CLUSTER ANALYSIS OF THE RESULTS OF INTRAOPERATIVE OPTICAL SPECTROSCOPIC DIAGNOSTICS IN BRAIN GLIOMA NEUROSURGERY

Osmakov I.A.[1], **Savelieva T.A.**[1,2], **Loschenov V.B.**[1,2], **Goryajnov S.A.**[3], **Potapov A.A.**[3]
[1]National Research Nuclear University MEPhI, Moscow, Russia
[2]Prokhorov General Physics Institute of the Russian Academy of Sciences, Moscow, Russia
[3]N.N. Burdenko National Scientific and Practical Center for Neurosurgery
of the Ministry of Healthcare of the Russian Federation, Moscow, Russia

Abstract

The paper presents the results of a comparative study of methods of cluster analysis of optical intraoperative spectroscopy data during surgery of glial tumors with varying degree of malignancy. The analysis was carried out both for individual patients and for the entire dataset. The data were obtained using combined optical spectroscopy technique, which allowed simultaneous registration of diffuse reflectance spectra of broadband radiation in the 500–600 nm spectral range (for the analysis of tissue blood supply and the degree of hemoglobin oxygenation), fluorescence spectra of 5-ALA induced protoporphyrin IX (Pp IX) (for analysis of the malignancy degree) and signal of diffusely reflected laser light used to excite Pp IX fluorescence (to take into account the scattering properties of tissues). To determine the threshold values of these parameters for the tumor, the infiltration zone and the normal white matter, we searched for the natural clusters in the available intraoperative optical spectroscopy data and compared them with the results of the pathomorphology. It was shown that, among the considered clustering methods, EM-algorithm and k-means methods are optimal for the considered data set and can be used to build a decision support system (DSS) for spectroscopic intraoperative navigation in neurosurgery. Results of clustering relevant to the pathological studies were also obtained using the methods of spectral and agglomerative clustering. These methods can be used to post-process combined spectroscopy data.

Keywords: optical spectroscopy, fluorescence, diffuse reflectance, 5-ALA, protoporphyrin IX, neurosurgery, gliomas, cluster analysis

For citations: Osmakov I.A., Savelieva T.A., Loschenov V.B., Goryajnov S.A., Potapov A.A. Cluster analysis of the results of intraoperative optical spectroscopic diagnostics in brain glioma neurosurgery, *Biomedical photonics,* 2018, vol. 7, no. 4, pp. 23–34. (in Russ.) 10.24931/2413–9432–2018–7–4–23–34.

Contacts: Osmakov I.A., e-mail: ilya.osmakov@gmail.com

КЛАСТЕРНЫЙ АНАЛИЗ РЕЗУЛЬТАТОВ ИНТРАОПЕРАЦИОННОЙ ОПТИЧЕСКОЙ СПЕКТРОСКОПИЧЕСКОЙ ДИАГНОСТИКИ В НЕЙРОХИРУРГИИ ГЛИАЛЬНЫХ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА

И.А. Осьмаков[1], **Т.А. Савельева**[1,2], **В.Б. Лощенов**[1,2], **С.А. Горяйнов**[3], **А.А. Потапов**[3]
[1]Национальный исследовательский ядерный университет «МИФИ», Москва, Россия
[2]Институт общей физики им. А.М. Прохорова Российской академии наук, Москва, Россия
[3]Национальный медицинский исследовательский центр нейрохирургии
имени академика Н. Н. Бурденко, Москва, Россия

Резюме

В работе представлены результаты сравнительного исследования методов кластерного анализа данных оптической интраоперационной спектроскопии при проведении операций по удалению глиальных опухолей различной степени злокачественности. Анализ проведен как для отдельных пациентов, так и для всей совокупности данных. Данные были получены методом комбинированной оптической спектроскопии, регистрирующим спектр диффузного отражения широкополосного излучения в диапазоне спектра 500–600 нм (с целью анализа кровенаполненности тканей и степени оксигенации гемоглобина), спектр флуоресценции индуцированного 5-аминолевулиновой кислотой протопорфирина IX (с целью анализа степени изменения тканей) и сигнал диффузно отраженного лазерного излучения, использовавшегося для возбуждения флуоресценции (с целью учета рассеивающих свойств тканей).

Для определения пороговых значений указанных параметров для опухоли, зоны инфильтрации и нормального белого вещества был проведен поиск естественных кластеров в имеющихся интраоперационных данных оптической спектроскопии и их сопоставление с результатами патоморфологической экспертизы. Было показано, что среди рассмотренных методов кластеризации EM-алгоритм и метод k-средних оптимальны для рассмотренного набора данных и могут быть использованы для построения системы поддержки принятия решений при спектроскопической интраоперационной навигации в нейрохирургии. Релевантные результатам патоморфологических исследований модели были также получены с помощью методов спектральной и агломеративной кластеризации. Эти методы могут быть использованы для постобработки данных комбинированной спектроскопии.

Ключевые слова: оптическая спектроскопия, флуоресценция, диффузное отражение, 5-АЛК, протопорфирин IX, нейрохирургия, глиомы, кластерный анализ.

Для цитирования: Осьмаков И.А., Савельева Т.А., Лощенов В.Б., Горяйнов С.А., Потапов А.А. Кластерный анализ результатов интраоперационной оптической спектроскопической диагностики в нейрохирургии глиальных опухолей головного мозга // Biomedical photonics. – 2018. – Т. 7, № 4 . – С. 23–34. doi: 10.24931/2413–9432–2018–7–4–23–34.

Контакты: Осьмаков И.А., e-mail: ilya.osmakov@gmail.com

Introduction

The incidence of cancer of the central nervous system is currently increasing steadily [1]. One of the main methods of their treatment is surgical removal of malignant tumors. However, the determination of the boundaries of glial tumors is a nontrivial task due to the peculiarities of their growth along the myelinated nerve fibers and vessels deep into the healthy white matter of the brain [2, 3], which leads to incomplete removal of the tumor and a high frequency of postoperative recurrences.

It is the infiltrating nature of the growth of glial tumors that necessitates the use of additional methods of their demarcation during surgery. At the same time, optical methods for determining the type and condition of biological tissues have a number of significant advantages: high speed, accuracy, non-invasive nature, a compact size of the working part of the tool. The most popular among optical methods is the registration of fluorescence markers of tumor changes of both endogenous and exogenous nature.

