

---



---

**МАТЕМАТИЧЕСКИЕ МЕТОДЫ  
И МОДЕЛИРОВАНИЕ В ПРИБОРОСТРОЕНИИ**

---



---

УДК 543.5; 543.26

© А. И. Белозерцев, О. В. Черемисина, С. З. Эль-Салим,  
В. В. Манойлов, И. В. Заруцкий

**АЛГОРИТМЫ ОБРАБОТКИ ДАННЫХ В ГАЗОАНАЛИТИЧЕСКИХ  
КОМПЛЕКСАХ С ПОЛУПРОВОДНИКОВЫМИ СЕНСОРАМИ  
ДЛЯ ОБНАРУЖЕНИЯ ПАРОВ ЯДОВИТЫХ ВЕЩЕСТВ  
В ОКРУЖАЮЩЕЙ СРЕДЕ**

В данной работе проведено описание алгоритмов обработки данных полупроводниковых сенсоров, входящих в состав аналитических комплексов для контроля паров ядовитых веществ в воздухе, жидкостях и твердых телах. Эти комплексы, в частности, востребованы для анализа компонент ракетных топлив (КРТ) как на наземных объектах космической инфраструктуры, так и в окружающей среде. Алгоритмы обработки данных предназначены для автоматического разделения образцов на чистые и на содержащие токсичные загрязнения, примеси или добавки (загрязненные образцы). В основе этих алгоритмов лежат методы, основанные на дискриминантном анализе многомерных данных.

*Кл. сл.:* алгоритмы обработки данных газового анализа, газочувствительные полупроводниковые сенсоры, газоаналитические приборные комплексы, охрана окружающей среды

### ВВЕДЕНИЕ

В настоящее время особую актуальность приобретают вопросы оценки воздействия на окружающую среду в результате ракетно-космической деятельности. Важное значение отводится задачам контроля в окружающей среде (в воздухе, водах, почвах и грунтах) несимметричного диметилгидразина (НДМГ), тетраоксида азота (АТ) и продуктов их разложения и взаимодействия. В работе [1] приведен обзор методов контроля НГМД с помощью различных приборных комплексов. Отмечены преимущества полупроводниковых сенсоров для контроля наличия компонент ракетных топлив (КРТ), описана технология изготовления таких сенсоров, приведены схематические и аппаратные решения для их применения в газоаналитических комплексах контроля в воздухе, растворах и твердых телах. В работе [1] не описаны алгоритмы обработки результатов измерений, предназначенные для количественного анализа компо-

нент, входящих в состав ядовитых веществ, а также для формирования информационных признаков, необходимых для повышения селективности обнаружения целевых веществ, других примесей или токсичных добавок. Целью данной работы является восполнение этого пробела.

### ГАЗОЧУВСТВИТЕЛЬНЫЕ СЕНСОРЫ И ЭТАПЫ ПРОВЕДЕНИЯ ИЗМЕРЕНИЙ С ИХ ПОМОЩЬЮ

Полупроводниковые датчики, применяемые для определения паров ядовитых веществ, состоят из газочувствительного слоя, подложки-диэлектрика и нагревательного элемента. Чувствительный элемент включает в себя полупроводниковый слой, нанесенный на подложку, с обратной стороны которой расположен нагреватель. В присутствии детектируемого газа изменяется проводимость чувствительного слоя, величина этого изменения зависит от типа и концентрации газа. Изменение проводимости основано на том, что при адсорбции

Термодесорбция	Термопауза	Выход в рабочий режим	Рабочий режим
20–40 с	10–60 с	10–30 с	30–60 с

**Рис. 1.** Временные интервалы различных этапов проведения измерений

молекул на поверхности газочувствительного слоя нарушается равновесное значение тока в зоне проводимости и происходит либо увеличение концентраций носителей заряда (в случае хемосорбции газа-донора), либо снижение их концентрации (в случае хемосорбции газа-акцептора) [2]. Изменения концентрации носителей определяют аналитический отклик датчика, который зависит от химического состава аналита и его количества — концентрации.

Применение полупроводниковых адсорбционных датчиков позволяет, во-первых, провести миниатюризацию газоаналитических приборов; во-вторых, расширить перечень определяемых веществ и, наконец, проводить количественный анализ.

Процесс измерений для определения концентрации газов с помощью таких сенсоров содержит следующие технологические этапы.

- Термодесорбция — очистка поверхности, подготовка поверхности к проведению измерений.
- Термопауза — контроль электрофизических характеристик, смещение запрещенной зоны до максимального значения.
- Выход в рабочий режим — оксидирование поверхности, подготовка к хемосорбции.
- Рабочий режим — проведение хемосорбции, измерение изменения электропроводности;

Временные интервалы этих этапов представлены на рис. 1.

На рис. 2 и 3 представлены результаты измерения напряжений от полупроводниковых сенсоров для различных типов анализируемых образцов. На рис. 2 представлен результат регистрации выходного сигнала сенсора для газа-восстановителя. На рис. 3 представлены результаты регистрации выходного сигнала сенсора для газа-окислителя (график 1) и для смеси двух газов — окислителя и восстановителя — (график 2).

## АЛГОРИТМЫ ОБРАБОТКИ ВЫХОДНЫХ СИГНАЛОВ СЕНСОРОВ

### Определение времени начала рабочего режима

Информационные параметры, которые свидетельствуют о типе и концентрации анализируемого газа могут быть определены только на этапе рабочего режима (рис. 1). Первой задачей программного обеспечения, которое реализует алгоритмы обработки сигналов полупроводниковых сенсоров, является определение времени начала рабочего режима. Эта задача решается двумя способами: вручную и автоматически. В ручном режиме указанное время определяется с помощью вертикального маркера, которым управляет оператор. Пример такого маркера показан на рис. 3.

