
**ИССЛЕДОВАНИЯ, ПРИБОРЫ, МОДЕЛИ
И МЕТОДЫ АНАЛИЗА**

УДК 681.327.8: 621.384

© В. В. Наумов, О. А. Гребенщиков

МЕТОДИКА ОЦЕНКИ ЭФФЕКТИВНОСТИ МОДЕЛЕЙ И ОБНАРУЖЕНИЯ МАСС-СПЕКТРОМЕТРИЧЕСКИХ ПИКОВ

В статье приведены результаты экспериментальных исследований выходных сигналов квадрупольного масс-спектрометра на базе автоматизированного комплекса, содержащего квадрупольный масс-спектрометр МХ-7306. Разработаны методика оценки эффективности моделей сигналов и оптимальный вариант корреляционного алгоритма сглаживания, обнаружения и идентификации масс-спектрометрических пиков; приведены аналитические выражения эффективности моделей пиков для различных аппроксимаций. Методика позволяет проводить выбор наилучшей с точки зрения вероятности правильного обнаружения аппроксимации масс-спектрометрического пика в каждом конкретном случае.

ВВЕДЕНИЕ

Одной из особенностей масс-спектрометрических анализаторов состава и свойств веществ является необходимость проведения достаточно сложной обработки выходного сигнала детектора для получения конечных результатов, которая в настоящее время проводится цифровыми методами. Для успешного решения задач обработки необходимо оценить эффективность моделей масс-спектрометрических пиков, найти оптимальные с точки зрения минимизации погрешностей и вероятности правильного обнаружения пиков алгоритмы обработки. В связи с этим представляют интерес исследования характеристик выходных сигналов серийных масс-спектрометров типов МХ-7304А, МХ-7306 и алгоритмов автоматизации их обработки, которые проводились с использованием автоматизированных комплексов, описанных в работах [1–3].

МЕТОДИКА ОЦЕНКИ ЭФФЕКТИВНОСТИ МОДЕЛЕЙ ПИКОВ

Авторы рассматривали шесть аппроксимаций форм пиков, аналитические выражения и графические представления которых сведены в таблицу.

Если в качестве основного критерия при выборе параметров корреляционного алгоритма использовать наименьшие значения амплитуд пиков, которые достоверно идентифицируются на фоне шума, то, ограничиваясь анализом задачи обнаружения в случае аддитивной смеси детерминированного сигнала и белого шума, алгоритм может быть записан в виде [4]:

$$S = \sum_{i=1}^k (I_i - I_{\phi}) \cdot f(t_i - t_0) \geq h_0, \quad (1)$$

где k — ширина окна функции свертки, I_i — регистрируемые отсчеты, I_{ϕ} — интенсивность фона, f — точки модели пика, h_0 — заданный порог.

Рассмотрим разработанный согласно этой формуле вариант корреляционного алгоритма сглаживания и обнаружения масс-спектрометрических пиков [3]. Уровень порога h_0 задается величиной, большей $(I_{\phi} + 2 \cdot \text{СКО}) k/2$, где, СКО — среднеквадратическое отклонение интенсивности фона. Алгоритм производит идентификацию пика, оценку его положения и максимального амплитудного значения, по которым производятся вычисление значений номера а.е.м. [1] и оценка интенсивности спектральных линий [5].

Для проведения анализа эффективности обнаружения пика рассмотрим аппроксимации формы пиков в виде, приведенном в таблице. В качестве критерия эффективности аппроксимаций можно воспользоваться значениями вероятностей правильного обнаружения пиков P_0 и ложной тревоги P_T , которые определяются по формулам [6]

$$P_0 = 0.5 (1 - \Phi(Z_2)), \quad P_T = 0.5 (1 - \Phi(Z_1)), \quad (2)$$

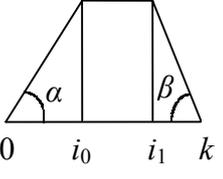
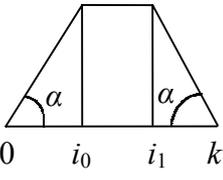
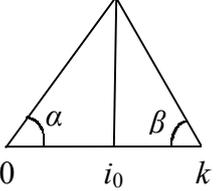
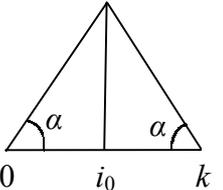
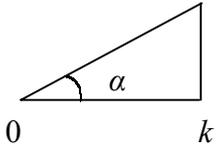
где $\Phi(Z_{1,2})$ — функция ошибок. Значения Z_1, Z_2 в зависимости от амплитуды (интенсивности) пика масс-спектрометрического сигнала и СКО шума $\sigma_{ш}$ определяются в соответствии со следующими выражениями [5]:

$$Z_1 = \frac{h_0}{\sigma_{ш} \sqrt{2F}}, \quad (3)$$

$$Z_2 = Z_1 - \frac{I_S}{\sigma_{ш} \sqrt{2}} \sqrt{F}, \quad (4)$$

где $F = \sum_{i=0}^k f_i^2$; I_S — амплитуда идентифицируемого пика; $f_i = f(t_i - t_0) = f(i \cdot \Delta t - t_0)$; Δt — период дискретизации сигнала.

