

УДК 621.391.26

© Л. В. Новиков

## МОДИФИЦИРОВАННЫЕ ВЕЙВЛЕТЫ В ОБРАБОТКЕ ДАННЫХ АНАЛИТИЧЕСКИХ ПРИБОРОВ. II. АЛГОРИТМЫ ОБРАБОТКИ

Дано описание алгоритмов обработки с применением модифицированных вейвлетов. Построены блоки фильтров, реализующие быстрые вычислительные алгоритмы. Приведены примеры применения для обработки сигналов различных аналитических приборов. Показано, что предложенный подход позволяет совместить вейвлетное подавление шума, деконволюцию или другие преобразования сигнала, в несколько раз повысить чувствительность и разрешающую способность прибора.

### ВВЕДЕНИЕ

В работе [1] рассмотрены вейвлеты, полученные модификацией известных ортонормированных вейвлетных базисов (вейвлетов-"прародителей")  $\{\varphi_{j,k}(t), \psi_{j,k}(t); j \in Z, k \in N\}$  некоторой функцией  $G(t)$ . Новые базисные функции, обозначаемые как  $\theta_{j,k}(t), \tilde{\theta}_{j,k}(t)$  — масштабирующие функции,  $\eta_{j,k}(t), \tilde{\eta}_{j,k}(t)$  — вейвлетные функции, позволяют, в частности, вычислить оценку полезного сигнала  $x(t)$ , искаженного аппаратной функцией  $H(t)$  и шумом  $u(t)$  в наблюдаемом сигнале  $y(t)$

$$y(t) = \int H(t - \tau) x(\tau) d\tau + u(t). \quad (1)$$

Эта оценка имеет вид

$$\tilde{x}(t) = \sum_k \tilde{c}_{j_0}(k) \varphi_{j_0,k}(t) + \sum_{j=J_0}^J \sum_k \tilde{d}_j(k) \psi_{j,k}(t), \quad (2a)$$

где  $j_0$  и  $J$  — самое тонкое и самое грубое значения масштаба соответственно. Показано, что получение оценки по формуле (2a) эквивалентно обработке наблюдаемого сигнала фильтром с частотным откликом  $\hat{G}(\omega)$

$$\hat{\tilde{x}}(\omega) = \hat{G}(\omega) \hat{y}(\omega) \quad (2b)$$

и вейвлетным подавлением шума по алгоритму, предложенному в работе [3].

Коэффициенты вейвлет-разложения наблюдаемого сигнала  $y(t)$  —  $\tilde{c}_{j_0}(k)$  и  $\tilde{d}_j(k)$  определяются как

$$\tilde{c}_{j_0}(k) = \int y(t) \tilde{\theta}_{j_0,k}(t) dt, \quad (3a)$$

$$\tilde{d}_j(k) = \int y(t) \tilde{\eta}_{j,k}(t) dt, \quad (3b)$$

где функции  $\tilde{\theta}_{j,k}(t)$  и  $\tilde{\eta}_{j,k}(t)$  должны удовлетворять следующим масштабирующим уравнениям:

$$\tilde{\theta}_{j,k}(t) = \sum_n h(n - 2k) \tilde{\theta}_{j+1,n}(t), \quad (4a)$$

$$\tilde{\eta}_{j,k}(t) = \sum_n g(n - 2k) \tilde{\theta}_{j+1,n}(t), \quad (4b)$$

или

$$\tilde{\theta}_j(t) = \sum_n \tilde{\alpha}_{j+1}(n) \varphi_{j+1,n}(t), \quad (5a)$$

$$\tilde{\eta}_j(t) = \sum_n \tilde{\beta}_{j+1}(n) \varphi_{j+1,n}(t), \quad (5b)$$

где  $h$  и  $g$  — масштабирующие и вейвлетные коэффициенты вейвлетов-"прародителей",  $\tilde{\alpha}$  и  $\tilde{\beta}$  — масштабирующие и вейвлетные коэффициенты новых вейвлетов, которые определяются по формулам:

$$\begin{aligned} \tilde{\alpha}_{j+1}(n) &= \\ &= \frac{1}{2^{j+1} 2\pi} \int_{-2^{j+1}\pi}^{2^{j+1}\pi} \hat{G}(\omega) \hat{h}(2^{-j-1}\omega) e^{i2^{-j-1}\omega n} d\omega, \end{aligned} \quad (6a)$$

$$\begin{aligned} \tilde{\beta}_{j+1}(n) &= \\ &= \frac{1}{2^{j+1} 2\pi} \int_{-2^{j+1}\pi}^{2^{j+1}\pi} \hat{G}(\omega) \hat{g}(2^{-j-1}\omega) e^{i2^{-j-1}\omega n} d\omega. \end{aligned} \quad (6b)$$

Обозначим

$$\tilde{\alpha}_{j+1}(\omega) = \hat{G}(\omega) \hat{h}(2^{-j-1}\omega), \quad (7a)$$

$$\widehat{\beta}_{j+1}(\omega) = \widehat{G}(\omega) \widehat{g}(2^{-j-1}\omega). \quad (76)$$

В настоящей части работы рассматриваются алгоритмы обработки наблюдаемых данных  $y(t)$  и приведены результаты их апробации на реальных сигналах аналитических приборов. Блок-схема алгоритма вейвлетной обработки наблюдаемого сигнала  $y(t)$  приведена на рис. 1. Наблюдаемый сигнал сначала подвергается дискретному вейвлетному преобразованию (ДВП), результатом которого являются коэффициенты вейвлет-разложения  $\check{c}_j(k)$  и  $\check{d}_j(k)$ . При удачном выборе вейвлетного базиса энергия сигнала оказывается сосредоточенной в небольшом числе коэффициентов, в то время как энергия шума рассредоточивается практически по всем коэффициентам вейвлет-разложения. Благодаря этому обстоятельству в результате пороговой обработки (ПО) почти все коэффициенты шума приравниваются нулю и сохраняются только те коэффициенты сигнала, которые оказались выше порогового уровня. В блоке обратного ДВП (ОДВП) происходит восстановление оценки полезного сигнала, а в блоке оценки параметров (ОП) — расчет параметров сигнала, содержащих информацию об анализируемом веществе.

### 1. БЫСТРЫЕ ВЫЧИСЛИТЕЛЬНЫЕ АЛГОРИТМЫ И БЛОКИ ФИЛЬТРОВ

#### Анализ

Чтобы получить быстрые алгоритмы для вычисления коэффициентов вейвлет-разложения

оценки полезного сигнала  $\check{x}(t)$ , подставим в (3а, 3б) вместо  $\check{\theta}_{j,k}$  и  $\check{\eta}_{j,k}$  их выражения из (4а) и (4б) соответственно. Получим:

$$\check{c}_j(k) = \sum_m h(m-2k) \check{c}_{j+1}(k), \quad (8а)$$

$$\check{d}_j(k) = \sum_m g(m-2k) \check{c}_{j+1}(k). \quad (8б)$$

Так же как и в случае традиционного КМА, мы получили рекурсивный алгоритм для вычисления коэффициентов, который требует знания начальных значений  $\check{c}_j(k)$  и  $\check{d}_j(k)$  на самом "тонком" масштабе  $J$ :

$$\check{c}_j(k) = \int y(t) \check{\theta}_{j,k}(t) dt, \quad (9а)$$

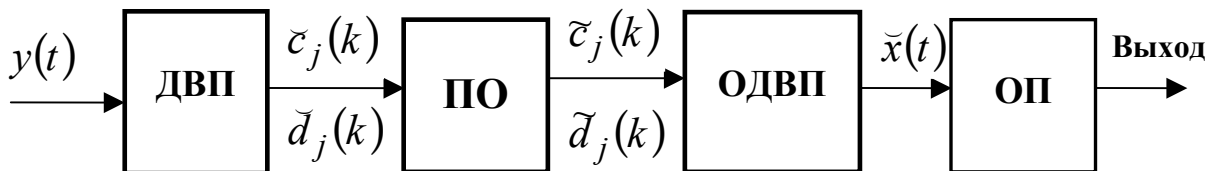
$$\check{d}_j(k) = \int y(t) \check{\eta}_{j,k}(t) dt. \quad (9б)$$

