Raman spectroscopy of blood plasma for cancer diagnosis

Cherkasova, O., Mankova, A., Vrazhnov, D.

O. Cherkasova, A. Mankova, D. Vrazhnov, "Raman spectroscopy of blood plasma for cancer diagnosis," Proc. SPIE 12086, XV International Conference on Pulsed Lasers and Laser Applications, 120861M (2 December 2021); doi: 10.1117/12.2613925

SPIE.

Event: XV International Conference on Pulsed Lasers and Laser Applications, 2021, Tomsk, Russian Federation

Raman spectroscopy of blood plasma for cancer diagnosis

O. Cherkasova [*a,b], A. Mankova[c], D. Vrazhnov [d,e]
[a]Institute of Laser Physics, Siberian Branch of the RAS, 15 B Ac. Lavrentiev's ave, 630090 Novosibirsk, Russia
[b]Institute on Laser and Information Technologies - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of RAS, 1 Svyatoozerskaya st., 140700 Shatura, Moscow Region, Russia
[c]Faculty of Physics, Lomonosov Moscow State University, 1 Leninskiye Gory, 119991 Moscow, Russia
[d] V.E. Zuev Institute of Atmospheric optics, Siberian Branch of the RAS, 1, Academician Zuev square, 634055 Tomsk, Russia
[e]Tomsk State University, 36 Lenin Ave., 634050 Tomsk, Russia

ABSTRACT

Raman spectra of blood plasma were studied in the dynamics of the experimental glioma. We used a DXR Raman Microscope (Thermo Scientific), excitation wavelengths of 532 nm, range 80–3200 cm^{-1}. Each sample of blood plasma was a droplet with a volume of 10 μL placed on a special aluminum plate. Machine learning methods were used to identify the most informative frequencies associated with cancer molecular markers. The most significant changes in the Raman spectra are observed in the 900–1700 cm^{-1} range.

Keywords: Raman spectra, blood plasma, U87 human glioblastoma, machine learning methods

INTRODUCTION

A Raman spectrum represents a molecular fingerprint of the sample and provides quantitative information regarding its chemical makeup. Biochemical changes in cells and tissues that may either be caused or are the cause of a disease can lead to significant changes in the Raman spectra. Raman spectroscopy is a powerful analytical technique that can be applied in the differential diagnosis of breast, skin and other types of cancer [1]. Raman spectroscopy and machine learning techniques were used to identify glioma margins and grades [2, 3], and biochemical changes induced by IDH1 mutation in human glioma [4, 5]. The Raman spectra included molecular vibrational fingerprints of carotenoids, tryptophan, amide I/II/III, proteins, and lipids. The cross-validated sensitivity, specificity, and accuracy were found to be 100%, 96.3%, and 99.6% to distinguish glioma tissues from normal brain tissues, respectively [6]. The analysis of blood can enable powerful, minimally invasive diagnostics for many diseases, including cancer [1, 7, 8]. Raman spectroscopy combined with advanced chemometrics is allowed the discrimination between samples of blood from breast cancer patients and healthy volunteer controls [8, 9], between samples of cirrhotic patients with and without hepatocellular carcinoma [10]. Using the principal component analysis, we previously assessed the general differences in the Raman spectra of the blood serum in rats in the dynamics of development of the experimental cholangiocarcinoma [11].

The purpose of our work is to study the Raman spectra of blood plasma in the dynamics of the development of experimental glioblastoma. Glioblastoma is among the deadliest neoplasms and represents approximately 49% of all brain and other central nervous system malignant tumors [12]. A positive prognosis in treatment is observed with the early detection of a malignant tumor. Early diagnostic methods have not yet been developed for these cancer types [13].