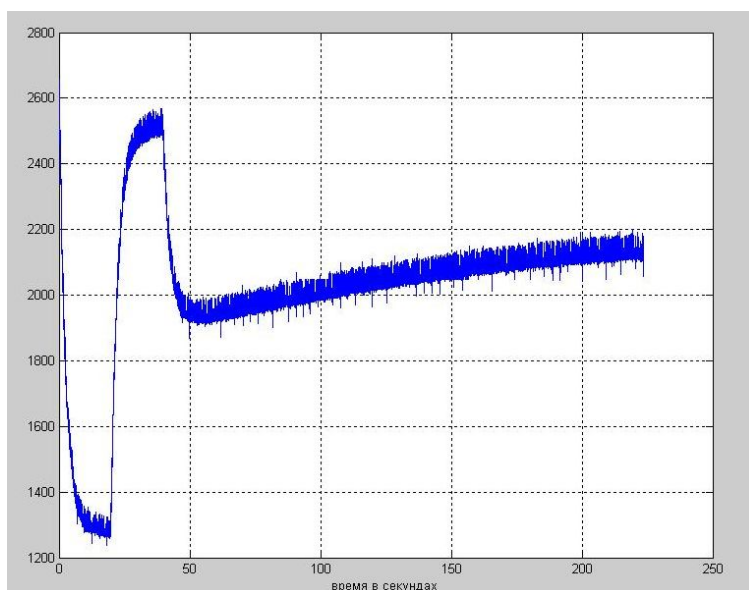
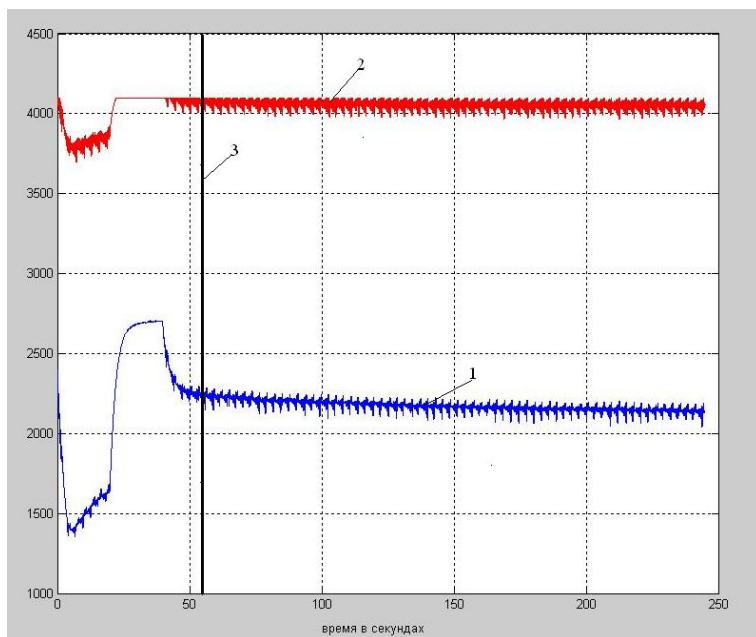


Рис. 2. Выходной сигнал сенсора для донора (газ-восстановитель)



**Рис. 3.** Выходной сигнал сенсора для акцептора (1) и смеси донора и акцептора (2). Маркер для определения времени начала рабочего режима (3)

В автоматическом режиме для определения этого времени используется алгоритм максимального правдоподобия [3]. Суть этого алгоритма заключается в следующем. В скользящем окне определяются параметры прямой линии, которой аппроксимируются измеряемые данные. Такая аппроксимация проводится методом наименьших квадратов, с помощью которого определяется угловой коэффициент этой прямой. Затем методом максимального правдоподобия сравниваются две гипотезы. Первая гипотеза заключается в том, что угловой коэффициент изменился и его величина соответствует заданному порогу. Вторая гипотеза заключается в том, что величина углового коэффициента не соответствует заданному порогу. Если оказывается верной первая гипотеза, то принимается решение о начале рабочего режима.

#### Сглаживание измеряемых данных

Для повышения воспроизводимости результатов анализа производится предварительное сглаживание исходных данных. Сглаживанию подвергаются все измеряемые данные, в том числе и те, которые не относятся к рабочему режиму. Были рассмотрены два хорошо известных алгоритма: первый — экспоненциальное сглаживание [4] и второй — сглаживание методом Савицкого—Голея [5]. Алгоритм экспоненциального сглаживания может быть модифицирован в удобную для реализации в микропроцессоре форму, при которой момент времени сглаженного значения сигнала отстает только на один шаг квантования от значения момента времени, соответствующего произ-

водимому измерению. Такая процедура проведения сглаживания в режиме on-line описана ниже.

#### Алгоритм экспоненциального сглаживания

Сглаженное значение  $s_s(t_i)$  сигнала  $s(t)$  в момент времени  $t_i$  равно

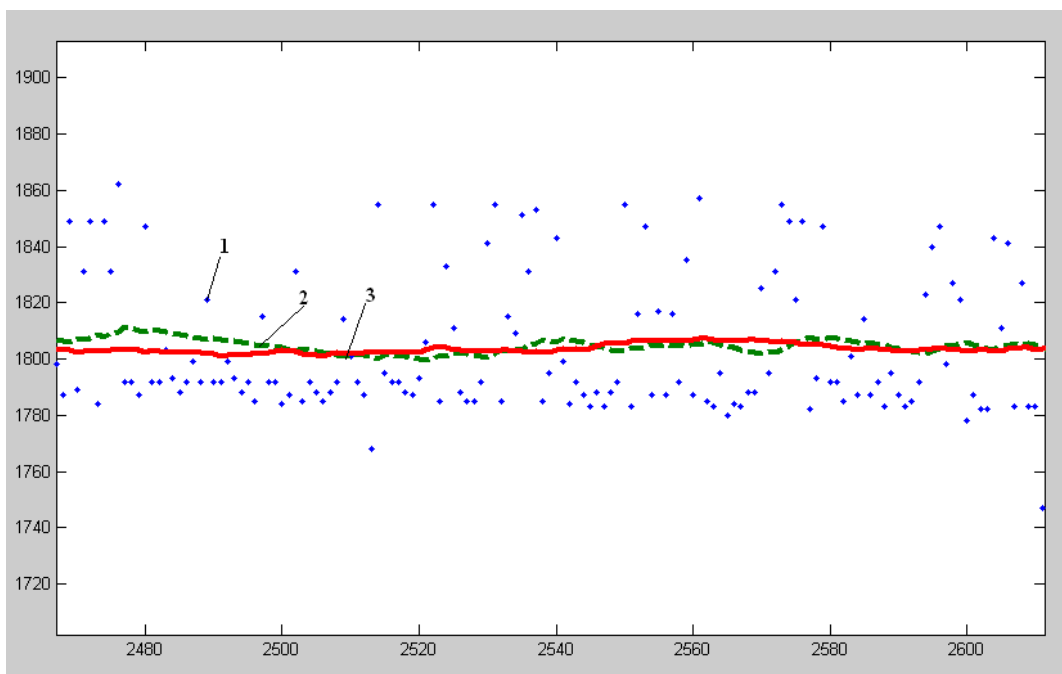
$$s_s(t_i) = ps(t_i) + (p-1)s_s(t_{i-1}). \quad (1)$$

В этой формуле  $s_s(t_{i-1})$  — сглаженное значение сигнала  $s(t)$  в момент времени  $t_{i-1}$ . Параметр  $p$  подбирался экспериментально и для большинства обрабатываемых данных был равен 0.03. На рис. 4 показан результат сглаживания по формуле (1) с указанным параметром  $p$ .