Чтобы вычислить эти коэффициенты, так же как и в случае традиционного КМА, необходимо воспользоваться отсчетами наблюдаемого сигнала  $y(t)$ . Подставляя в (9а, 9б) выражения для  $\check{\theta}_{j,k}$  и  $\check{\eta}_{j,k}$  из (5а, 5б) и выполняя преобразования, получим

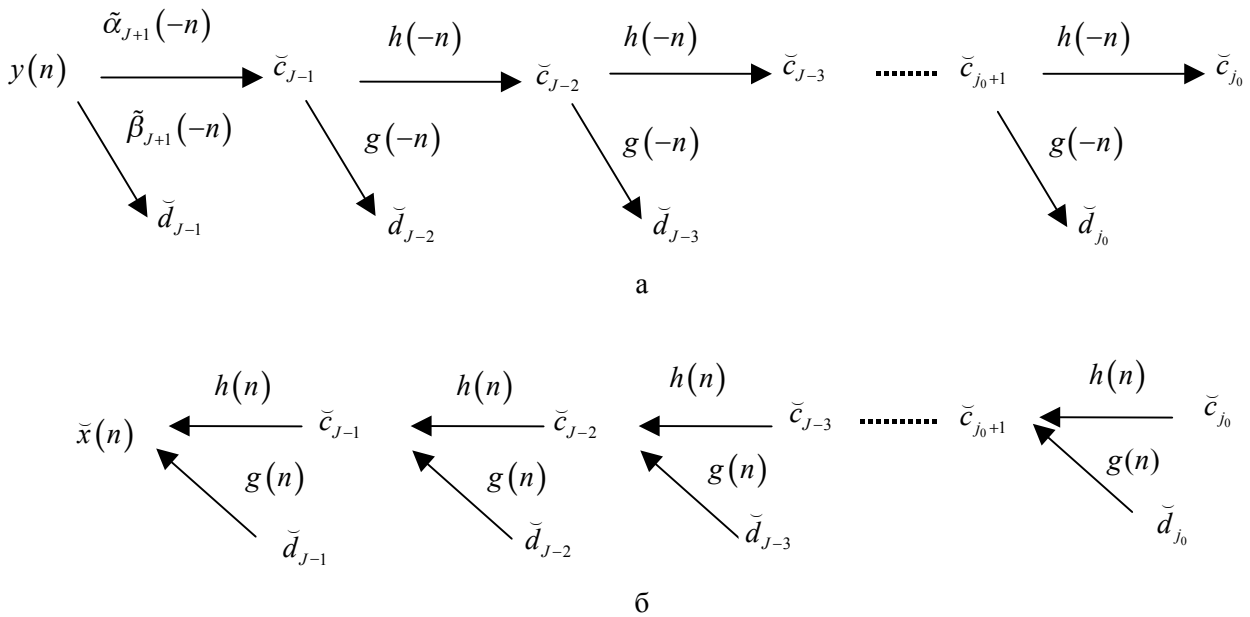
$$\begin{aligned} \check{c}_j(k) &= \sum_m \check{\alpha}_{j+1}(m-2k) \int y(t) \phi_{j+1,m}(t) dt \approx \\ &\approx \sum_m y(m) \check{\alpha}_{j+1}(m-2k) \end{aligned} \quad (10а)$$

и

$$\check{d}_j(k) \approx \sum_m y(m) \check{\beta}_{j+1}(m-2k). \quad (10б)$$



**Рис. 1.** Вейвлетное подавление шума. ДВП — дискретное вейвлетное преобразование; ПО — пороговая обработка; ОДВП — обратное ДВП; ОП — оценка параметров;  $y(t)$  — наблюдаемый сигнал;  $\check{c}_j(k)$ ,  $\check{d}_j(k)$  — коэффициенты вейвлет-разложения наблюдаемого сигнала;  $\check{c}_j(k)$ ,  $\check{d}_j(k)$  — коэффициенты вейвлет-разложения наблюдаемого сигнала после пороговой обработки;  $\check{x}(t)$  — оценка полезного сигнала



**Рис. 2.** Схема 4-ступенчатого анализирующего (а) и синтезирующего (б) блоков фильтров.  $\tilde{\alpha}_{j+1}$ ,  $h$  и  $\tilde{\beta}_{j+1}$ ,  $g$  — коэффициенты низкочастотного и высокочастотного фильтров вейвлетов

По формулам (8\*) и (10\*) можно построить вычислительную схему алгоритма анализа наблюдаемых данных, которая показана на рис. 2, а. В отличие от схемы [1, рис. 3, а] отсчеты наблюдаемого сигнала в первом каскаде подвергаются обработке низкочастотным и высокочастотным фильтрами с другими коэффициентами, а именно  $\tilde{\alpha}_{j+1}$  и  $\tilde{\beta}_{j+1}$  соответственно. Коэффициенты фильтров на последующих ступенях анализа остаются прежними.

**Синтез**

Учитывая что  $\tilde{U}_{j+1} = \tilde{U}_j \oplus \tilde{M}_j$  (см. [1, п. 3]), базисную функцию  $\tilde{\theta}_{j+1,k} \in \tilde{U}_{j+1}$  можно представить в виде

$$\tilde{\theta}_{j+1,m}(t) = \sum_k \langle \tilde{\theta}_{j+1,m}, \theta_{j,k} \rangle \tilde{\theta}_{j,k}(t) + \sum_m \langle \tilde{\theta}_{j+1,m}, \eta_{j,k} \rangle \tilde{\eta}_{j,k}(t)$$

или с учетом формул (26) и (29) в [1]:

$$\tilde{\theta}_{j+1,m}(t) = \sum_k h(m-2k)\tilde{\theta}_{j,k}(t) + \sum_m g(m-2k)\tilde{\eta}_{j,k}(t). \quad (11)$$

Умножая далее (11) скалярно слева и справа на  $y(t)$ , получим

$$\tilde{c}_{j+1}(m) = \sum_k h(m-2k)\tilde{c}_j(k) + \sum_m g(m-2k)\tilde{d}_j(k). \quad (12)$$

Мы получили, таким образом, рекурсивную формулу восстановления коэффициентов вейвлет-разложения искомой оценки  $\tilde{x}(t)$  "от грубого к тонкому разрешению" с использованием, как и в анализе, фильтров-"прародителей"  $h$  и  $g$ . Вычислительная схема восстановления отсчетов оценки  $\tilde{x}(t)$  ничем не отличается от схемы традиционного КМА (ср. [1, рис. 3, б] и рис. 2, б).

Покажем теперь, что схема анализа—синтеза, показанная на рис 2, обеспечивает получение оценки полезного сигнала  $\tilde{x}(t)$  по наблюдаемому сигналу  $y(t)$ , если значение масштаба  $J$  выбрано равным  $-1$ , т. е. при  $J = -1$ . Чтобы положить коэффициенты  $c_0(k)$ , равными отсчетам наблюдаемого сигнала  $y(k)$ , т. е.  $c_0(k) = y(k)$ , как уже отмечалось в [1], спектр масштабирующей функции  $\tilde{\varphi}_0(\omega)$  в пределах спектральной полосы  $\tilde{y}(\omega)$

должен быть максимально плоским. Это означает, что при использовании в качестве "прародителей" вейвлетов Добеши, имеющих плоскую частотную характеристику вплоть до частоты Найквиста, частота отсчетов наблюдаемого сигнала должна быть меньше  $\pi$ . Тогда из формулы (10) на первом шаге анализа необходимо выполнить:

$$\check{c}_{-1}(k) = \sum_m \check{\alpha}_0(m-2k)y(m), \quad (12a)$$

$$\check{d}_{-1}(k) = \sum_m \check{\beta}_0(m-2k)y(m). \quad (12б)$$

Рассмотрим работу первого каскада анализа и последнего каскада синтеза блока фильтров, который должен восстанавливать искомый сигнал в его дискретных точках  $\check{x}(k)$  (рис. 3). Можно показать [2], что вход и выход такого фильтра связаны выражением

$$\begin{aligned} \check{x}(\omega) = & \frac{1}{2} \left[ \overline{\check{\alpha}_0(\omega)} \hat{h}(\omega) + \overline{\check{\beta}_0(\omega)} \hat{h}_1(\omega) \right] \hat{y}(\omega) + \\ & + \frac{1}{2} \left[ \overline{\check{\alpha}_0(\omega + \pi)} \hat{h}(\omega) + \overline{\check{\beta}_0(\omega + \pi)} \hat{h}_1(\omega) \right] \hat{y}(\omega + \pi). \end{aligned}$$

