ISSN 0868–5886 НАУЧНОЕ ПРИБОРОСТРОЕНИЕ, 2012, том 22, № 2, с. 21–26

УДК 621.391.26

© Ю. А. Титов, В. В. Манойлов, А. Г. Кузьмин

ВЫДЕЛЕНИЕ ИНФОРМАЦИОННЫХ ПРИЗНАКОВ В МАСС-СПЕКТРОМЕТРИЧЕСКИХ СИГНАЛАХ ПРОБ ВОЗДУХА

В работе предлагается использование нового метода выделения информационных признаков из обрабатываемых сигналов — обобщенного спектрального анализа в приспособленном базисе. Проведены исследования функциональной структуры и адаптируемой математической основы системы классификации сигналов различных типов, в том числе масс-спектров воздуха, полученных на квадрупольном анализаторе. Разработаны алгоритмы синтеза приспособленных к обрабатываемым сигналам базисных функций ортогональных преобразований. В качестве примера использования данного метода в аналитическом приборостроении рассматривается обработка масс-спектров, полученных на квадрупольном масс-спектрометре, для анализа чистоты воздуха.

Кл. сл.: методы обработки сигналов, масс-спектрометрия, оценка параметров масс-спектров воздуха, анализ алгоритмов обработки данных

ВВЕДЕНИЕ

Для выделения информационных признаков из масс-спектрометрических сигналов проб воздуха развивается новый метод обработки сигналов на основе применения обобщенного спектрального анализа в приспособленном базисе. На основе усредненных масс-спектров проб воздуха, не содержащих примесей, строится приспособленный базис для обобщенного спектрального анализа. Такой базис используется в качестве обучающего эталона. Далее приспособленный базис строится для масс-спектра исследуемого воздуха. После чего происходит сравнение двух этих базисов. Применение предлагаемого метода позволяет производить анализ проб воздуха и обнаруживать пробы воздуха, содержащие примеси.

Задача по выявлению информативных признаков в масс-спектрометрических сигналах может быть представлена как задача нахождения вектора $\mathbf{Y}$

$$\mathbf{Y} = F\mathbf{X},$$

где $\mathbf{X}$ — вектор исходных данных преобразования; F — оператор преобразования; $\mathbf{Y} = \left[y_1, y_2, ..., y_L\right]^{\mathrm{T}}$ — вектор признаков, который должен адекватно характеризовать вектор исходного описания с точки зрения задачи оценки параметров масс-спектра.

При решении этой задачи возникает вопрос выбора класса оператора преобразования F. Известно, что общий подход к отысканию процедур оценки признаков основан на линейных преобразованиях. Среди них наибольшее распространение получили ортогональные преобразования. При использовании ортогональных преобразований имеется возможность выбора систем базисных функций, наиболее приспособленных к решаемой задаче. При этом может обеспечиваться высокая вычислительная эффективность таких преобразований, которые могут выполняться по быстрым алгоритмам.

ОРТОГОНАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ. ОСНОВНЫЕ СВОЙСТВА

В рамках теории цифровой обработки сигналов понятие ортогонального преобразования может быть представлено в виде матричного уравнения

$$\mathbf{Y} = \mathbf{HX}, \tag{1}$$

где $\mathbf{H}$ — матрица ортогонального преобразования; $\mathbf{X} = \left[x_1, x_2, ..., x_N\right]^{\mathrm{T}}$ — вектор анализируемого сигнала; $\mathbf{Y} = \left[y_1, y_2, ..., y_N\right]^{\mathrm{T}}$ — вектор, спектральных коэффициентов (является отображением $\mathbf{X}$ в спектральной области).

Оператор преобразования $\mathbf{H}$ включает в себя систему базисных функций и имеет структуру

$$\mathbf{H} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix} = \begin{bmatrix} \left[h_{11} \cdots h_{1N}\right] \\ \left[h_{21} \cdots h_{2N}\right] \\ \vdots \\ \left[h_{M1} \cdots h_{MN}\right] \end{bmatrix},$$

где $\left[h_{i1}, h_{i2}, \cdots, h_{iN}\right]$ — вектор дискретных значений i-й базисной функции.

