Automated Extraction of Data for Construction of Parkinson's Disease Experimental Model

I. Gurevich[1)], I. Koryabkina[1)], E. Kozina[2)], A. Myagkov[3)], H. Niemann[4)], M. Ugrumov[2)], and V. Yashina[1)]

1) Dorodnicyn Computing Centre of the Russian Academy of Sciences, Vavilov st. 40, 119333 Moscow, Russian Federation, igourevi@ccas.ru, werayashina@gmail.com

2) Koltzov Institute of Developmental Biology of the Russian Academy of Sciences, Vavilov st. 40, 119334 Moscow, Russian Federation, e.a.Kozina@gmail.com, michael.ugrumov@mail.ru

3) Lomonosov Moscow State University, Leninskie Gory GSP-1, 119992 Moscow, Russian Federation, aem.istranet@gmail.com

4) Friedrich-Alexander University of Erlangen-Nuremberg, Martensstr. 3, Erlangen, 91058, Germany, niemann@informatik.uni-erlangen.de

Abstract: *This paper is devoted to the description of automated method for experimental data acquisition which is required to fill the model of Parkinson's Decease preclinical stage. Digital images of the immunostained brain sections of experimental animals are used as a data source. Proposed method: 1) is based on the following mathematical morphology operations: opening, grayscale reconstruction, closing, bot-hat transformation, morphological gradient, watershed transformation; 2) enables: to smooth heterogeneous complex background; to select small objects on images depending on given sizes and gray values; to eliminate out-of-focus objects; to separate close objects; to calculate features of selected objects; 3) is intended for automatic extraction of dopaminergic neurons terminals on striatum frontal section images. Experimental investigations confirmed that the developed method supports automated processing and analysis of section images*

***Keywords*:** image mining, mathematical morphology, grayscale reconstruction, segmentation, Parkinson's disease.

1. INTRODUCTION

Now and in foreseeable future an image is one of the main tools to represent information in scientific researches, particularly in medicine and biology. Thus development and application of modern mathematical apparatus for automation of image mining becomes one of the breakthrough challenges for theoretical computer science. Paper authors developed theoretical basis [2] and elements of information technology [5] for automated morphological analysis of lymphoid cell nuclei of diseased hemoblastoses, which were fundamentals for creation of system for automated diagnostics of oncological blood diseases. This paper is devoted to development of mathematical tools and elements of information technology that supports automatic extraction of experimental data for filling a model of preclinical stage of Parkinson's disease (PD) [1].

The disease is characterized by a progressive degeneration of dopaminergic (DA-ergic) neurons [6, 8] in the substantia nigra pars compacta (SN) leading to a dopamine (DA) depletion in the striatum. As a result, parkinsonian patients lose the ability to control their movements. Construction of the experimental models is crucial for the research of neurodegenerative disease pathogenesis.

Application of the developed method allows one to estimate quantitatively a) a degeneration of DA-ergic neurons in the substantia nigra and their axons in the striatum after specific dopaminergic neurotoxin – 1-methyl-4phenyl-1,2,3,6-tetrahydropyridine (MPTP) administration and also b) a functional condition of dopaminergic neurons and axons remained after MPTP administration. The model shows differences of DA-ergic neurons features between experimental (the group of animals that were injected by the toxin) and control (the group of animals that were not injected by the toxin) groups.

The initial data is digital images of the immunostained sections of various brain areas. DA-ergic neurons were labeled on serial sections (a thickness of 20 microns) of the substantia nigra (Fig. 1),

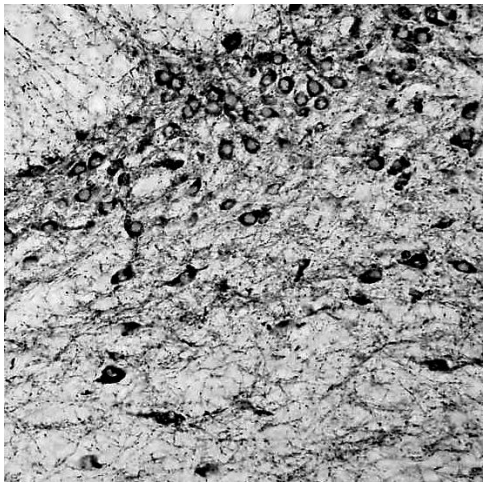

Fig. 1. Neurons

and their fibers (axons) on sections of the striatum (a thickness of 12 microns) (Fig. 2) by immunohystochemestry for tyrosinehydroxylase (TH) (TH is the specific enzyme of DA synthesis). Experimental data has been received from digital images of distal parts of axons (terminals).

