FRACTAL CHARACTERISTICS OF CELL NUCLEUS CHROMATIN DENSITY USING FOR THE INTERPHASE CONDITION DETERMINATION

Alexandr SHUMEYKO*, Gennadiy SHEVCHENKO, Stanislav GERASIMENKO****

* City Kamianske, Dniprovsk Technical University, Computer Science Department
** City Dnipro, Noosphere company, Scientific Center

Summary. Medical images digital analysis is gaining popularity due to the increase in information presented in the form of images and the automating possibility of their processing at the computer technology development current level. Cells characteristics classification is a powerful non-invasive tool for determining diseases associated with changes in the cellular structure, which requires researchers to introduce new features most accurately reflecting the studied disease into the analysis procedure. Cell nuclei chromatin distribution fractal characteristics are a fairly accurate interphase state indicator which detection is an important stage for subsequent analysis and diagnosis.

Keywords: Medical images analysis, diseases early diagnosis, fractal dimension, digital image processing, extraction of cell characteristics.

1. INTRODUCTION

Fractal properties studies (including those for the purpose of early cancer diagnosis) of chromatin distribution in human cell nuclei have recently become of specialists' interest [1,2,11]. This subject is of interest due to the possibility of diagnostic process automation that comes to the corresponding human cells digital images analysis [4,5]. One of the key elements of this approach is the selection of the digital image area containing the cell nucleus for the extracted fragment subsequent analysis. The study of the chromatin density distribution in the cell nucleus is complicated by the fact that this value is different in different periods of cell life (see, for example, [4]). The process of protein biosynthesis occurs in the cell and all the major cell structures are doubled during the interphase. An exact original chromosome copy from the chemical compounds existing in the cell is synthesized along the original chromosome and the DNA molecule is doubled. The doubled chromosome consists of two halves that are chromatids. Each of the chromatids contains one DNA molecule. On average the interphase lasts 10-20 hours in the plant and animal cells. Then comes mitosis that is the process of cell division. Mitosis is a continuous process, but it is divided into four stages for convenience, depending on the nucleus chromatin look at that time. Prophase, metaphase, anaphase and telophase are distinguished in mitosis process. Chromosomes shorten and thicken due to their spiralization during the prophase. The fission spindle formation is completed and it consists of two microtubules

types: chromosomal microtubules that are bound to chromosome centromeres, and centrosome microtubules (pole) that are extended from the pole to the pole of the cell during the metaphase. Chromatids doubled in chromosomes interphase diverge to the poles of the cell during the anaphase. There are two diploid chromosomes sets in the cell at that moment. During the telophase the processes that are reverse to those observed in the prophase occur: the chromosomes despiralization (unwinding) begins. Nuclear envelope is formed from the cytoplasm membrane structures around the chromosomes at each pole and small nuclei arise in nuclei. The fission spindle is destroyed. The cytoplasmis divides (cytotomy occurs) and two cells are formed during the telophase. As a result of mitosis, two daughter cells appear from the mother cell with the same set of chromosomes.

2. FRACTAL DIMENSIONALITY USING FOR CELL NUCLEUS STUDYING

The chromatin density distribution changes significantly at each stage of mitosis. If chromatin is evenly distributed in the cell nucleus during the interphase, then during the prophase it is gathered into "lumps" which during the menafase and anaphase are broken up into parts and pressed against the opposite sides of the cell border.

Considering the problem of chromatin density distribution studying we are interested only in the interphase period of the cell life cycle. Thus, the method for cell nuclei chromatin density distribution studying should take this feature into account. In our case the input data are presented in the full-color images forms obtained as a result of cytological research.

The fractal halftone images analysis is an effective method of this kind of research [1, 2,4,11].

First of all, these are the following methods.

The BOX COUNTING method. This method is applied to the image of any structure on the plane and can be adapted to 3D objects [3,7]; it is related to the similarity dimension and allows to determine the fractal dimension of not strictly self-similar objects such as, for example, metal structures [9]. For the box-dimensionality estimation the Euclidean space containing the object image is separated by a grid with a cell of size r and non-empty squares quantity $N(r)$ occupied by the investigated object are counted. Then the size r is reduced, and the number of non-empty fields $N(r)$ is counted again. The slope of the graph in the logarithmic scale $N(r)$ on $1/r$ corresponds to the dimension value. As can be seen from the definition, it is necessary to carry out the segmentation of the image for this method that substantially worsens the obtained fractal characteristic.

