

УДК 519.2:534.83

ВЕЙВЛЕТ-ФИЛЬТРАЦИЯ ШУМОВ РАЗЛИЧНОЙ СТАТИСТИЧЕСКОЙ ПРИРОДЫ

^{1,2}Воскобойников Ю.Е., ²Крысов Д.А.¹ФГБОУ ВО «Новосибирский государственный архитектурно-строительный университет (Сибстрин)», Новосибирск, e-mail: voscob@mail.ru;²ФГБОУ ВО «Новосибирский государственный технический университет», Новосибирск, e-mail: voscob@mail.ru, tamahouk@sibnet.ru

Большинство используемых алгоритмов вейвлет-фильтрации носят пороговый характер: коэффициент разложения, меньший по абсолютной величине некоторой пороговой величины, приравнивается к нулю. В противном случае коэффициент подвергается некоторому (чаще всего нелинейному) преобразованию, которое задается применяемой пороговой функцией. В настоящее время используются как однопараметрические, так и двухпараметрические пороговые функции. При этом ошибка вейвлет-фильтрации зависит как от вида пороговой функции и ее параметров, так и статистической природы фильтруемого шума. Однако в опубликованных исследованиях, связанных с выбором как пороговой функции, так и величины порога, рассматривается только нормальное распределение погрешностей измерений. На практике же часто встречаются погрешности измерений, распределения которых отличаются от нормального. Примерами может служить равномерное распределение, распределение Пуассона и другие. Поэтому возникает ряд вопросов, связанных как с выбором пороговых функций, так и ее параметров при других (отличных от нормального) распределениях погрешностей измерений. В данной работе выполнен ряд исследований, позволяющих дать ответы на эти вопросы, особенно связанные с оценением оптимальных пороговых величин (минимизирующих среднеквадратическую ошибку фильтрации) при трех распределениях погрешностей измерений. По результатам этих исследований даны практические рекомендации.

Ключевые слова: вейвлет-фильтрация, пороговые функции, распределение погрешностей измерений, выбор наилучшей пороговой функции, оценивание оптимальных пороговых величин

WAVELET-FILTRATION OF NOISE DIFFERENT STATISTICAL NATURE

^{1,2}Voskoboynikov Yu.E., ²Krysov D.A.¹The Novosibirsk State University of Architecture and Civil Engineering (Sibstrin), Novosibirsk, e-mail: voscob@mail.ru;²The Novosibirsk State Technical University, Novosibirsk, e-mail: voscob@mail.ru, tamahouk@sibnet.ru

The majority of the used wavelet-filtration algorithms have threshold character: decomposition coefficients smaller on an absolute value of some threshold size are equated to zero. Otherwise the coefficient is exposed to some transformation (most often nonlinear). The applied threshold function is set the transformation. Currently both single-parametrical and two-parametrical threshold functions are used. At the same time, the error of wavelet-filtration depends as on a type of threshold function and its parameters, and the statistical nature of the filtered noise. However, in the published researches, closed with the choice of both threshold function, and size of thresholds, only measurement's normal distribution of errors is considered. In practice, errors of measurements, which distributions differs from normal, often meet. Uniform distribution, Poisson's distribution and others can be examples. Therefore there is a number of questions, closed as with the choice of threshold functions, as her parameters at others measurement's distributions of errors (other than normal). In this paper, a number of research have been carried out that make it possible to provide answers to these questions. Especially questions related to the estimation of optimal threshold values (minimizing the root-mean-square filtering error) for three distributions of measurement errors. Based on the results of these researches practical recommendations are given.