Аппроксимации формы и параметры моделей пиков

№	Графическое представление	Аналитическое выражение	Параметры формы пиков	$F = \sum f_i^2$
1		$f_i = \begin{cases} \alpha \frac{i}{k}, & 0 \leq i \leq i_0 \\ 1, & i_0 \leq i \leq i_1 \\ \beta \frac{k-i}{k}, & i_1 \leq i \leq k \end{cases}$	$\begin{aligned} &\alpha \neq \beta, \quad i_0 \neq i_1, \\ &\alpha > 0, \quad \beta > 0, \\ &i_0 \neq 0, \quad i_1 \neq k \end{aligned}$	$k + \frac{\alpha + \beta}{6k} - \frac{2k}{3} \left(\frac{1}{\alpha} + \frac{1}{\beta} \right)$
2		$f_i = \begin{cases} \alpha \frac{i}{k}, & 0 \leq i \leq i_0 \\ 1, & i_0 \leq i \leq i_1 \\ \alpha \frac{k-i}{k}, & i_1 \leq i \leq k \end{cases}$	$\begin{aligned} &\alpha = \beta, \quad i_0 \neq i_1, \\ &\alpha > 0, \quad \beta > 0, \\ &i_0 \neq 0, \quad i_1 \neq k \end{aligned}$	$k + \frac{\alpha}{3k} - \frac{4k}{3\alpha}$
3		$f_i = 1, \quad 0 \leq i \leq k$	$i_0 = 0, \quad i_1 = k$	k
4		$f_i = \begin{cases} \alpha \frac{i}{k}, & 0 \leq i \leq i_0 \\ 1, & i = i_0 \\ \beta \frac{k-i}{k}, & i_1 \leq i \leq k \end{cases}$	$\begin{aligned} &\alpha \neq \beta, \quad i_0 = i_1, \\ &\alpha > 0, \quad \beta > 0, \\ &i_0 \neq 0, \quad i_1 \neq k \end{aligned}$	$\frac{k}{3} + \frac{\alpha + \beta}{6k}$
5		$f_i = \begin{cases} \alpha \frac{i}{k}, & 0 \leq i \leq i_0 \\ 1, & i_0 \leq i \leq i_1 \\ \alpha \frac{k-i}{k}, & i_1 \leq i \leq k \end{cases}$	$\begin{aligned} &\alpha = \beta, \quad i_0 = i_1, \\ &\alpha > 0, \quad \beta > 0, \\ &i_0 \neq 0, \quad i_1 \neq k \end{aligned}$	$\frac{k}{3} + \frac{\alpha}{3k}$
6		$f_i = \frac{i}{k}, \quad 0 \leq i \leq k$	$i_0 = i_1 = k$	$\frac{k}{3} + \frac{1}{6k} + \frac{1}{2}$

В тех случаях, когда трудно принять решение о соотношении удельных весов ошибок ложного обнаружения и пропуска пика, можно считать их равноценными и оценивать эффективность аппроксимации только вероятностью правильного обнаружения P_O [6]. При этом условия из (2)–(4) с учетом того, что если $P_T = 1 - P_O$, то $\Phi(Z_1) = \Phi(Z_2)$ и, следовательно, $-Z_1 = Z_2$, следует, что:

$$h_0 = (I_{sm} / 2)F; \quad (5)$$

$$Z_2 = -0.35 \frac{I_s \sqrt{F}}{\sigma_u}, \quad (6)$$

где I_{sm} — амплитуда наименьшего идентифицируемого пика.

Таким образом, чтобы оценить P_O по заданным значениям порога h_0 и отношения сигнал/шум I_s/σ_u в соответствии с (2)–(4) или выбрать уровень порога по заданным значениям P_O и I_{sm} в соответствии с (5), необходимо воспользоваться выражениями суммы F для каждой из приведенных в таблице аппроксимаций пика. Покажем получение выписанных в таблице выражений и использованные методики.

