Diagnostics of infectious and bronchial-pulmonary diseases using photoacoustic spectroscopy of the man-expired air

B.G. Ageev,[1] Yu.V. Kistenev,[2] E.P. Krasnozhenov,[2] O.Yu. Nikiforova,[1] E.S. Nikotin,[2] G.S. Nikotina,[1,2] Yu.N. Ponomarev,[1] and V.A. Fokin[2]

[1]V.E. Zuev Institute of Atmospheric Optics, Siberian Branch of the Russian Academy of Sciences, Tomsk
[2]Siberian State Medical University, Tomsk

Received August 1, 2008

The aim of this work is the investigation of gas emission of various pathogenic bacterium cultures using the photoacoustic method of laser spectroscopy, as well as the estimation of possibility of noninvasive express diagnostics of infectious and bronchial-pulmonary diseases by the spectroscopic analysis of the man-expired air. Absorption spectra were recorded with an intracavity photoacoustic sensor on the base of a waveguide CO_2 laser with discrete frequency tuning. The investigation of spectra was carried out by the data mining method. It was established that, in spite of similar, on the whole, behavior of bacterium gas emission spectra, available distinctions are sufficient to select some bacterium kinds. Measurements were carried out with a test group consisting of 100 persons and a group of patients (60 persons) affected by different bronchopulmonary pathologies. It was shown that absorption spectra allow the differentiation of patients according to their pathology.

Introduction

Methods for noninvasive diagnostics of states of biological objects are world-wide in their distribution in last decades, in particular, spectral gas analysis of gas exchange between biosystems and environment. In this case, the object of the study is, e.g., the man-expired air.[1–4]

Along with main components, the man-expired air can contain molecules-biomarkers of endogenous origin, the specificity of emission of which is determined by pathogenic processes in organism caused by different infections. A large group of infections concerns with the exposure to pathogenic bacteria, vital functions of which are accompanied by both their own gas emission and the action on the gas exchange of the human organism.

In this work, we present the results of investigation of the gas emission by a number of pathogenic bacteria by laser spectroscopy methods and the estimate of capability of noninvasive instant diagnosis of infectious and bronchial-pulmonary diseases on the base of spectroscopic analysis of the man-expired air. When solving the above problems, the following aspects were taken into account.

1. The specificity of the gas emission by bacterium cultures is caused by a low gas-exchange rate of bacteria in comparison with macroobjects and small doses of emitted gas-markers. In this case, the use of PA laser gas analysers is most optimal.

2. A large number of absorbing components in the recorded spectra essentially embarrass the use of traditional spectroscopic gas analysis methods for multicomponent mixtures; therefore, we used additionally the method of experimental data mining.

1. Technique for measurement and analysis of spectroscopic data

Absorption spectra of gas emission of bacterium colonies and samples of the man-expired air were obtained with the use of an intracavity laser PA sensor ILPA.[5,6] The sensor has been designed for real-time detection of gas admixtures in air with absorption bands of the 9.2–10.8 μm spectral range. The list of gases detectable by sensors of such type consists of several tens of names.[7]

The principle of operation of the sensor is based on the PA effect, arising when absorbing laser radiation by gases. A waveguide CO_2 laser[7] used in ILPA emits at 70 lines of the ground-state isotope of the carbonic acid molecule ($^{12}C^{16}O_2$). A resonance differential flow PA detector is mounted inside the cavity of the CO_2 laser, which allows an exciting radiation power of about 100 W to be attained at certain lasing lines.

The measured PA signal A at the frequency ν is defined by the equation[7]

$$A(\nu) = U / W = \alpha\Big[\sum C_i k_i(\nu) + \beta(\nu)\Big],$$

where U and W are the microphone acoustic signal and laser radiation power; α is the sensitivity of the PA receiver; C_i and $k_i(\nu)$ are the concentration and absorptivity of the ith absorbing gas at the frequency ν; β is the background absorptivity.