Keywords: wavelet filtration, threshold functions, distributions of measurement errors, choice of the best threshold function, estimation of optimum threshold sizes

Напомним, что качество вейвлет-фильтрации зашумленного сигнала в основном определяется алгоритмами обработки коэффициентов вейвлет-разложения. Большинство применяемых на практике алгоритмов носят *пороговый характер*: коэффициент разложения меньший по абсолютной величине некоторой пороговой величины зануляется, в противном случае коэффициент сохраняется или подвергается некоторому (в общем случае нелинейному) преобразованию [1–3]. Распространение на практике получили как *однопараметрические поро-*

говые функции, зависящие от одного «управляющего» параметра, называемого *порогом* (например, часто используемые «жесткая» и «мягкая» функции), так и двухпараметрические *пороговые функции*, в определенной степени свободные от недостатков, присущих однопараметрическим функциям [4].

Опубликованные исследования по выбору вида пороговой функции и величины порога выполнялись при предположении, что шум измерения имеет нормальное распределение. Это предположение является традиционным практически для всех пу-

бликаций (как отечественных, так и зарубежных), посвященных пороговым алгоритмам вейвлет-фильтрации. В работе [5] проведены комплексные исследования влияния четырех плотностей распределений (нормальное, равномерное, пуассоновское, импульсное) на ошибку фильтрации одномерных дискретных сигналов при различных уровнях погрешностей измерений. На основе анализа введенного в работе коэффициента эффективности были выбраны оптимальные (по минимуму среднеквадратической ошибки фильтрации) пороговые функции для каждого вида плотности распределения шума измерения. Однако эти результаты были получены на основе вычислительных экспериментов, когда задавался «точный» сигнал, т.е. не искаженный погрешностями измерений. Поэтому открытым остается вопрос: каким образом (или, иначе, каким алгоритмом) выбрать наилучшее значение пороговых величин этих оптимальных пороговых функций? Неудачный выбор величины порога может свести на нет преимущества той или иной пороговой функции. Ответ на этот важный для практики вопрос дается в настоящей работе.

Пороговые алгоритмы вейвлет-фильтрации сигналов

Напомним, что алгоритмы вейвлет-фильтрации основаны на *многомасштабном* (multiresolution) представлении функции $f(x)$ в базисе вейвлет-функции [1–3]:

$$\hat{f}(x) = \sum_k a_{j_0+J,k} \Phi_{J,k}(x) + \sum_j \sum_k d_{j,k} \Psi_{j,k}(x).$$

Функции $\{\Phi_{j,k}(x)\}$ называют масштабирующими (или отцовскими), а функции – вейвлет-функциями (или материнскими). Коэффициенты разложения $a_{j,k}$ называют

аппроксимирующими, $d_{j,k}$ – детализирующими, и они определяются выражениями

$$a_{j,k} = \int_R f(x) \Phi_{j,k}(x) dx,$$

$$d_{j,k} = \int_R f(x) \Psi_{j,k}(x) dt,$$

где R – интервал определения функции $f(t)$. Переменная j характеризует уровень разложения, и ее часто называют коэффициентом масштаба, j_0 – начальный уровень разложения, J – максимальный уровень разложения, а переменная k – временной сдвиг той или иной базисной функции. Системы функций $\{\Phi_{j,k}(x)\}$, $\{\Psi_{j,k}(x)\}$ составляют ортогональный базис пространств вейвлет-функций (подробнее см. [1–3]). Обозначим коэффициенты разложения зашумленного сигнала $\tilde{f}(x)$ как $\tilde{a}_{j,k}$, $\tilde{d}_{j,k}$. Заметим, что относительные погрешности аппроксимирующих коэффициентов на порядок и более меньше погрешностей коэффициентов $\tilde{d}_{j,k}$ (см. [4]). Поэтому на практике обработке подвергаются только детализирующие коэффициенты $\tilde{d}_{j,k}$ с использованием различных пороговых функций (например, [1, 4, 6]). Суть такой обработки заключается в том, что коэффициенты $\tilde{d}_{j,k}$ меньшие по модулю задаваемой пороговой величины принимаются равными нулю, в противном случае эти коэффициенты подвергаются преобразованию, чаще всего нелинейному (подробнее см. [7, 8]).