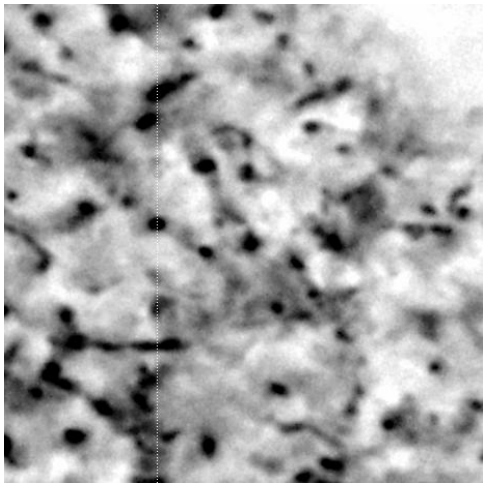

Fig. 2. Terminals

The major characteristic of PD model is the number of DA-ergic axons, which innervate the striatum, when various schemes of MPTP administration (a dose, quantity of injections, intervals between injections) are used. The extent of degeneration is defined as a difference between the number of terminals of DA-ergic axons in control and in experimental groups. DA-ergic neurons and axons remained after MPTP administration are supposed to have to increase its functional activity to compensate DA deficiency. One of the indicators of increased functional activity of neurons and their fibers is their size increases. TH (DA synthesis key enzyme) concentration increase is supposed to be another specific indicator of functional activity of DA-ergic axons and neurons.

The development of PD preclinical stage model is a complex screening analysis which is being done cooperatively by physiologists, biochemists and morphologists. The morphological research for constructing the model requires processing and analysis of a great many serial brain sections images. Studies of each section require quantitative and qualitative feature measurements and analysis of several thousand neurons and axons.

Creating of mathematical methods providing automation of specified image analysis is a requirement for an effective construction and research of adequate PD preclinical stage models. Automated extraction of information from medical images is based on the joint use of image processing techniques, mathematical theory of image analysis and mathematical theory of pattern recognition. The designed methods are given by the specialized algorithmic schemes, that include the following principal stages of automated extraction of information from images: 1) preprocessing (image quality enhancement, irrelevant details and artifacts elimination, statistical and logical filtering); 2) image analysis (object detection and their edge extraction; segmentation; the choice and the estimation of features describing the structure and the content of the image et al.); 3) construction of object representations; 4) classification of images and objects presented on the images; 5) recognition. Application of these schemes allows one to speed up significantly the research of PD at the cost of automatization of experimental data model filling and automatization of model investigation by means of computer experiments.

The proposed method provides automated analysis of digital images of distal parts of axons and allows one to estimate the number and the significant features of terminals. The detailed description of the method is given below in Sec. 2. Section 3 reviews the results of the method application.

2. METHOD DESCRIPTION

The proposed method is designed for the isolation of the small informative elongated objects on the frontal striatum section images and for located objects feature calculation. Images and objects represented on that images were characterized as follows: a) initial image resolution is 0.0117 $\mu m^2/pixel^2$; b) terminals (Fig. 2) are rounded objects with area varying from 0.6 - 0.7 μm^2 up to 2.5 – 3 μm^2; c) terminals color differs from background color; d) terminals can have oval, round, prolate or irregular shape. The method provides segmented object contours and object features. The automated segmentation method include 7 following stages: 1. opening by reconstruction; 2. bot-hat transformation by dual reconstruction; 3. closing by dual reconstruction; 4. H-dome elimination transformation; 5. object and background marker extraction; 6. morphological gradient image modification; 7. watershed segmentation.