To eliminate the effect of the background absorption (absorption by other gases and nonselective absorption by walls and windows of the PA cell) in gas concentration measurements, the differential absorption method is used, i.e., measurements are carried out at two laser lines: lines with the frequency

0235-6880/08/12 969-06 $02.00

ν_{on}, where absorption by the gas under study is maximal, and lines with the frequency ν_{off}, where such absorption is negligible and the contribution of the background in the recorded signal is maximal. If the lines are chosen sufficiently close and $\beta(\nu_{on}) = \beta(\nu_{off})$, then the concentration C of the gas under study can be obtained by the equation[7]

$$C = \frac{1}{\alpha} \frac{A(\nu_{on}) - A(\nu_{off})}{k(\nu_{on}) - k(\nu_{off})}.$$

To determine the sensitivity, 35 scans for the control gas mixture of 5000 ppm CO_2 in nitrogen were recorded during 2 hours. Optimal couples of spectral lines were chosen with the use of computer code LPM[8]: $10R(18)$–$10R(32)$ and $10R(22)$–$10R(34)$ for the $10R$-branch with an expectable error of CO_2 concentration determination of 27–29%, as well as $10P(22)$–$10P(30)$ and $10P(18)$–$10P(8)$ for the $10P$-branch with an expectable error of 37%.

Calculated sensitivities of the ILPA sensor for four chosen line couples are shown in Fig. 1.

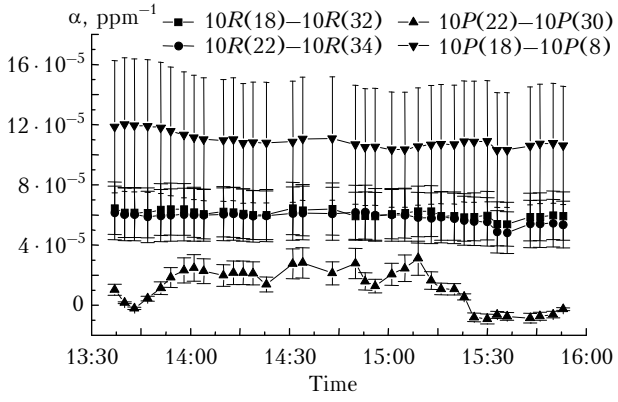

Fig. 1. Sensor sensitivity calculated by the measurements for four couples of CO_2 lasing lines in the 10 μm band.

The quite large sensitivity error is caused by the relatively "smooth" carbonic acid spectrum, therefore absorptivities at lines ν_{on} and ν_{off} differ only by several percents.

It is seen that more stable in time and compatible values of α are obtained when using line couples of the $10R$-branch. Figure 2 shows α calculated only for this branch.

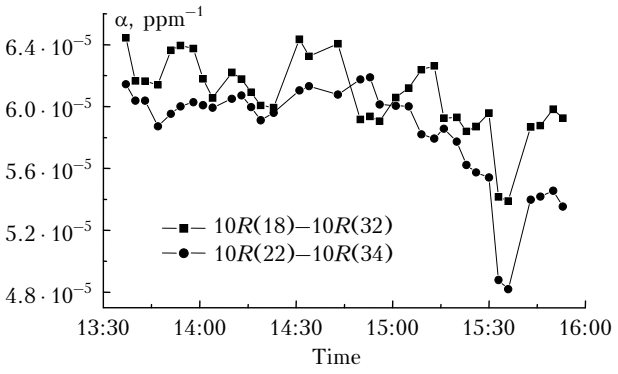

Fig. 2. Sensor sensitivity calculated by the measurements for two couples of CO_2 lasing lines for the $10R$-branch.

After deleting the values corresponding to the measurement time 15:33 and 15:36 and averaging, the ILPA sensitivity $\alpha = 6.1 \cdot 10^{-5}$ ppm^{-1} for the line couple $10R(18)$–$10R(32)$ and $\alpha = 5.9 \cdot 10^{-5}$ ppm^{-1} for the couple $10R(22)$–$10R(34)$ were obtained. The $\alpha = 6 \cdot 10^{-5}$ ppm^{-1} was used later on in calculations of gas concentrations in mixtures.

To estimate the spectrum specificity, the data mining method was used, based on the integral criterion of estimation of the studied object state.[9]

According to this method, a quantitative estimate of the biosystem state was carried out by the specificity of the gas emission absorption spectrum for the object under study with respect to the absorption spectrum, characterizing its reference state. The estimation technique was as followed.