#### Сравнение алгоритма экспоненциального сглаживания с алгоритмом Савицкого—Голея

На рис. 4 представлены результаты сравнения алгоритма экспоненциального сглаживания по формуле (1) и сглаживание по алгоритму Савицкого—Голея. В алгоритме Савицкого—Голея использовалась линия регрессии второго порядка в скользящем окне из 141 точки. Общее количество точек в исходном сигнале в рабочем режиме было 5317. Как видно из рисунка, на участке рабочего режима линии 2 и 3 почти совпадают.

Для количественной оценки результатов сравнения алгоритмов сглаживания были вычислены три евклидова расстояния в евклидовой метрике (корень квадратный из суммы квадратов невязок): 1 — между исходными данными и полиномом



**Рис. 4.** Результат сглаживания исходного сигнала в рабочем режиме. Исходный сигнал — множество точек 1; результат экспоненциального сглаживания по формуле (1) — пунктирная линия 2; результат сглаживания по алгоритму Савицкого—Голея — сплошная линия 3

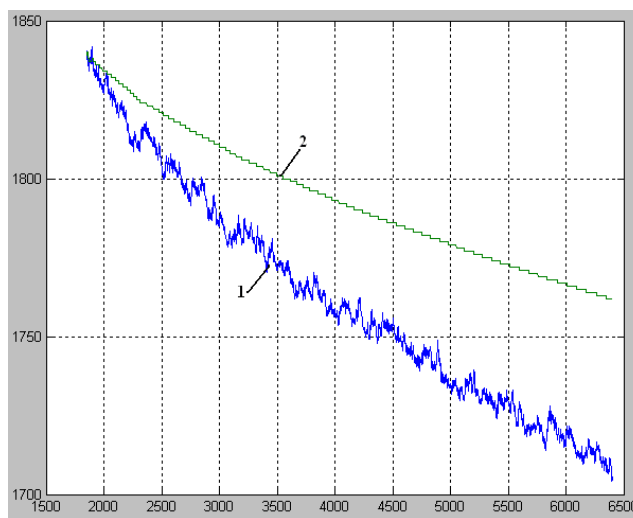
второй степени, с помощью которого проводилась аппроксимация исходных данных; 2 — между сглаженными данными, полученными по формуле (1), и полиномом; 3 — между сглаженными данными, полученными по алгоритму Савицкого—Голея, и полиномом. Для данных, представленных на рис. 4, эти расстояния оказались:

- 1 — 1768,
- 2 — 218,
- 3 — 226.

Аналогичные результаты получались для других исходных данных. Таким образом, как видно из представленных данных, оба алгоритма пригодны для сглаживания. Однако алгоритм Савицкого—Голея [5] не может использоваться в режиме on-line, поскольку сглаженное значение сигнала сильно отстает от текущего отсчета. Сглаживание по формуле (1) лучше подходит для работы on-line, и по этой причине ему отдано предпочтение.

#### Вычисление Z-сигнала

Следующим этапом в проведении первичной обработки исходных данных является вычисление так называемого Z-сигнала [6]. По сути эта характеристика обрабатываемого сигнала является слагаемым в известной формуле оценки дисперсии —



**Рис. 5.** Сглаженный исходный сигнал — 1 и накопленный  $\langle s(k) \rangle$  — 2.

По горизонтальной оси — номера отсчетов

второго момента. Последующие этапы первичной обработки выполняются с Z-сигналом. Эти этапы позволяют выделить информационные характеристики, необходимые для проведения идентифика-

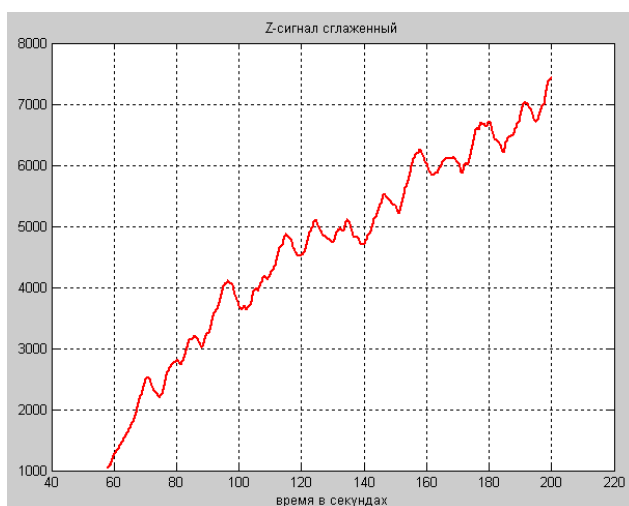


Рис. 6. Z-сигнал, сглаженный в рабочем режиме

ции газа в исследуемом образце.

Текущее значение Z-сигнала в точке  $k$  определяется по формуле

$$Z(k) = [s(k) - \langle s(k) \rangle]^2. \quad (2)$$

В этой формуле  $s(k)$  — текущее значение измеренного сигнала в рабочем режиме,  $\langle s(k) \rangle$  — среднее значение  $k$  отсчетов измеренного сигнала в рабочем режиме.  $\langle s(k) \rangle$  вычисляется путем суммирования  $k$  отсчетов, сумма затем делится на  $k$ . Вычисление среднего для очередной точки измеряемого сигнала дает возможность проследить его изменение в течение всего рабочего режима и затем использовать для идентификации, т. к. такое изменение для каждого исследуемого газа разное. Общее количество точек в рабочем режиме может выбираться с помощью маркера или путем вычисления времени с момента начала рабочего режима. Как правило, продолжительность рабочего режима 30–60 с.

На рис. 5 представлены следующие сигналы: сглаженный исходный сигнал — линия 1; накопленный  $\langle s(k) \rangle$  — линия 2 в рабочем режиме. По горизонтальной оси отложены номера отсчетов.

На рис. 6 представлен сглаженный Z-сигнал в рабочем режиме.