В работе [5] показано, что для рассматриваемых здесь трех распределений погрешностей измерений (нормальное, равномерное, пуассоновское) оптимальной функцией является двухпараметрическая пороговая функция $T_{HEB}(\tilde{d}, \lambda_1, \lambda_2)$ вида

$$T_{HEB}(\tilde{d}, \lambda_1, \lambda_2) = \begin{cases} 0, & \text{если } |\tilde{d}| \leq \lambda_1; \\ \frac{\tilde{d}}{e-1} \cdot \left(e^{\frac{|\tilde{d}|-\lambda_1}{\lambda_2-\lambda_1}} - 1 \right), & \text{если } \lambda_1 < |\tilde{d}| \leq \lambda_2, \\ \tilde{d}, & \text{если } |\tilde{d}| > \lambda_2, \end{cases} \quad (1)$$

зависящая от двух пороговых величин λ_1, λ_2 (имеющая средний коэффициент оптимальности $\bar{k}_{HEB} = 1$). Крайне мало этой функции проигрывает (средний коэффициент оптимальности $\bar{k}_{HYP} = 0,990$) однопараметрическая пороговая функция $T_{HYP}(\tilde{d}, \lambda)$ вида

$$T_{HYP}(\tilde{d}, \lambda) = \begin{cases} \text{sign}(\tilde{d}) \cdot \sqrt{\tilde{d}^2 - \lambda^2}, & \text{если } |\tilde{d}| \geq \lambda, \\ 0, & \text{если } |\tilde{d}| < \lambda. \end{cases} \quad (2)$$

Учитывая несущественный проигрыш по коэффициенту оптимальности и тот факт, что оценивание одного оптимального порога существенно проще, чем двух (подробнее см. [4]), в дальнейшем будем рассматривать только функцию (2). На рис. 1 показаны графики функции (1) (сплошная кривая) и функции (2) (точечная кривая), которые хорошо иллюстрируют пороговый характер обработки коэффициентов разложения ($\lambda_1 = 1$, $\lambda_2 = 2$). Очевидно, что пороговые величины существенно влияют на точность вейвлет-фильтрации и поэтому перейдем к задаче оценивания оптимального порога λ_{opt} для функции (2).

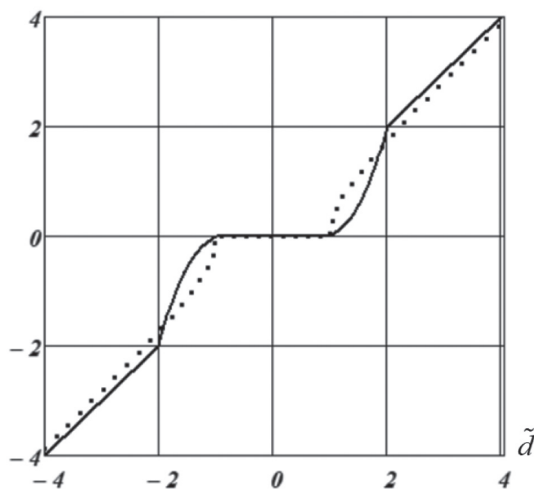


Рис. 1. Графики пороговых функций

Оптимальным порогом будем называть величину λ_{opt} , минимизирующую ошибку фильтрации, определяемую выражением

$$\Delta(\lambda) = \|\hat{f}_\lambda - f\|^2, \quad (3)$$

где $\hat{f}_\lambda$ – вектор, составленный из значений отфильтрованного сигнала, f – вектор, составленный из значений точного (не зашумленного сигнала) сигнала, $\|\cdot\|$ – евклидова норма вектора. Вычислить λ_{opt} из условия минимума (3) невозможно из-за незнания значений «точного» сигнала (вектор f). Поэтому на практике используют различные алгоритмы выбора λ , позволяющие лучше или хуже оценить λ_{opt} . Подробный обзор и сравнения наиболее распространенных на практике алгоритмов выбора для нормального распределения погрешностей измерений приведен в публикациях [4]. Было показано, что наиболее точно оценивает λ_{opt} алгоритм, построенный на основе критерия оптимальности. Попробуем использовать этот алгоритм выбора и в случае распреде-

лений, отличных от нормального. Кратко изложим суть этого алгоритма (подробнее см. [4, 9]).