Let S_0 and S be the preset reference and estimated states, characterized by the spectrum sets $\{\mathbf{b}_j \mid j \in N_{S_0}\}$ and $\{\mathbf{b}_i \mid i \in N_S\}$, respectively. Here $\mathbf{b}$ is the vector with absorption intensities at preset frequencies as coordinates; N_{S_0} and N_S are the numbers of objects of the reference and estimated states (sample volumes). The value of quantitative estimate of the object state $\mathbf{b}_i \in S$ can be characterized by its measure of closeness to the reference state S_0; in this case, relative positions of objects representing the reference system state in the attribute space are to be taken into account. Considering this, the criterion of integral estimate (IE) of the closeness of the object, characterized by the spectrum $\mathbf{b}_i$, to the state S_0 can be defined as

$$I_{S_0}(\mathbf{b}_i) = \frac{1}{2m} d(\mathbf{b}_i, S_0),$$

where $d(\mathbf{b}_i, S_0)$ is calculated as an average distance from the spectrum $\mathbf{b}_i$ to spectra from S_0; m is the dimension of the attribute space.

When using criteria based on multivariate methods for data analysis, the main problem is small volume of samples, characterizing the reference state. This results in significant variability of estimates, obtained with their use. To obtain effective estimates of some state, we used the Monte Carlo method for reference data modeling. To do this, we modeled M sampling sets X_k ($k = \overline{1,M}$) of sufficiently large volume, corresponding to statistical characteristics of the reference state S_0 represented by the initial set of objects $X : \{\mathbf{b}_i \mid i \in \overline{1,N_{S_0}}\}$. Then, estimates $I_{S_0,k}(\mathbf{b})$ were calculated for each set X_k, distribution of them was used further to obtain quantitative estimates of the state under study. I_{S_0} were estimated with the use of the developed computer program[10] by means of modeling samples, corresponding to volumes of $N_{S_0} = 1000$ observations and $M = 200$ under the assumption that parameters of the reference sample satisfy the multinormal distribution law.

The fragment of file with results of the gas mixture spectrum measurement is shown in Fig. 3. Here "LineFreq" is the frequency, cm^{-1}, "mic" is the current value of the microphone signal, dB; "pyro" is the current value of the pyroreciever signal, dB;

"absorbtion" is the absorbtion strength (the ratio of the microphone signal to the pyroreceiver one); "difference" is the difference between the current and background absorptivities.

LineFreq	mic	pyro	absorbtion	difference
931.27	9.11095e-06	9.35643e-05	0.097413	0.097413
933.18	0.0234707	0.464001	0.0505834	0.0505834
935.19	0.0247729	0.509786	0.0485947	0.0485947
937.03	0.0311982	0.613976	0.0508134	0.0508134
938.97	0.0296406	0.39835	0.0744086	0.0744086
940.82	0.0354193	0.346142	0.102326	0.102326
943.13	0.0532001	0.523279	0.101667	0.101667
944.47	0.0731994	0.552094	0.132584	0.132584
946.25	0.0422652	0.33654	0.125588	0.125588
947.96	0.0718149	0.317219	0.226389	0.226389
949.76	0.0559656	0.318392	0.175776	0.175776
951.47	0.0515613	0.24358	0.211681	0.211681
953.11	0.0345685	0.171841	0.201166	0.201166
954.84	0.0230406	0.176352	0.130652	0.130652
956.48	0.00557059	0.0550645	0.101164	0.101164
958.04	1.33444e-05	9.01116e-05	0.148179	0.148179
965.06	1.17192e-05	7.91985e-05	0.148131	0.148131
966.56	0.0136815	0.0898731	0.152232	0.152232
967.96	0.0375461	0.192856	0.194686	0.194686

Fig. 3. The fragment of file with ILPA-sensor measurements.

