DIAGNOSTIC ANALYSIS USING TEXTURAL FEATURES OF THE LACHRYMAL FLUID CRYSTAL IMAGES

Alexander Kupriyanov, Nataly Ilyasova, and Alexander Malapheyev

Image Processing Systems Institute of Russian Academy of Sciences
Molodogvardeyskaya st. 151, 443001, Samara, Russia
phone: + (7) 8462 325622, fax: + (7) 8462 325620, email: akupr@smr.ru

ABSTRACT

The application of the direction field method and the statistical textural analysis for crystallograms classification is proposed. As global features, the features the expert uses for the crystallogram-based classification of the eye pathology were taken: unidirectedness of the crystal rays, relative area of crystal domains with clear-cut rays, ray density, crystal transparency. As textural features, the characteristics of the second-order distribution were taken. Experimental studies were conducted on the lachrymal fluid crystallograms.

1. INTRODUCTION

Pathological conditions cause multiple changes in the molecular composition of tissue and biological fluid. Many authors believe that biological fluids [1] (blood, saliva, urine, and others) are indicative of metabolism impairment caused by the pathology in a human organ.

Crystallographic studies (CS) are used as an integrated method that allows one to make implicit conclusions about the matter structure. The CS of biological fluids can provide information that would allow a more accurate diagnostic of inflammatory, cancer, dystrophic, and allergic diseases. In clinical practice, crystallogram photographs are analyzed. It is very difficult if not impossible, to visually single out the critical pathological signs.

In this connection, computerized methods for crystallogram image processing are becoming important tools of scientific research and enhancement of early diagnostics. Our studies aim to develop methods for the automated analysis of crystallograms, investigate their diagnostic value, and generate quantitative integrated estimates of pathology probability on the basis of the crystallogram classification features.

According to the crystallographic analysis method used in the clinical setting, the normal fluid crystallogram is transparent and comprises thin, mostly unidirectional, clear-cut rays originating from a common crystallization center. Pathological crystals feature a great variety of directions and irregular contours. The pathological crystal is opaque, with numerous ray fractures and bulges. A distinctive feature of pathology is the large density of crystal rays in some areas (fig. 1). Thus, by analyzing the crystallograms the ophthalmologist had classified as those with and without pathology, we were able to extract the global features the expert uses for the crystallogram-based classification of the eye pathology: unidirectedness of the crystal rays; relative area of domains with clear-cut rays of the crystal; ray density; and crystal transparency [2].

2. FORMATION OF CRYSTALLOGRAM GLOBAL DIAGNOSTIC FEATURES

The quasiperiodic structure is an important feature of the crystallogram images[3]. Because of this, most classification features we discuss are based on the notion of the complex direction field [4] derived from the function of image intensity $I(x,y)$:

$$\dot{\psi}(x,y) = w(x,y)\exp\big(i2\psi(x,y)\big),\ 0 \le w(x,y) \le 1. \quad (1)$$

$$\operatorname{tg}\psi(x,y) = -\frac{\partial I(x,y)/\partial x}{\partial I(x,y)/\partial y},\ 0 \le \psi(x,y) < \pi. \quad (2)$$

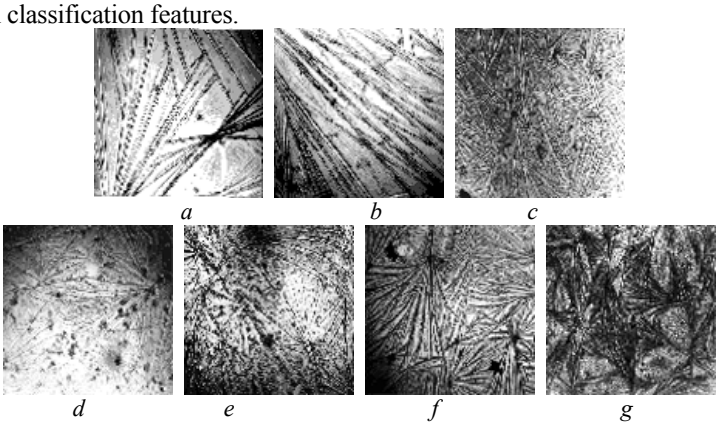


Figure 1. Crystallogram samples in normal condition (a-b) and pathological condition (c-g)