[*]o.p.cherkasova@gmail.com; phone +7 9659231192

XV International Conference on Pulsed Lasers and Laser Applications, edited by Anton V. Klimkin, Maxim V. Trigub, Victor F. Tarasenko, Proc. of SPIE Vol. 12086, 120861M · © 2021 SPIE · 0277-786X · doi: 10.1117/12.2613925

METHODS

We used a DXR Raman Microscope (Thermo Scientific) with a magnification of 10x, excitation wavelengths of 532 nm, range 80–3200 cm^{-1}. Each sample of blood plasma was a droplet with a volume of 10 μL placed on a special aluminum plate [11] (Figure 1). This plate has identical holes in the form of a funnel with a diameter of 5 and a depth of 2 mm. We can assume that the distribution of samples in each well and the film thickness of the sample after drying will be the same. The measurements were carried out after complete drying of the sample, which took place for 10 min at room temperature. The Raman spectrum was recorded at three points in the center of the well and averaged. Each spectrum was averaged over 300 scans.

Figure 1. Special aluminum plate, top view.

The model of orthotopic xenotransplantation of the U87 human glioblastoma cells into immunodeficient mice SCID line was used [14]. The U87 cell suspension (5 μL, 500000 cells per animal) was introduced in the subcortical brain structures through a hole in the mouse cranium. The control (healthy) mice were injected with the same volume of saline. Animals were removed from the experiment on days 7, 14, 21 and 28 after tumor inoculation. Each group consisted of 10 individuals. The dynamics of the tumor growth was estimated by means of T2-weighted images [15]. The experiments were carried out in compliance with the principles of humanity according to directive 86/609/EEC of the European Community.

Two tasks were considered in the framework of the analysis of Raman spectra using machine learning methods. First, the difference between the control group and the comparison group was studied with the informative feature selection that has the greatest impact on group separability. Second, the dynamics of changes in Raman spectra for glioma groups with an increase in the duration of tumor development were considered. The question of the influence of the background subtraction procedure on the choice of informative features by the method of principal components analysis (PCA), and Support Vector Machines (SVM) was investigated separately. Details about PCA and SVM can be found in our previous work [16-22]. All computations were made by Python 3.8.2 with sklearn 0.22.2 (PCA and machine learning), rampy 0.4.6 (baseline correction), and scipy 1.4.1 (Savitsky-Golay filtering) libraries.

We used the following data processing pipeline. First, baseline correction was made. Second, PCA with standard scaler and loadings matrix analysis/SVM was applied. For better visualization of loadings matrix, Savitsky-Golay filter was applied with the following parameters window size 51, and polynomic of order 7 selected. SVM was used with default parameters. Results were tested by the Stratified K Fold procedure with 3 splits. Class weights were balanced to make groups of equal size.

RESULTS

Figure 2 shows the dynamics of tumor size growth. Significant tumor growth took place from 14 to 21 days after cell inoculation. The intensity of a number of bands in the Raman spectrum changes with an increase in tumor volume. Figure 3 shows the averaged Raman spectra of blood plasma for each group. The most significant changes in the Raman spectra are observed in the 900–1700 cm^{-1} range. Raman spectra are high-dimensional data. PCA reduces the dimensionality of the data, reveals informative sections of the spectra, and makes it possible to visualize the results,

which makes it possible to evaluate the possibility of creating a data model for differentiating representatives of various experimental groups.

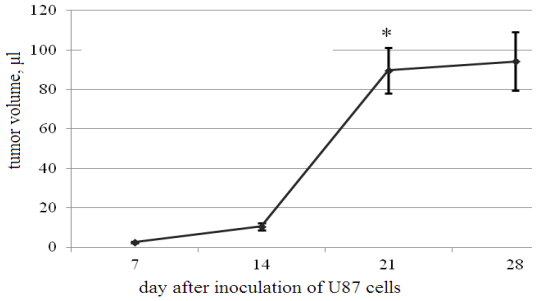

Figure 2. U87 glioblastoma volume versus days after inoculation. * p<0.001, one-way ANOVA, LSD-test.