The absorption strength values at CO_2 lasing lines within the 932–954 and 966–982 cm^{-1} ranges were used as coordinates of vectors **x** of the biosystem state. Such choice of data is caused by the best reproducibility of the absorption signal in the above ranges. Two integral estimates I_1 and I_2 of the biosystem state, obtained in such a way and corresponding to the above frequency ranges, determined its state in the form of point in the two-dimensional attribute space.

2. Investigation results of gas emission by bacteria at CO_2 lasing lines

The investigation process was as follows: a pure bacterium culture was inoculated in tubes with nutrient medium, which then were placed in a thermostat (37°C); then, at a certain stage of the colony growth, a small amount of air (2–5 cm^3) from a tube was taken into a receiver of ILPA sensor; later on, the absorption spectrum of the taken sample was scanned.

The recorded spectra of gas samples for several bacterium kinds are shown in Fig. 4, and the results of analysis of the spectra, obtained by the data mining method, are shown in Fig. 5.

It is seen from Figs. 4 and 5 that the recorded scans of air samples over colonies of different bacteria are similar in general; nevertheless, there is some difference for different kinds of bacteria both in values of the sensor signal and in the behavior of the spectral dependence. In general, despite the similar character of the spectra, the existing differences are enough to confidently distinguish at least some kinds (Proteus and Klebsiella) by the ratio of IE criterion values in the 932–954 and 966–982 cm^{-1} ranges.

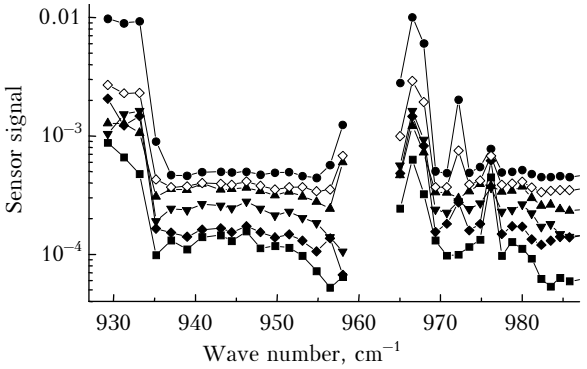

Fig. 4. Recorded spectra of gas samples for Klebsiella (–•–); Staphylococcus (–◇–); Pseudomonas aeruginosa (–▲–); Escherichia coli (–▼–); Proteus (–◇–); and nutrient broth (–■–).

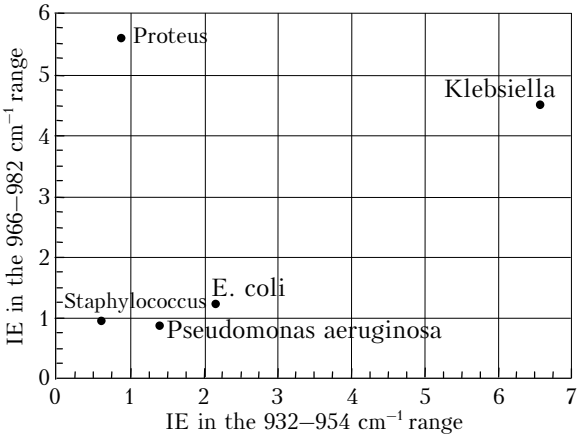

Fig. 5. Integral estimate of variations of gas emission content for different bacteria.

ILPA sensor-recorded scans of gas sample spectra are characterized by significant irregularity. Figure 6 shows the recorded scan of control gas mixture (5000 ppm CO_2 + N_2) and the absorption spectrum of this mixture calculated on the base of the HITRAN-2004 database.

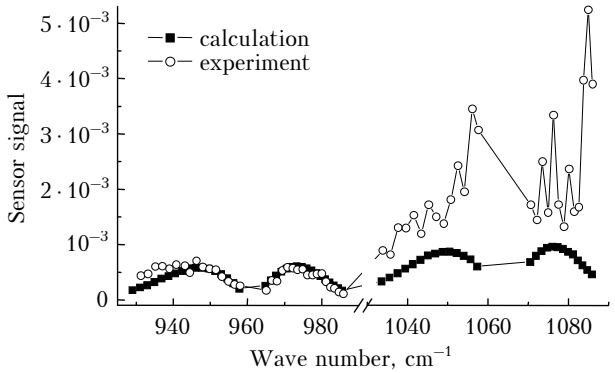

Fig. 6. The scan of the control gas mixture spectrum recorded on March 24, 2008 at 13:47 and the absorption spectrum of this mixture calculated on the base of the HITRAN-2004 database.