At the moment, the working tool based on the principle of optical detection of fluorescence used in neurosurgery is the Opmi Pentero microscope with Blue400 mode, which allows to observe the level of fluorescence (excited in the purple range of the spectrum) accumulation of protoporphyrin IX (PPIX) induced by 5-aminolevulinic acid (5-ALA) in tumor cells. The main disadvantage of this method is the subjectivity of the evaluation of the recorded signal by a neurosurgeon. It is the doctor who, largely at his/her own discretion, determines which brightness of fluorescence shall be considered to be a subthreshold level where tumor destruction must be stopped. And the tactics of surgeons in this regard may differ. Therefore, it is essential to use a quantitative approach to intraoperative analysis of the type of fabric which is supported by optical spectrum analysis. Moreover, in the case of a sufficiently large sample of morphologically verified conclusions, it is preferable to use not only numerical values but also the preliminary conclusions about the type of tissues.

In the case of one parameter characterizing the degree of tissue malignancy (fluorescence PPIX), determining the type of tissue only by the value of the parameter is a fairly trivial task. However, as the number of parameters increases, the task becomes more complex and the use of statistical methods of data analysis for a preliminary conclusion is required. Machine learning methods are ideally suited for solving this problem.

This work is devoted to the preliminary cluster analysis of spectroscopic data for individual patients, as well as for the entire data set, with the use of the built-in libraries in Python language. It describes an optical spectroscopy method that uses the analysis of fluorescence spectra of 5-ALA-induced PPIX and diffuse reflection spectra of tissues, with subsequent extraction of information from them on the absorbers in tissues and their light scattering properties. To determine the threshold values for the tumor, the infiltration zone and the normal level, it is necessary to search for natural clusters in the available intraoperative optical spectroscopy data and compare them with the results of pathomorphological examination.

Materials and methods

Intraoperative optical spectroscopy method

A device was developed for simultaneous registration of diffuse reflection spectra and laser-induced fluorescence, consisting of a spectroanalyzer (LASA-01-BIOSPEC), two radiation sources (helium-neon laser λ=632.8 nm and a halogen lamp), fiber-optic transfer system for the delivery of radiation to and from the tissue, as well as a personal computer with special software for registration and analysis of spectra in real time. The device uses a cross-filter system that allows the separation of the visible range of the spectrum into two areas: the registration of

the diffuse reflection spectrum and the fluorescence spectrum of PPIX.

During the measurements, the distal end of the fiber optic probe was brought closer to the tissue to the degree of contact without pressure. As a result of the measurement, the input of the spectrometer receives fluorescent, as well as broadband and laser radiation which is diffusely reflected by the tissue. The recorded spectral dependences are subjected to mathematical processing in accordance with the algorithms described in [4], in a real-time mode.

The scattering properties of the tissues were estimated by the intensity of the backward scattered laser radiation and are given in comparison with the doubled value of the unchanged cortex (since, according to the literature, the diffuse reflection signal from white matter in the visible range of the spectrum is on average twice higher than from gray matter). The fluorescence intensity was calculated as the ratio of PPIX fluorescence intensity in the range of 690–730 nm to the intensity of the backward scattered laser radiation. Fluorescent contrast was determined as the ratio of the fluorescence intensity of the tissue studied to the fluorescence intensity of the normal cortex. The examples of recorded spectra are shown in Fig. 1.

The calculation of the parameters for the analysis was made according to the following formulas:

$$FI_i = \frac{S_{[690..730],i}}{S_{[625..640],i}} \qquad FC_i = \frac{FI_i}{FI_{norm}}$$

$$ScC_i = \frac{S_{[625...640],i}}{k*S_{[625..640],norm}} \qquad Hb_{total,i} = [Hb]_i + [HbO_2]_i$$

$$HbC_i = \frac{Hb_{total,i}}{Hb_{total,norm}}$$

$$Sat(Hb)_i = \frac{HbO_{2i}}{Hb_{total,i}} \qquad Sat(Hb)C_i = \frac{Sat(Hb)_i}{Sat(Hb)_{norm}}$$

where *S* is the area under the graph in the range indicated in the lower index; *i* is the fluorescence intensity calculated from the current spectrum; *norm* is the fluorescence intensity calculated on the basis of the normal tissue spectrum (usually from the cortex at some distance from the tumor projection); *FI* is the fluorescence intensity; *FC* is the contrast of the tissue under study compared with the normal fluorescence intensity; *ScC* is the contrast of the tissue under study compared to the tissue which has normal light scattering level; *k* is the coefficient of the fluorescence intensity with due account for the differences in light scattering for white and gray matter (k=2 when used as the norm of gray matter, k=1 when used as the norm of white matter); *[Hb]* is the concentration of reduced hemoglobin; *[HbO_2]* is the concentration of oxygenated hemoglobin; Hb_{total} is the total concentration of hemoglobin in the tissue (blood filling); *Sat(Hb)* is the degree of hemoglobin oxygenation (oxygen saturation).