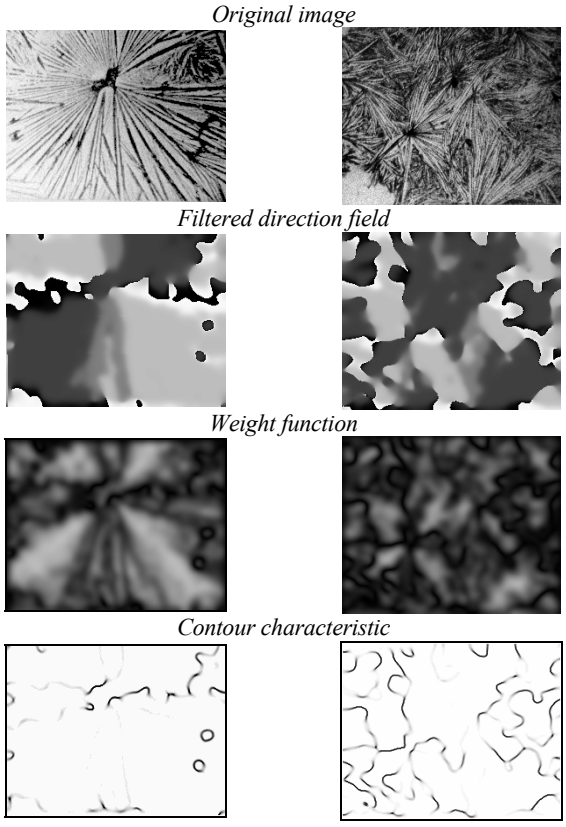
The direction field $\psi(x,y)$ represents the tangent angle to the level lines of the intensity function; the weight function $w(x,y)$ stands for the certainty (reliability) in the determination of the direction field at a given point.

The coefficient of the crystals unidirectedness is defined as:

$$K_1 = \frac{1}{|D|}\iint_D \gamma(x,y)dxdy\,,$$

$$\gamma(x,y) = \left(\frac{\partial \sin\psi(x,y)}{\partial x}\right)^2 + \left(\frac{\partial \sin\psi(x,y)}{\partial y}\right)^2 \cdot \left(\frac{\partial \cos\psi(x,y)}{\partial x}\right)^2 + \left(\frac{\partial \cos\psi(x,y)}{\partial y}\right)^2$$

Table 1. Characteristic crystallogram images and the direction fields in normal condition (left column) and pathological condition (right column)



The contour characteristic of the direction field of the first pathological crystallogram is depicted in table 1.

To quantify the domains with the pronounced unidirectedness of lines, we use the coefficient of clear-cut lines K_2 defined as the ratio of the total area S_p of domains with the greatest the weight values of the direction field to the entire image area S : $K_2 = S_p / S$

Quantitatively, the line density feature in the crystallogram is found to be based on the frequency properties of the image intensity function. As a classification criterion, we take here the mean value of the ray density over the image domain D wherein the weight function takes its greatest value and the value of spectral frequency is certain.

The image intensity function is considered to be locally periodic and admitting the following approximation:

$$I(x,y) = A\sin(\omega_x x + \omega_y y + \varphi) + B\,,$$

where ω_x and ω_y are the spatial projections. The coefficient of the line density K_3 is defined as the mean value of the squared spatial frequency of the crystallogram intensity function:

$$K_3 = \frac{1}{|D|}\sum_D \omega^2\,, \quad \omega^2 = \omega_x^2 + \omega_y^2\,.$$

The crystallogram transparency is characterized through the probability distribution of the intensity function. The "transparent" crystallogram features a positive shift of the mean value of intensity $\overline{I}$ with respect to the midpoint $I_c = (I_{\max} + I_{\min})/2$ of the intensity range. This criterion can be quantified by the coefficient: $K_4 = (\overline{I} - I_c)/\overline{I}$.

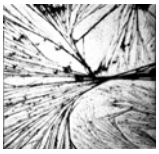
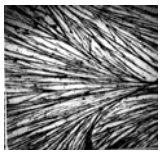
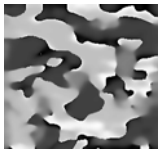
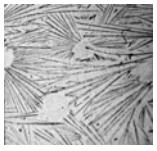
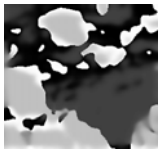
The probability of crystallogram's being pathology-free was used as a criterion for the independent classification by each feature. In particular, the probability is equal to one if the value of the criterion is greater than the threshold of norm, and it is equal to zero if the criterion is smaller than the threshold of pathology. In the intermediate range the dependence is linear. For each feature, the threshold of norm and pathology is chosen from the condition of the minimum classification error under the given criterion. The final estimate of the pathology probability depends on the partial estimates of the pathology probabilities derived from each criterion.