Вторая сумма в квадратных скобках равна нулю, и, следовательно, равна нулю "элайзинговая" составляющая синтезируемого сигнала. Действительно, имеем с учетом (7\*) при  $j = -1$

$$\begin{aligned} \overline{\check{\alpha}_0(\omega + \pi)} \hat{h}(\omega) + \overline{\check{\beta}_0(\omega + \pi)} \hat{h}_1(\omega) = \\ = \hat{G}(\omega + \pi) \overline{\hat{h}(\omega + \pi)} \hat{h}(\omega) + \\ + \hat{G}(\omega + \pi) \overline{\hat{h}_1(\omega + \pi)} \hat{h}_1(\omega) = \end{aligned}$$

$$\begin{aligned} = \hat{G}(\omega + \pi) \times \\ \times \left[ \hat{h}(\omega) \overline{\hat{h}(\omega + \pi)} + \exp(i\pi) \hat{h}(\omega) \overline{\hat{h}(\omega + \pi)} \right] = 0. \end{aligned}$$

Первая сумма в квадратных скобках равна

$$\begin{aligned} \overline{\check{\alpha}_0(\omega)} \hat{h}(\omega) + \overline{\check{\beta}_0(\omega)} \hat{h}_1(\omega) = \\ = \hat{G}(\omega) \overline{\hat{h}(\omega)} \hat{h}(\omega) + \hat{G}(\omega) \overline{\hat{h}_1(\omega)} \hat{h}_1(\omega) = \\ = \hat{G}(\omega) \left[ |\hat{h}(\omega)|^2 + |\hat{h}_1(\omega)|^2 \right] = 2\hat{G}(\omega). \end{aligned}$$

Следовательно,

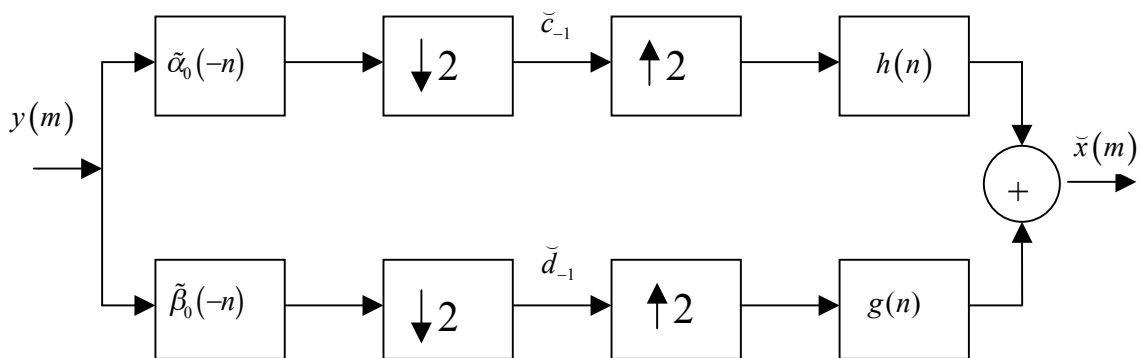
$$\check{x}(\omega) = \hat{G}(\omega) \hat{y}(\omega).$$

Сравнение последнего выражения с формулой (2б) показывает, что применение вычислительной схемы рис. 2 для обработки наблюдаемых данных  $y(t)$  при соответствующем выборе  $G(t)$  приводит к желаемому результату получения оценки полезного сигнала  $\check{x}(t)$ .

## 2. ОПИСАНИЕ АЛГОРИТМОВ

Таким образом, практическая реализация обработки данных с использованием модифицированных вейвлетов включает выполнение следующих этапов:

а) выбор или синтез функции  $G(t)$  в зависимости от круга задач решаемых прибором;



**Рис. 3.** Первая ступень блока фильтров анализа-синтеза:  $y(m)$  — наблюдаемый сигнал;  $\check{x}(m)$  — оценка сигнала;  $\downarrow 2$  и  $\uparrow 2$  — прореживание и интерполяция с коэффициентом 2;  $\check{c}_{-1}$  и  $\check{d}_{-1}$  — коэффициенты анализа сигнала первой ступени

- б) выбор вейвлета-"прародителя" (коэффициентов  $h(n)$  и  $g(n)$ );
- в) вычисление коэффициентов  $\tilde{\alpha}_0(n)$  и  $\tilde{\beta}_0(n)$ ;
- г) ДВП;
- д) пороговая обработка вейвлет-коэффициентов;
- е) ОДВП;
- ж) оценивание параметров пиков.

Рассмотрим более подробно алгоритмы, реализуемые на каждом этапе обработки.

### Выбор функции $G(t)$

**Деконволюция.** Как уже отмечалось, выбор функции  $\hat{G}(\omega)$ , равной  $\hat{H}^{-1}(\omega)$ , приводит к большому уровню выходного шума. С целью подавления эффекта бесконечного усиления на практике применяют окна  $W(\omega)$  или используют восстанавливающий фильтр Винера—Тихонова [4], регуляризирующий усиление на верхних частотах,

$$\hat{G}(\omega) = \frac{\overline{\hat{H}(\omega)}}{|\hat{H}(\omega)|^2 + \beta R(\omega)}, \quad (14)$$

где  $\beta$  — регуляризирующий параметр,  $R(\omega)$  — некоторый четный полином.

Выбирая константу  $\beta$  и полином  $R(\omega)$ , можно легко корректировать ход частотной характеристики  $\hat{G}(\omega)$ .

**Дифференцирование.** Чтобы получить оценку  $p$ -й производной сигнала  $x(t)$ , искаженного прибором и шумами, функция  $\hat{G}(\omega)$  будет иметь вид

$$\hat{G}(\omega) = (i\omega)^p \frac{\overline{\hat{H}(\omega)}}{|\hat{H}(\omega)|^2 + \beta R(\omega)}. \quad (15)$$

Если же необходимо получить оценку производной полезного сигнала без учета аппаратной функции, то можно воспользоваться дифференцирующими фильтрами Савицкого—Голэя [5]. Тогда функция  $\hat{G}(\omega)$  будет частотным откликом цифрового дифференцирующего фильтра, т. е.

$$\hat{G}(\omega) = \hat{G}(\exp(i\omega)) = G_0(z), \quad (16)$$

где  $z = \exp(i\omega)$ .

**Интегрирование.** При оценке текущего значения интеграла  $\int_{t-T}^t \tilde{x}(\tau) d\tau$  с учетом искажений, вносимых прибором и шумами, функция  $\hat{G}(\omega)$  будет иметь вид

$$\hat{G}(\omega) = (i\omega)^{-1} \frac{\overline{\hat{H}(\omega)}}{|\hat{H}(\omega)|^2 + \beta R(\omega)}. \quad (17)$$

**(Карунена—Лоэва)-подобные разложения (декорреляция).** Известно, что разложение Карунена—Лоэва приводит к некоррелированным коэффициентам разложения [6]. Однако этот же результат можно получить с помощью предлагаемых вейвлетов. Предположим, что на входе системы  $\hat{H}(\omega)$  действует белый шум с единичной дисперсией  $x(t)$ . Тогда на ее выходе он оказывается коррелированным с корреляционной функцией

$$K(\tau) = \frac{1}{2\pi} \int \hat{K}(\omega) \exp(i\omega\tau) d\omega,$$

где  $\hat{K}(\omega)$  — спектральная плотность мощности и

$$\hat{K}(\omega) = |\hat{H}(\omega)|^2.$$

Выберем  $\hat{G}(\omega) = \hat{K}(\omega)^{-1/2}$  и покажем, что коэффициенты разложения  $\tilde{c}_j(k)$  и  $\tilde{d}_j(k)$  при этом будут некоррелированными. Действительно, с учетом (3а) получим

$$\begin{aligned} E[\tilde{c}_j(m)\tilde{c}_j(k)] &= \\ &= \iint E[y(t)y(\tau)]\tilde{\theta}_{j,k}(t)\tilde{\theta}_{j,k}(\tau) dt d\tau = \\ &= \iint K(t-\tau)\tilde{\theta}_{j,k}(t)\tilde{\theta}_{j,k}(\tau) dt d\tau = \\ &= \frac{1}{2\pi} \int \hat{K}(\omega) |\tilde{\theta}_j(\omega)|^2 \exp(i\omega(m-n)) d\omega = \\ &= \frac{1}{2\pi} \int |\varphi(\omega)|^2 \exp(i\omega(m-n)) d\omega = \delta(m-n), \end{aligned}$$

где символ  $E$  обозначает статистическое усреднение.