The method stage description is outlined according to the following scheme: 1) general description of the applied transformation or algorithm; 2) mathematical content; 3) the significance of the particular transformation while task solving. Some stages of the method are supplied either by plots of intensity function for a particular column or by particular step resulting image. Intensity function corresponding to particular stage is plotted with solid line, this of previous step (if any) is depicted by dashed line and dash-dot line is used to demonstrate the partial result while applying of transformation. In the equations presented below the grayscale reconstruction [9] of I from marker J is denoted by $\rho_I(J)$ (similar the dual reconstruction [9] is denoted by $\rho_I^*(J)$).

2.1 OPENING BY RECONSTRUCTION

Grayscale opening by reconstruction, denoted by $I \circ \rho B$, consists of the following: an initial image I is eroded [7] by a flat structuring element B , then the obtained image is used as marker to reconstruct the initial image.

$$I \circ_\rho B = \rho_I(I \ominus B) \quad (1)$$

In the general case one can use opening by reconstruction operation [7, 9] for intensity narrow peaks elimination while preservation of the background mean intensity values and wider peaks, with the help of an appropriately sized structuring element B. Note the difference between this transformation and ordinary opening. While in the opening, the dilation of the eroded image by the same structuring element only partially recovers intensity values of regions, not entirely smoothed by the erosion, the opening by reconstruction allows one to avoid such drawbacks.

Thereby, the first method step is intended for

elimination of initial image narrow peaks, corresponding to the background. The results carrying out the present transformation are depicted on Fig. 3.

The erosion was done by flat disk structuring element with radius that is greater than the narrowest terminal width and that is smaller than the most prolate terminal length. This step is essential for the reduction of the background regions. There are lots of intensity local minima in these regions, that are used as markers of the objects in the next.

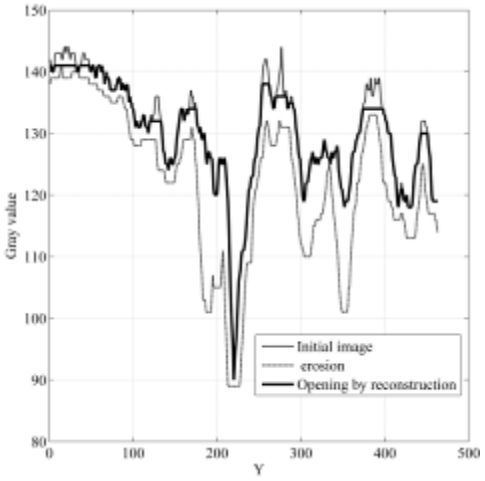

Fig. 3. Opening by Reconstruction. Plot of intensity function. X=453

2.2 BOT-HAT TRANSFORMATION BY DUAL RECONSTRUCTION

Grayscale bot-hat transformation by dual reconstruction [11, 9] is the subtraction of the input image from closed by dual reconstruction image (see Sec. 2.3).

$$BotHat_{\rho}^{B}(I) = \rho_I^{\star}(I \oplus B) - I \qquad (2)$$

Bot-hat by dual reconstruction is used for heterogeneous complex background removal, when grey values of objects are less than those of the background. Choosing of an appropriately sized structuring element, closing by dual reconstruction allowed one to mark narrow troughs while not marking wider troughs in the image. It produces acceptable background approximation. And the subtraction partially corrects the uneven background. The use of dual reconstruction is necessary to preserve intensity values in regions that are not entirely smoothed by the dilation.

The main goal of this step is the correction of complex heterogeneous background of the initial image. Fig. 4 reveals the results of bot-hat by dual reconstruction application.

While applying this transformation the inner structure of terminals remains unchanged. It is achieved by using of the dual reconstruction and by the fact, that the used structuring element is grater than almost all terminals.