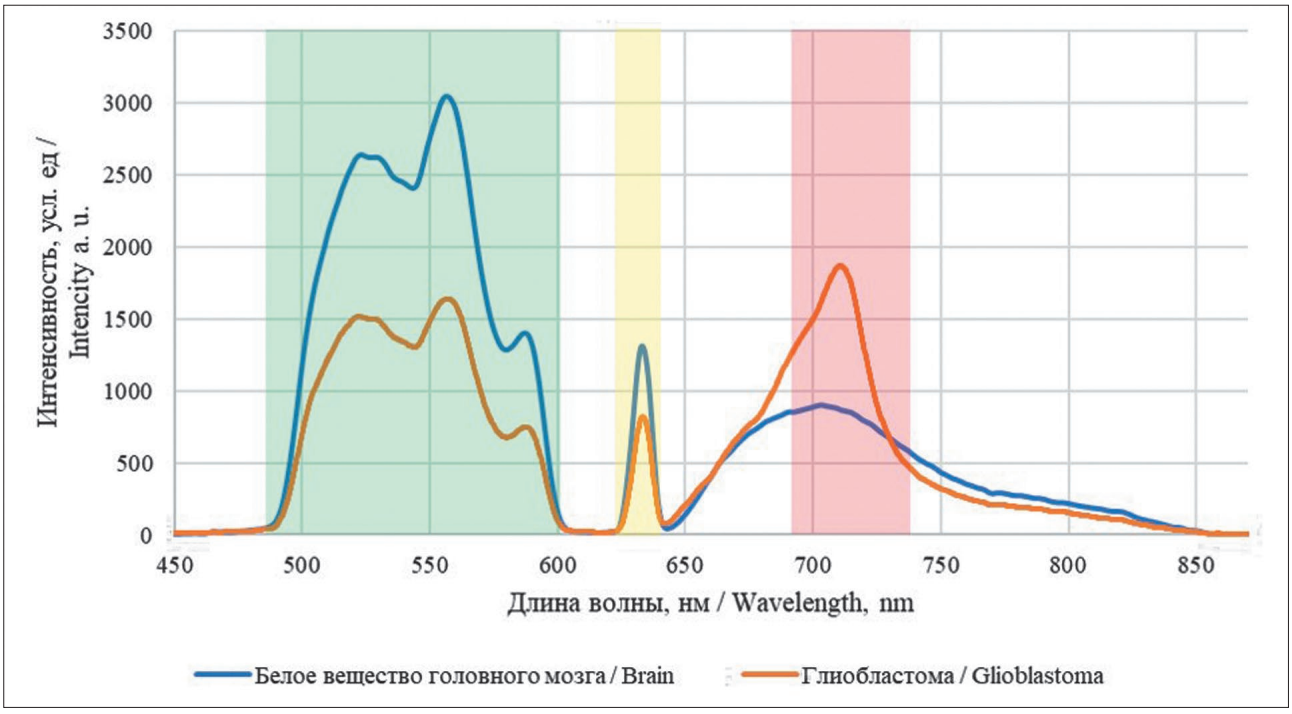


Рис. 1. Пример спектров различных типов тканей: зеленым цветом обозначена область оценки степени оксигенации, желтым – диффузного отражения лазерного излучения, красным – флуоресценции
Fig. 1. Example of spectra characteristic for different types of tissue: green – spectral range used for evaluation of oxygenation level, yellow – diffuse reflectance of laser light, red – fluorescence spectrum

Clinical data

The study retrospectively used the data from 13 patients. Three patients diagnosed with glioblastoma/astrocytoma from the sample underwent separate research. The training of the clustering algorithm was carried out on each patient separately, and then on the aggregate of the three, in order to compare the quality metrics on the test sample, which included the remaining patients. Thus, the algorithm was tested on the objects not included in the training sample. The patients were orally administered a solution of the hydrochloride of 5-aminolevulinic acid (Alasens product, manufacturer: FSUE «SRC «NIOPIK», Russia) calculated as 25 mg/kg body weight, 2–4 hours before tumor removal. Videofluoroscopy intraoperative navigation was performed with the use of operating microscope (Opmi Pentero, Carl Zeiss, Germany) with fluorescence module simultaneously with spectroscopic navigation device LESA-01-Biospec (OOO «BIOSPEC», Russia). 2 to 11 tissue samples were taken from each patient for subsequent histological analysis and comparison of its results with the data of spectroscopic examination. Each tissue sample corresponded to a number of spectra (from 1 to 10). Thus, 77 tissue samples and 876 spectra were analyzed, of which 335 were verified by histological conclusions. A scatter diagram of all verified objects is shown in Fig. 2.

Working with missing data

The specifics of the collected data is that the technical methods for simultaneous registration of all the parameters were not used in the early development and use of intraoperative optical spectroscopy method. It was only possible to measure the following pairs: the total concentration of hemoglobin in the tissues and the degree of its oxygenation or fluorescence intensity and the area under the peak of the echo signal. Due to this fact, some data was missing.

Missing data refers to empty parameter values of the objects. Their processing is a separate section of statistics and independent research work. In this study, the following standard methods of their processing were considered: removal, in which the sample was reduced 2 to 2.5 times, which is an impractical method; data zeroing led to the appearance of a set of objects with different histological labels at point 0; averaging by parameters, in which the algorithms obtained low quality metrics.

These unsatisfactory results led to the creation of a multi-step data processing strategy which included:

1. The division of data into complete and incomplete
2. The division of the complete data into the training sample and the test sample
3. The classification of incomplete data by diagnosis
4. The separation of data broken down by diagnoses in accordance with the types of tissues
5. Averaging by each type
6. The combination of training data and data averaged by type.

There were also patients who had only one pair of parameters. In such patients, the missing parameters were averaged by type with the parameters of all patients with the same diagnosis. Thus, the test sample included the objects with true parameters, and the training sample was made as large as possible. This strategy, in comparison with other methods of missing data processing, proved to have the highest quality metrics, the

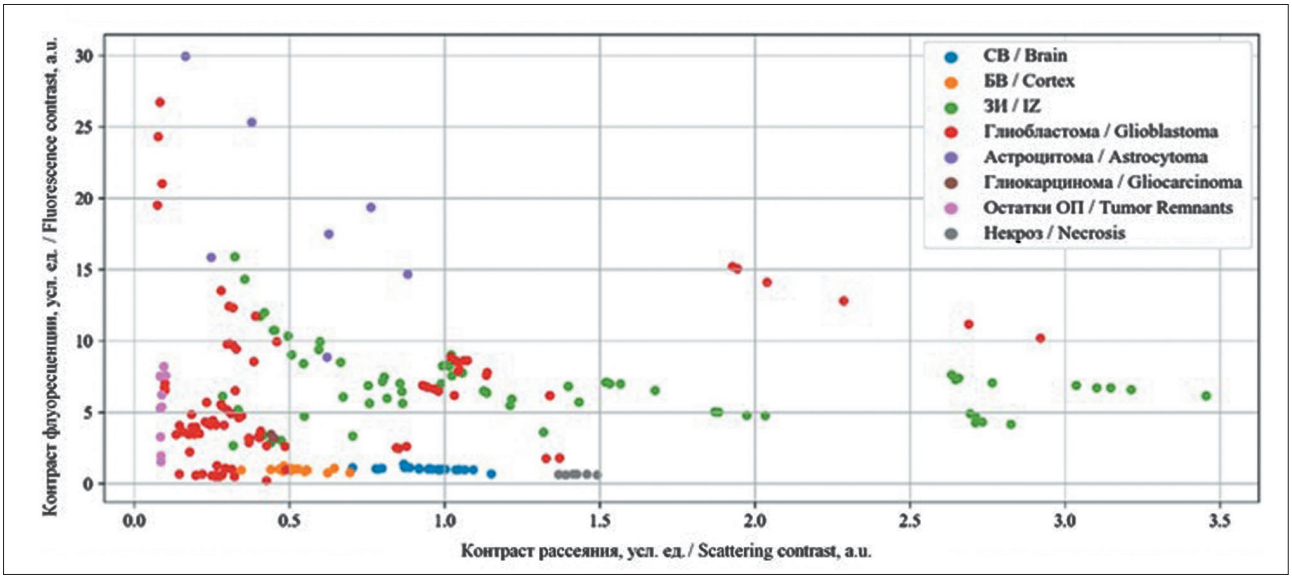


Рис. 2. Точечная диаграмма верифицированных данных, где СВ – серое вещество головного мозга, БВ – белое вещество головного мозга, ЗИ – зона инфильтрации, ОП – опухоль
Fig. 2. Scatter plot of verified data, where IZ – Infiltration Zone

largest number of successful algorithms, the preservation of most objects and the highest degree of interpretation.