Аналогично можно показать, что

$$E[\tilde{d}_j(m)\tilde{d}_j(n)] = \delta(m-n).$$

Из формул [1, (19а, 21а)] при  $k=0$  можно получить интересный результат:

$$\hat{K}(\omega)\hat{\theta}_j(\omega) = \tilde{\theta}_j(\omega),$$

или

$$\int K(t-\tau)\tilde{\theta}_j(\tau)d\tau=\theta_j(t),$$

указывающий на связь масштабирующих функций модифицированных вейвлетов.

Однако так же, как в случае с деконволюцией, выбор  $\widehat{G}(\omega)=\widehat{K}(\omega)^{-1/2}$  может оказаться некорректным вследствие слишком большого усиления на высоких частотах, поэтому  $\widehat{G}(\omega)$  необходимо строить по формуле (14), положив  $\widehat{H}(\omega)=\widehat{K}(\omega)^{1/2}$ .

**Коррекция частотной характеристики системы.** Чтобы скорректировать частотную характеристику системы, необходимо умножить функцию  $\widehat{G}(\omega)$  на подходящий полином от частоты  $\omega$ . Например, чтобы поднять верхние частоты, достаточно умножить эту функцию на полином вида  $1+a_0\omega^2$ .

**Выбор модифицируемых вейвлетов.** Как было показано в [1], функции  $\theta, \tilde{\theta}, \eta, \tilde{\eta}$  порождают КМА, если в качестве модифицируемого выбран вейвлет с компактным носителем в частотной области (вейвлет Мейера или вейвлет Шеннона). Однако эти вейвлеты во временной области затухают очень медленно, определяются большим числом коэффициентов  $h$  и  $g$  и неудобны для приложений. Поэтому в качестве модифицируемых вейвлетов необходимо использовать другие вейвлеты, например Добеши высоких порядков, имеющих хорошую крутизну фронтов частотного спектра. Возникающая при этом систематическая ошибка может быть оценена численно в каждом конкретном случае.

**Вычисление коэффициентов  $\tilde{\alpha}_0$  и  $\tilde{\beta}_0$ .** Эти коэффициенты определяются формулами (6а) и (6б) при  $j=-1$ :

$$\tilde{\alpha}_0(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\widehat{G}(\omega)\widehat{h}(\omega)e^{i\omega n}d\omega, \quad (18a)$$

$$\tilde{\beta}_0(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\widehat{G}(\omega)\widehat{g}(\omega)e^{i\omega n}d\omega. \quad (18б)$$

Для вычислений по этим формулам функции  $\widehat{G}(\omega)$ ,  $\widehat{h}(\omega)$  и  $\widehat{g}(\omega)$  должны быть определены на дискретном множестве (сетке) частоты  $\omega$

$$\omega=(N_1:N_1-1)\pi/N_1,$$

где  $N_1$  выбирается в пределах 100÷200, что дает вполне удовлетворительный результат. Тогда ко-

эффициенты  $\tilde{\alpha}_0$  и  $\tilde{\beta}_0$  вычисляются по формуле обратного дискретного преобразования Фурье (ДПФ), которая для  $\tilde{\alpha}_0$  имеет вид

$$\tilde{\alpha}_0(n)=\frac{1}{2N_1}\sum_{m=-N_1}^{N_1-1}\widehat{G}(m)\widehat{h}(m)e^{i\pi mn/N_1}$$

для всех  $n=(M_0:M_1)$ , где  $M_0$  и  $M_1$  выбираются такими, чтобы  $\tilde{\alpha}_0(n)\approx 0$  при  $M_0\geq n\geq M_1$ .

По аналогичной формуле вычисляются коэффициенты  $\tilde{\beta}_0$ .

Если приборная функция  $H(t)$ , которая в ряде задач определяет  $G(t)$ , задана на дискретной сетке  $t=(-N_0:N_0)$ , то ее ДПФ вычисляется по формуле

$$\widehat{H}(m)=\sum_{k=0}^{2N_0-1}H(k)e^{-i\pi mk/N_1}$$

для всех  $m=(-N_1:N_1-1)$ .

Если же функция  $\widehat{G}(\omega)$  определена по формуле (16) и так как  $h(n)$  и  $g(n)$  являются КИХ-фильтрами, коэффициенты  $\tilde{\alpha}_0(n)$  и  $\tilde{\beta}_0(n)$  могут быть получены путем умножения полиномов по степеням  $z$ :

$$\tilde{\alpha}_0(z)=(\widehat{G}_0*h)(z), \quad (19a)$$

$$\tilde{\beta}_0(z)=(\widehat{G}_0*h_1)(z). \quad (19б)$$

Тогда искомые коэффициенты будут коэффициентами полиномов  $\tilde{\alpha}_0(z)$  и  $\tilde{\beta}_0(z)$ .

**Быстрые дискретные вейвлетные преобразования (ДВП).** ДВП имеют целью вычисление коэффициентов  $\tilde{c}_j(k)$  и  $\tilde{d}_j(k)$  для значений масштабов  $0\geq j\geq j_0$ . Вычислительные процедуры и реализующие их блоки фильтров рассмотрены в [1, раздел 1] и здесь в разделе 1. Из формул (8\*) и (10\*) следует, что в основе структуры ДВП лежит блок фильтров, показанный в [1, рис. 4]. Их последовательное соединение по схеме рис. 2, а реализует процедуру анализа наблюдаемого сигнала  $y(t)$ . В отличие от традиционного КМА, как показано на рис. 2, а, для достижения требуемого результата первый каскад блока фильтров должен быть снабжен коэффициентами фильтров  $\tilde{\alpha}_0(n)$  и  $\tilde{\beta}_0(n)$ . Напомним, что операция децимации на каждом шаге анализа понижает число коэффициентов вдвое. Поэтому, чтобы достигнуть наиболь-

шей глубины анализа по масштабу, выбирают число отсчетов наблюдаемого сигнала  $y(n)$  кратным степени двойки, т. е.  $N = 2^{J_0}$ . Тогда формально можно выполнить максимально  $J_0$  шагов анализа ( $J_0 = \log_2(N)$ ), получив в результате  $N-1$  коэффициентов  $\check{d}_j$  и один коэффициент  $\check{c}_{j_0}$ . На практике часто применяют ДВП с избыточным числом коэффициентов (избыточное (redundant) ДВП), при котором иногда достигается лучшее подавление шума. В этом преобразовании в результате децимации коэффициенты не отбрасываются, а сохраняются. Тогда на первом шаге анализа получим пару векторов  $\check{d}_j(k)$ :  $\check{d}_j^A(k)$  и  $\check{d}_j^B(k)$  длиной  $N/2$  каждый — и пару векторов  $\check{c}_j(k)$ :  $\check{c}_j^A(k)$  и  $\check{c}_j^B(k)$  длиной  $N/2$  каждый. Векторы  $\check{c}_j^A(k)$  и  $\check{c}_j^B(k)$  на втором шаге подвергаются аналогичной обработке, в результате которой получаем четыре вектора  $\check{d}_j(k)$  длиной  $N/4$  каждый и четыре вектора  $\check{c}_j(k)$  длиной  $N/4$  каждый, и т. д.; выполнив  $J_0$  шагов, получим  $N \log_2(N)$  коэффициентов  $\check{d}_j$  и  $\check{c}_{j_0}$ .

Программы, реализующие ДВП и избыточное ДВП, содержатся в последних версиях МАТЛАБ, а также их можно найти на различных сайтах ИНТЕРНЕТ, например [7].

**Пороговая обработка коэффициентов вейвлет-разложения.** Пороговая обработка имеет целью удаление шума из наблюдаемых данных. Процедура пороговой обработки осуществляется только для коэффициентов  $d_j(k)$ , которые сравниваются с порогом  $\text{Tr}$ , обнуляются, если этот коэффициент ниже порога, или сохраняют свое значение в противоположном случае. Очевидно, что идеальное удаление шума будет в том случае, если энергия сигнала в результате ДВП оказывается сосредоточенной в минимальном числе коэффициентов, а энергия шума размыта по максимальному их числу. Этого можно достичь в том случае, если базис будет хорошо коррелирован с сигналом, что и происходит для сигналов, имеющих небольшую продолжительность (для пиков), при применении вейвлетного базиса. Очевидно, что наибольшая корреляция будет в том случае, если масштабирующие и вейвлетные функции совпадают с формой сигнала.

