [in Russian]

2. Paringer RA, Kupriyanov AV. The Method for Effective Clustering the Dendrite Crystallogram Images. 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding "OGRW 2014", Electronic on-site Proceedings. University of Koblenz-Landau, 2014.
3. Paringer RA, Kupriyanov AV. The method of forming features for image classification dendritic crystallograms. Information technologies and mathematical science and management, Irkutsk, Russia, 2015; 1: 50-55. [in Russian]
4. Paringer RA, Kupriyanov AV. Research methods for classification of the crystallogram images. Proceedings of the 12th international conference "PRIP'2014", Minsk, Belarus, 2014: 231-234.
5. Paringer RA, Kupriyanov AV. Developing improve image quality classification methods dendritic crystallograms. Collection of scientific works of the international scientific technical conference "PIT-2014", Samara, Russia, 2014: 282-286. [in Russian]
6. Kravtsova NS, Paringer RA, Kupriyanov AV. Assessing local features space-governmental spectrum for classification of texture images. Samara: Proceedings of the international youth-term scientific conference "XIII Korolyovskie chteniya", Samara, Russia. SSAU, 2015; 2: 134-135. [in Russian]
7. Kravtsova NS, Paringer RA, Kupriyanov AV. Research of algorithms of classification of texture images based on the analysis of the spatial. Methods of computer diagnostics in biology and Medicine, Saratov, Russia. Publishing House of Saratov source, 2015: 190-193. [in Russian]
8. Kravtsova NS, Paringer RA, Kupriyanov AV. Developing technologies detection of informative segment in the spectral region images for classification crystallograms. Collection of scientific works of the international scientific technical conference "PIT-2016", Samara, Russia, 2016: 104-106. [in Russian]
9. Fukunaga K. Introduction to statistical pattern recognition. San Diego: Academic Press, 1990; 592 p.
10. Ilyasova NYu, Kupriyanov AV, Paringer RA. Formation of features for improving the quality of medical diagnosis based on discriminant analysis method. Computer Optics, 2014; 38(4): 751-756. [in Russian]
11. Ilyasova NYu, Kupriyanov AV, Paringer RA. The Discriminant Analysis Application to Refine the Diagnostic Features of Blood Vessels Images. Optical Memory & Neural Networks (Information Optics), 2015; 24(4): 309-313.