Cluster analysis

For cluster analysis, the unsupervised learning approach was used. In this paper, we have considered the following methods of clustering: the k-means method, spectral clustering, expectation-maximization method (EM-algorithm), agglomerative clustering, and density clustering, the iterations of which are described below.

Before cluster analysis, the data underwent preliminary standardization, as it is necessary before such processing. This is to ensure that the weights selected in the algorithms are not operated with the parameters of different orders.

The k-means method. This is one of the most common methods used for primary data processing, which gained particular popularity after the publication of McQueen's study [5]. It involves choosing n-random clustering centers. Then, each object is compared to each center, and the object is assigned to the cluster to the center of which the object is the closest. Finally, the centers are calculated.

Spectral clustering method. In this method, similarity matrices are defined for the objects. Next, the two nearest objects are combined according to the similarity matrix so that the objects within the cluster are as different as possible from the objects of other clusters [6].

The EM-algorithm. The method is to maximize likelihood. It is based on the fact that the density of distribution probability for the objects in a sample is a weighted sum of the densities of probability in each cluster. All clusters are selected from a certain family of distributions, which are often families of normal distributions [7].

Agglomerative clustering method. In this method, pairwise distances between objects are sorted in ascending order, and each is assigned to its own cluster. Then a pair of the nearest clusters is selected and combined into one. (The search for the closest clusters can be performed with the use of various combination methods). After that, the number of centers is calculated.

Density clustering method. In this method, there must be a certain number of other points near the object within a certain radius; if this condition is not met, the object is labeled as noise.

From the specifics of the use of the clustering algorithms considered, it is possible to conclude that such methods as k-means and EM-algorithm can produce the output model of data clustering which can then be used to predict new objects.

The input parameters were chosen in the way that healthy objects were separated as much as possible from the rest of the sample into a separate cluster, but the number of clusters did not exceed 8. This is due to the fact that the number of histologically different objects may not exceed 8.

Quality metrics

In order to assess the quality of clustering results, various quality metrics are used. Such estimates must not depend on the label values themselves but only on the sample partition as such. In addition, true labels of objects are not always known, so it is necessary to have estimates that make it possible to evaluate the quality of clustering based on only an unlabeled sample.

There are external and internal quality metrics. The external metrics use the information about true clustering, while internal metrics use no external information and evaluate the quality of clustering only on the basis of the dataset. The optimal number of clusters is usually determined with the use of internal metrics.

Adjusted Rand Index (ARI). It is assumed that the true labels of the objects are known. This measure does not depend on the label values as such but only on the partitioning of the sample into clusters. Let n be the number of objects in the sample, then a is the number of pairs of objects that have the same labels and are in the same cluster, and b is the number of pairs of objects that have different labels and are in different clusters. The Rand Index then is:

$$RI = \frac{2(a+b)}{n(n-1)}$$

That is, it is the share of objects for which these partitions (initial and resulting from clustering) are «approved». Rand Index (*RI*) expresses the similarity of two different clusterings of the same sample. For this index to give values close to zero for random clustering with any n and any number of clusters, it is necessary to normalize it. This is how the Adjusted Rand Index is determined:

$$ARI = \frac{RI - E[RI]}{\max(RI) - E[RI]}$$

This measure is symmetric and does not depend on the values of the labels and their swapping. Thus, this index is a measure of the distance between various sample partitions. *ARI* takes values in the range [-1,1]. Negative values correspond to «independent» cluster partitions; values close to zero correspond to random partitions, and positive values indicate that two partitions are similar (coincide at $ARI = 1$).

Adjusted Mutual Information (AMI). This measure is very similar to *ARI*. It is also symmetric and does not depend on the values of the labels and their swapping. It is determined with the use of the entropy function, with the interpretation of sample splits as discrete distributions (the probability of the assignment to a cluster is equal to the share of objects in the cluster). The *AMI* index is defined as the mutual information for two distribu-

tions corresponding to the sample-to-cluster partitions. Intuitively, mutual information measures the proportion of information common to both partitions: how much the information on one of them reduces the uncertainty in respect of the other.

AMI index is determined in the way which is similar to the determination of *ARI*, making it possible to avoid the growth of the *AMI* index with the increase in the number of classes. It takes values in the range of [-1,1]. The values close to zero indicate the independence of the partitions, and those close to one, their similarity (coincidence at *ARI* = 1).

Homogeneity, completeness, V-measure. Formally, these measures are also defined with the use of entropy and conditional entropy functions, with the consideration of sample partitions as discrete distributions:

$$h = 1 - \frac{H(C|K)}{H(C)}$$

$$c = 1 - \frac{H(K|C)}{H(K)}$$

here *K* is the result of clustering, *C* is the true division of the sample into classes. Thus, *h* measures the degree to which each cluster consists of objects of the same class, and *c* measures the degree to which the objects of the same class belong to the same cluster. These measurements are not symmetrical. Both take on values in the range of [0,1], and larger values correspond to more accurate clustering. These measures are not normalized like *ARI* or *AMI* and, therefore, they depend on the number of clusters. Random clustering will not produce zero values in case of a large number of classes and a small number of objects. In these cases, it is preferable to use *ARI*. However, if the number of objects is more than 1000 and the number of clusters is less than 10, this problem is not so pronounced and can be ignored.

To account for both values, *h* and *c*, a *V*-measure is also introduced as their harmonic mean:

$$v = 2\frac{hc}{h + c}$$

It is symmetric and shows how much the two clusterings are similar to each other.

Silhouette. In contrast to the above metrics, this coefficient does not imply the knowledge of the true labels of objects and makes it possible to assess the quality of clustering with the use of only the (unlabeled) selection and the result of clustering. First, the silhouette is defined separately for each object. *a* is the average distance from this object to the objects from the same cluster, *b* is the average distance from this object to the objects from the nearest cluster (different from the one in which the object itself is). Then the silhouette of the object is the value:

$$s = \frac{b - a}{\max(a, b)}$$

The silhouette of a selection is the average value of the silhouette of the objects in that selection. Thus, a silhouette shows how the average distance to the objects of the same cluster differs from the average distance to the objects of other clusters. This value is in the range of [-1,1]. Values close to -1 correspond to the clustering variant with a high spread, values close to zero mean that clusters intersect and overlap, and values close to 1 correspond to «dense» clearly outlined clusters. Thus, the larger the silhouette, the more clearly the clusters are outlined, and they are compact, tightly grouped clouds of points.