#### Нахождение свертки производной исходного сигнала с Z-сигналом и производной от этой свертки

В работе [7] отмечается, что свертка производной исходного сигнала и Z-сигнала несет информацию, необходимую для идентификации ядови-

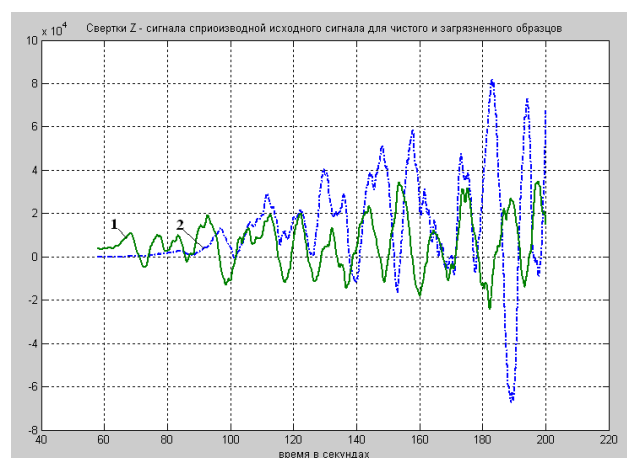


Рис. 7. Результат свертки Z-сигнала с производной исходного сигнала в рабочем режиме.

Для чистого образца — 1 (сплошная линия) и загрязненного образца — 2 (пунктир с точками)

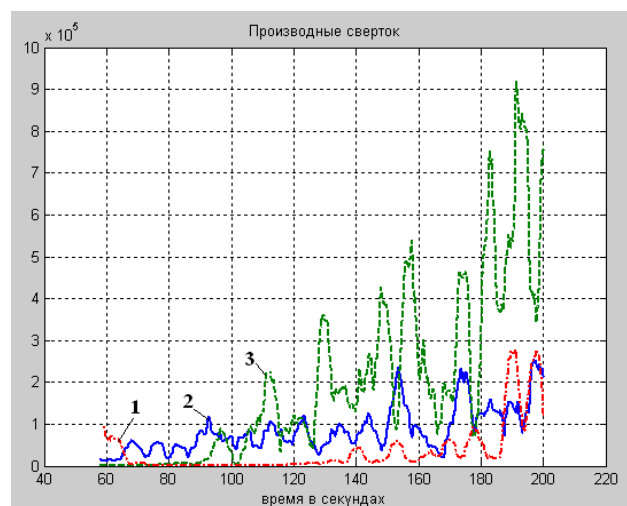


Рис. 8. Производные сигналов свертки Z-сигнала с производной исходного сигнала в рабочем режиме для двух чистых образцов (1 и 2) и загрязненного образца (3) — пунктирная линия

тых веществ в исследуемом образце. Производная этого сигнала-свертки определяет изменения спектральной составляющей во времени и позволяет связать Z-сигнал с кинетикой (скоростью реакции хемосорбции) и динамикой (переходом в равновесное состояние), что полностью определяет состояние гетерогенной системы (адсорбат—адсорбент). В нашем случае: адсорбат — молекулы примеси в газовой фазе, адсорбент — газочувствительный слой сенсора. На рис. 7 представлен сигнал свертки производной исходного

сигнала с Z-сигналом для разных образцов исследуемых газов. Цифрами 1 и 2 на рисунке обозначены номера образцов исследуемого газа: 1 — чистый образец, 2 — загрязненный образец. Показанный на рисунке результат дает возможность качественно заметить различие сигналов разных образцов исследуемых газов. Это различие может быть использовано для того, чтобы отличить загрязненный образец от чистого на этапе вторичной обработки. На рис. 8 представлен результат дифференцирования сигнала-свертки.

Показанный на рис. 8 результат качественно показывает еще более сильное различие сигналов чистых образцов и загрязненного образца.

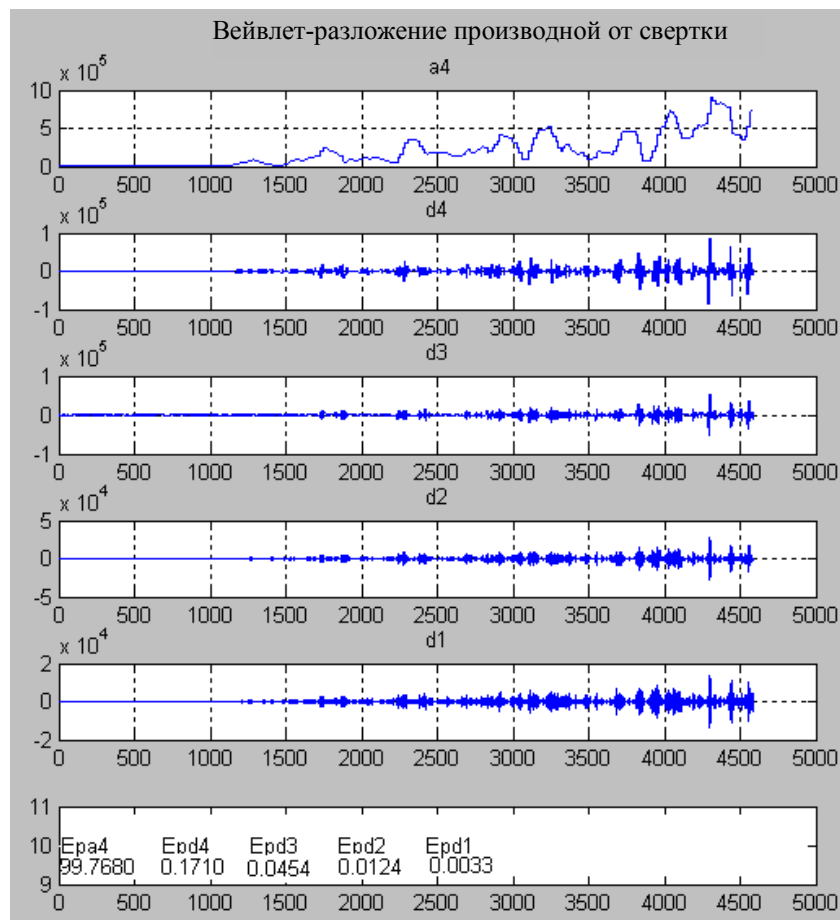
Для количественной оценки отличий сигналов чистых образцов и сигнала загрязненного образца вычислялись евклидовы расстояния между рассмотренными сигналами. Евклидово расстояние между сигналами чистых образцов 1 и 2 оказалось равным 862, а между сигналом 2 и сигналом загрязненного образца 3 оказалось равным 3120.