Для обеспечения асимптотической оптимальности пороговой величины введем на каждом j -м уровне разложения «уровне-зависимую» пороговую величину λ_j в виде

$$\lambda_j = \beta \cdot \sigma \cdot \sqrt{2 \cdot \ln(N_j)}, \quad (4)$$

где N_j – количество обрабатываемых коэффициентов на соответствующем уровне разложения, σ^2 – дисперсия погрешностей измерений. На рис. 2 показаны типичные зависимости относительной ошибки филь-

трации $\frac{\|\hat{f}_\beta - f\|}{\|f\|}$ от параметра β для трех

распределений (нормальное – сплошная кривая, равномерное – точечная, пуассоновское – штриховая кривая). Видно наличие оптимального значения β_{opt} , которое минимизирует ошибку фильтрации. При этом для каждого распределения погрешностей имеется «свое» значение β_{opt} .

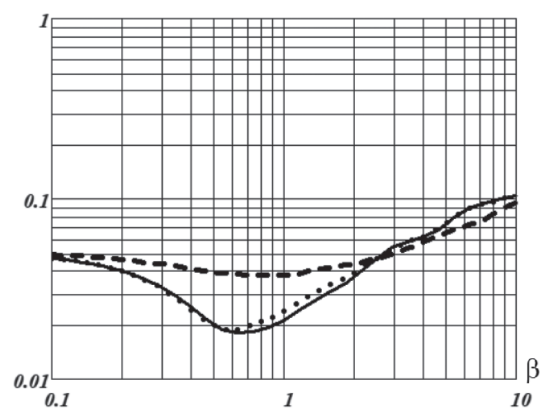


Рис. 2. Зависимость относительной ошибки фильтрации от параметра β

Точное значение β_{opt} невозможно определить из-за незнания «точного» сигнала, и поэтому рассмотрим алгоритм оценивания оптимального коэффициента β_{opt} из условия минимума (3). Определим статистику

$$\rho_w(\beta) = \frac{1}{\sigma^2} \cdot \sum_{i=1}^N \tilde{f}_i \cdot e_i(\beta),$$

где $e(\beta) = \tilde{f} - \hat{f}_\beta$ – вектор невязки, $\hat{f}_\beta$ – результат вейвлет-фильтрации с параметром β . В качестве оценки для β_{opt} принимается величина β_w для которой выполняется условие

$$\vartheta_{\gamma/2, N} \leq \rho_w(\beta_w) \leq \vartheta_{1-\gamma/2, N},$$

где $\vartheta_{\gamma/2, N}, \vartheta_{1-\gamma/2, N}$ – квантили χ^2 – распределения с числом степеней свободы

N уровней $\gamma/2, 1-\gamma/2$ соответственно, N – число значений фильтруемой функции, $\gamma = 0,05$ – вероятность ошибки первого рода при проверке статистической гипотезы об оптимальности параметра β (подробнее см. [4, 9]). Подставляя оценку β_w в (4), получаем оценку $\lambda_j^w = \beta_w \cdot \sigma \cdot \sqrt{2 \cdot \ln(N_j)}$ для оптимальной пороговой величины на j -м уровне разложения. Напомним, что ранее исследования такого алгоритма оценивания оптимальных порогов были выполнены только для погрешностей, распределенных по нормальному закону [4, 9].

Ответим на вопрос о точности оценивания этим алгоритмом оптимальных порогов в случае погрешностей, не подчиняющихся нормальному распределению на основе анализа результатов проведенных вычислительных экспериментов.