Система функций $\{h_i\}$ является базисной, если все функции ортогональны друг другу и образуют полную систему, для которой не существует ни одной другой функции, ортогональной ко всем остальным. При выполнении условия полноты имеет место равенство $M = N$. Система ортогональных базисных функций является нормированной (или ортонормированной), если норма каждой базисной функции в пространстве L_2 равна единице, т. е. для любого $1 \le i \le N$.

Выбор той или иной базисной системы производится по требованиям конкретной задачи анализа сигналов[1, 2]. При решении задачи оценки наличия примесей в масс-спектрах воздуха базис должен обеспечивать получение пространства информативных признаков невысокой размерности и выражаться в факторизованной форме, которая ведет к возможности оперативного выполнения ортогонального преобразования по быстрому алгоритму. Сокращение размерности при этом заключается в том, что при ортогональном преобразовании основная информативная (релевантная) часть исходной информации концентрируется в k спектральных коэффициентах $k << N$. За счет исключения остальных спектральных коэффициентов как неинформативных размерность вектора информативных признаков существенно уменьшается. Таким образом, следует выбирать ту базисную систему, которая обеспечивает максимальное уменьшение размерности вектора информативных признаков при минимальных информационных потерях. В работе [2] предложен новый подход к построению ортогональных преобразований с параметрически перестраиваемыми по форме базисными функциями с сохранением их ортогональности и полноты. Рассмотрим аппарат таких преобразований более подробно.

ПЕРЕСТРАИВАЕМЫЕ ОРТОГОНАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ

Существо метода перестраиваемых ортогональных преобразований состоит в таком факторизованном представлении матрицы спектрального оператора, в котором ненулевые элементы факторизованных матриц взаимосвязаны условиями ортонормированности и полноты и имеют степени свободы, обеспечивающие параметрическое формирование множества базисов с алгоритмами быстрых преобразований [4, 5]. В основе построения алгоритма лежит представление оператора преобразования H в виде произведения слабозаполненных (факторизованных) далее неразложимых матриц:

$$\mathbf{H} = \mathbf{G}_n \mathbf{G}_{n-1} \cdots \mathbf{G}_1,$$

$\mathbf{G}_i(\phi_{i1}, \cdots, \phi_{i\frac{N}{2}})$ — факторизованные (слабозаполненные далее неразложимые) матрицы, ненулевые элементы которых зависят от параметров ϕ_{ij}, $i = 1, \cdots, n$; $j = 1, \cdots, \frac{N}{2}$. Здесь n — число факторизованных матриц $\mathbf{G}_i$, в частности при $N = 2^n$ равно $\log_2 N$ (N — размерность оператора H). При этом каждая из матриц $\mathbf{G}_i$ содержит $\frac{N}{2}$ элементарных матричных блоков, обобщенная модель которых имеет параметрическую форму и названа в силу неразложимости спектральным ядром. Для операторов размерности $N = 2^n$ факторизованные матрицы $\mathbf{G}_i$ составляются из параметрических элементов ядер

$$\mathbf{V}_{ij} = \begin{bmatrix} \alpha_{ij} & \cdots & \gamma_{ij} \\ \beta ij & \cdots & \delta_{ij} \end{bmatrix}. \tag{2}$$

Пример структуры факторизованной матрицы $\mathbf{G}_i$ имеет вид:

$$\mathbf{G}_i = \begin{bmatrix} \begin{bmatrix} \alpha_{i1} & 0 & \cdots & 0 & \gamma_{i1} \\ \beta_{i1} & 0 & & 0 & \delta_{i1} \end{bmatrix} & 0 & \cdots & 0 \\ 0 \quad \begin{bmatrix} \alpha_{i2} & 0 & \cdots & 0 & \gamma_{i2} \\ \beta_{i2} & 0 & & 0 & \delta_{i2} \end{bmatrix} & & \cdots & 0 \\ \ddots & & \ddots & \\ 0 \quad \cdots \quad 0 & \begin{bmatrix} \alpha_{iN/2} & 0 & \cdots & 0 & \gamma_{iN/2} \\ \beta_{iN/2} & 0 & & 0 & \delta_{iN/2} \end{bmatrix} & & \end{bmatrix}. \tag{3}$$