Such irregularity of the recorded spectra can result in large errors in determination of gas concentrations by the differential technique; therefore, gas concentrations were calculated by the single-wave technique in the following way. The carbonic acid absorbs at all CO_2 lasing lines; therefore, first, the correlation coefficients of the microphone signal (column "mic" in Fig. 3) with the laser radiation power (column "pyro") were calculated for all lasing lines. Three lines were chosen in centers of 10*P*- and 10*R*-branches, to which sufficiently high correlation coefficients answered, and the carbonic acid concentration was determined by averaging the values obtained for each line. Then it was assumed that the background signal was caused mainly by the carbonic acid absorption: $\beta(\nu) = C_{CO_2} k_{CO_2}(\nu)$; therefore, the contribution of carbonic acid was preliminary subtracted when calculating concentrations of water vapor and ammonia: $C_{H_2O,\,NH_3} = [A(\nu)/\alpha - \beta(\nu)]/k_{H_2O,\,NH_3}$. The concentration of water vapor was calculated at 10*R*(20) (975.93 cm^{-1}) and of ammonia – at 10*R*(6) and 10*R*(8) (966.25 and 967.71 cm^{-1}), where contribution of these gases in absorption is large.

Concentrations of components were calculated for mixtures of three types. First, concentrations of three components under study were estimated in the control gas mixture (5000 ppm $CO_2 + N_2$). When studying different bacteria, 5 scans were recorded for each air sample and the results were averaged; therefore, concentrations of components for the control gas mixture were calculated after averaging the results obtained for five consequently recorded scans.

Thus, 7 values of component concentrations were obtained for the whole 2-hour measurement series, each value was a 15–20 min average. In addition, about 30 scans of the same air sample were recorded during 2 h in one of the measurement series above the surface of the distilled water, and this experiment was repeated in another series. Correspondingly, 6 values of average concentrations were obtained for a 2-hour measurement run. Finally, air samples were studied above the surface of nutrient broth and colonies of different bacteria. Three-time air intake was carried out from each tube; 5 scans were recorded for each air sample. The calculated concentrations of components in each sample are shown in Figs. 7–9.

The calculation error of the carbonic acid concentration in the control gas mixture did not exceed 10%, the concentration of water vapor in samples above the distilled water corresponded to 70–80% of the air humidity at the room temperature. However, two circumstances should be noted. First, when analyzing one air sample, the trend is observed toward the decrease in concentrations of the components under study (Figs. 7 and 8). Second, negative values of concentrations of water vapor and ammonium were obtained in studying the control gas mixture.

For quantitative gas analysis, the mixture composition is to be known;[7] hence, concentrations of H_2O and NH_3 have been determined incorrectly. However, when analyzing biogenic air samples, it is desirable at least to have possibility to estimate concentrations of a fixed set of components, even if the mixture composition is not known surely.

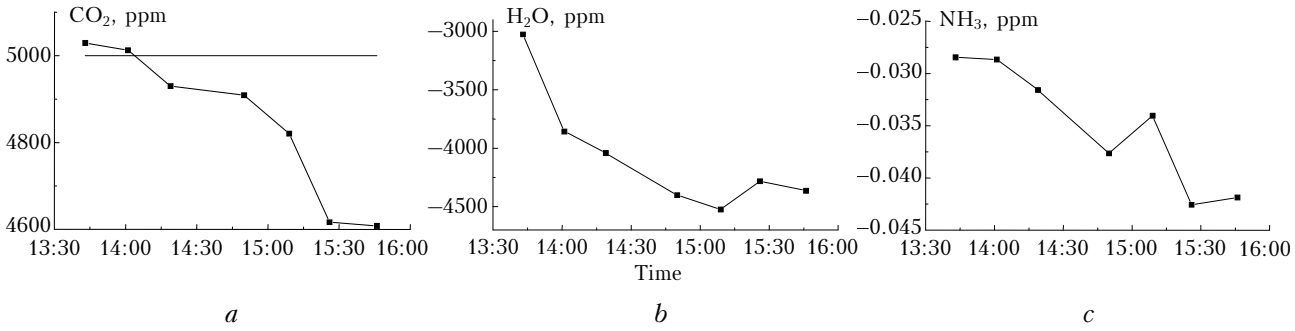

Fig. 7. Concentrations of carbonic acid (*a*), water vapor (*b*), and ammonia (*c*), obtained for the control gas mixture (5000 ppm $CO_2 + N_2$).