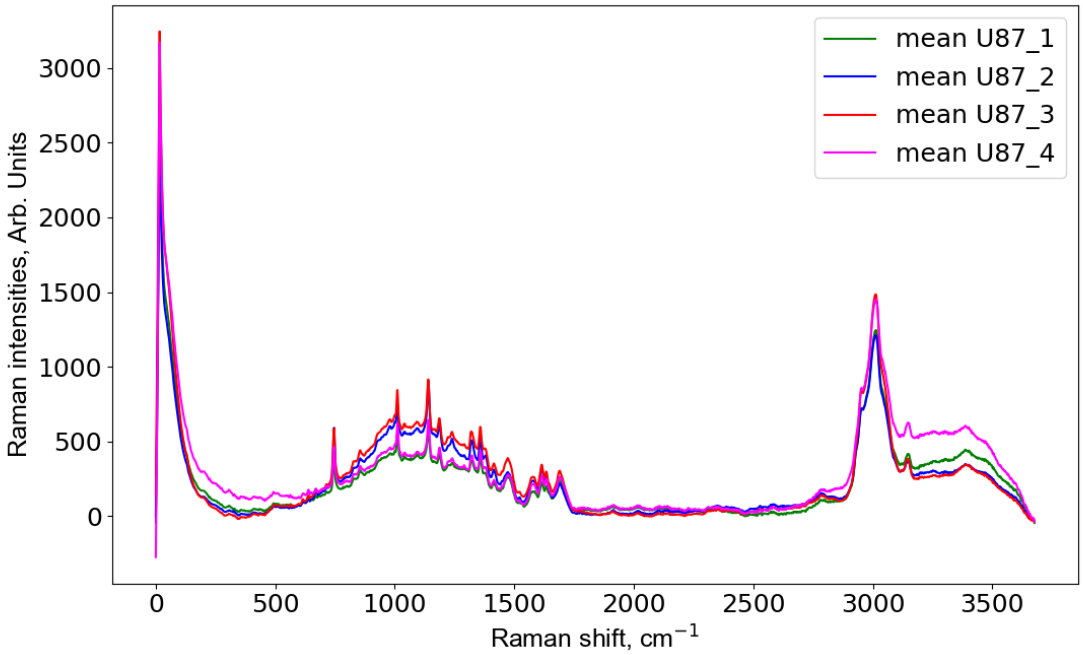

Figure 3. Mean Raman spectra of U87 glioblastoma experimental groups.

A significant increase in tumor size is observed on day 21 (3 weeks) after the injection of tumor cells. We will analyze the Raman spectra of this group in more detail. The first four components, which make up 97.8% of the variance, are the most informative and significant for this group. After clarifying the number of principal components containing the largest share of information, possible combinations of the projections of the four principal components onto the plane were constructed (see Figure 4.). As can be seen from Figure 4, the class separability is carried out according to the second principal component. Therefore the main contribution to the group separability will be made by the frequencies corresponding to the second principal component shows that the ranges of 450-500, 1850-1950, 2400-2500 cm^{-1} will be informative (see Figure 5).