В этом случае параметрические элементы матриц-ядер (2) вычисляются следующим образом:

$$\mathbf{V}_{ij} = \begin{bmatrix} \alpha_{ij} & \cdots & \gamma_{ij} \\ \beta_{ij} & \cdots & \delta_{ij} \end{bmatrix} = \begin{bmatrix} \cos(\phi_{i,j}) & \cdots & w_{i,j}\sin(\phi_{i,j}) \\ \sin(\phi_{i,j}) & \cdots & -w_{i,j}\cos(\phi_{i,j}) \end{bmatrix},$$

$w_{i,j} = \exp(j\theta_{i,j})$, $\phi \in [0, 2\pi]$, $\theta \in [0, 2\pi]$.

Задавая параметры элементов ядер $\phi_{i,j}$ и $\theta_{i,j}$, можно формировать спектральные операторы H.

Система базисных функций строится на основе обучающего эталона $\hat{\mathbf{R}}_{эт}$ в соответствии с условием (2). Элементы матриц находятся по итерационной процедуре. Для $N = 2$ имеем:

$$\begin{bmatrix} c_1 & s_1 \\ s_1 & -c_1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \Rightarrow \begin{cases} c_1 r_1 + s_1 r_2 = y_1, \\ s_1 r_1 - c_1 r_2 = y_2, \end{cases}$$

где $\begin{cases} c_1 = \cos(\phi_1), \\ s_1 = \sin(\phi_1). \end{cases}$

Равенство $\mathbf{Y}_1 = \mathbf{Y}_ц$ ($\mathbf{Y}_ц$ — целевой вектор) выполняется при $y_2 = s_1 r_1 - c_1 r_2 = 0$, откуда получаем

$$\frac{s_1}{c_1} = \frac{r_2}{r_1} \Rightarrow \phi_1 = \operatorname{arctg}\left(\frac{r_2}{r_1}\right).$$

Для $N = 4$ имеем $\mathbf{G}_2 \mathbf{G}_1 \hat{\mathbf{R}}_{эт} = Y$.

Первый шаг в процедуре — произведение $\mathbf{G}_1 \hat{\mathbf{R}}_{эт} = \mathbf{Y}_1$

$$\begin{bmatrix} c_{11} & 0 & s_{11} & 0 \\ s_{11} & 0 & -c_{11} & 0 \\ 0 & c_{12} & & s_{12} \\ 0 & s_{12} & & -c_{12} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \end{bmatrix},$$

где $\begin{cases} y_{11} = c_{11} r_1 + s_{11} r_3, \\ y_{12} = s_{11} r_1 - c_{11} r_3, \\ y_{13} = c_{12} r_2 - s_{12} r_4, \\ y_{14} = s_{12} r_2 - c_{12} r_4. \end{cases}$

Элементы y_{11} и y_{12} этого вектора определяются произведением первого ядра матрицы $\mathbf{G}_1$ на элементы r_1 и r_3

$$\begin{bmatrix} c_{11} & s_{11} \\ s_{11} & -c_{11} \end{bmatrix} \begin{bmatrix} r_1 \\ r_3 \end{bmatrix} = \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix},$$

и соответственно y_{13} и y_{14} находятся аналогично через произведение второго ядра на элементы r_2 и r_4

$$\begin{bmatrix} c_{12} & s_{12} \\ s_{12} & -c_{12} \end{bmatrix} \begin{bmatrix} r_2 \\ r_4 \end{bmatrix} = \begin{bmatrix} y_{13} \\ y_{14} \end{bmatrix}.$$

Параметры — $\phi_{11} = \operatorname{arctg}\left(\frac{r_1}{r_3}\right)$, $\phi_{12} = \operatorname{arctg}\left(\frac{r_2}{r_4}\right)$; результирующий вектор — $\mathbf{Y}_1 = [y_{11}, 0, y_{13}, 0]^{\mathrm{T}}$.