Experimental studies [2] have shown that the above-considered features have different weights upon the crystallogram diagnostics. Thus the weight coefficients for each criterion were taken to be proportional to the quality of the classification.

The classification results are shown in Table 2. In the table, the column Type indicates the a priori estimate of an image by the ophthalmologist (N - norm, P - pathology); P1 through P4 are the probabilities of norm according to the corresponding classification criteria; R1, R2 are the resulting estimates of the probability of norm obtained via different techniques for combining the classification criteria (R2 is for the optimal combination); C1 and C2 show whether the classification result corresponds to the a priori estimate, provided that threshold is 0.6.

The global diagnostics on a series of samples (150 crystallograms) has made it possible to extract from a variety of crystallograms the normal and pathological groups and quantify the classification features. The error in the pathology recognition in the crystallograms with quasiperiodic structures did not exceed 3-5%. A more detailed processing based on a series of local features will make it possible in the future to go to the differential diagnostics, thus diagnosing separate groups of diseases: tumors, dystrophic and inflammatory diseases. The objective of the next section is studies and formalization of these diagnostics features.

Table 2. The results of classification on the learning sample.

Image	Direction field	Type	P1	P2	P3	P4	R1	C1	R2	C2
		N	0.634	0.4	0.964	1	0.72	+	0.736	+
		N	1	1	0.44	0.294	0.72	+	0.783	+
		P	0	0	1	0	0.15	+	0.319	+
		N	0.863	1	0.476	0.824	0.83	+	0.751	+
		P	0.614	0.525	0.456	1	0.67	-	0.576	+

3. FORMATION OF CRYSTALLOGRAM LOCAL TEXTURAL DIAGNOSTIC FEATURES

The image texture is analyzed to provide a series of features for the classification of the eye fluid crystallograms according to the familiar types of pathologies. The textural features were formed on the basis of human visual perception, so the aim was to extract the information that a human interpreter associates with the texture.

The different images represent a particular texture for each class of the crystallograms, which is a global representation of the crystal. A clinical expert extracted seven main classes according to the severity of pathology. The first two classes form a norm group (fig. 1a-1b). The last five classes form a pathology group (fig. 1c-1g). The crystallograms available for this study came from 70 patients with different types of pathologies. The texture analysis was carried out on the images of lachrymal crystallograms using the second-order statistics of the gray levels. The gray-level-cooccurrence features [5] have proven to be very successful in the extraction of textural information [6].

To describe the gray-level cooccurrence (GLC) matrix, we need the following definitions and symbols: D is the image field containing $M \times N$ pixels, $x_{m,n}$ is the gray value of pixel coordinates $(m,n) \in D$, G is the number of the gray levels in the image.

The indicator-function (3) shows whether two neighboring pixels at the distance d have the determined levels:

$$f_{i,j}(x_{m,n}, x_{m+k,n+l}) = \begin{cases} 1; & x_{m,n} = i \ , x_{m+k,n+l} = j \\ 0; & x_{m,n} \neq i \quad or \quad x_{m+k,n+l} \neq j \end{cases}, \quad i,j = 0,\ 1,\ ...,\ \ G-1. \tag{3}$$

The normalized values of the GLC matrix are defined as

$$P_{k,l}(i,j) = C_{k,l}(i,j) \Big/ \sum_i \sum_j C_{k,l}(i,j) \ ,$$

$$C_{k,l}(i,j) = \sum_{(m,n)\in D} \sum_{(m+k,n+l)\in D} f_{i,j}(x_{m,n}, x_{m+k,n+l}) \ ,$$

$$k,l = 0,\ \pm 1,\ \pm 2,\ ... \quad .$$

The dimension of matrix $P_{k,l}$ is $G \times G$. The distinction between the opposite directions was disregarded. Therefore, the symmetrical matrices $P^s_{k,l}$ were generated as follows: $P^s_{k,l} = \left(P_{k,l} + P_{-k,-l}\right) \Big/ 2$. For the present purposes, we chose to avoid the characterization of texture in a given direction. Each calculated matrix P^s_d is the average of four matrices calculated in the four directions (0°, 45°, 90° and 135°): $P^s_d = (P^s_{d,0} + P^s_{d,d} + P^s_{0,d} + P^s_{-d,d}) / 4$.