Рассмотренные в данном разделе статьи алгоритмы первичной обработки выходных сигналов

сенсоров реализованы в программном обеспечении, которое выполняет обработку данных в реальном масштабе времени (режим on-line). Кроме обработки данных по рассмотренным выше алгоритмам в режиме on-line производятся также вычисления функции плотности распределения сигналов, быстрое фурье-преобразование рассмотренных выше сигналов свертки и сигналов производных свертки для различных компонент исследуемой газовой смеси. Полученные в результате этих вычислений и преобразований данные также содержат информацию о различии чистых и загрязненных образцов исследуемых газов.

#### Дискретное вейвлет-преобразование с использованием материнских функций Хаара

Суть дискретного многоступенчатого дискретного вейвлет-преобразования состоит в том, что сначала на первой ступени с помощью вейвлет-фильтров нижних и верхних частот исходный сигнал разделяется на две части. В первой части содержатся составляющие нижних частот исходного сигнала. Сигнал с такими составляющими



**Рис. 9.** Вейвлет-преобразование сигнала-производной от свертки исходного сигнала с Z-сигналом для измерения загрязненного образца

называется "аппроксимирующий сигнал  $a1$ ". Во второй части преобразованного сигнала содержатся верхние частоты исходного сигнала. Сигнал с такими составляющими называется "детализирующий сигнал  $d1$ ". Частота среза для нижних и верхних частот на первой ступени равна частоте Найквиста, что соответствует половине частоты дискретизации. Аппроксимирующий сигнал  $a1$  и детализирующий сигнал  $d1$  содержат одинаковое количество отсчетов, которое равно половине количества отсчетов в исходном сигнале. На втором этапе преобразования (второй ступени преобразования) аналогичной фильтрации подвергается только аппроксимирующий сигнал, полученный на первой ступени. Детализирующий сигнал записывается в память и будет использован на этапе реконструкции, о которой будет сказано ниже. На втором этапе преобразования частота среза равна четверти частоты дискретизации. На этом этапе получают также два сигнала: аппроксимирующий сигнал  $a2$  и детализирующий сигнал  $d2$ . Аппроксимирующий сигнал  $a2$  и детализирующий сигнал  $d2$  содержат одинаковое количество отсчетов, которое равно четверти количества отсчетов в исходном сигнале. Для обработки данных полупроводниковых сенсоров используются 4 ступени. Таким образом, получают 5 различных сигналов: один аппроксимирующий сигнал  $a4$  и четыре детализирующих сигнала. Полученные сигналы имеют разное количество отсчетов, для выравнивания количества отсчетов используется операция реконструкции [8, 9].

На рис. 9 показаны пять сигналов, рассмотренных выше. Вейвлет-преобразованию подвергался сигнал-производная от свертки исходного сигнала

с  $Z$ -сигналом для загрязненного образца. В нижнем окне показано распределение энергий этих сигналов в процентах от суммарной энергии. По горизонтальной оси указаны номера отсчетов сигналов, измеренных в рабочем режиме. На рис. 9 введены следующие обозначения:  $a4$  — аппроксимирующий сигнал4;  $d1$  — детализирующий сигнал1;  $d2$  — детализирующий сигнал2;  $d3$  — детализирующий сигнал3;  $d4$  — детализирующий сигнал4.

Дискретному вейвлет-преобразованию подвергались также другие сигналы:

- 1) свертка  $Z$ -сигнала с производной исходного сигнала;
- 2) производная свертки  $Z$ -сигнала с производной исходного сигнала;
- 3) производная исходного сигнала;
- 4) вторая производная исходного сигнала.

Более объективную информацию о различии сигналов чистого и загрязненного образцов дает сравнение распределений энергий этих сигналов в процентах от суммарной энергии всех сигналов.

В табл. 1 и 2 представлены результаты вычисления энергий аппроксимирующих и детализирующих вейвлет-коэффициентов, полученных от сигналов чистого образца (канал 1) и загрязненного образца (см. табл. 2, канал 2).

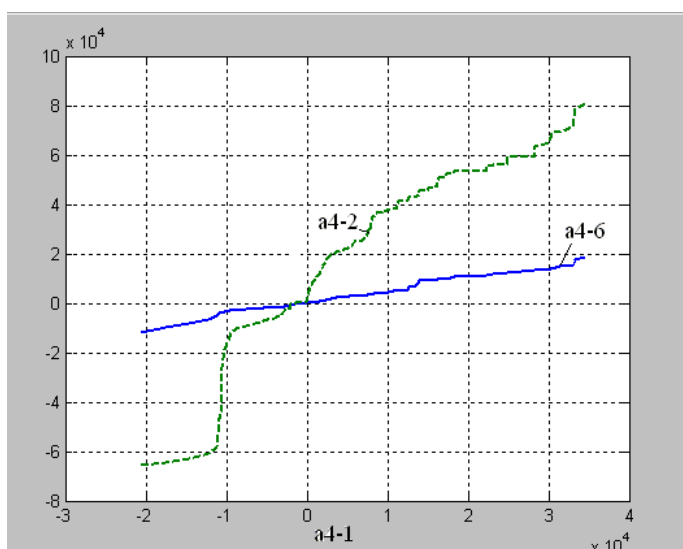
Вейвлет-фильтры нижних и верхних частот были синтезированы на основе материнских вейвлетов Хаара. Полученные результаты вейвлет-преобразований для каждого компонента исследуемой смеси газов отличаются друг от друга и могут служить информационными критериями для проведения классификации методами многомерного статистического анализа.

**Табл. 1.** Распределение энергий аппроксимирующего ( $a4$ ) и детализирующих вейвлет-коэффициентов ( $d2$ ,  $d3$ ,  $d4$ ,  $d1$ ) для сигнала чистого образца (канал 1)

Вид преобразования	Энергии вейвлет-разложения канала 1 ( в процентах)				
	$E_{a4}$	$E_{pd4}$	$E_{pd3}$	$E_{pd2}$	$E_{pd1}$
Свертка	99.280	0.528	0.145	0.037	0.010
Производная свертки	99.834	0.117	0.036	0.010	0.003
Первая производная	82.834	8.452	4.647	2.595	1.472
Вторая производная	99.367	0.463	0.126	0.032	0.008

**Табл. 2.** Распределение энергий аппроксимирующего ( $a_4$ ) и детализирующих вейвлет-коэффициентов ( $d_2, d_3, d_4, d_1$ ) для сигнала загрязненного образца (канал 2)