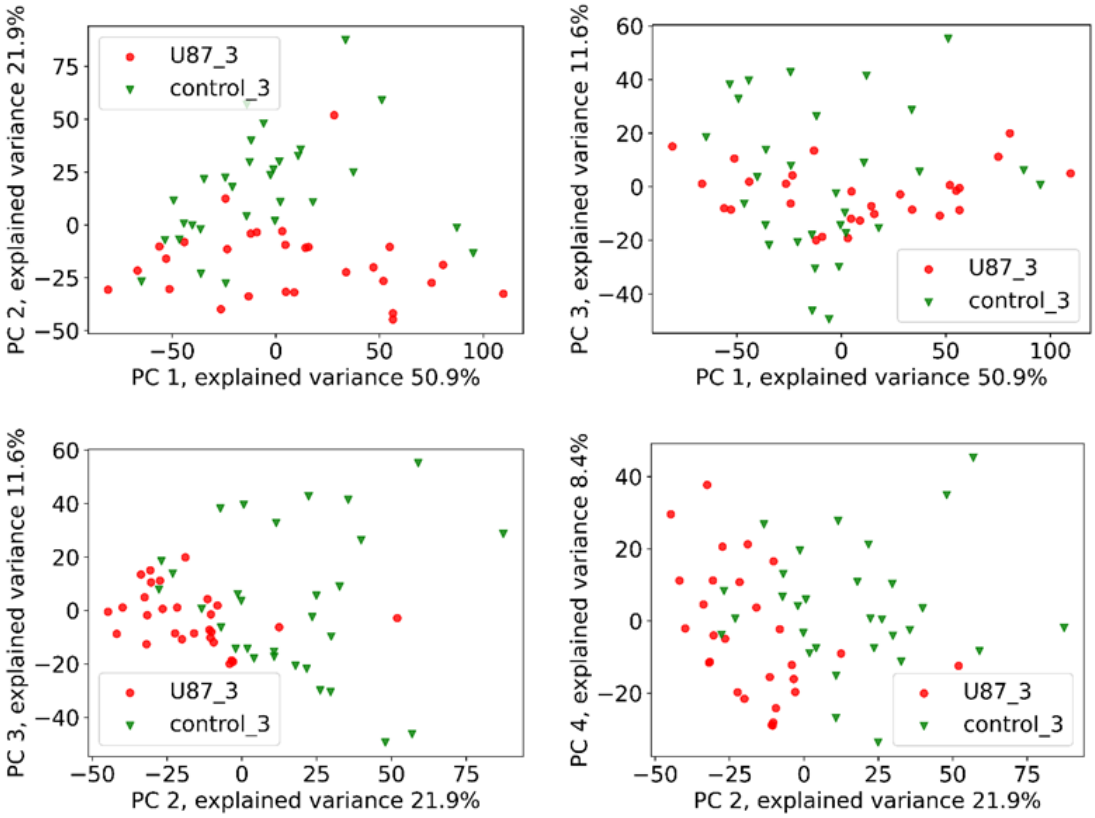

Figure 4. The projections of the various principal components on a plane for the third week of the experiment

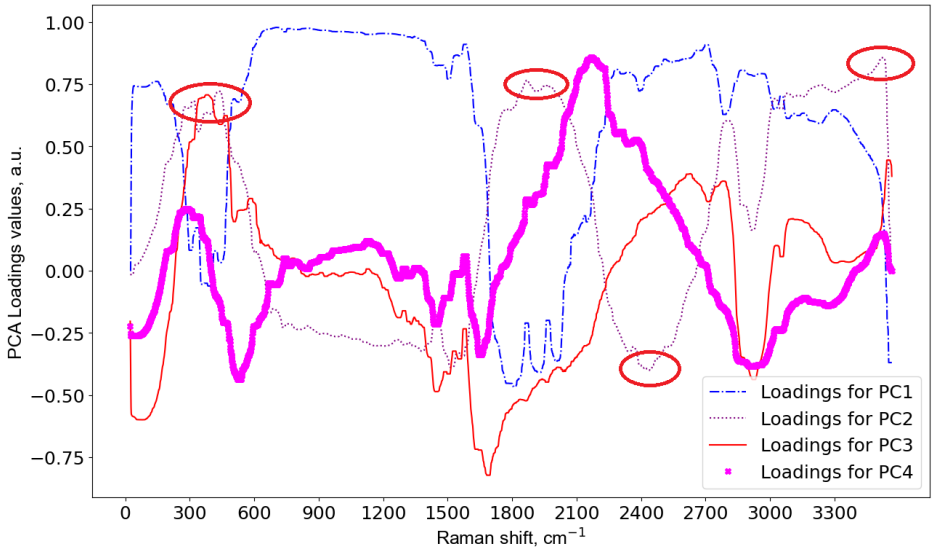

Figure 5. Analysis of the load matrix for the third week of the experiment.