Для второго шага произведения по аналогии имеем $\mathbf{G}_2 \mathbf{Y}_1 = \mathbf{Y}_2$

$$\begin{bmatrix} c_{21} & 0 & s_{21} & 0 \\ s_{21} & 0 & -c_{21} & 0 \\ 0 & c_{12} & & s_{12} \\ 0 & s_{12} & & -c_{12} \end{bmatrix} \begin{bmatrix} y_{11} \\ 0 \\ y_{13} \\ 0 \end{bmatrix} = \begin{bmatrix} y_{21} \\ y_{22} \\ y_{23} \\ y_{24} \end{bmatrix},$$

$$\begin{bmatrix} c_{21} & s_{21} \\ s_{21} & -c_{21} \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{13} \end{bmatrix} = \begin{bmatrix} y_{21} \\ y_{22} \end{bmatrix}, \begin{bmatrix} c_{21} & s_{21} \\ s_{21} & -c_{21} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} y_{21} \\ y_{22} \end{bmatrix}.$$

Параметры — $\phi_{21} = \operatorname{arctg}\left(\frac{y_{11}}{y_{13}}\right)$, $\phi_{22} = 0$, и достигается искомое равенство $\mathbf{Y}_2 = \mathbf{Y}_ц = [y_{21}, 0, 0, 0]^{\mathrm{T}}$.

ПРИМЕНЕНИЕ ПРИСПОСОБЛЕННОГО ПРЕОБРАЗОВАНИЯ К АНАЛИЗУ МАСС-СПЕКТРОВ ВОЗДУХА

При анализе масс-спектров, полученных в различных экспериментах одной пробы и от различных проб, существуют ситуации, при которых необходим количественный сравнительный анализ масс-спектров. В ряде случаев применение приспособленных преобразований может решить задачу количественных сравнений масс-спектров проб чистого воздуха и проб воздуха с примесями.

Критериями для принятия решения о наличии различий в сравниваемых масс-спектрах является сравнение норм и квадратов максимальных значений начальных компонент в векторах, полученных в результате выполнения преобразования соответственно эталона и тестируемых сигналов. В качестве обучающего эталона используется усредненный масс-спектр чистого воздуха. Данный подход обработки сигналов был исследован для изучения возможностей его применения в масс-спектрометрии. Систему приспособленных базисных функций, отвечающих требованиям ортонормированности и полноты, предлагается строить на основе обучающего эталона $\bar{\mathbf{R}}_{эт}$. Оператор ортогонального преобразования считается приспособленным к $\bar{\mathbf{R}}_{эт}$, если выполняется условие

$$\frac{1}{N}\mathbf{H}_n\bar{\mathbf{R}}_{эт} = \mathbf{Y}_ц = \left[y_{ц1}, 0, \ldots, 0\right]^{\mathrm{T}}.$$

Здесь $\mathbf{H}_n$ — приспособленный оператор; $\mathbf{Y}_ц$ — целевой вектор, принимаемый в качестве критерия настройки оператора; $\bar{\mathbf{R}}_{эт} = \left[r_{эт,1}, \ldots, r_{эт,N}\right]^{\mathrm{T}}$ — вектор обучающего эталона. Элементы матриц $\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_i, \mathbf{G}_{i+1}$ находятся по итерационной процедуре, описанной в работах [2, 3, 6]. На каждом шаге итерации заполненность матриц уменьшается в 2 раза. На $(n-1)$-м шаге мы получаем диагональную матрицу, а на n-м шаге — одно число. После построения приспособленного оператора преобразования $\mathbf{H}$ вектор преобразованного сигнала получается по формуле $\mathbf{Y} = \mathbf{H}\mathbf{F}^{\mathrm{T}}$, где $\mathbf{F}$ — исходный сигнал. Критерием (целевым вектором) для принятия решения о наличии примеси является сравнение норм и квадратов максимальных значений начальных компонент в векторах, полученных в результате выполнения преобразования соответственно стандартного образца и тестируемых сигналов.