Результаты исследования и их обсуждение

В качестве тестового сигнала в вычислительных экспериментах был принят «контрастный» сигнал с резким изменением амплитуд (искаженные значения показаны на рис. 3). Спектр сигнала является широкополосным, что существенно усложняет его фильтрацию из-за совпадений высокочастотных составляющих сигнала и погрешностей измерений. Точные значения (количество отсчетов $N = 2048$) этого сигнала искажались шумами измерений, имеющими следующие плотности распределений: нормальное распределение; равномерное распределение; пуассоновское распределение. Дисперсии этих распределений определялись таким образом, чтобы обеспечить задаваемый относитель-

ный уровень шума, определяемый соотношением $\delta_\eta = \frac{\|\tilde{f} - f\|}{\|f\|}$ (в экспериментах он

был равен 0,05; 0,10; 0,15). На рис. 3, а, показаны значения сигнала, искаженные равномерным шумом, а на рис. 3, б – искажение пуассоновским шумом (относительный уровень шума $\delta_\eta = 0,10$).

Далее, в качестве характеристики точности оценивания оптимальной величины β_{opt} введем коэффициент эффективности E_w , определяемый соотношением [4]:

$$E_w = \frac{\|\hat{f}_{\beta_{opt}} - f\|}{\|\hat{f}_{\beta_w} - f\|}. \text{ Очевидно, что } 0 < E_w \leq 1$$

и чем больше E_w отклоняется от 1 к нулю, тем больше проигрыш по точности алгоритма вейвлет-фильтрации с параметром β_w . Величина E_w является случайной величиной – меняется в зависимости от вектора погрешностей η , который искажает точные значения сигнала. Поэтому в качестве неслучайной характеристики берем выборочное среднее значение $\bar{E}_w$, которое будет вычисляться по выборке $\{\hat{f}_{\beta_w}^{(l)}\}$ объемом 50, где $\hat{f}_{\beta_w}^{(l)}$ – результат фильтрации l -й реализации зашумленного сигнала.

В таблице приведены средние значения относительной ошибки фильтрации

$$\delta_f = \frac{\|\hat{f}_{\beta_w} - f\|}{\|f\|} \text{ (объем выборки также равен$$

50) и средние значения коэффициента эффективности $\bar{E}_w$ для трех уровней погрешностей, имеющих разные распределения.