The 2D Variation [6] method is designed for processing binary and grayscale images. In this case, the gray colour minimum and maximum levels are found within each square cell with the side r. Thus, two-dimensional minimum and maximum functions are defined for each grid spacing.

Then, the volume difference between the minimum and maximum functions is

determined for the entire image. Then, $V(r) = r^s const$ where *V* is the "difference" volume, *r* is the grid spacing, and *S* is the fractal dimension.

A similar procedure is repeated for all pixels of the image that are sufficiently far from the border, then *r* is increased and the procedure is repeated. The slope *s* of the graph *V(r)* on a logarithmic scale is used to determine the fractal dimension:

$D_{2D} = 3 - \frac{s}{2}$

This method allows us to evaluate both the dimension of each RGB color channel separately, and the dimension of the image "brightness".

Another method for estimating the grayscale images fractal dimension is the Brownian dimension [8], which is proportional to the Hurst exponent *D=3-H*. The parameter *H* can be calculated statistically from the dispersion law using the properties of the Brownian motion:

$H = -\frac{log(\sigma_{\Delta R}(\Delta I))}{log(\Delta R)}$

$\sigma_{\Delta R}(\Delta I)$- Root-mean-square increments deviation, i.e. brightness difference in points $R + \Delta R$ and R, $\Delta I = I(R + \Delta R) - I(R)$, *R* - coordinates of a point.

The algorithm of calculations is as follows: for all points of the image (that are sufficiently far from the edge), brightness increments ΔI are calculated in the ΔR- neighborhood, the root-mean-square deviation is calculated for the resulting array of increments $\sigma_{\Delta R}(\Delta I)$, then we change ΔR and repeat the previous steps. The graph slope of the $\sigma_{\Delta R}(\Delta I)$ on ΔR dependence in logarithmic coordinates gives us the value of *H*.

2.1. THE BUILDING OF THE CHROMATIN DENSITY DISTRIBUTION CHANGE CHARACTERISTICS IN A GIVEN DIRECTION

For the solved problem the complexity of these characteristics using is that the obtained values characterize the image fractality as a whole, but it is important to determine the local unevenness of the color distribution (chromatin density) or at least the unevenness in one direction that is typical for cell mitosis.

Let the monochromatic image of the cells nuclei be given.

$K_k = \{i = 0, 1, ..., n_k - 1\}(k = 1, 2, ..., N)$

Using the pixel color value as the corresponding material point mass, we can find the center of mass

$X_k = \frac{1}{n_k}\sum_{i=0}^{n_k-1} x_i^k, Y_k = \frac{1}{n_k}\sum_{i=1}^{n_k-1} y_i^k$

And we draw a bundle of straight lines through this point

$y = tg\varphi(x - X_k)$, where $\varphi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

For each fixed φ, the points masses of the set K_k that are on the line $y = tg\varphi\left(x - X_k\right)$ represent a one-dimensional array $P(\varphi) = \left\{P_0, P_1, ..., P_{N(\varphi)}\right\}$.

Let $\Delta P(\varphi) = \frac{1}{N(\varphi)}\sqrt{\sum_{i=0}^{N_v-1}\left|P_i - P_{i+1}\right|^2}$. If the chromatin in the cell nucleus is uniformly distributed, then the density change $\Delta P(\varphi)$ in any direction is approximately the same, that is, if $\Pi(\Delta P(\varphi))$ is a Hardy permutation $\Delta P(\varphi)$, then it contains a large horizontal (or almost horizontal) section. We approximate $\Pi(\Delta P(\varphi))$ to a broken line $\lambda\left(\varphi, \left\{\tau_i\right\}_{i=0}^{3}, t\right)$ with three links and boundary conditions (to remove the emissions, both large and small)

$$\lambda\left(\varphi, \left\{\tau_i\right\}_{i=0}^{3}, \tau_0\right) = 0, \lambda\left(\varphi, \left\{\tau_i\right\}_{i=0}^{3}, \tau_3\right) = 1 \text{ where } \tau_0 = 0,\ \tau_3 = N_v.$$

So that for

$$\left\|\Pi(\Delta P(\varphi)) - \lambda\left(\varphi, \left\{\tau_i\right\}_{i=0}^{3}\right)\right\|_2 \to min \text{ by } \tau_1, \tau_2 \qquad (1)$$

We need additional constructions to find this problem solution.