Вектор $\mathbf{Y}$, полученный в результате преобразования исходного сигнала в приспособленном базисе, сравнивается с вектором $\mathbf{Y}_{эт}$ чистого воздуха, полученным также в результате преобразования в приспособленном базисе. Эти два вектора информационно одинаковы, если

$$\begin{cases} \left\|\mathbf{Y} - \mathbf{Y}_{эт}\right\| \le \delta, \\ \left|Y(1)^2 - Y_{эт}(1)^2\right| \le \varepsilon. \end{cases} \tag{4}$$

В противном случае масс-спектры, которые сравниваются, информационно различны. Параметры δ и ε выбираются в процессе обучения с использованием информации о величине среднего квадратичного отклонения (СКО) шума. По величинам нормы разности векторов $\left\|\mathbf{Y} - \mathbf{Y}_{эт}\right\|$ и разности квадратов главных компонент $\left|Y(1)^2 - Y_{эт}(1)^2\right|$ можно судить о количественном информационном отличии масс-спектров, которые сравниваются. Полученные в результате сравнения числовые величины не обязательно могут свидетельствовать о наличии загрязнений в воздухе, но могут дать информацию об отличиях масс-спектров измеряемой пробы от усредненного масс-спектра чистого воздуха.

Практическое применение

Рассмотрим применение данного метода в масс-спектрометрии. На рис. 1 представлен масс-спектр воздуха, который получен путем многократных измерений газовой смеси заранее известного состава. Такой масс-спектр был выбран в качестве эталонного сигнала $\hat{\mathbf{R}}_{эт}$. На рис. 2 представлен масс-спектр воздуха содержащего примеси. Легко заметить, что содержание азота (масса 28) во втором масс-спектре выше, чем в первом.

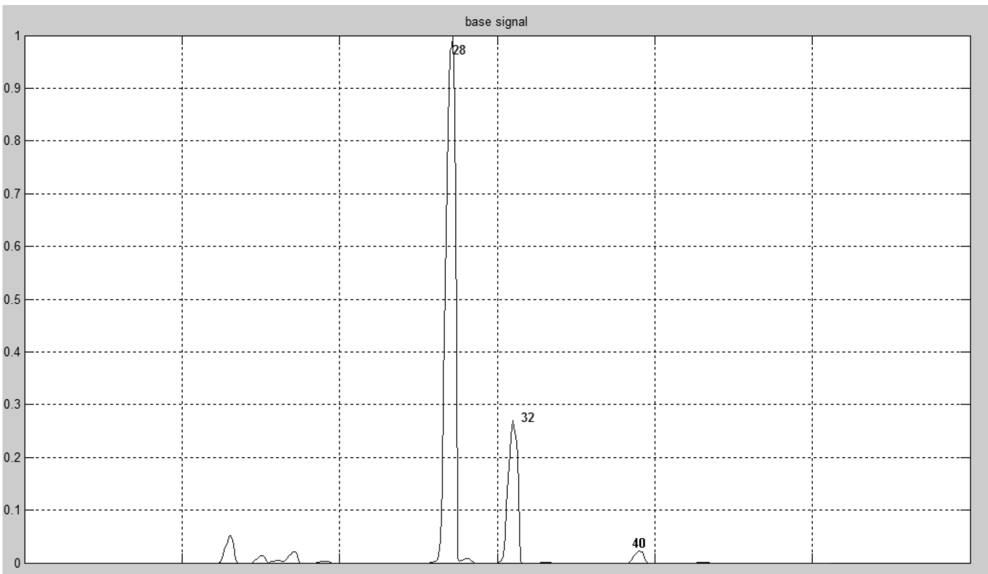

Рис. 1. Спектр воздуха, используемый в качестве эталонного сигнала $\hat{\mathbf{R}}_{эт}$.
Над спектральными линиями указаны соответствующие массы