2.3 CLOSING BY DUAL RECONSTRUCTION

Grayscale closing by dual reconstruction, denoted by $I \bullet_{\rho} B$, consists of the following: an initial image I is dilated [7] by a flat structuring element B , after that the obtained image $I \oplus B$ is used as marker to reconstruct the initial image.

$$I \bullet_{\rho} B = \rho_I^{\star}(I \oplus B) \qquad (3)$$

In the general case, with the help of an appropriately sized structuring element B, one can use closing by dual reconstruction operation [7, 9] for intensity narrow troughs elimination while preservation of mean intensity values and wider troughs. Note the following. While in the ordinary grayscale closing, the erosion of the dilated image by the same structuring element only partially recovers intensity values of regions, retained after the dilation. The closing by dual reconstruction allows one to avoid such drawbacks.

This step of the method is aimed to nonuniform regions smoothing in the interior of the terminals.

The structuring element was chosen to be completely contained in all possible terminals. This step is essential for providing robust marking of terminals procedure. Whereas terminal initially has many intensity local minima, the marker extraction procedure will not give appropriate results without the use of current operation.

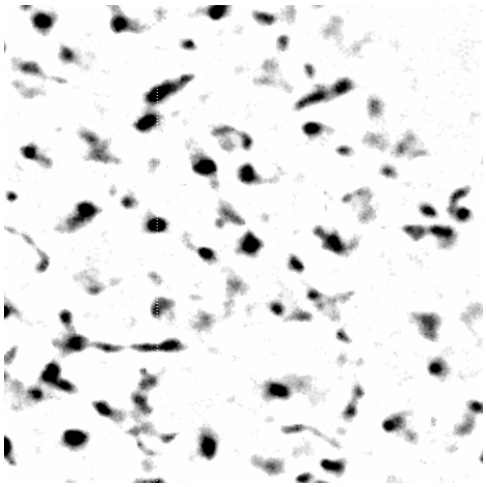

Fig. 4. Bot-Hat by Dual Reconstruction

2.4 H-DOME ELIMINATION TRANSFORMATION

Reconstruction proved to be a very efficient technique to extract regional maxima and minima from grayscale images. Moreover, the method extends to the determination of such “maximal structures”, which are called h-domes and h-basins [9]. As it was shown in [9] the binary image (mask) M(I) of the regional maxima of I is given by:

$$M(I) = I - \rho_I(I - 1) \qquad (4)$$

Then, the h-dome $D_h(I)$ image of the h-domes of a grayscale image I was defined as follows:

$$D_h(I) = I - \rho_I(I - h) \qquad (5)$$

And consequently h-dome elimination is the subtraction of an h-dome image from the initial image.

In contrast to the top-hat transformations, the h-dome elimination transformation removes light structures without involving any size or shape criterion. The only parameter h is related to the height of these structures.

Terminals are located at different depth on 12 microns thickness section of the striatum. While taking photos of a section the only section plane is in the microscope focus, some terminals can have intensity values greater than the others. In addition, sections can be nonuniformly stained, that also involves differences in terminal intensity values.

But at first, it is nessesary to detect in-and out-of- focus objects. And it is the aim of the current step of the method. The image with eliminated h-domes is presented on Fig. 5.

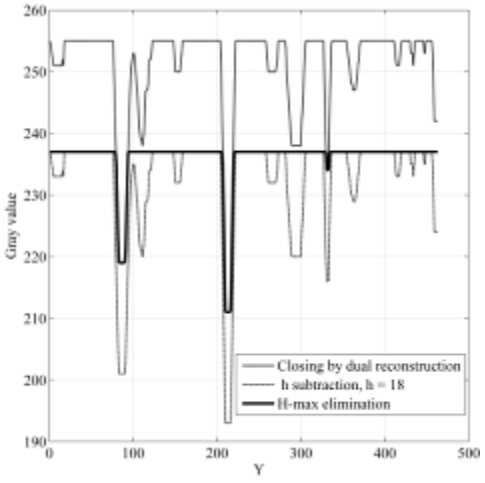

Fig. 5. H-dome elimination Plot of intensity function. X=149

A technique for h-parameter estimation was offered for automation of the segmentation procedure. It proceeds on the idea of the selected marker intensity values clustering into two groups. It was noticed that the closer an object to microscope's focus the greater difference between the maximum and the minimum of the intensity values inside the object. As a result of previous transformations, maxima of all remained objects became equalized, and the focus closeness can be measured by minimal intensity values of the objects which are revealed as regional minima. The initial values for clusters' centres is assigned with minimum and maximum intensity values of the previous step resulting image.