Различают жесткую и мягкую пороговые обработки. Жесткая пороговая обработка строится по алгоритму [8]

$$d_j(k) = \begin{cases} d_j(k), & \text{если } |d_j(k)| > \text{Tr}; \\ 0, & \text{если } |d_j(k)| \leq \text{Tr}, \end{cases}$$

а мягкая пороговая обработка:

$$d_j(k) = \begin{cases} d_j(k) - \text{Tr}_j, & \text{если } |d_j(k)| \geq \text{Tr}_j; \\ d_j(k) + \text{Tr}_j, & \text{если } |d_j(k)| \leq -\text{Tr}_j; \\ 0, & \text{если } |d_j(k)| < \text{Tr}_j, \end{cases}$$

где порог для каждого значения масштаба  $j$  определяется как

$$\text{Tr}_j = \sigma_j \sqrt{2 \ln N}. \quad (20)$$

Дисперсия шума  $\sigma_j^2$  определяется по формуле

$$\sigma_j^2 = \sigma^2 \sum_n d_{0,j}^2(n), \quad (21)$$

где  $d_{0,j}$  — коэффициенты разложения функции  $G(t)$  по вейвлетным функциям  $d_{0,j} = \int G(t) \check{\eta}_{j,k}(t) dt$ ,  $\sigma^2$  — дисперсия шума  $u(t)$ . Эта дисперсия оценивается по наблюдаемым данным  $y(t)$  путем вычисления медианы коэффициентов  $d_j(k)$  на самой тонкой шкале ( $j=0$ ). Если эта медиана равна  $M_d$ , то оценка  $\sigma$  будет

$$\sigma = \frac{M_d}{0.6745}. \quad (22)$$

**Обратные дискретные вейвлетные преобразования (ОДВП).** ОДВП имеют целью восстановление дискретных значений оценки сигнала  $\check{x}(n)$  по коэффициентам вейвлет-разложения наблюдаемого сигнала  $y(t)$ , прошедшим пороговую обработку. Вычислительная схема ОДВП показана на рис. 2, б.

**Оценка параметров пиков.** Параметрами пиков, содержащими информацию об анализируемом веществе, являются их амплитуда, положение максимума (или центра тяжести) и площадь. Для определения этих параметров обработке подвергаются отсчеты  $\check{x}(n)$ , которые являются случайными величинами, подчиняющимися, как правило, закону нормального распределения. Чтобы оценить перечисленные выше параметры пика, нужно прежде всего его обнаружить. Для этого используются такие критерии, как интенсивность

отсчетов  $\tilde{x}(n)$ , крутизна фронтов и ширина пика. Все эти показатели вычисляются в текущем окне данных из трех и более отсчетов и сравниваются с эмпирически выбранными порогами. Площадь определяется как сумма отсчетов в пределах от начала до конца пика, интенсивность — как величина максимального отсчета в этих же пределах, а положение пика — как положение максимума. Для более точного определения параметров отсчеты в окне данных могут быть аппроксимированы, например, квадратичным сплайном.

### 3. ПРИМЕРЫ ПРИМЕНЕНИЙ

В практике предварительной обработки сигналов решаются задачи оценки функции  $\tilde{x}(t)$ , ее  $p$ -й производной или текущего значения интеграла

$\int_{t-T}^t \tilde{x}(\tau) d\tau$  при условии минимизации величины соответствующего среднеквадратического отклонения (СКО):

$$\|\tilde{x}(t) - x(t)\| < \varepsilon_0,$$

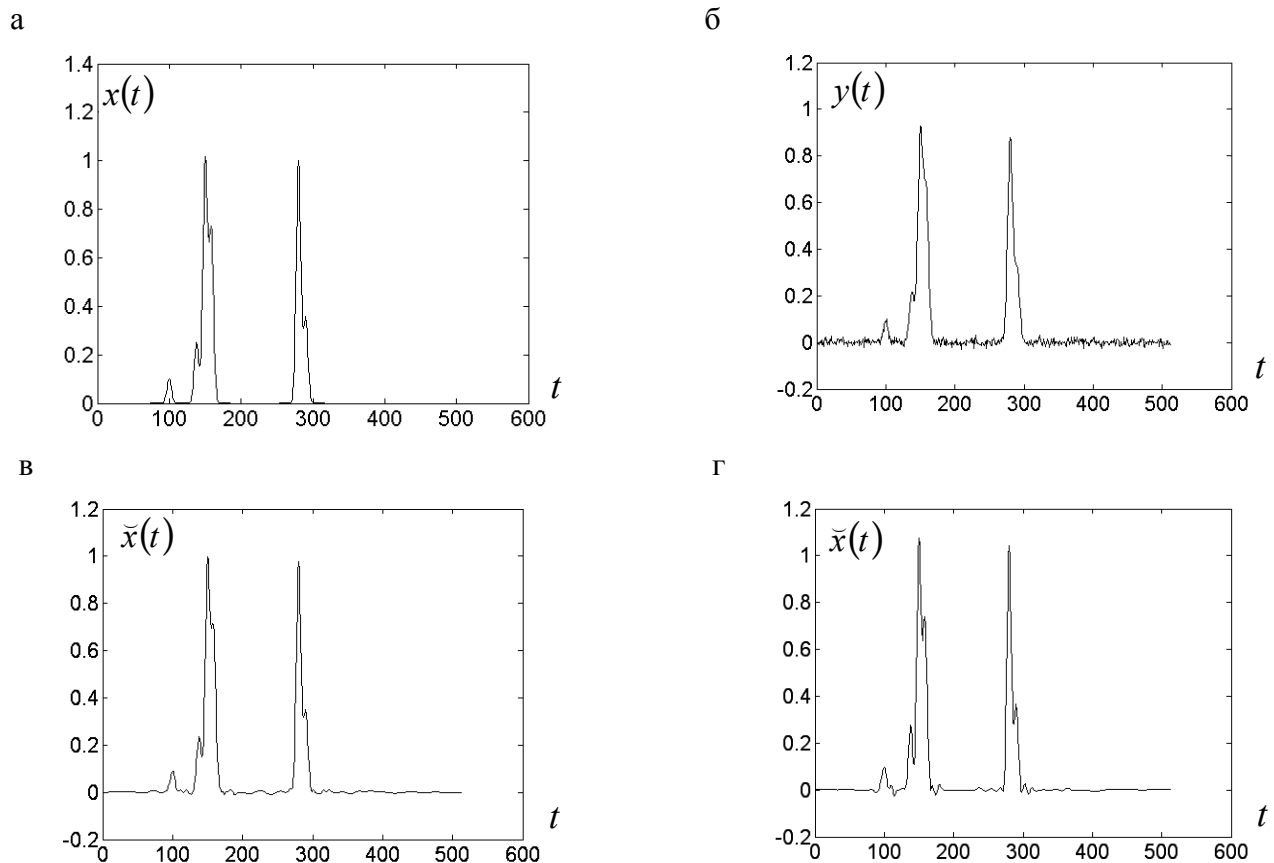
$$\|\tilde{x}^{(p)}(t) - x^{(p)}(t)\| < \varepsilon_p,$$

$$\left\| \int_{t-T}^t \tilde{x}(\tau) d\tau - \int_{t-T}^t x(\tau) d\tau \right\| < \varepsilon_T,$$

где  $\varepsilon_* > 0$  — некоторая заданная ошибка.