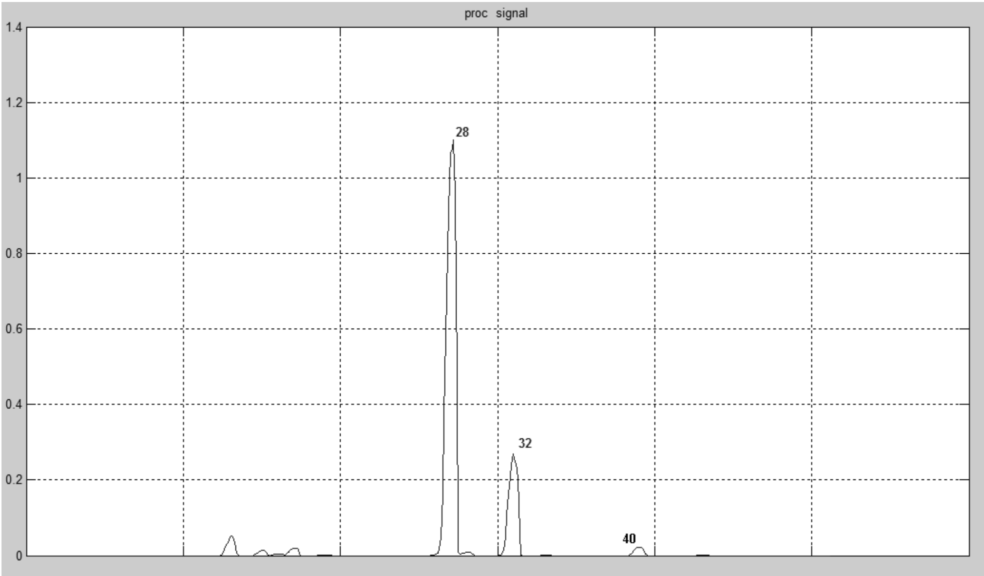

Рис. 2. Спектр воздуха, в котором увеличено содержание азота (масса 28)

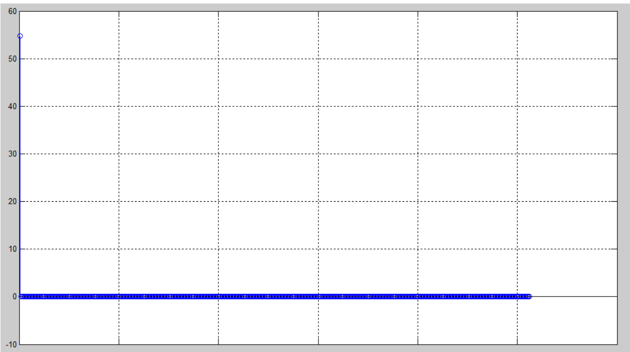

Рис. 3. Спектр эталонного сигнала $\hat{\mathbf{R}}_{эт}$ в приспособленном базисе. Этот спектр содержит один элемент, отличный от нуля

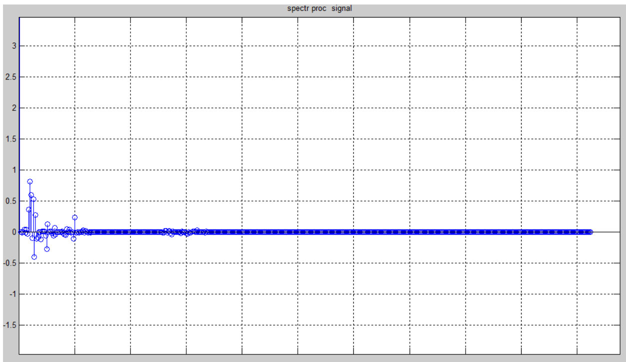

Рис. 4. Спектр сигнала, представленного на рис. 2, в приспособленном базисе

На рис. 3 представлен спектр эталонного сигнала $\hat{\mathbf{R}}_{эт}$, полученный по формуле (1) в приспособленном базисе. Как видно, этот спектр содержит один элемент, отличный от нуля. На рис. 4 представлен спектр воздуха с примесями. Нетрудно заметить, что в этом случае имеется ряд элементов, отличных от нуля.

Принятие решения о чистоте воздуха по его масс-спектру осуществляется двумя путями.