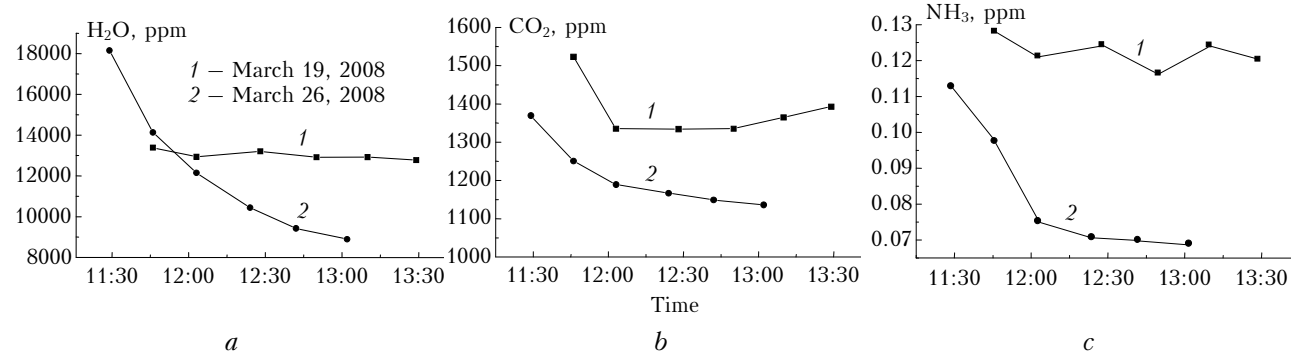

Fig. 8. Concentrations of water vapor (*a*), carbonic acid (*b*), and ammonia (*c*), obtained for an air sample above the surface of distilled water.

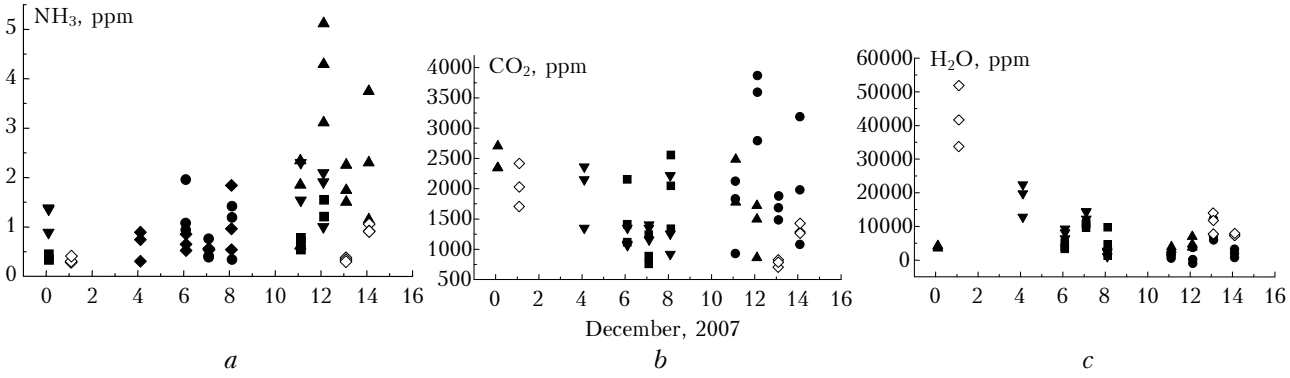

Fig. 9. Concentrations of ammonia (*a*), carbonic acid gas (*b*), and water vapor (*c*) in air samples above the bacterium colonies: ■ – nutrient broth; ● – Pseudomonas aeruginosa; ▲ – Klebsiella; ▼ – Escherichia coli; ◆ – Proteus; ◇ – Staphylococcus.