Вид преобразования	Энергии вейвлет-разложения канала 2 ( в процентах)				
	$E_{a4}$	$E_{pd4}$	$E_{pd3}$	$E_{pd2}$	$E_{pd1}$
Свертка	99.346	0.478	0.133	0.034	0.009
Производная свертки	99.768	0.171	0.045	0.012	0.003
Первая производная	82.972	7.859	4.706	2.737	1.726
Вторая производная	99.544	0.334	0.0916	0.0236	0.006

**Рис. 10.** Зависимости аппроксимирующих коэффициентов  $a_{4-2}$  для сигналов загрязненного образца от аппроксимирующих коэффициентов  $a_{4-1}$  сигнала чистого образца (кривая  $a_{4-2}$ ) и аппроксимирующих коэффициентов  $a_{4-6}$  для сигналов образца с возможными загрязнениями от аппроксимирующих коэффициентов сигналов чистого образца  $a_{4-1}$  (кривая  $a_{4-6}$ )

В книге [10] отмечается важность наглядного представления различий многомерных переменных перед проведением классификаций. В этой книге приведен также ряд интересных графических представлений этих различий. Одним из простых, но в тоже время достаточно наглядным является представление различий многомерных данных на двумерном графике следующим методом. На одной оси откладываются значения упорядоченных (отсортированных от меньшего к большему) значений сигнала одного образца, а на другой — сигналов другого образца. На рис. 10 представлены две кривые. Кривая  $a_{4-2}$  представляет собой зависимость аппроксимирующих коэффициентов сигнала загрязненного образца от аппроксимирующих коэффициентов чистого образца  $a_{4-1}$ . Кривая  $a_{4-6}$  представляет собой зависимость аппроксимирующих коэффициентов образца с возможным загрязнением от аппроксимирующих коэффициентов сигнала чистого образца  $a_{4-1}$ . Чем больше отличие указанных кри-

вых, тем сильнее различаются измеренные сигналы. Кроме того, чем чище измеряемые объекты, тем ближе к прямой линии указанные зависимости.

Информация, представленная на рис. 10, дает качественное представление различий. Для объективного анализа различий необходимо вычислить евклидовы расстояния для результатов измерений сигналов чистых образцов и сигналов загрязненных образцов, а также необходимо использовать методы классификации, основанные, например, на дискриминантном анализе [11, 12].

## КЛАССИФИКАЦИЯ ПРОБ

### Евклидовы расстояния для аппроксимирующих коэффициентов

Вычисления евклидовых расстояний для аппроксимирующих коэффициентов сигналов-производных свертки  $Z$ -сигнала с производной исходно-



го сигнала показало, что значения этих расстояний практически совпадают с евклидовыми расстояниями вычисленными для таких же сигналов без вейвлет-преобразований. Такой результат объясняется тем, что больше 98 % энергии содержится в аппроксимирующих коэффициентах этих сигналов.

#### Формирование матрицы исходных данных для проведения классификации

Рассмотренные алгоритмы обработки сигналов дают возможность сформировать матрицу данных для разбиения обрабатываемых сигналов на два класса: класс сигналов чистых образцов и класс сигналов загрязненных образцов.

Столбцами такой матрицы являются данные энергий аппроксимирующего (a4) и детализирующих вейвлет-коэффициентов (d2, d3, d4, d1) (см. рис. 10).

Таким образом, формируется матрица, размер которой по строкам определяется количеством исследуемых образцов газов с добавлением строк, необходимых для проведения обучения. Каждая строка матрицы содержит 20 элементов (в матрице 20 столбцов), заполненных числами, подобными представленным в табл. 1 и 2.

#### Проведение классификации чистых и загрязненных компонент газовых смесей методом дискриминантного анализа

Дискриминантный анализ является разделом многомерного статистического анализа [11], который включает в себя методы классификации многомерных наблюдений по принципу максимального сходства анализируемых наблюдений с наблюдениями, отнесенными к определенным классам (группам) по результатам обучения. В данной работе метод дискриминантного анализа применен для классификации чистых и загрязненных образцов исследуемых газовых смесей для разделения их на два класса: чистые образцы и загрязненные образцы. Для проведения классификации используется матрица данных, упомянутая выше.

#### Преобразование обрабатываемых измерений методом главных компонент и выполнение дискриминантного анализа с квадратичной дискриминантной функцией

Широко применяемый в прикладной многомерной статистике метод главных компонент (МГК) позволяет существенно сократить размерность обрабатываемых данных. Реальное количество главных компонент, или иначе говоря размерность обрабатываемых данных, с помощью которых представляются наблюдаемые данные, можно найти, анализируя вектор собственных значений ковариационной матрицы  $\text{CovM}$  исходной матрицы данных  $A$ . Для этого выполним следующие вычисления:

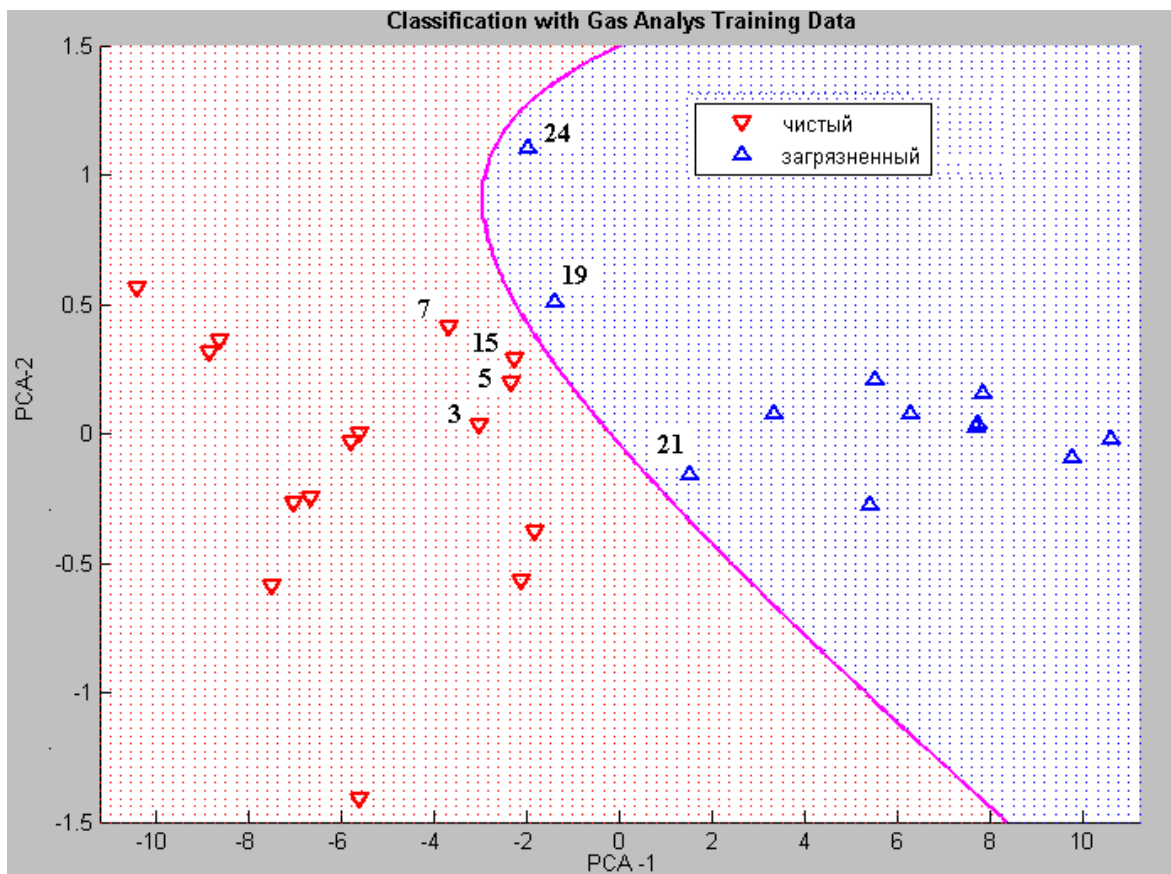
$$\text{CovM} = \text{cov}(A), \quad (3)$$