At the next stage, we used Supervised Learning methods to estimate the linear separability of the data, for example, by

using linear kernel SVM [17]. To validate estimation of the data separability, we used a standard approach for Supervised Learning: initial data were randomly divided into 3 splits with the same distribution of the data among classes as in the original dataset. Linear SVM was trained on each split with class weight balancing. Obtained quality metrics such as sensitivity, specificity, accuracy, and precision were averaged. The averaged quality metrics: sensitivity was equal to 0.86±0.04; specificity was equal to 0.83 ± 0.05; accuracy was equal to 0.84 ± 0.04. Also, Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC) analyses were performed for each data split and the averaged ones. The mean ROC curve and AUC analysis for the linear SVMs is shown in Figure 6. The results demonstrate the potential of developed models for the separation of control and experimental groups.

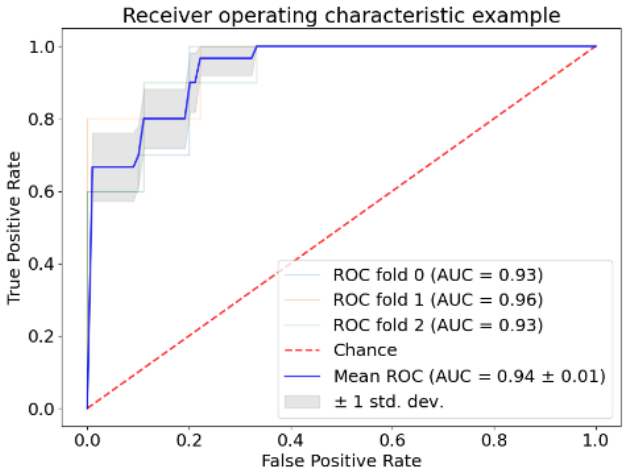

Figure 6. Mean ROC curve for linear SVMs for the 3 splits.

Also, linear SVM allows obtaining the most informative features based on the proximity of the feature vector to the separating hyperplane. An averaged relative features' importance was estimated for the three train/test splits and linear SVMs. Figure 7 illustrates the informative frequencies for this classifier. These are 700, 1000, 1180, 1500, 1770, 3000 $см^{-1}$.

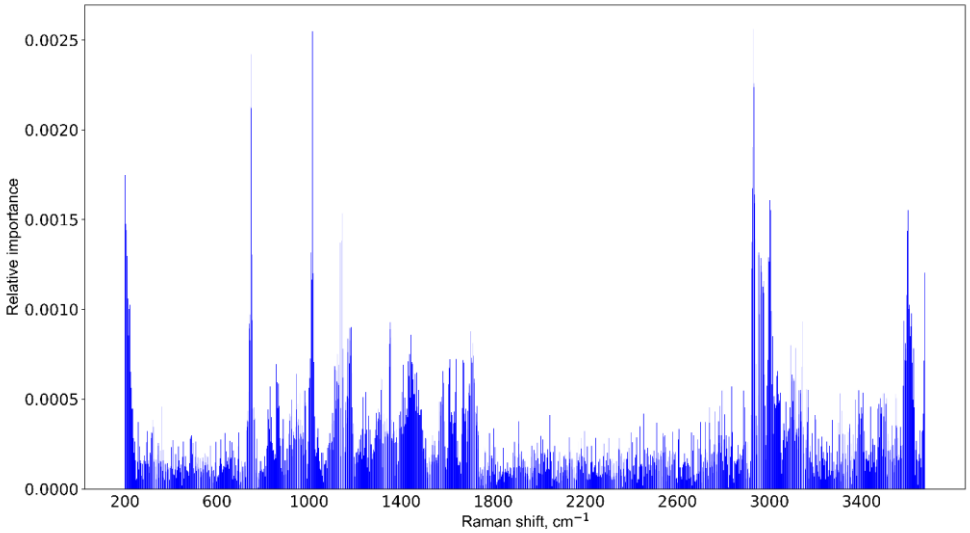

Figure 7. Mean the informative frequencies for the third week of the experiment.