To sum up, h-dome elimination corresponds to out-of-focus objects removal. H-parameter estimation technique agrees with hand-selected in-focus objects.

2.5 OBJECT AND BACKGROUND MARKERS EXTRACTION

All the previous steps were intended to avoid oversegmentation when applying watershed transformation to morphological gradient image (see Sec. 2.7). The oversegmentation is caused by noise or other local irregularities on the gradient image. A very effective way to reduce oversegmentation is based on the idea of markers [4, 11]. A marker is a connected component of pixels of the image. Objects (inner) and background (outer) markers are distinguished. Typically the outer markers surrounds inner ones. Then they are used for gradient image transformation. Applying the watershed segmentation algorithm to the modified gradient, only marked objects are selected.

Object markers are extracted as regional minima of the previous stage resulting image. Background markers are estimated from the distance transform [3] of object markers binary image as given below. The distance transform assigns each pixel of the image the distance to the closest object pixel. Then the watershed segmentation algorithm is applied to received image. Watershed lines correspond to pixels that are located at the maximum distance from the closest objects markers.

2.6 MORPHOLOGICAL GRADIENT IMAGE MODIFICATION

After marker extraction, grayscale reconstruction is used to modify the gradient image G into an image G" [9] so that:

- its only minima are located on the extracted markers,
- its watershed lines separating markers are preserved.

Morphological gradient, denoted by G, is obtained as the difference between the dilation and the erosion of the Step 3 (Sec. 2.3) resulting image by the same structuring element.

$$G(p,q) = (I \oplus B)(p,q) - (I \ominus B)(p,q) \qquad (6)$$

Let m be the maximal value of the pixels of G. Then the binary marker image M, revealed at the previous step, is used to modify G in the following way:

$$G' = \rho^{*}_{\min\{G+1,(m+1)M\}}((m+1)M) \qquad (7)$$

In this process, pixels located on markers are given value 0 in G' and non-marked catchment basins get filled up. It is shown on Fig. 6.

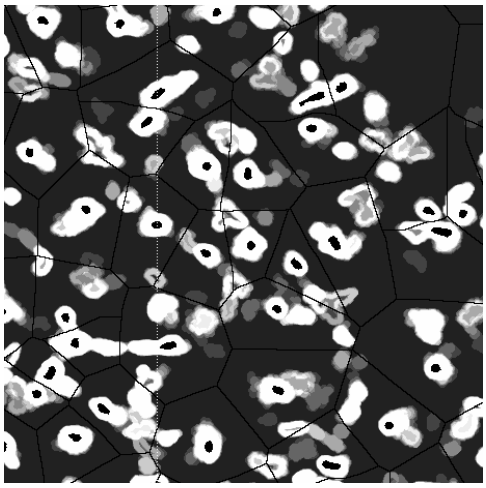

Fig. 6. Modification of Gradient Image

2.7 WATERSHED SEGMENTATION

Watershed notion [3, 10] is based on the interpretation of an image as the 3D surface. In such a "topographic" interpretation points of three types are regarded:

1. local minima;
2. points, located on slopes, i. e. the ones, from which water is slide into the same local minimum;
3. points, located on peaks, i. e. the ones, from which water is slide into one of the local minima with equal possibility.

Concerning a particular local minimum, a set of points, satisfying second condition, is called catchment basin (or spillway). Sets of points obeyed the third condition form highest crest-lines and are called watersheds.

In practice watershed segmentation algorithm is applied to a gradient image. In such a case basin local minima are well matched with low intensity values of the gradient image. This is often corresponds to the objects, whereas, watershed lines, agreed with high intensity values, are close to object boundaries.

So, the modified gradient is processed with watershed segmentation algorithm and object contours are retrieved.