A set of statistic features $F = \left(f_1^d, \ldots, f_6^d\right)$ was calculated to summarize the GLC matrix. These features are textural features[7] (Table 3). *Variance* describes the degree of image homogeneity. *Contrast* describes the degree of image contrast. *Inertia* describes the presence of sharp edges. *Correlation* describes the degree of statistical dependence of pixels. *Shade* describes the degree of equiprobable appearance of dark and bright areas in the image (near-object shadows, etc.). *Entropy* is the measure of image disorder.

The K-nearest neighbours (KNN) method was applied to a series of images. The classifier was developed using randomly selected 50% of the data set, with the testing performed with the remaining 50% of data. With the KNN method, each pattern of the training set is stored as a prototype. The class of a new pattern is directly obtained from the computation of the distance between this pattern and each prototype in database. Among the KNN, the majority class is ascribed to the unknown pattern.

Table 3. The average nonnormalised values of each feature for major groups - norm and pathology

Feature	Norm	Pathology
Variance $f_1^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\left[P_d^s(i,j)\right]^2$	0,06	0,03
Contrast $f_2^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\lvert i-j\rvert P_d^s(i,j)$	0,08	0,14
Inertia $f_3^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)^2 P_d^s(i,j)$	2,1	4,2
Correlation $f_4^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-M_x)(j-M_x)P_d^s(i,j)$	25	40
Shade $f_5^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i+j-M_x)^3 P_d^s(i,j)$	1,5	3,7
Entropy $f_6^d = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\ln\left[P_d^s(i,j)\right]P_d^s(i,j)$	7,6	9,1

As the number of samples in each class was relatively small, the classification was conducted ten times with various randomly selected testing and training sets. The classification rates presented in Table 4 correspond to the average results obtained in the experiments [7].

These results indicate that the method of textural analysis can be used to identify the class of the crystallogram, and, hence, determine the severity and type of pathology, with a relatively small probability of false miss errors.

Table 4. The classification results of the 1144 crystallograms samples

Group	Norm	Pathology
Correctly classified	334	810
Percentage of correct classification	87%	98%

4. CONCLUSION

Methods of the direction field and statistical textural analysis have been used to construct a classifier that allows the lachrymal fluid crystallogram type to be identified. A conceptual possibility to use the technique for disease diagnostics has been proved. Some experiments yielded an accuracy of near 95%.

It is possible to construct an expert system to diagnose the pathology type of biological liquid crystallograms. Additionally, the effectiveness and informativeness of the features were studied using the discriminant analysis method [7]. Since certain features are highly correlated with others the classification quality can be further improved. In future research, we propose to use a combination of textural and direction field analysis.

Acknowledgements: This work was financially supported by the U.S. Civilian Research and Development Foundation (CRDF) and the Ministry of Education of the Russian Federation international grant REC-SA-014-02 under Basic Research and Higher Education (BRHE) program.

References

1. R. J. Morris Statistical pattern recognition for macromolecular crystallographers, Acta Cryst. (2004). D60, 2133-2143
2. Ilyasova, N.Yu., Dvoryanova, T.P., Ustinov A.V., Khramov, A.G.: Computer-aided Analysis of the Tear-Crystal Image. Proceedings of 13-th biennal international conference Biosignal-96, Brno, (1996) 29-31.
3. D.S. Raicu, J.D. Furst, D. Channin, D.H. Xu, & A. Kurani, A Texture Dictionary for Human Organs Tissues' Classification. Proceed. of the 8th World Multiconf. on Syst., Cyber. and Inform., July 18-21, 2004.
4. Soifer, V.A., Kotlyar, V.V., Khonina, S.N., Khramov, A.G.: The Method of the Directional Field in the Interpretation and Recognition of Images with Structure Redundancy. Pattern Recognition and Image Analysis, v.6, No.4, (1996) 710-724.
5. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans.Syst.Man Cybern. v.3, (1973), 610-621.
6. O.Basset, F.Dupont, A.Herandez, C.Odet, Texture image analysis: application to the classification of bovine muscles from meat slice images // Optical. Engineering, v.38(11), p.1950-1959, 1999.
7. Kupriyanov, A.V., Khramov, A.G., Ilyasova, N.Yu.: Analysis of Features of Texture Images for Crystallogram Identification and Classification. Optical Memory & Neural Networks, vol.11, No.1, (2002) 19 – 28.