CONCLUSION

We studied the blood plasma of mice in the dynamics of the development of U87 glioblastoma. Machine learning methods made it possible to conduct an analysis to separate the experimental and control groups in terms of sensitivity, specificity, and accuracy. The best models were obtained for a period of 21 days after inoculation of tumor cells when there is the greatest difference in the Raman spectra of control and experimental animals. The analysis of load matrices was carried out, and the most informative frequencies associated with specific metabolites, the concentration of which changes in the blood in the dynamics of tumor growth, were determined.

ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basic Research (grant # 19-52-55004), the Ministry of Science and Higher Education of the Russian Federation within the Agreement No. 075-15-2019-1950, within the State assignment FSRC "Crystallography and Photonics" RAS, by the Interdisciplinary Scientific and Educational School of Moscow University "Photonic and Quantum Technologies. Digital Medicine". The research was carried out with the support of a grant under the Decree of the Government of the Russian Federation No. 220 of 09 April 2010 (Agreement No. 075-15-2021-615 of 04 June 2021).

REFERENCES

[1] Kong,K., Kendall, C., Stone, N. and Notingher, I., "Raman spectroscopy for medical diagnostics - From in-vitro biofluid assays to in-vivo cancer detection", Adv Drug Deliv Rev 89, 121–134 (2015).

[2] Galli, R., Meinhardt, M., Koch, E., Schackert, G., Steiner, G., Kirsch, M., and Uckermann, O., "Rapid Label-Free Analysis of Brain Tumor Biopsies by Near Infrared Raman and Fluorescence Spectroscopy—A Study of 209 Patients", Front. Oncol. 9, 1165 (2019).

[3] Kopec, M., Błaszczyk, M., Radek, M. and Abramczyk, H., "Raman imaging and statistical methods for analysis various type of human brain tumors and breast cancers", Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 262, 120091 (2021).

[4] Sciortino, T., Secoli, R., d'Amico, E., Moccia, S., Conti Nibali, M., Gay, L., Rossi, M., Pecco, N., Castellano, A., De Momi, E. et al., "Raman Spectroscopy and MachineLearning for IDH Genotyping of Unprocessed Glioma Biopsies", Cancers 13, 4196 (2021).

[5] Uckermann, O., Yao, W., Juratli, T. A., Galli, R., Leipnitz, E., Meinhardt, M., Koch, E., Schackert, G., Steiner, G. and Kirsch, M., "IDH1 mutation in human glioma induces chemical alterations that are amenable to optical Raman spectroscopy", J. Neurooncol. 139(2), 261-268 (2018).

[6] Zhou, Y., Liu, C.-H., Wu, B. et al., "Optical biopsy identification and grading of gliomas using label-free visible resonance Raman spectroscopy," J. Biomed. Opt. 24(9), 095001 (2019).

[7] Qi, D. and Berger, A. J. " Chemical concentration measurement in blood serum and urine samples using liquid-core optical fiber Raman spectroscopy", Appl. Opt. 46, 1726-1734 (2007).

[8] Atkins, C. G., Buckley, K., Blades, M. W. and Turner, R. F. B. "Raman Spectroscopy of Blood and Blood Components", Applied Spectroscopy 71(5) 767–793 (2017).

[9] Pichardo-Molina, J. L., Frausto-Reyes, C., Barbosa-Garcia, O. et al, "Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients", Lasers. Med. Sci. 22, 229-236 (2007).

[10] Byrne, H. J., Bonnier, F., McIntyre, J. and Parachalil, D. R., " Quantitative analysis of human blood serum using vibrational spectroscopy", Clin. Spectrosc. 2, 100004 (2020).