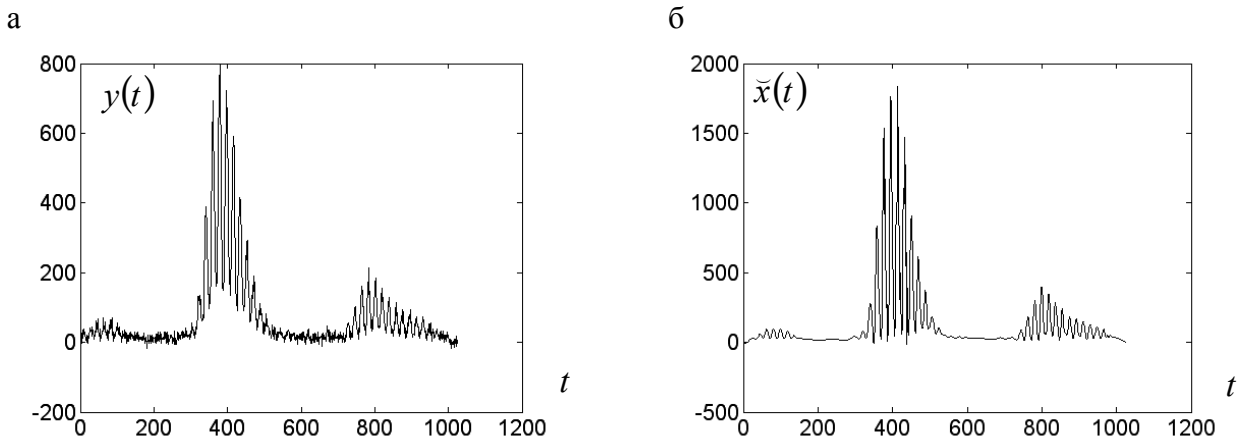
Чаще всего используют относительное СКО, определяемое как

$$\text{RMSE}(x - \tilde{x}) = \left\{ \varepsilon^2 / \int |x(t)|^2 dt \right\}^{\frac{1}{2}}. \quad (23)$$



**Рис. 4.** Пример деконволюции модельного сигнала. а — оригинальный сигнал  $x(t)$ ; б — наблюдаемый сигнал  $y(t)$  с шумом при отношении сигнал/шум 10 ( $\text{RMSE}(x - y) = 0.1$ ); в — оценка сигнала  $\tilde{x}(t)$  с шумом ( $\text{RMSE}(x - \tilde{x}) = 0.03$ ,  $\beta = 0.02$ ,  $\hat{R}(\omega) \equiv 1$ ); г — то же, что и (в), но с усилением верхних частот с помощью полинома  $1 + \omega^2$





**Рис. 5.** Фрагмент спектра инсулина времяпролетного масс-спектрометра.  
а — наблюдаемый сигнал, б — оценка сигнала, полученная методом обратной свертки

В дальнейшем будем пользоваться величиной относительного СКО, чтобы оценить степень зашумленности сигналов, качество подавления шума, качество полученной оценки и т. п. Отношение сигнал/шум в приведенных ниже примерах будем оценивать как отношение амплитуды компонента сигнала самой минимальной интенсивности к среднеквадратическому значению шума.

**Деконволюция.** Применим сначала предложенный подход для решения задачи деконволюции. Задача деконволюции является некорректной, т. к. не существует физически реализуемых способов точного восстановления искаженного линейной системой сигнала. Приближенное решение задачи возможно путем применения функции  $\hat{G}(\omega)$ , заданной формулой (14). Выберем модельный сигнал с белым шумом, определенный на множестве точек  $N = 2^9$ , состоящий из шести пиков гауссовой формы  $a_k \exp\left(-\frac{(t-t_k)^2}{2\mu_0^2}\right)$ , где

$$a_k = [0.1 \ 0.25 \ 1.0 \ 0.7 \ 1.0 \ 0.35],$$

$$t_k = [100 \ 138 \ 150 \ 159 \ 280 \ 290], \quad \mu_0 = 3.2.$$

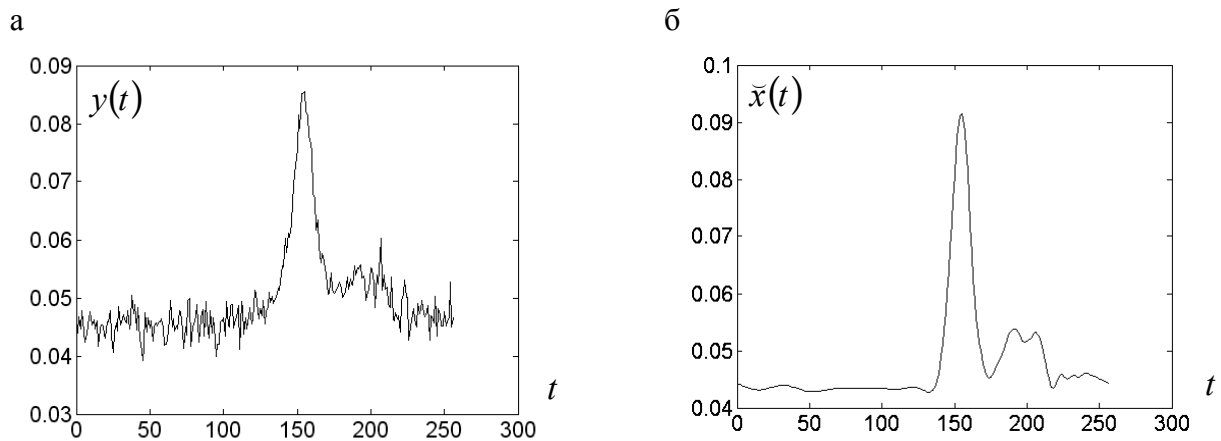
Приборная функция принята равной  $H(t) = \exp\left(-\frac{t^2}{2\mu_0^2}\right)$ ,  $\hat{R}(\omega) \equiv 1$ , и коэффициенты  $\tilde{\alpha}_0(n)$  и  $\tilde{\beta}_0(n)$  определялись по формулам (18\*).

Результаты приведены на рис. 4. Как следует из рис. 4, г, определенный положительный эффект дает усиление верхних частот умножением  $\hat{G}(\omega)$  на полином  $1 + \omega^2$ .

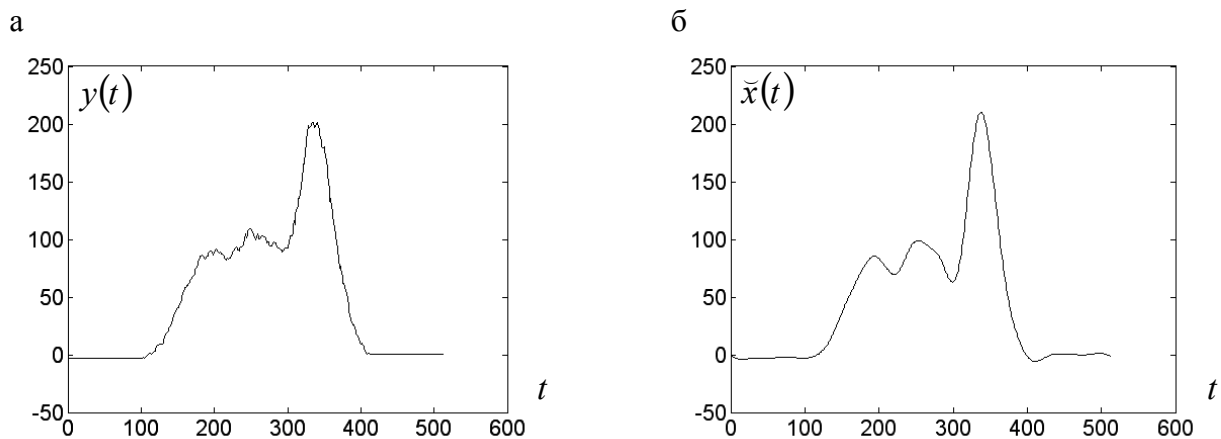
На рис. 5–8 приведены примеры обработки реальных сигналов времяпролетного масс-спектрометра, спектрометра Мессбауэра, электронного спектрометра и секвенатора НАНОФОР 3. Приборные функции в первых двух случаях определялись по форме одиночного пика этих приборов, во втором и третьем случаях выбиралась в виде гауссовой кривой. Фильтр  $\hat{G}(\omega)$  строился аналогично предыдущему примеру.

**Дифференцирование.** Операция дифференцирования применяется в тех случаях, когда необходимо оценить, например, скорость или ускорение движения, выделить изменения, происходящие в наблюдаемом сигнале, локализовать скачки в уровне сигнала и оценить их величину и др. Дифференцирование в присутствии шумов приводит к значительному увеличению их интенсивности и снижению отношения сигнал/шум, особенно при вычислении производных второго и более высоких порядков. Предлагаемый метод, совмещающий операцию дифференцирования и вейвлетного подавления шума, позволяет в несколько раз уменьшить влияние шума на оценку производной.