Let $\Delta_n: 0 = t_0 < t_1 < ... < t_{n-1} < t_n = T$ be an arbitrary partition of the $[0, T]$ segment and let $S_r\left(\Delta_n\right)$ be the set of all splines of the r order of the minimal defect in the Δ_n partition, that is, the set of functions with continuous *(r-1)*th derivative on the $[0, T]$ interval that coincide on each interval $\left(t_{i,n}, t_{i+1,n}\right)(i = 0, 1, ..., n-1)$ with an algebraic polynomial of at most r degree.

We denote by $P\left(\Delta_n\right)$ the operator mapping C^{ρ} in $S_r\left(\Delta_n\right)$, $\rho \geq r - k$. For a fixed r, the sequence $\left\{P^*(\Delta_n^*)\right\}_{n=1}^{\infty}$ will be called the asymptotically best spline regression for the function $x(t)$ if for $n \to \infty$

$$\int_a^b \left(x(t) - P^*\left(\Delta_n^*, t\right)\right)^2 dt = \inf inf\left\{P\left(\Delta_n\right) \in S_r\left(\Delta_n\right)\right\}(1 + o(1)) \qquad (2)$$

Let on the $[0, T]$ segment at the nodes of the partition Δ_n be given values of the function x_i, $i = 0, 1, ..., n$. A cubic spline $s_3\left(x, \Delta_n\right) = s_3\left(\left\{x_i\right\}_{i=0}^{n}, \Delta_n\right)$ is called an interpolation with boundary conditions $s_3^{(i)}\left(x, \Delta_n, 0\right) = x(0), i = 0, 1$

and $s_3^{(i)}\left(x, \Delta_n, T\right) = x(T), i = 0, 1$ if

$$s_3\left(x, \Delta_n, t_i\right) = x_i, i = 0, 1, ..., n$$

If $x_i = \frac{T_i}{n}, i = 0, 1, 2, ..., n$ then instead of $s_3^{''}\left(x, \left\{\frac{T_i}{n}\right\}_{i=0}^{n}, t\right)$ we write $s_3\left(\left\{x_i\right\}_{i=0}^{n}\right)$.

The following assertion is proved in [11].

Theorem 1. Let $x \in L_2^3$ be such that $x^{''}$ vanishes on a finite number of segments or points from $[0, T]$ and $X(t)$ is a second antiderivative $x(t)$ such that $X^{'}(0) = 0$ and $X(0) = 0$. We choose M_i from the conditions

$$\lambda_i M_{i-1} + 2M_i + \mu_i M_{i+1} = \frac{6}{h_{i-1}+h_i}\left(\frac{X_{i+1}-X_i}{h_i} - \frac{X_i - X_{i-1}}{h_{i-1}}\right), i = 1, 2, ..., n-1$$

where $\mu_i = \frac{h_{i-1}}{h_{i-1}+h_i}, \lambda_i = 1 - \mu_i$.

Then

$$inf\left\{s_1(\Delta_n) \in S_1(\Delta_n)\right\} = \left\| x - s_3^{''}\left(\left\{M_i\right\}_{i=0}^{n}, \Delta_n\right)\right\|_2$$

We choose the nodes t_i^* of the partition Δ_n^* from condition

$$\int_0^{t_i^*}\left(\left|s_3^{''}\left(\left\{M_i\right\}_{i=0}^{n}, t\right)\right| + \frac{1}{n^\gamma}\right)^\alpha dt = \frac{i}{n}A_n, \; A_n = \int_0^T\left(\left|s_3^{''}\left(\left\{M_i\right\}_{i=0}^{n}, t\right)\right| + \frac{1}{n^\gamma}\right)^\alpha dt,$$

then for $n \to \infty$

$$inf\left\{s_1(\Delta_n) \in S_1(\Delta_n)\right\} = \left\| x - s_3^{''}\left(\left\{M_i\right\}_{i=0}^{n}, \Delta_n^*\right)\right\|_2 (1 + o(1)) == \frac{\theta}{n^2}\left\|x^{''}\right\|_\alpha + o\left(\frac{1}{n^2}\right),$$

where $\theta = \frac{1}{2!}\left(\int_0^1 t^2(1-t)^2 dt\right)^{1/2}$ and $\alpha = 2/5$.