[11] Mankova, A. A., Cherkasova, O. P., Lazareva, E. N., et. al., "Study of Blood Serum in Rats with Transplanted Cholangiocarcinoma Using Raman Spectroscopy", Opt. Spectrosc. 128 (7), 964-971 (2020).

[12] Miller, K. M., Ostrom, Q. T., Kruchko, K., et al., "Brain and other central nervous system tumor statistics", CA Cancer J Clin 0, 1–26 (2021).

[13] Cherkasova, O., Peng, Y., Konnikova, M., Kistenev, Y., Shi, C., Vrazhnov, D., Shevelev, O., Zavjalov, E., Kuznetsov, S. and Shkurinov, A., "Diagnosis of Glioma Molecular Markers by Terahertz Technologies", Photonics 8, 22 (2021).

[14] Zavjalov, E. L., Razumov, I. A., Gerlinskaya, L. A. and Romashchenko, A. V., "In vivo MRI Visualization of U87 Glioblastoma Development Dynamics in the Model of Orthotopic Xenotransplantation to the SCID Mouse", Russ. J. Genet. Appl. Res. 6 (4), 448–453 (2016).
[15] Shevelev, O. B., Seryapina, A. A., Zavjalov, E. L., Gerlinskaya, L.A., Goryachkovskaya, T. N., Slynko, N. M., Kuibida, L.V., Peltek, S. E., Markel, A. L. and Moshkin, M. P., "Hypotensive and neurometabolic effects of intragastric Reishi (Ganoderma lucidum) administration in hypertensive ISIAH rat strain", Phytomedicine 41, 1–6, (2018).
[16] Kistenev, Y. V., Karapuzikov, A. I., Kostyukova, N. Y., Starikova, M. K. et al., "Screening of patients with bronchopulmonary diseases using methods of infrared laser photoacoustic spectroscopy and principal component analysis", J. Biomed. Opt. 20, 065001 (2015).
[17] Konnikova, M. R., Cherkasova, O. P., Nazarov, M. M., Vrazhnov, D. A., et al., "Malignant and benign thyroid nodule differentiation through the analysis of blood plasma with terahertz spectroscopy," Biomed. Opt. Express 12, 1020-1035 (2021).
[18] Kistenev, Y. V., Borisov, A. V. and Shapovalov, A. V., "Determination of component concentrations in models of exhaled air samples using principal component analysis and canonical correlation analysis", Proc. SPIE 9810, International Conference on Atomic and Molecular Pulsed Lasers XII, 98101Z (2015).
[19] Kistenev, Y. V., Shapovalov, A. V., Borisov, A. V., Vrazhnov, D. A., Nikolaev, V. V., Nikiforova O. Y., "Applications of principal component analysis to breath air absorption spectra profiles classification", Proc. SPIE 9810, International Conference on Atomic and Molecular Pulsed Lasers XII, 98101Y (2015).
[20] Kistenev, Y. V., Karapuzikov, A. I., Kostyukova, N. Yu., et al. "Screening of patients with bronchopulmonary diseases using methods of infrared laser photoacoustic spectroscopy and principal component analysis", J. Biomed. Opt. 20(6), 065001 (2015).
[21] Kistenev, Y. V., Kuzmin, D. A., Vrazhnov, D. A., Borisov, A. V., "Classification of patients with broncho-pulmonary diseases based on analysis of absorption spectra of exhaled air samples with SVM and neural network algorithm application", Proc. SPIE 10035, 22nd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, 1003507 (2016).
[22] Kistenev, Y.V., Bukreeva, E.B., Bulanova, A.A. et al., "Laser spectroscopy and chemometric study of the specific features of air exhaled by patients with lung cancer and chronic obstructive pulmonary disease", Phys. Wave Phen. 22, 210–215 (2014).