With the silhouette, you can select the optimal number of clusters *k* (if it is not known in advance) and select the number of clusters that maximize the value of the silhouette. Unlike the previous metrics, the silhouette depends on the shape of the clusters, and reaches larger values on the more convex clusters obtained by algorithms based on the restoration of the distribution density.

To assess the quality of clustering, clusters were manually merged in such a way that healthy objects were in a separate cluster, and all other objects were combined into a cluster of pathology (not healthy ones). Thus, the obtained metrics will evaluate how well the used method distinguishes the healthy objects from the sick ones.

Results and discussion

Results of the analysis of data for individual patients

Patient G. Diagnosis: diffuse astrocytoma with pronounced polymorphism.

Patient G. had a sample of 12 objects. The quality metrics are presented in Table 1. The resulting models in the visualization can not be correlated with true representations. However, k-means and agglomerative clustering methods were able to group healthy tissues into a separate cluster, but the very model of data clustering turned out to have very broad boundaries, which allow fluorescence intensity above 7.5, which is not typical for a healthy brain area. Based on the results of the analysis of the data of this patient, it is obvious that these methods of processing should be used on sufficiently large samples.

For patient G., the methods that have the highest quality estimates are density clustering, k-means, and agglomerative clustering.

Patient S. Diagnosis: glioblastoma Grade IV.

As can be seen from Table 2, all algorithms except the density method showed equally good results in the

division of normal and pathologic samples into separate clusters. This is due to its features, which result in marking some objects as noise, greatly reducing its quality metrics. However, this method allocates objects that are close in time into separate clusters. This feature can be useful further on for the averaging of such objects in order to prevent them from making large weights in the measurements.

Patient S., in comparison with patient G., had a sample of 82 objects, that is, it was almost 6 times more numerous. Healthy tissues were well grouped into a separate cluster, as it can be seen by the example of the visualization of the agglomerative clustering results (Fig. 3*a*). The input parameters of clustering were selected in such a way that allows to differentiate healthy tissues from all others.

It is also worth noting that the obtained straight boundaries in the k-means method (Fig. 3*b*) are not relevant to the complex boundary between clusters found in the experiment, which is not apparent in the EM-algorithm (Fig. 3*c*). However, the EM algorithm did not have a gradient transition between clusters, which would be typical for infiltration zones and for the diffuse nature of glioblastomas. In addition, it is difficult to interpret the resulting model of data division into clusters, because healthy tissues were included in a large cluster with characteristics that differ from those for healthy tissues.

Patient B. Diagnosis: Glioblastoma**.**

Patient B. had a sample of 59 objects; the quality metrics of the obtained models are shown in Table 3. The best methods were EM-algorithm, spectral clusterization, k-means, agglomerative clustering. However, the values of these metrics are not high enough to use the resulting clustering models due to insufficient sample size.

Results of the analysis of the data set of all patients

To begin with, an analysis was carried out of the data set of those patients who were earlier considered separately.

Patients B. + G. + C.

The results of the processing of the integrated data of patients B., G. and S. are shown in Fig. 4 and in table 4. The high-quality metrics and the more predictable nature of the healthy tissue model make it possible to say that the increase of the sample has a positive impact on the comprehensive assessment of the results.

The clustering density method obtained an abnormally large number of clusters, more than 10. A more detailed study of this phenomenon showed that most clusters consist of objects which are close-standing in terms of the registration time, that is, very likely, these spectra corresponded to the same small area of tissue.

The quality assessment of B. + G. + C. models in other patients

Since the processing of the total data set of the three patients produced potentially plausible models, the data

Таблица 1
Метрики качества пациента Г. на отложенной выборке
Table 1
Quality metrics of held-out set for patient G.

Название Metric	AMI	ARI	Гомогенность Homogeneity	Полнота Completeness	V-мера V-measure	Силуэт Silhouette
ЕМ-алгоритм EM-algorithm	0,4666	0,6409	0,5183	0,6468	0,5754	0,2609
СК SC	0,4666	0,6409	0,5183	0,6468	0,5754	0,3502
ПК DC	9,8715	0,0000	9,8715	1,0000	1,9743	0,3705
k-средних k-means	1,0000	1,0000	1,0000	1,0000	1,0000	0.4551
АК AC	1,0000	1,0000	1,0000	1,0000	1,0000	0,4551

СК – спектральная кластеризация, ПК – плотностная кластеризация, АК – агломеративная кластеризация
SC – spectral clustering, DC – density-based clustering, AC – agglomerative clustering

were then used to evaluate the quality metrics on their basis, in respect of the data of patients for whom no clustering was performed.

Evaluation of the quality of predictions for patient D. (127 objects, 41 verified ones), patient L. (30 objects, 23 verified ones) and the remaining group of 9 patients (422 objects, 93 verified ones), on models obtained in patients B., G., S., is presented in table 5.

It can be seen from the obtained metrics that in all patients the EM-algorithm coped with the task almost perfectly, given that the test sample included some types of tissues that the algorithm had not dealt with before,

Таблица 2
Метрики качества пациента С. на отложенной выборке
Table 2
Quality metrics of held-out set for patient S.

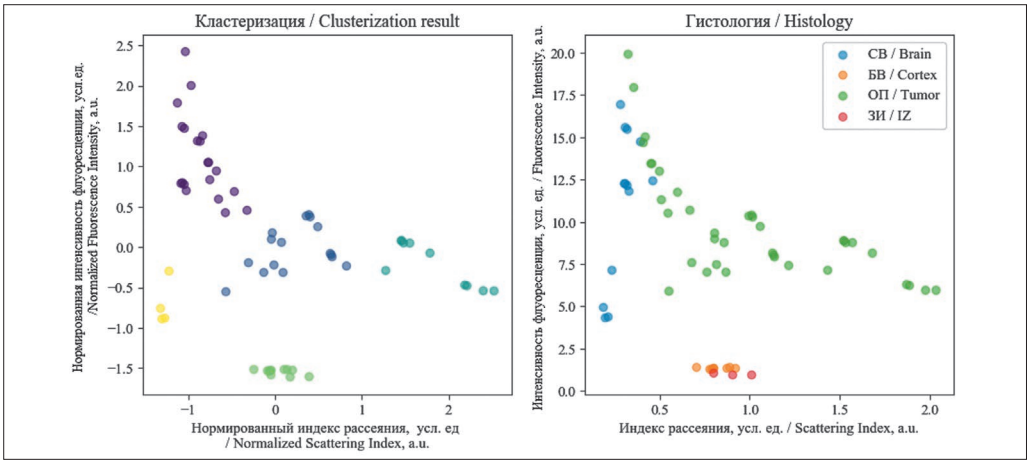
Название Metric	AMI	ARI	Гомогенность Homogeneity	Полнота Completeness	V-мера V-measure	Силуэт Silhouette
ЕМ-алгоритм EM-algorithm	1,0000	1,0000	1,0000	1,0000	1,0000	0,2446
СК SC	1,0000	1,0000	1,0000	1,0000	1,0000	0,2446
ПК DC	0,4464	0,4411	1,0000	0,4548	0,6252	0,1153
k-средних k-means	1,0000	1,0000	1,0000	1,0000	1,0000	0,2446
АК AC	1,0000	1,0000	1,0000	1,0000	1,0000	0,2446