Using the Theorem 1 results, we solve problem (1). The solution are the values of τ_1 and τ_2 that realize the extremum. Then the value of

$$\delta = \frac{\lambda\left(\varphi, \left\{\tau_i\right\}_{i=0}^3, \tau_2\right) - \lambda\left(\varphi, \left\{\tau_i\right\}_{i=0}^3, \tau_1\right)}{\tau_2 - \tau_1}$$

will characterize the uniformity of the chromatin distribution; the closer it is to zero, the

more uniformly (regardless of the direction) it is distributed, and the value of $\theta = \lambda(\varphi, \{\tau_i\}_{i=0}^{3}, \tau_1)$ characterizes the density value, that is, the greater θ, the more chromatin color jumpings, and the value of $\vartheta = \tau_2 - \tau_1$ characterizes the presence or absence of emissions; the closer ϑ to $N(\varphi)$, the less the data emissions.

Note that a large value of δ is typical of the menafase and anaphase states, a small value θ is typical for cells in the prophase state, and a small value of ϑ can indicate that we observe the nucleus telophase.

Thus, on the one hand, the obtained characteristics using allows to discard those cells that are in the mitosis state, and on the other hand, to obtain a characteristic of the chromatin density.

Moreover, we can use the fractal dimension in a given direction for density estimation. We use the following construction.

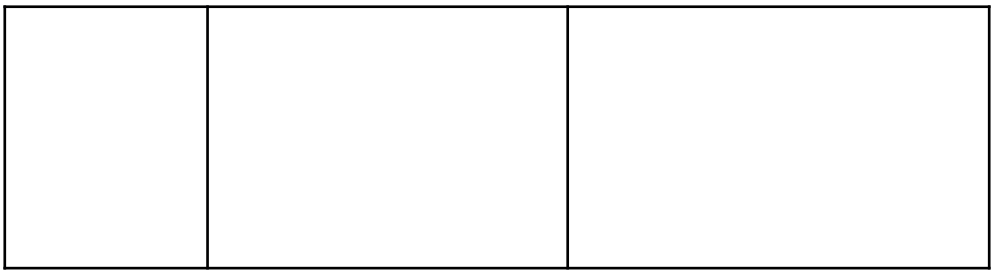
Let $r \ll N(\varphi)$ and $P_i = 0$ for $i > N(\varphi)$, then for $i = 0, 1, ..., [N(\varphi)/r]$

$m_i = min\{P_{ir,} P_{ir+1}, ..., P_{(i+1)r}\}$, $M_i = max\{P_{ir,} P_{ir+1}, ..., P_{(i+1)r}\}$ and $V_i(r) = (M_i - m_i)r$.

Further, let $V(r) = \sum_{i=0}^{[N(\varphi)/r]} V_i(r)$. A similar procedure is repeated for all image pixels that are sufficiently far from the border, then r is increased and the procedure is repeated. The slope s of the graph $V(r)$ on a logarithmic scale is used for the fractal dimension determination.

As in the previous case, if the chromatin is distributed in the core uniformly, then the fractal dimension is the same in any direction.

Figure 1 shows the cell nucleus fractal characteristics in the interphase state, and Figure 2 shows the cell in the mitosis state.