The values of CO_2 and NH_3 concentrations, being essentially higher than the background values (350 and $5 \cdot 10^{-4}$ ppm), which are characteristic of the atmospheric air, were obtained in analysis of air samples above the surface of the distilled water. This can be caused both by the essential difference between the air compositions in the room, where the experiment was conducted, and clear air and by a non-adequate technique for determination of the concentration.

According to results of the quantitative analysis of air samples above the colonies of bacteria, note the following: the highest concentration of ammonia is characteristic for Klebsiella, while the lowest one (at the level of absorption by the nutrient broth) – for staphylococcus. The results are similar for three other kinds of bacteria.

Note that according to results of data mining other bacterium kinds were distinguished, i.e., Proteus and Klebsiella. This can well reflect the fact that data mining gives a general integral estimate of an object state, which cannot be always connected with the presence of some components in air samples. Note also that the obtained results are to be considered as preliminary and requiring further investigations both in the measuring procedure and the processing technique.

3. The results of the study of man-expired air at CO_2 lasing lines

The main components of atmospheric air (oxygen and nitrogen) do not absorb at CO_2 lasing lines. Absorbing components are mainly carbonic acid, contributing to the signal at all radiation wavelengths, and water vapor, the contribution of which prevails over other components only at several wavelengths. The noticeable absorption can be also caused by ozone, methane, ammonia, hydrogen peroxide, sulfur dioxide, nitric oxides, ethylene, ethane, and bensole.

The man-expired air is not homogeneous. An adult healthy person expires 0.5 l of air or more for each breathing cycle. The first 150 ml represent upper air passages, where gas exchange is absent. To be informative, air is to be expired from the lung depth. Therefore, we used for analysis an expiratory reserve volume, exhaled by a patient into a syringe after a corresponding instruction and under the control of a medical man. At the first stage of investigation of the man-expired air (spectroscopic analysis), gas samples without any selection of gas components were used to avoid the information loss.

To sample the man-expired air, a calibrating gas syringe of a commercial spirometer of 1 l in volume was used. The mouthpieces part was removable to avoid transmission of infection from one patient to another.

Measurements were carried our in a control group of healthy people and in a group of patients with different diseases.

Figure 10 shows spectra of the man-expired air, obtained with the ILPA sensor in a group of patients with different diseases (from the 3rd Tomsk City Hospital). The control group consisted of 100 persons.

The estimation results of the measure of variability of the man-expired air composition are shown in Fig. 11.

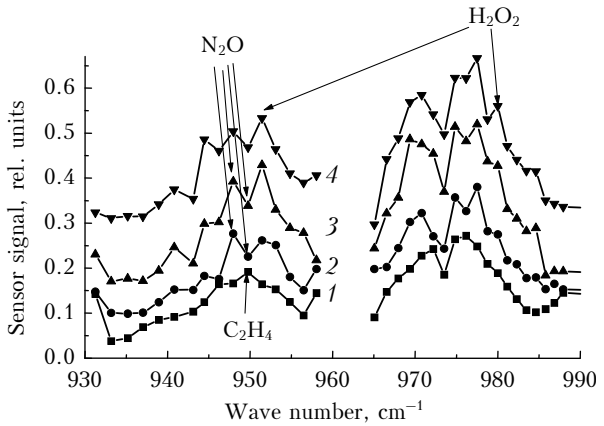

Fig. 10. Scans of the man-expired air of several patients of the 3rd Tomsk City Hospital with stomach ulcer (*1*), bronchial asthma (*2*), COPD (*3*), and cardio ischemia (*4*).

The reference state consisted of samples of the air, expired by healthy persons ($n = 31$). The states under estimation are samples corresponding to some bronchial-pulmonary diseases: BA – bronchial asthma ($n = 11$), COPD – chronic obstructive pulmonary disease ($n = 16$), P – pneumonia ($n = 13$), and CI –

cardio ischemia ($n = 11$); the parameters of the reference sample modeling were $N_{S_0} = 1000$ and $M = 200$.