1. Сравнение с заранее установленной константой δ разности вектора $\mathbf{Y}$, полученного путем преобразования в приспособленном базисе проверяемого сигнала, и вектора $\mathbf{Y}_{эт}$ — преобразованного эталонного сигнала $\hat{\mathbf{R}}_{эт}$ (верхняя строка формулы (4)).

2. Сравнение с заранее установленной константой ε модуля разности квадратов первых элементов в векторах $\mathbf{Y}$ и $\mathbf{Y}_{эт}$.

Если оба неравенства в формуле (4) выполняются, то измеренный масс-спектр принадлежит чистому воздуху. В противном случае в измеренном с помощью масс-спектрометра воздухе содержатся примеси.

ЗАКЛЮЧЕНИЕ

Рассмотренные алгоритмы оценки наличия примесей в масс-спектрах воздуха на основе ортогональных преобразований в приспособленном базисе имеют следующие основные преимущества.

1. Независимость от формы исходного сигнала.
2. Простота реализации.
3. Возможность автоматизированной оценки наличия загрязнений в окружающем воздухе.

Дальнейшее изучение возможностей метода преобразований в приспособленном базисе для оценки чистоты воздуха по его масс-спектру требует выполнения следующих работ:

- формирования сигналов, которые могут быть приняты за эталонные;
- выбора констант δ и ε для того чтобы принять достоверное решение о загрязнении воздуха по его масс-спектру при наличии в нем шумов и помех;
- проведения экспериментальных работ для оценки чувствительности метода и воспроизводимости его результатов.

СПИСОК ЛИТЕРАТУРЫ

1. *Солодовников А.И., Спиваковский А.М.* Основы теории и методы спектральной обработки информации. Учебн. пособие. Л.: Изд-во ЛГУ, 1986. 272 с.
2. *Абденнби А.* Разработка и исследование метода и алгоритмов классификации сигналов на основе приспосабливаемых спектральных ортогональных преобразований. Дис. … канд. техн. наук. СПбГЭТУ, 2005. 120 с.
3. *Неймарк Ю.И., Басин Ю.Г.* Алгоритмы приспособленного базиса в задачах распознавания образов // Техническая кибернетика. № 2, 1970. С. 145–161.
4. *Ахмед Н.Д., Рао К.Р.* Ортогональные преобразования при обработке цифровых сигналов. М.: Связь, 1980. 248 с.
5. *Солодовников А.И.* Синтез полных ортонормированных функций, имеющих алгоритм быстрого преобразования // Вопросы теории систем автоматического управления. Межвуз. сборник. Л.: Изд-во ЛГУ, 1978. Вып. 4. С. 94–105.
6. *Манойлов В.В.* Развитие методов обработки информации в масс-спектрометрии для изотопного и элементного анализа, Дис. … докт. техн. наук по специальности 01.04.01 "Приборы и методы экспериментальной физики". Санкт-Петербург, 2008. 263 с.

Институт аналитического приборостроения РАН, г. Санкт-Петербург

Контакты: *Манойлов Владимир Владимирович*, manoilov_vv@mail.ru

Материал поступил в редакцию 5.10.2011.

ALLOCATION OF INFORMATION ATTRIBUTES IN THE MASS SPECTROMETRIC SIGNALS OF AIR

Yu. A. Titov, V. V. Manoilov, A. G. Kuzmin

Institute for Analytical Instrumentation of RAS, Saint-Petersburg

The paper deals with a new method of isolating informational attributes from the processed signals, the generalized spectral analysis in the adapted basis. The studies of the functional structure of mathematical foundations and adaptable classification system of different signal types, including the quadrupole mass spectra of air, were carried out. The algorithms for synthesis of basic functions adapted to process signals of orthogonal transforms were developed. Mass spectra which was obtained by a quadrupole mass spectrometer processing, obtained on quadrupole mass spectrometer, for the analysis of air purity was used as an example.

Keywords: signal processing methods, mass spectrometry, parameter estimation of the mass spectra of air, analysis of the data processing algorithms