Сравним этот метод с существующим, например использующим дифференцирующие фильтры второго порядка Савицкого—Голэя (СГ) [5], на модельном сигнале рис. 4, а (см. рис. 9).



**Рис. 6.** Мессбауэровский спектр продуктов реакции образования SnS (олово—сера). а — наблюдаемый сигнал; б — оценка сигнала, полученная методом обратной свертки с усилением верхних частот с помощью полинома  $1+10\omega^2$



**Рис. 7.** Фотоэлектронный спектр. а — наблюдаемый сигнал; б — оценка сигнала, полученная методом обратной свертки с усилением верхних частот с помощью полинома  $1+150\omega^2$

Ошибка дифференцирования определялась по величине СКО от истинного значения производной  $x''$ , вычисленной по известным правилам дифференцирования. Если при отсутствии шума (рис. 9, б) этой ошибкой можно пренебречь ( $RMSE(x'' - \tilde{x}'') \approx 10^{-12}$ ), то в присутствии шума (рис. 9, в) метод СГ приводит к величине СКО  $\sim 28\%$ . Более того, производная от пика минимальной интенсивности (первый пик) оказалась полностью скрытой в шумах. Предлагаемый метод (рис. 9, г) приводит к величине СКО, равной  $\sim 8\%$ , и к четкому выделению производной минимального пика на фоне шумов.

Вычисление коэффициентов  $\tilde{\alpha}_0(n)$  и  $\tilde{\beta}_0(n)$

в этом случае производилось по формулам (19\*).

**(Карунена—Лозва)-подобные разложения (декорреляция).** В формуле (14) положим  $\hat{R}(\omega) \equiv 1$  и  $\beta = 3.5 \cdot 10^{-8}$ . В качестве модельного сигнала  $x(t)$  выберем белый гауссовский шум,  $N = 2^{10}$  и фильтр

$$\hat{H}(\omega) = \frac{1}{\sqrt{1 + \gamma^2 \omega^2}}$$

при  $\gamma^2 = 10$ .

Тогда случайный процесс на выходе фильтра

будет описываться экспоненциальной корреляционной функцией  $K(\tau) = \exp\left\{-\left|\frac{\tau}{\gamma}\right|\right\}$ . Коэффициенты

$\tilde{\alpha}_0(n)$  и  $\tilde{\beta}_0(n)$  вычисляются по формулам (19\*). Результаты обработки приведены на рис. 10.

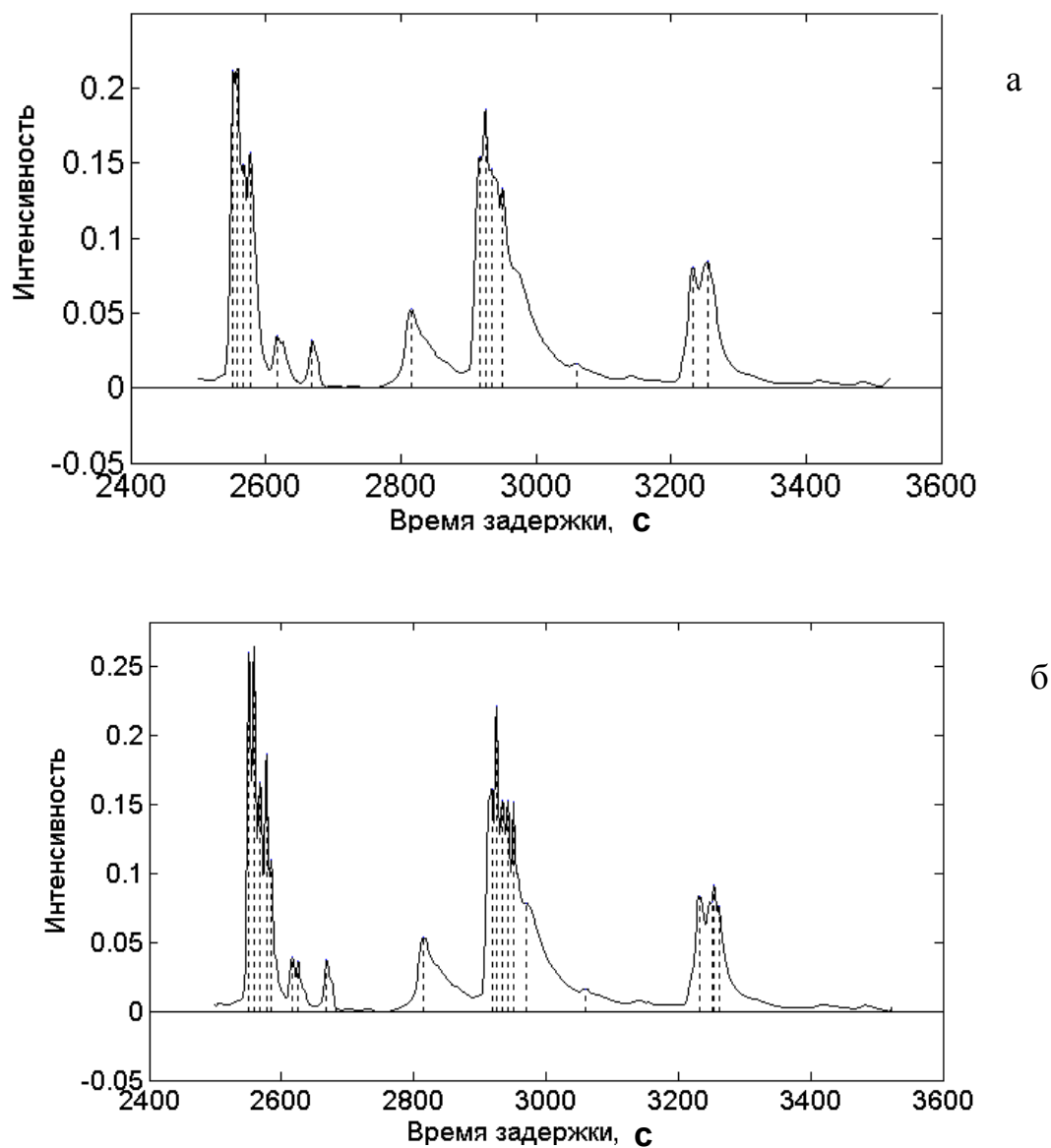
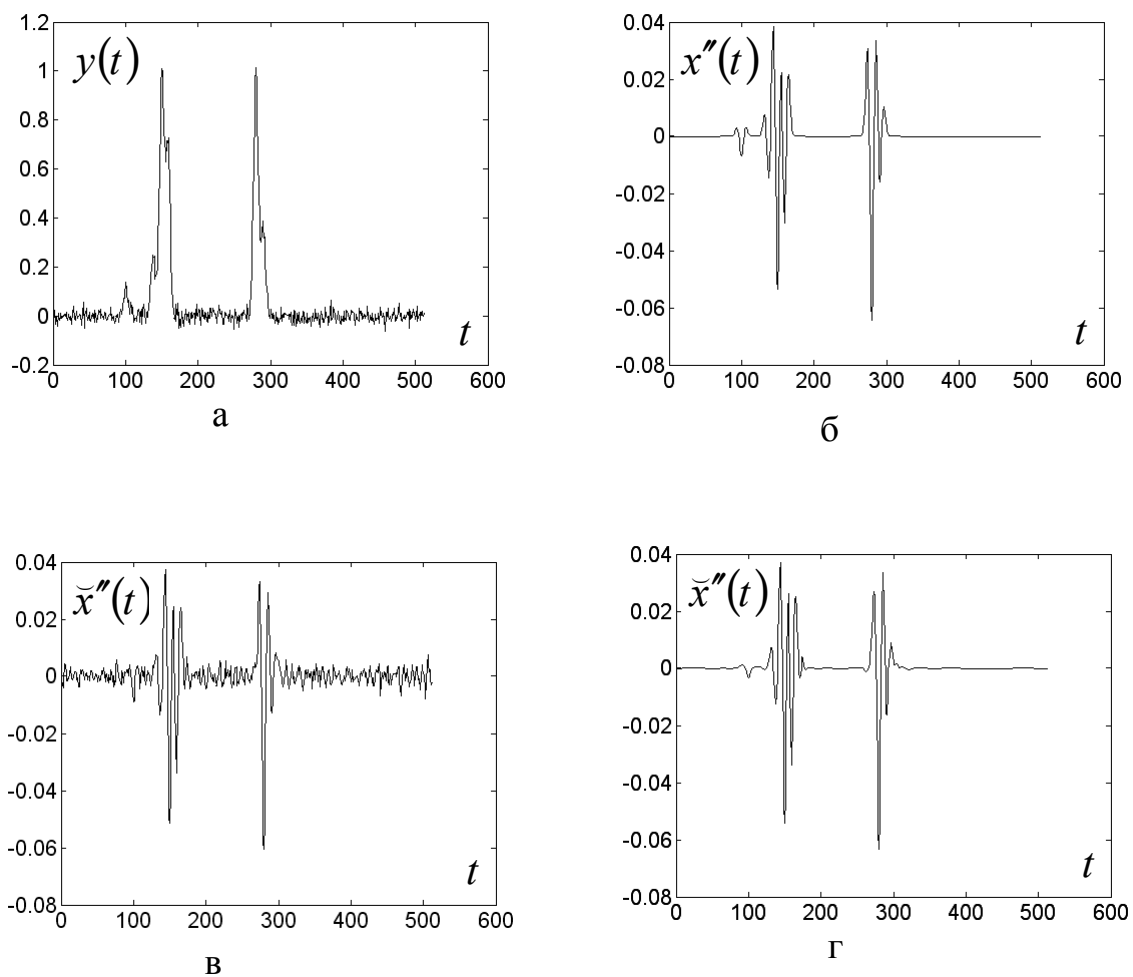