СК – спектральная кластеризация, ПК – плотностная кластеризация, АК – агломеративная кластеризация
SC – spectral clustering, DC – density-based clustering, AC – agglomerative clustering

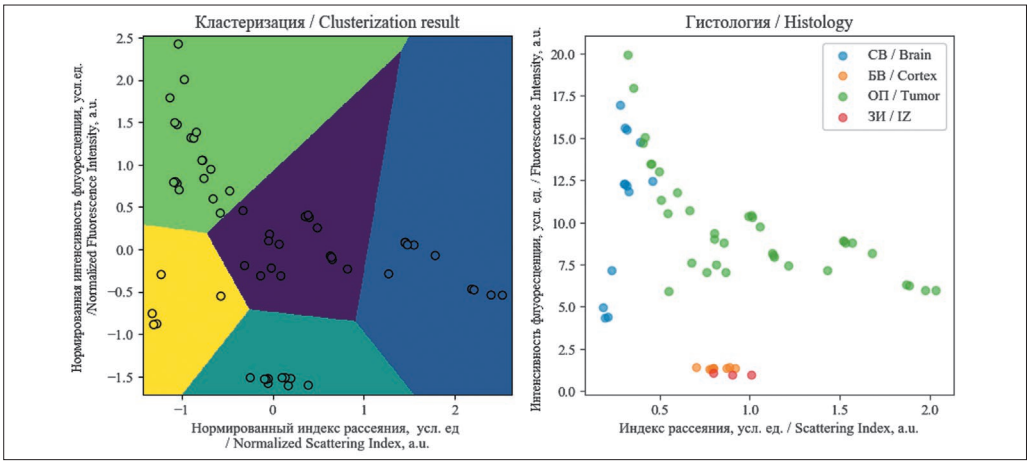
Таблица 3
Метрики качества пациента Б. на отложенной выборке
Table 3
Quality metrics of held-out set for patient B.

Название Metric	AMI	ARI	Гомогенность Homogeneity	Полнота Completeness	V-мера V-measure	Силуэт Silhouette
ЕМ-алгоритм EM-algorithm	0,8133	0,9150	0,8710	0,8174	0,8434	0,4313
СК SC	0,8133	0,9150	0,8710	0,8174	0,8434	0,4313
ПК DC	0,3053	0,2838	0,7268	0,3192	0,4436	0,0815
k-средних k-means	0,8133	0,9150	0,8710	0,8174	0,8434	0,4313
АК AC	0,8133	0,9150	0,8710	0,8174	0,8434	0,4313

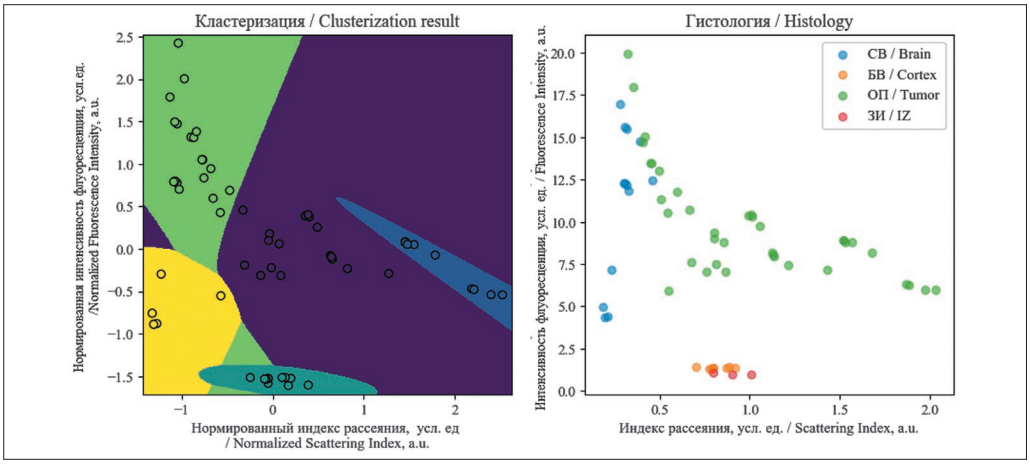
СК – спектральная кластеризация, ПК – плотностная кластеризация, АК – агломеративная кластеризация
SC – spectral clustering, DC – density-based clustering, AC – agglomerative clustering



a



b



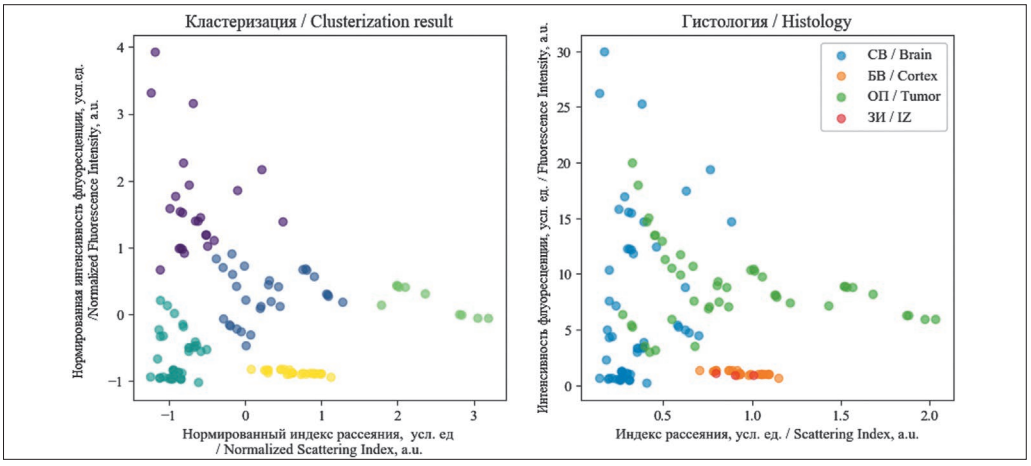
c

Рис. 3. Визуализация результатов пациента С. с применением различных методов кластеризации, в сравнении с реальным распределением объектов (справа), где СВ – серое вещество головного мозга, БВ – белое вещество головного мозга, ОП – опухоль, ЗИ – зона инфильтрации:

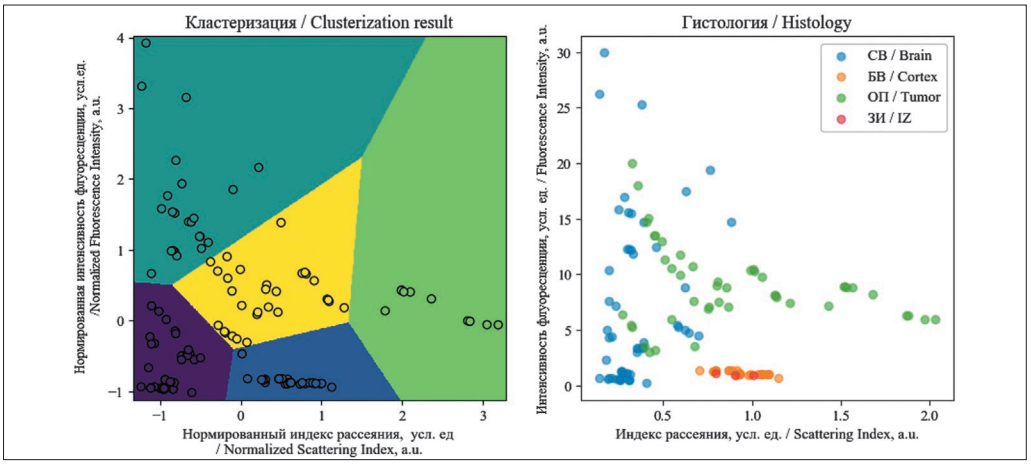
a – агломеративная кластеризация;
b – кластеризация k-средних;
c – кластеризация EM-алгоритмом

Fig. 3. Visualization of clusterization results (left) compared to actual distribution (right) for patient S., where IZ – infiltration zone:

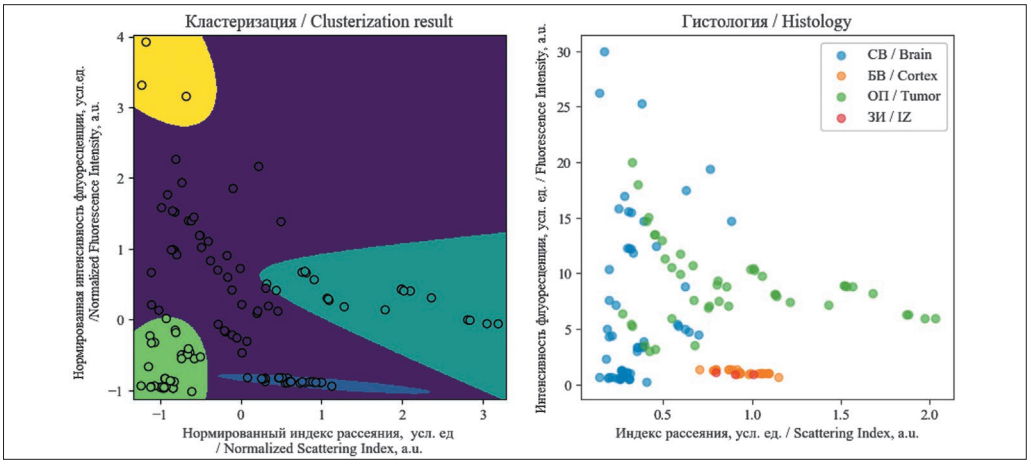
a – aglomerative clusterization;
b – k-means clusterization;
c – EM-algorithm clusterization



a



b



c

Рис. 4. Визуализация результатов пациентов Б.+Г.+С. с применением различных методов кластеризации, в сравнении с реальным распределением объектов (справа), где СВ – серое вещество головного мозга, БВ – белое вещество головного мозга, ОП – опухоль, ЗИ – зона инфильтрации:

a – плотностная кластеризация;
b – кластеризация k-средних;
c – кластеризация EM-алгоритм

Fig. 4. Visualization of clustering results (left) compared to actual distribution (right) for patients B.+G.+S., where IZ – infiltration zone:

a – density-based clustering;
b – k-means clustering;
c – EM-algorithm clustering

Таблица 4
Метрики качества пациентов Б.+Г.+С. на отложенной выборке
Table 4
Quality metrics of held-out set for patients B.+G.+S.

Название Metric	AMI	ARI	Гомогенность Homogeneity	Полнота Completeness	V-мера V-measure	Силуэт Silhouette
ЕМ-алгоритм EM-algorithm	1,0000	1,0000	1,0000	1,0000	1,0000	0,2366
СК SC	1,0000	1,0000	1,0000	1,0000	1,0000	0,2366
ПК DC	0,0191	0,1026	0,0452	0,0544	0,0494	0,2397
k-средних k-means	0,6442	0,8291	0,8058	0,6548	0,7225	0,2366
АК AC	1,0000	1,0000	1,0000	1,0000	1,0000	0,2397

СК – спектральная кластеризация, ПК – плотностная кластеризация, АК – агломеративная кластеризация
SC – spectral clustering, DC – density-based clustering, AC – agglomerative clustering

Таблица 5
Метрики качества предсказаний пациентов на полученных на пациентах Б., Г., С. моделях
Table 5
Quality metrics of patient predictions based on models obtained from patients B., G. and S.

Пациент Patient	Название Metric	AMI	ARI	Гомогенность Homogeneity	Полнота Completeness	V-мера V-measure	Силуэт Silhouette
Д. D.	ЕМ-алгоритм EM-algorithm	1,0000	1,0000	1,0000	1,0000	1,0000	0,1420
	k-средних k-means	0,6860	0,8479	0,8091	0,6976	0,7449	0,1420
Л. L.	ЕМ-алгоритм EM-algorithm	1,0000	1,0000	1,0000	1,0000	1,0000	0,8165
	k-средних k-means	1,0000	1,0000	1,0000	1,0000	1,0000	0,8165
Все All	ЕМ-алгоритм EM-algorithm	1,0000	1,0000	1,0000	1,0000	1,0000	–0,1452
	k-средних k-means	0,5468	0,7529	0,5587	0,7425	0,6376	–0,1452

whereas the k-means method showed a relatively worse result in most cases.