$$\lambda = \text{eig}(\text{CovM}). \quad (4)$$

В формулах (3) и (4):  $\text{cov}$  — вычисление ковариационной матрицы и  $\text{eig}$  — операция вычисления собственных чисел. Начиная с определенного номера элемента этого вектора, его элементы приближаются к нулю. Если выбрать из этих элементов только превышающие некоторый заданный порог, то таким образом мы выберем число главных компонент, которых достаточно для представления наших исходных данных по координатам главных компонент. В работе [12] мы ограничились шестью компонентами и получили приемлемый результат. Попробуем провести дискриминантный анализ, используя в качестве координат нашего многомерного пространства только две главные компоненты. В табл. 3 представлены собственные числа ковариационной матрицы  $A$ . В этой таблице должно быть 20 столбцов по числу переменных, т. е. значений энергий сигналов, полученных в результате дискретного вейвлет-преобразования, однако в результате расчета собственных чисел отличными от нуля оказались только 7 первых собственных чисел, которые представлены в таблице.

Табл. 3. Собственные числа ковариационной матрицы исходных данных

Номер главной компоненты											
1	2	3	4	5	6	7	8	9	10	11	12
41.7	0.190	0.09	0.06	0.009	0.006	0.001	0.0	0.0	0.0	0.0	0.0



**Рис. 11.** Проведение дискриминантного анализа для двух переменных, представляющих собой главные компоненты. Цифрами обозначены номера образцов наиболее близких к разделительной линии

На рис. 11 показан результат проведения дискриминантного анализа с квадратичной дискриминантной функцией, в котором были использованы только две главные компоненты: самая главная PCA-1 и следующая за ней PCA-2

Для получения результатов, приведенных на рис. 11 был проанализирован массив данных, состоящий из обучающей выборки для чистых образцов и загрязненных образцов. Результатом проведения обучения является квадратичная дискриминантная функция, показанная на рис. 11 в виде параболы. В контрольной выборке содержались результаты измерений 15 чистых образцов и 12 загрязненных. Обработка данных такого массива показала, что количество чистых образцов, отнесенных к классу загрязненных, равно нулю. Количество загрязненных образцов, отнесенных к классу чистых, также равно нулю.

Данный результат показывает принципиальную возможность использования двух главных компонент для разделения чистых и загрязненных образцов исследуемых газов.

## ВЫВОДЫ

1. Рассмотренные алгоритмы преобразования сигналов полупроводниковых сенсоров позволяют различать сигналы чистых и загрязненных образцов исследуемой смеси без использования дополнительных операций подготовки проб.

2. Наиболее информативным сигналом для выявления различий между чистыми и загрязненными образцами является сигнал свертки  $Z$ -сигнала с производной исходного сигнала.

3. Применение дискретного вейвлет-преобразования дает возможность количественно оценить различия сигналов чистых образцов от сигналов загрязненных образцов.

4. В качестве информационных критериев количественной оценки различий сигналов чистых и загрязненных образцов использованы энергии аппроксимирующих и детализирующих коэффициентов вейвлет-преобразований.

5. Полученные количественные характеристики

различий чистых и загрязненных образцов целесообразно использовать для автоматической классификации методами многомерной статистики.

### СПИСОК ЛИТЕРАТУРЫ

1. Белозерцев А.И., Черемисина О.В., Эль-Салим С.З., Манойлов В.В. Развернутые газоаналитические приборные комплексы для обнаружения компонент ракетных топлив в окружающей среде (обзор) // Научное приборостроение. 2017. Т. 27, № 2. С. 91–102. URL: <http://213.170.69.26/mag/2017/full2/Art12.pdf>.
2. Волькенштейн Ф.Ф. Электронные процессы на поверхности полупроводников при хемосорбции. М.: "Наука", Гл. ред. физ.-мат лит., 1987. 432 с.
3. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. М.: Финансы и статистика, 1983. 471 с.
4. Грешилов А.А., Стакун В.А., Стакун А.А. Математические методы построения прогнозов. М.: Радио и связь, 1997. 112 с.
5. Savitzky A., Golay M.J.E. Smoothing and differentiation of data by simplified least squares procedures // Analytical Chemistry. 1964. Vol. 36, no. 8. P. 1627–1639.
6. Белозерцев А.И., Черемисина О.В., Эль-Салим С.З., Манойлов В.В. Количественное определение несимметричного диметилгидразина в растворах методом спектроскопии комбинационного рассеяния // Научное приборостроение. 2017. Т. 27, № 2. С. 47–56. URL: <http://213.170.69.26/mag/2017/full2/Art6.pdf>.
7. Николаев А.В., Эль-Салим С.З. Классификация сигналов в процессе анализа газовых смесей. URL: <http://rl-omega.ru/docs/Classification.pdf>.
8. Обработка сигналов и изображений. URL: <http://matlab.exponenta.ru/wavelet/book1/14/wavemngr.php>.
9. Дьяконов В.П. Вейвлеты. От теории к практике. Серия: Полное руководство пользователя. Изд. Солонпресс, 2010. 400 с.
10. Krzanowski W.J. Principles of multivariate analysis: a us-

er's perspective. New York: Oxford University Press, 1988.