3. EXPERIMENTAL CHECK OF PROPOSED METHOD

The initial image with white color marked object boundaries, revealed during the proposed method application, is presented at Fig. 7. Fig. 8 depicts manually extracted objects for the same image.

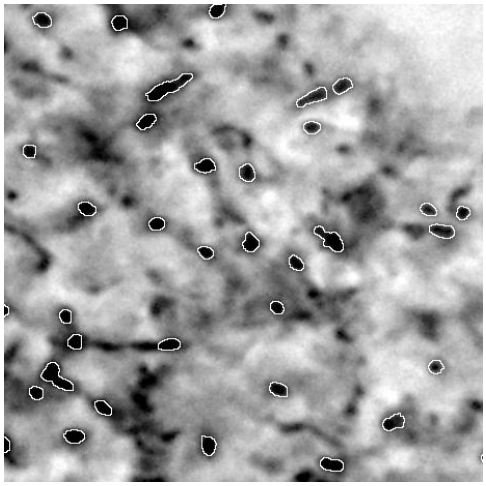

Fig. 7. Automatic extraction

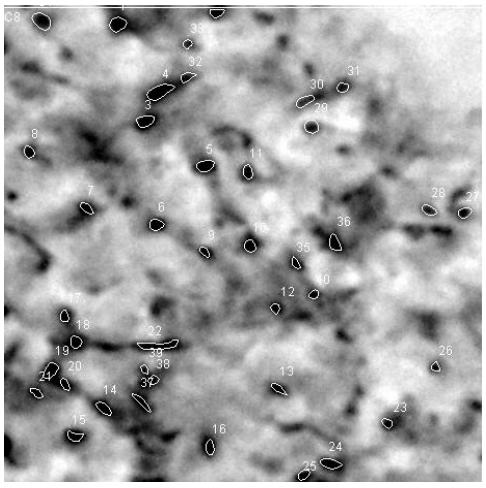

Fig. 8. Manual extraction

With the aid of the developed method a ten striatum frontal section images (five belong to experimental group and five – to control group) were analyzed. Main subtasks of this analysis were the following: automated and manual extracted objects comparison; terminal features calculating; detecting the differences in experimental and control groups. The following data was found for each image and for each group of images:

1. selected objects numeric features;
2. averaged numeric features;
3. results of hypothesis tests to compare the distributions of feature values in manual and automated calculations;
4. explanatory material permitting manual and automated selected object areas and mean gray values comparison:
   a. plots for comparison and evaluation of area distribution fractiles;
   b. plots for comparison and evaluation of mean value distribution fractiles;
   c. plots of joint area and mean gray value distribution.

Table 1 contains means and standard deviations of terminal area and mean gray value distributions, when different ways of object selection techniques are used. In this table coincident objects are those, that were the same while extracting them manually and automatically.

Table 1. Area and mean gray value distribution average characteristics in «experiment» and «control»

Average characteristic	Estimation method	Area		Mean gray value	
		exp.	cont.	exp.	cont.
Mean	manual estimation	1.37	1.30	96.23	82.93
	coincident objects	1.56	1.70	95.79	83.60
	automated estimation	1.75	1.73	100.68	88.61
Standard deviation	manual estimation	0.69	0.52	9.93	10.77
	coincident objects	0.56	0.70	10.99	11.57
	automated estimation	0.69	0.71	12.25	14.70

Such a considerable difference in the mean area becomes clear if we take into account the fact that the human vision is not so perfect in detecting precise boundaries of the objects and the fact that morphologists extract not all objects presented on an image, but only those they believe to be in-focus terminals. Furthermore the manual terminal extraction was done with computer mouse and it is not always possible to control hand and mouse movements totally. But, in spite of all written above the hypothesis tests allow us to conclude that there is no reliable difference between area distributions in manual and automated estimation. The results also were satisfactory for PD experts. Concerning a little difference in manually estimated area it can be guessed that the method extracts more essential objects. Hypothesis tests on area distributions give the following results: there is no reliable difference between control and experimental groups in manual estimation and there is such a difference in automated estimation.