ИСПОЛЬЗОВАНИЕ МЕТОДИКИ

Вычисление суммы F проведем для несимметричной трапецеидальной аппроксимации, т. к. соответствующим заданием значений параметров α , β , i_0 , i_1 могут быть получены все остальные. Например, при $\alpha = \beta = 3$ получается равнобокая трапеция; при $i_0 = 0$, $i_1 = k$ — прямоугольник и т. д. Итак, сумма равна

$$F = \sum_{i=0}^{i_0-1} \left(\frac{\alpha i}{k} \right)^2 + \sum_{i=i_1+1}^k \left(\frac{\beta(k-i)}{k} \right)^2 + i_1 - i_0 + 1 = \\ = \frac{\alpha^2}{k^2} \sum_{i=0}^{i_0-1} i^2 + \frac{\beta^2}{k^2} \sum_{i=0}^{k-i_1-1} i^2 + i_1 - i_0 + 1.$$

Далее, учитывая, что $i_0 = k/\alpha$, $i_1 = k - k/\beta$, и используя формулу для суммы квадратов, получаем

$$F = k + \frac{\alpha + \beta}{6k} - \frac{2k}{3} \left(\frac{1}{\alpha} + \frac{1}{\beta} \right). \quad (7)$$

Полученные соотношения позволяют записать формулу для вычисления Z_1 , Z_2 , h_0 в соответствии с (3)–(5) для всех рассматриваемых аппроксимаций пиков.

В частном случае, когда $\alpha = \beta = 3$, получим $F = 5k/9 + 1/k$, что отличается от приведенного в [6] соотношения на значение второго слагаемого. Таким образом, формула (7) позволяет более точно вычислить значения Z_1 , Z_2 , h_0 , что особенно существенно при небольшом количестве используемых отсчетов k .

Для проведения сравнительного анализа эффективности рассматриваемых аппроксимаций достаточно проанализировать значения Z_2 , вычисленные в соответствии с (6), поскольку $\Phi(Z)$ — монотонно возрастающая функция. Полученные результаты показывают, что величина Z_2 для всех аппроксимаций форм пиков одинаково зависит от отношения сигнал/шум. Следовательно, для проведения сравнительного анализа достаточно исследовать выражения, приведенные в таблице, в зависимости от значений параметров α , β , k .

Полученная классификация моделей полезного сигнала позволяет проводить выбор наилучшей с точки зрения вероятности правильного обнаружения аппроксимации в каждом конкретном случае в зависимости от значений параметров α , β , k . Например, при $\alpha = 1.7$, $\beta = 2.43$, $k = 20$, что имеет место в автоматизированном комплексе [3], прямоугольная аппроксимация лучше несимметричной трапецеидальной и лучше несимметричной треугольной.

СПИСОК ЛИТЕРАТУРЫ

1. Наумов В.В., Белозеров А.В., Гребенщиков О.А. Экспериментальная оценка динамики дрейфа пиков квадрупольного масс-спектрометра и определение массовых чисел // Научное приборостроение. 2001. Т. 11, № 1. С. 36–39.
2. Анишкевич Н.Н., Белозеров А.В., Наумов В.В. Автоматизированный масс-спектрометрический комплекс для анализа газов // Приборы и техника эксперимента. 1988. № 5. С. 235–236.
3. Анишкевич Н.Н., Белозеров А.В., Наумов В.В. и др. Автоматизация масс-спектрометрического анализа водорастворенных газов в реальном времени // Управляющие системы и машины. 1990. № 1. С. 88–92.
4. Русинов Л.А., Новиков Л.В. Спектральный подход к первичной обработке сигналов аналитических приборов. Л.: ЛГУ, 1984. 160 с.
5. Наумов В.В., Белозеров А.В., Гребенщиков О.А. Система автоматического поиска и слежения за глобальным максимумом пика масс-спектра // Научное приборостроение. 2001. Т. 11, № 1. С. 60–64.
6. Гуревич А.А., Русинов Л.А., Могильницкий А.М. и др. Автоматизация обработки масс-спектрометрической информации. М.: Энергия, 1978. 182 с.

Институт электроники НАН Беларуси, Минск

Материал поступил в редакцию 6.02.2007.

THE TECHNIQUE FOR ESTIMATING THE MODEL EFFICIENCY AND DETECTING MASS SPECTRUM PEAKS

V. V. Naumov, O. A. Grebenshchykov

Institute of Electronics, Belarus National Academy of Sciences, Minsk

The paper presents the results of experimental investigation of quadrupole mass spectrometer output signals by using a computerized system comprising quadrupole mass spectrometer MX-7306. A technique for estimating the efficiency of model signals and an optimum version of the correlation algorithm designed to smooth, detect and identify mass spectrum peaks have been developed. Analytical expressions for peak models efficiency for various approximations are represented. In each specific case, the technique considered allows one to choose the best mass-spectrum peak approximation (i. e., approximation ensuring high probability of correct detection).