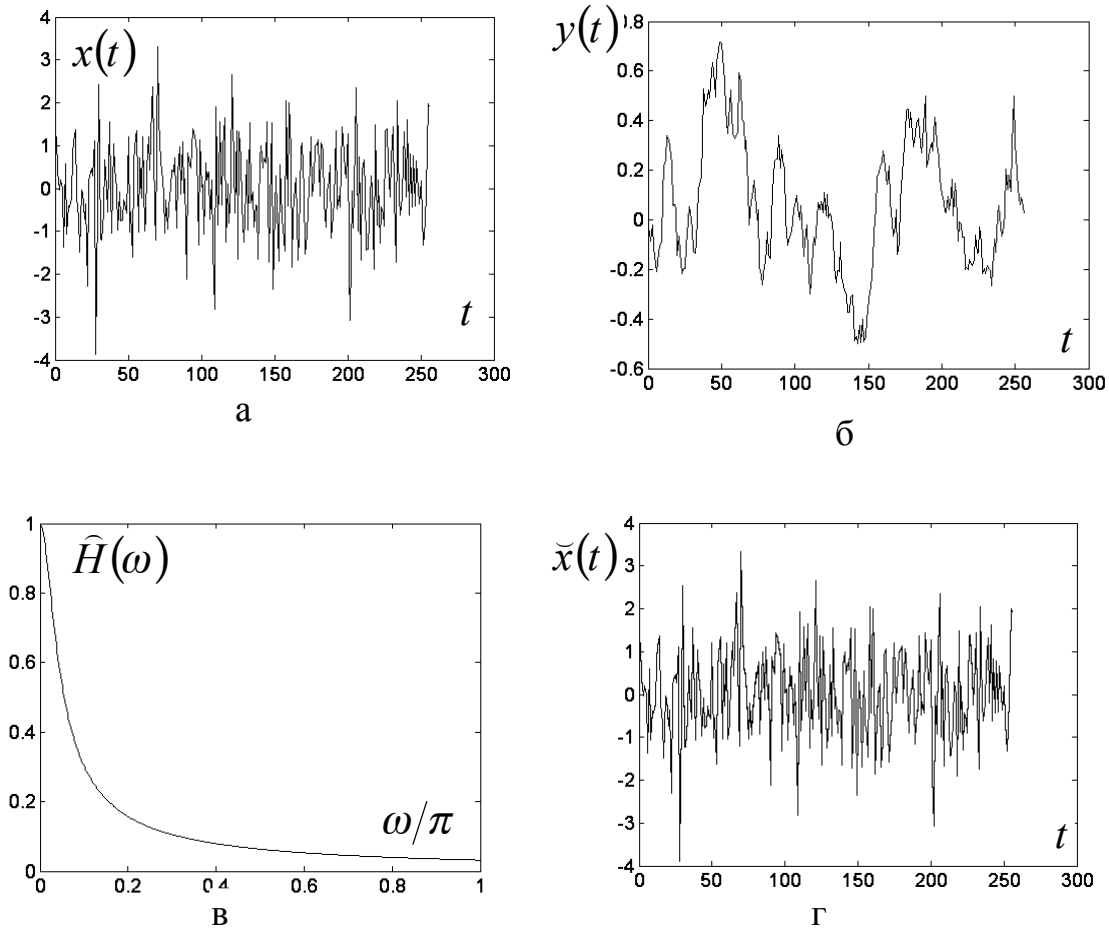
Рис. 8. Фрагмент сигнала ДНК, полученного на приборе НАНОФОР 3.

а — наблюдаемый сигнал; б — оценка сигнала, полученная методом обратной свертки



**Рис. 9.** Пример вычисления второй производной модельного сигнала с использованием девятиточечного третьего порядка фильтра Савицкого—Голэя.

а — наблюдаемый сигнал  $y(t)$  с шумом при отношении сигнал/шум 5 ( $\text{RMSE}(x - y) = 0.1$ ); б — оценка второй производной без шума ( $\text{RMSE}(x'' - \tilde{x}'') = 2 \cdot 10^{-12}$ ,  $\beta = 0.004$ ,  $\hat{R}(\omega) \equiv 1$ ); в — оценка второй производной без подавления шума ( $\text{RMSE}(x'' - \tilde{x}'') = 0.28$ ,  $\beta = 0.04$ ,  $\hat{R}(\omega) \equiv 1$ ); г — оценка второй производной с подавлением шума ( $\text{RMSE}(x'' - \tilde{x}'') = 0.087$ ,  $\beta = 0.04$ ,  $\hat{R}(\omega) = 1$ )



**Рис. 10.** Декорреляция (отбеливание) шума.  
 а — фрагмент тест-сигнала ( $N=1024$ ); б — наблюдаемый сигнал  $y(t)$ , в — частотная характеристика системы  $\hat{H}(\omega) = \frac{1}{\sqrt{1+\gamma^2\omega^2}}$ ; г — декоррелированный сигнал ( $\text{RMSE}(x-\tilde{x})=0.028$ )

### СПИСОК ЛИТЕРАТУРЫ

1. Новиков Л.В. Модифицированные вейвлеты в обработке данных аналитических приборов. I. Основы теории // Научное приборостроение. 2006. Т. 16, № 1. С. 3–15.
2. Добеши И. Десять лекций по вейвлетам. М.: РХД, 2002. 464 с.
3. Donoho D., Johnstone I. Ideal spatial adaptation via wavelet shrinkage // Biometrika. 1994. V. 81. P. 425–455.
4. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1979. 288 с.
5. Savitzky A., Golay M.J.E. Smoothing and differentiation of data by simplified least squares procedures // Analytical Chemistry. 1964. V. 36, N 8. P. 1627–1639.
6. Френкс Л.Е. Теория сигналов. М.: Советское радио, 1974. 343 с.
7. ([www-dsp.rice.edu/software/rwt.shtml](http://www-dsp.rice.edu/software/rwt.shtml)); ([www-stat.stanford.edu/~wavelab](http://www-stat.stanford.edu/~wavelab)).
8. Mallat S. A wavelet tour of signal processing. San Diego: Academic Press, 1999. 637 p.

*Институт аналитического приборостроения РАН,  
 Санкт-Петербург*

Материал поступил в редакцию 7.07.2005.

**MODIFIED WAVELETS IN DATA PROCESSING  
FOR ANALYTICAL INSTRUMENTS.  
II. DATA PROCESSING ALGORITHMS**

**L. V. Novikov**

*Institute for Analytical Instrumentation RAS, Saint-Petersburg*

The paper describes data processing algorithms using modified wavelets. The fast computational algorithms are implemented with filter banks. The application is illustrated by the examples of signal processing for various analytical instruments. It is shown that the proposed approach combines wavelet based noise rejection and deconvolution or other signal transformations, provides several times higher sensitivity and resolution of the instrument.