The number of objects extracted manually and automatically on the initial image fragment and automatically on the whole image are given in Tab. 2, where in image names «c» stands for control and «e» for experiment.

Table 2. Terminals number

Image	1–c	2–e	3–c	4–e	5–c	6–e	7–c	8–e	9–c	10–e
Automated segmentation (whole)	1664	900	1623	891	1423	917	1632	899	1980	1002
Manual extraction (fragment)	40	11	29	20	36	14	35	11	33	12
Automatic segmentation (fragment)	36	16	34	20	35	15	43	12	34	13

4. CONCLUSION

Experimental investigations confirmed, that the developed method supports automated processing and analysis of immunostained striatum frontal section images and helps to define characteristics, which are essential for preclinical stage PD model construction. The designed method for automatic extraction and feature calculation of dopaminergic neurons terminals on section images can also provide the results with precision comparable those of manual object feature estimation.

Experiments revealed the following: 1) terminal numbers of DA-ergic axons in experimental group are considerably decreased in comparison to terminal numbers in control group; 2) DA-ergic neurons change functional activity after neurotoxin administration.

Software implementation of the developed mathematical methods allows one to speed up significantly the research of PD at the cost of automation of experimental data model filling and automation of model investigation by means of computer experiments. The obtained results constitute the important stage of researches of dopaminergic nigrostriatal system at developing PD, which will allow research of compensatory mechanisms. Hereafter it will allow compensatory mechanisms studying with the purposes of its management.

The future investigations include cauterization of selected objects in an extended feature space, aiming to support automated assignment of extracted objects to selected clusters. The same methods can be developed for similar problems solving. In particular, for estimation of DA-ergic neurons degeneration in the substantia nigra after MPTP administration and for estimation of dopaminergic neurons, remained after MPTP administration, functional condition.

5. ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research Grants Nos. 07-07-13545, 08-01-90022 and by the project of the Program of the Presidium of the Russian Academy of Sciences "Fundamental Sciences to Medicine 2009".

6. REFERENCES

[1] Albin R. L., Young A. B., Penney J. B. The functional anatomy of basal ganglia disorders, Trends Neurosci 12 (1989). p. 366–75.

[2] Gurevich, I.B., Yashina, V.V., Koryabkina, I.V., Niemann, H., Salvetti, O. Descriptive approach to medical image mining: An algorithmic scheme for analysis of cytological specimens. Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications 18(4) (2008). p. 542-562.

[3] Gonsales, R.C., Woods, R.E.: Digital Image Processing. 2 edn. Pearson Education, Inc (2002) publishing as Prentice Hall.

[4] Gonsales R. C., Woods R. E., Eddins S. L. Digital Image Processing using MAT-LAB. 1 edn. Pearson Education, Inc (2004) publishing as Prentice Hall.

[5] Gurevich, I., Harazishvili, D., Jernova, I., et al.: Information technology for the morphological analysis of the lymphoid cell nuclei. The 13th Scandinavian Conference on Image Analysis. Volume 2749 of LNCS. (2003). p. 541-548.

[6] N. Ogawa, K. Mizukawa, Y. Hirose et al. Mptp-induced parkinsonian model in mice: biochemistry, pharmacology and behavior, Eur Neurol.. 26(1) (1987). p. 16–23.

[7] Soille, P. Morphological Image Analysis: Principles and Applications. Springer, Berlin (2004).

[8] Tipton, K.F., Singer, T.P.: Advances in our understanding of the mechanisms of the neurotoxicity of mptp and related compounds. J Neurochem 61 (1993). p. 1191-1206.

[9] Vincent, L. Morphological grayscale reconstruction in image analysis, Applications and efficient algorithms. IEEE Transactions on Image Processing 2 (1993). p. 176-201

[10] Vincent, L., Soille, P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Machine Intell 6(12) (June 1991). p. 583-598

[11] Wu, Q., Merchant, F., Castleman, K. MICROSCOPE IMAGE PROCESSING. Elsevier Inc. (april 2008).