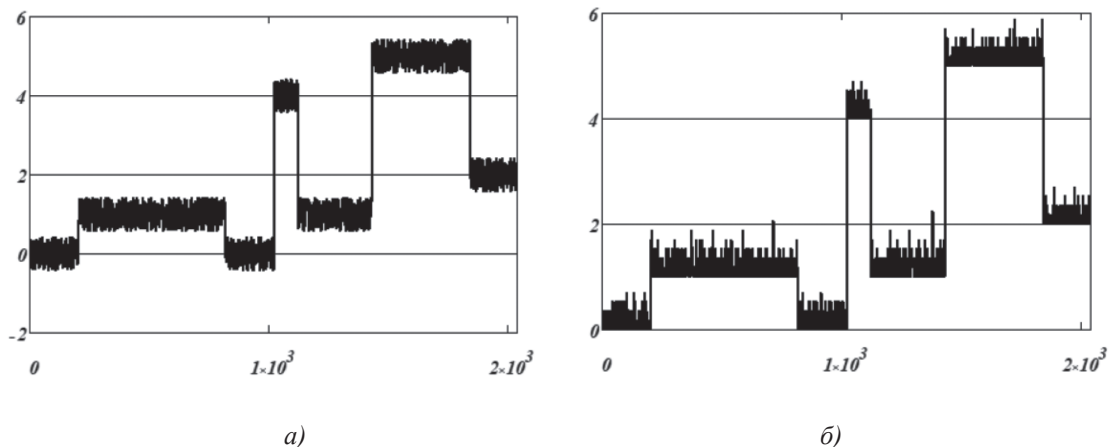


Рис. 3. Зашумленные значения тестового сигнала

Средние значения ошибки фильтрации и коэффициента эффективности

δ_η	Нормальное распределение		Равномерное распределение		Распределение Пуассона	
	$\bar{\delta}_f$	$\bar{E}_w$	$\bar{\delta}_f$	$\bar{E}_w$	$\bar{\delta}_f$	$\bar{E}_w$
0,05	0,020	0,982	0,019	0,984	0,038	0,996
0,10	0,038	0,968	0,037	0,976	0,076	0,993
0,15	0,055	0,961	0,053	0,968	0,114	0,990

Выводы

Анализ результатов, приведенных в таблице, позволяет сделать следующие выводы:

1. Алгоритм выбора пороговой величины, основанный на критерии оптимальности, позволяет с приемлемой точностью оценить β_{opt} не только в случае погрешностей, имеющих нормальное распределение, но и в случае фильтрации погрешностей, имеющих равномерное и пуассоновское распределение (высокое значение $\bar{E}_w \approx 1$ для всех распределений).

2. Самая большая ошибка фильтрации наблюдается для пуассоновского распределения погрешностей (как следствие ненулевого математического ожидания погрешностей). Высокие значения коэффициента эффективности $\bar{E}_w$ объясняются пологим характером кривой относительной ошибки фильтрации в области ее минимума (см. рис. 2). Поэтому для такого распределения погрешностей применение вейвлет-фильтрации оказывается неэффективным, даже при оптимальных значениях пороговых величин.

3. Изложенный подход к выбору пороговой величины достаточно просто реализуется на практике, где в случае неизвестной величины дисперсии σ^2 используется ее оценка, приведенная в [4, с. 66].

Список литературы

1. Mallat S.A. Wavelet tour of signal processing: the sparse way / S. Mallat. – N.Y.: Academic Press, 2008. – 621 p.
2. Kovacevic J. Fourier and wavelet signal processing / J. Kovacevic, V.K. Goyal, M. Vetterli. – Lausanne: EPFL, 2013. – 268 p.
3. Дьяконов В. П. Вейвлеты. От теории к практике. Изд. 2-е, перераб. и доп. / В.П. Дьяконов. – М.: СОЛОН-Р, 2010. – 400 с.
4. Воскобойников Ю.Е. Алгоритмы вейвлет-фильтрации с примерами в пакете MathCAD / Ю.Е. Воскобойников. – Deutschland: Palmarium Academic Publishing, 2016. – 196 с.
5. Воскобойников Ю.Е. Исследование потенциальной точности пороговых алгоритмов вейвлет-фильтрации шумов различной статистической природы / Ю.Е. Воскобойников, Д.А. Крысов // Автоматика и программная инженерия. – 2016. – № 4 (18). – С. 65–76.
6. Ясин А.С. Цифровая вейвлет-фильтрация зашумленных данных: влияние порогового уровня и выбора вейвлета / А.С. Ясин, А.Н. Павлов, А.Е. Храмов // Радиотехника и электроника. – 2016. – Т. 61, № 1. – С. 1–7.
7. Hedaoo P. Wavelet thresholding approach for image denoising / P. Hedaoo, S.S. Godbole // International journal of network security & its applications. – 2011. – vol. 3. – № 4. – P. 16–22.
8. Gupta T. Contemplation of thresholding in wavelet denoising / T. Gupta, P.K. Bhatia // Journal of basic and applied engineering research. – 2014. – vol. 1. – № 8. – P. 83–89.
9. Voskoboinikov Yu.E. Estimating optimum threshold sizes in a wavelet-filtration algorithms / Yu.E. Voskoboinikov, A.V. Gochakov // Optoelectronics, instrumentation and data processing. – 2011. – vol. 47. – № 2. – P. 3–14.