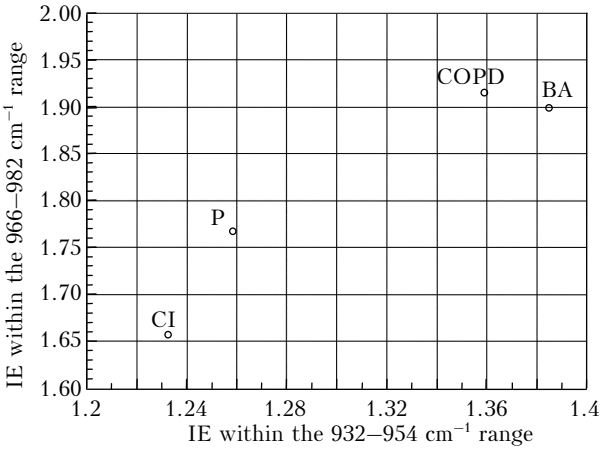

Fig. 11. Integral estimate of variations of the man-expired air at some deceases.

Conclusion

As a result of the performed investigations, we have stated that in spite of similar on the whole behavior of the bacterium gas emission spectra, available distinctions are sufficient to surely select at least several bacteria kinds by the data mining method.

It has been also shown that the recorded absorption spectra of the man-expired air allow differentiation of patients by their diseases. The obtained integral estimates of the man-expired air composition at pathologies under study can be medically interpreted. Cardio ischemia is the pathology, less affecting the bronchopulmonary system, which is reflected in a less deviation of the IE of the composition of the man-expired air from the norm as compared with other analyzed states. At the same time, bronchial asthma and chronic obstructive pulmonary diseases are chronic processes close in aetiology. In Fig. 11, points, answering these states, are close to each other and significantly farther from the reference state than pheumonia. This can be interpreted as the appearance of system shifts in organism functioning when some process becomes chronical.

Acknowledgements

This work was fulfilled under the financial support of Russian Foundation for Basic Research (Grant No. 08-02-99031-r_ofi) and State Contract No. 217/3 of 07.06.2008.

Reference

1. E.V. Stepanov and V.A. Milyaev, Usp. Fiz. Nauk **170**, No. 4, 458–462 (2000).
2. E.V. Stepanov and V.A. Milyaev, Quant. Electron. **32**, No. 11, 987–992 (2002).
3. T.H. Risby and S.F. Solga, Appl. Phys. B **85**, 421–426 (2006).
4. M.R. McCurdy, Yu. Bakhirkin, G. Wysocki, R. Lewicki, and F.K. Tittel, J. Breath Res. **1**, 014001 (12 pp), doi: 10.1088/1752-7155/1/1/014001 (2007).
5. Intracavity laser photo-acoustic sensor ILPA-1. Pasport. Technical specification. Service manual (LCS Facility Management Ltd., Novosibirsk).
6. B.G. Ageev, V.A. Kapitanov, Yu.N. Ponomarev, and V.A. Sapozhnikova, Atmos. Oceanic Opt. **20**, No. 9, 726–729 (2007).
7. V.I. Kozintsev, M.L. Belov, V.A. Gorodnichev, and Yu.V. Fedotov, *Laser Photo-acoustic Analysis of Multicomponent Gas Mixtures* (Publishing House of Bauman MSTU, Moscow, 2003), 352 pp.
8. V.V. Zuev, A.A. Mitsel', M.Yu. Kataev, I.V. Ptashnik, and K.M. Firsov, Comput. Phys. **9**, No. 6, 649–656 (1995).
9. S.I. Karas', Yu.V. Kistenev, O.Yu. Nikiforova, Ya.S. Pekker, V.A. Fokin, and A.V. Shapovalov, *Nonlinear Analysis of Biomedical Data* (Phublishing House of Tomsk Polytechnical University, Tomsk, 2006), 118 pp.
10. V.A. Fokin, I.S. Khakimov, and O.Yu. Nikiforova, *Computer program "StatSys"*, Inventor's application No. 2006613281; Appl. 09.29.2006; Publ. 11.22.2006.