11. Манойлов В.В., Титов Ю.А., Кузьмин А.Г., Заруцкий И.В. Алгоритмы дискриминантного анализа для классификации масс-спектров выдыхаемых газов // Научное приборостроение. 2017. Т. 27, № 3. С. 33–42. URL: <http://213.170.69.26/mag/2017/full3/Art5.pdf>.
12. Манойлов В.В., Титов Ю.А., Кузьмин А.Г., Заруцкий И.В. Методы обработки и классификации масс-спектров выдыхаемых газов с использованием дискриминантного анализа // Научное приборостроение. 2016. Т. 26, № 3. С. 50–57. URL: <http://213.170.69.26/mag/2016/full3/Art7.pdf>.

*АО "Научно-исследовательский институт физических измерений", г. Пенза (Белозерцев А.И.)*

*Санкт-Петербургский горный университет (Черемисина О.В.)*

*ООО "Омега", Санкт-Петербург (Эль-Салим С.З.)*

*Институт аналитического приборостроения РАН, Санкт-Петербург (Манойлов В.В., Заруцкий И.В.)*

*Университет ИТМО, Санкт-Петербург (Манойлов В.В.)*

Контакты: Манойлов Владимир Владимирович, [manoilov\\_vv@mail.ru](mailto:manoilov_vv@mail.ru)

Материал поступил в редакцию 3.11.2017

# ALGORITHMS FOR DATA PROCESSING IN GAS ANALYTICAL COMPLEXES WITH SEMICONDUCTOR SENSORS FOR DETECTING CONTAMINANTS IN ROCKET FUEL COMPONENTS

A. I. Belozertsev<sup>1</sup>, O. V. Cheremisina<sup>2</sup>, S. Z. El Salim<sup>3</sup>, V. V. Manoylov<sup>4,5</sup>, I. V. Zarutskiy<sup>4</sup>

<sup>1</sup>Research Institute of Physical Measurements. Penza, Russian Federation

<sup>2</sup>Saint-Petersburg Mining University, Russian Federation

<sup>3</sup>Ltd "Omega", Saint-Petersburg, Russian Federation

<sup>4</sup>Institute for Analytical Instrumentation of RAS, Saint-Petersburg, Russia

<sup>5</sup>ITMO University, Saint-Petersburg, Russian Federation

In this paper, we describe the algorithms for processing data from semiconductor sensors that are part of the analytical means for controlling the vapors of rocket fuel components (RFC) in air, liquids, and solids. Data processing algorithms are designed for the quantitative analysis of individual components of RFC, as well as for the formation of information signs necessary for the classification of detected components. A method based on wavelet analysis, which allows to automatically distinguish a pure sample of the investigated gas from a sample containing impurities, impurities or additives is proposed.

*Keywords:* algorithms for processing gas analysis data, gas sensitive semiconductor sensors, components of rocket fuels, gas analytical instruments, environmental protection

## REFERENCES

1. Belozertsev A.I., Cheremisina O.V., El-Salim S.Z., Manoylov V.V. [Deployed gas analytical instrumentation systems for detection of rocket fuel components in the environment (review)]. *Nauchnoe Priborostroenie* [Scientific Instrumentation], 2017, vol. 27, no. 2, pp. 91–102. (In Russ.). Doi: 10.18358/np-27-2-i91102.
2. Volkenshetein F.F. *Elektronnye processy na poverhnosti poluprovodnikov pri hemosorbtsii* [Electronic processes on the surface of semiconductors under chemisorption]. Moscow, Nauka Publ., 1987. 432 p. (In Russ.).
3. Ajvazyan S.A., Enyukov I.S., Meshalkin L.D. *Prikladnaya statistika. Klassifikatsiya i snizhenie razmernosti. Spravochnoe izdanie* [Applied statistics. Classification and decrease in dimension. Reference book]. Moscow, Finansy i statistika Publ., 1983. 471 p. (In Russ.).
4. Greshilov A.A., Stakun V.A., Stakun A.A. *Matematicheskie metody postroeniya prognozov* [Mathematical methods for constructing predictions]. Moscow, Radio and Communication Publ., 1997. 112 p. (In Russ.).
5. Savitzky A., Golay M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 1964, vol. 36, no. 8, pp. 1627–1639. Doi: 10.1021/ac60214a047.
6. Belozertsev A.I., Cheremisina O.V., El-Salim S.Z., Manoylov V.V. Quantitative determination of asymmetric dimethylhydrazine in solutions by the Raman spectroscopy method. *Nauchnoe Priborostroenie* [Scientific Instrumentation], 2017, vol. 27, no. 2, pp. 47–56. (In Russ.). Doi: 10.18358/np-27-2-i4756.
7. Nikolaev A.V., El-Salim S.Z. *Classification of signals during the analysis of gas mixtures*. URL: <http://rl-omega.ru/docs/Classification.pdf>. (In Russ.).
8. *Matlab. Exponenta*. URL: <http://matlab.exponenta.ru/wavelet/book1/14/wavemngr.php>. (In Russ.).
9. D'yakonov V.P. *Vejvlety. Ot teorii k praktike. Seriya: Polnoe rukovodstvo pol'zovatelya* [Wavelets. From theory to practice. Series: A comprehensive user guide]. Moscow, Solon-press Publ., 2010. 400 p. (In Russ.).
10. Krzanowski W.J. *Principles of Multivariate Analysis: A User's Perspective*. New York: Oxford University Press, 1988.
11. Manoylov V.V., Titov Yu.A., Kuzmin A.G., Zarutskiy I.V. [Algorithms of discriminant analysis for classification of mass spectra of exhaled gases]. *Nauchnoe Priborostroenie* [Scientific Instrumentation], 2017, vol. 27, no. 3, pp. 33–42. (In Russ.). Doi: 10.18358/np-27-3-i3342.
12. Manoilov V.V., Titov Yu.A., Kuzmin A.G., Zarutskii I.V. [Methods for processing and classifying mass spectra of exhaled gases using discriminant analysis]. *Nauchnoe Priborostroenie* [Scientific Instrumentation], 2016, vol. 26, no. 3, pp. 50–56. (In Russ.). Doi: 10.18358/np-26-3-i5056.

Contacts: Manoylov Vladimir Vladimirovich,  
manoilov\_vv@mail.ru

Article received in edition 3.11.2017