Conclusion

The following most universal models can be distinguished from the visualized models and measured quality metrics: EM-algorithm, k-means method, spectral clustering and agglomerative clustering. However, the last two methods do not provide ready-made models that can evaluate new data, which excludes them in the creation of decision-making assistance systems, but they are suitable for post-processing of the data.

When the number of obtained clusters is greater than the number of the types of labels, it creates practical difficulties in iterating and merging clusters for the evaluation of models. Sensitivity to the sample size can be seen in the quality metrics and the nature of the model boundaries in patient C. and the integrated patients B. + G. + C., who had 12 objects and 41 objects, respectively. In most cases, and with a sufficient sample, almost all algorithms perfectly coped with the task in individual patients, and the method of density clustering, which obtained, on average, poor quality metrics, was found to be special in the identification of objects close in time, which can help in further research.

The drawbacks listed, except the lack of operability on insufficient samples, can be mitigated with the use of other machine learning methods, namely, supervised learning, where the model will be trained on specific answers, which are labels of a class represented by histological findings.

The results of the study of spectroscopic data make it possible to identify correlations between several parameters numerically, with the use of machine learning methods, determined by spectra and histological conclusions about the presence of tissue malignancy signs.

In comparison with the method of statistical data processing presented earlier, the method of intraoperative registration of combined spectra described in the article [8], the sensitivity increased, on average, from 88% to 90%, and the specificity from 82% to 91%.

The results presented in the article were obtained with the use of research equipment of the Core Facilities Center «Technological and Diagnostic Center for the Production, Research and Certification of Micro and Nanostructures» of the Federal State Budgetary Institution of Science A. M. Prokhorov General Physics Institute of the Russian Academy of Sciences.

The study was done with the financial support of the Ministry of Education and Science of the Russian Federation (agreement RFMEFI60717X0183).

REFERENCES

1. De Robles P., Fiest K.M., Frolkis A.D., Pringsheim T., Atta C., St Germaine-Smith C., Day L., Lam D., Jette N. The worldwide incidence and prevalence of primary brain tumors: a systematic review and meta-analysis, *Neuro-Oncology*, 2015, vol. 17(6), pp. 776–783. doi:10.1093/neuonc/nou283
2. Claes A., Idema A.J., Wesseling P. Diffuse glioma growth: a guerilla war, *Acta Neuropathol.*, 2007, vol. 114, pp. 443–458. doi:10.1007/s00401–007–0293–7
3. Sutter M., Eggspuehler A., Grob D., Jeszenszky D., Benini A., Porchet F., Mueller A., Dvorak J. The validity of multimodal intraoperative monitoring (MIOM) in surgery of 109 spine and spinal cord tumors, *Eur Spine J.*, 2007, vol. 16, suppl. 2, pp. 197–208.
4. Savel'eva T.A., Loshchenov V.B., Goryainov S.A., Shishkina L.V., Potapov A.A. A spectroscopic method for simultaneous determination of protoporphyrin IX and hemoglobin in the nerve tissues at intraoperative diagnosis, *Russian Journal of General Chemistry*, 2015, vol. 85, no. 6, pp. 1549–1557.
5. MacQueen J. Some methods for classification and analysis of multivariate observations. In *Proc. 5th Berkeley Symp. on Math. Statistics and Probability*. 1967. pp. 281–297
6. Jianbo S., Jitendra M. Normalized Cuts and Image Segmentation, *IEEE Transactions on PAMI,* 2000, vol. 22(8) , pp. 888–905.
7. Jordan M.I., Xu L. *Convergence results for the EM algorithm to mixtures of experts architectures: Tech. Rep. A.I. Memo No. 1458.* MIT, Cambridge, MA, 1993. 33 p.
8. Potapov A.A., Goriaĭnov S.A., Loshchenov V.B., Savel'eva T.A., Gavrilov A.G., Okhlopkov V.A., Zhukov V.Iu., Zelenkov P.V., Gol'bin D.A., Shurkhaĭ V.A., Shishkina L.V., Grachev P.V., Kholodtsova M.N., Kuz'min S.G., Vorozhtsov G.N., Chumakova A.P. Intraoperative Combined Spectroscopy (Optical Biopsy) of Cerebral Gliomas, *N. N. Burdenko Journal of Neurosurgery*, 2013, vol. 2, pp. 3–10.

ЛИТЕРАТУРА

1. De Robles P., Fiest K.M., Frolkis A.D., et al. The worldwide incidence and prevalence of primary brain tumors: a systematic review and meta-analysis // Neuro-Oncology. – 2015. – Vol. 17(6). – P. 776–783. doi:10.1093/neuonc/nou283
2. Claes A., Idema A.J., Wesseling P. Diffuse glioma growth: a guee rilla war // Acta Neuropathol. – 2007. – Vol. 114. – P. 443–458. doi:10.1007/s00401–007–0293–7
3. Sutter M., Eggspuehler A., Grob D., et al. The validity of multir modal intraoperative monitoring (MIOM) in surgery of 109 spine and spinal cord tumors // Eur Spine J. – 2007. – Vol. 16, Suppl. 2. – P. 197–208.
4. Savel'eva T.A., Loshchenov V.B., Goryainov S.A., et al. A spectrov scopic method for simultaneous determination of protoporphyo rin IX and hemoglobin in the nerve tissues at intraoperative diagł nosis // Russian Journal of General Chemistry. – 2015. – Vol. 85, No. 6. – P. 1549–1557.
5. MacQueen J. Some methods for classification and analysis of mulQ tivariate observations. In Proc. 5th Berkeley Symp. on Math. Statisa tics and Probability. – 1967. – P. 281–297
6. Jianbo S., Jitendra M. Normalized Cuts and Image Segmentation // IEEE Transactions on PAMI. – 2000. – Vol. 22(8). – pp. 888–905.
7. Jordan M.I., Xu L. Convergence results for the EM algorithm to mixtures of experts architectures: Tech. Rep. A.I. Memo No. 1458. – MIT, Cambridge, MA, 1993. – 33 p.
8. Potapov A.A., Goriaĭnov S.A., Loshchenov V.B., et al. Intraoperative Combined Spectroscopy (Optical Biopsy) of Cerebral Gliomas // N. N. Burdenko Journal of Neurosurgery. – 2013. – Vol. 2. – P. 3–10.