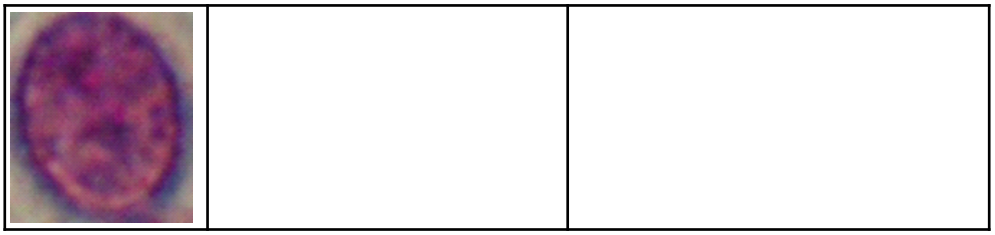


Fig. 1. Cell characteristics in the interphase state

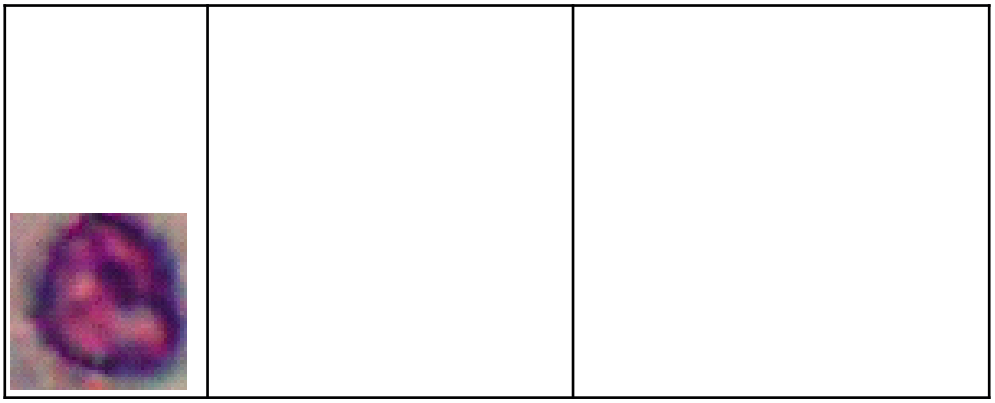


Fig. 2. Cell characteristics in the mitosis state

2.2. CELLS DETERMINATION CHARACTERISTICS BUILDING IN THE INTERPHASE STATE

The problem is simplified a little bit in the case when the cells isolation is required in the interphase state (which is the most often case in point).

As can be seen from the chromatin density graph behavior, if these data are described by quadratic regression

$$\sum_{i=0}^{100} \left(\Pi\left(\Delta P\left(\varphi_i\right)\right) - ai^2 - bi - c\right)^2 \rightarrow$$

then the coefficients a, b, c will characterize the chromatin density change (For the cells shown in Figure 1 and Figure 2, the corresponding coefficients are given in Table 1).

Table 1

Parabolic regression coefficients of chromatin density

	a	*b*	*c*
Figure 2a	0.0015835	0.2163586	1.5100646
Figure 2b	0.0037519	0.1356285	4.4715336

The best characteristic is the leading coefficient. For a test sample this value ranges from 0.000804 to 0.0061442.

The coefficients permutation characterizing fractality is well described by the cubic

regression model (see Tab. 2).

Table 2
Cubic regression coefficients of chromatin fractality characteristics

	a	*b*	*c*	*d*
Figure 2a	0.000000284	0.0000076194	0.0109206	-0.2296844
Figure 2b	0.0000003088	0.0000083126	0.0117783	-0.2497191

Thus, the quadratic regression leading coefficients of the sorted chromatin densities array and the cubic regression leading coefficients of the sorted array of coefficients characterizing fractality can indicate the cell interphase state.

3. CONCLUSIONS

Using the change characteristics of the chromatin density distribution in the cell nucleus allows obtaining efficient algorithms for cells isolating in the interphase state on digital images.

4. REFERENCES

[1] Acharya R. et al.: *Analysis of cardiac health using fractal dimension and wavelet transformation,* ITBM-RBM, Vol. 26, Issue 2, 2005, p. 133-139.

[2] Einstein A. J. et al. *Fractal characterization of chromatin appearance for diagnosis in breast cytology,* The Journal of Pathology, Vol. 185, Issue 4, 1998, p. 366-381.

[3] Falconer K. *Fractal Geometry. Mathematical foundations and applications,* John Wiley & Sons, New York, 1990.

[4] Klyushin, D. A. et al. *A new breast cancer screening method based on interphase bucal epithelium nuclei fractal analysis,* Journal of Computational and Applied Mathematics, No. 1, 2011.

[5] Klyushin D. A. et al. *Screening of breast cancer using Peano curve*, Journal of Computational and Applied Mathematics, No. 4, 2012, p. 117-121.

[6] Kraft R. *Estimating the Fractals and Dimensions,* Journal of agricultural, nutritional and environmental sciences, No. 10, 1996.

[7] Kronover R.M. *Fractals and chaos in dynamic systems,* POSTMARKT, Moscow, 2000.

[8] Ligun A.A., Shumeiko A.A. *Asymptotic methods for reconstructing curves,* Institute of Mathematics of the National Academy of Sciences of Ukraine, Kiev, 1997.

[9] Mihalev A.I., Derevyanko A.I., Pomulev V.V. *Fractal estimation of metal structures granularity,* System technologies. Regional intercollegiate collection of scientific works, Issue 2 (10), Dnepropetrovsk 2000, p. 104-112.

[10] Shumeyko A.A., Shumeyko E.A. *About the asymptotically optimal piecewise linear regression construction*, Informatics and Mathematical Methods in Modeling, Vol. 1, No.2, 2011, p.99-106.

[11] Smith T. G. et al. *A fractal analysis of cell images*, Journal of Neuroscience Methods,

Vol. 27, Issue 2, 1989, p. 173-180.