

# РЕДАКТИРОВАНИЕ ШУМА В ИЗОБРАЖЕНИИ НА ОСНОВЕ Пороговой нейронной сети

асп. Аврамов В.В., доц. Петровский Н.А.

Белорусский государственный университет информатики и радиоэлектроники

В работе рассматриваются методы адаптивного шумоподавления на основе пороговых нейронных сетей в вейвлет-домене. Обучение данного типа нейронных сетей, в отличие от классических нейронных сетей, где обучение заключается в подстройке весовых коэффициентов, направлено на подстройку пороговой функции активации, с целью достижения оптимального порога усечения. Лучших результатов, в сравнении с адаптивными субполосными порогами, позволяют достичь методы, основанные на подборе порога для каждого отдельно взятого вейвлет-коэффициента. Предлагается пороговая нейронная сеть с 2D неразделимым параунитарным банком фильтров на основе умножителей кватернионов.

**1. Задача редактирования шума и пороговые методы.** Шумы и артефакты в методах визуализации ухудшают качество изображения, что влияет на качество изображения и, как следствие, на их интерпретацию человеком, а также на точность компьютерных методов в случае медицинской визуализации. Следовательно, редактирование шумов в изображении становится главным требованием для многих практических приложений. В общем случае задача редактирования шумов заключается в получении восстановленной копии полезного сигнала  $\mathbf{x}$  из его искаженной версии  $\mathbf{y} = \mathbf{x} + \mathbf{n}$ , где  $\mathbf{n}$ , например, аддитивный Гауссовый белый шум. Таким образом, целью обработки является получение на выходе системы редактирования шумов  $f$  восстановленного сигнала  $\hat{\mathbf{x}} = f(\mathbf{y})$  как можно более близкого к полезному сигналу  $\mathbf{x}$ , что аналитически может быть сформулировано следующим образом: как минимизация функции потерь  $E(\mathbf{x}, \hat{\mathbf{x}})$  (среднеквадратического отклонения)

$$\min_{\mathbf{x}, \hat{\mathbf{x}} \in \mathbb{R}} E(\mathbf{x}, \hat{\mathbf{x}}) = \min_{\mathbf{x}, \hat{\mathbf{x}} \in \mathbb{R}} \|\mathbf{x} - \hat{\mathbf{x}}\|^2 . \quad (1)$$

При оценке сигнала шума  $\mathbf{n}$  может и должна использоваться любая доступная априорная информация о распределении энергии сигнала и шума, которая может быть получена путем перевода исходного сигнала в другое пространство, на основе линейного ортогонального преобразования  $T$ , как показано в [1] (рисунок 1). Применение такого преобразования для искаженной версии сигнала позволяет изменить исходное распределение сигнала, что обеспечивает важную информацию для осуществления разделения сигнала и шума. Таким образом, полезный сигнал может быть извлечен из его искаженной версии путем обработки области преобразования некоторой пороговой функцией, сохраняющей энергию коэффициентов преобразования  $T(\mathbf{x} + \mathbf{n})$  попадающих в область распределения полезного сигнала, и подавляющей энергию шумовых коэффициентов, не попавших в данную область. На этой концепции и основываются пороговые методы редактирования шумов.

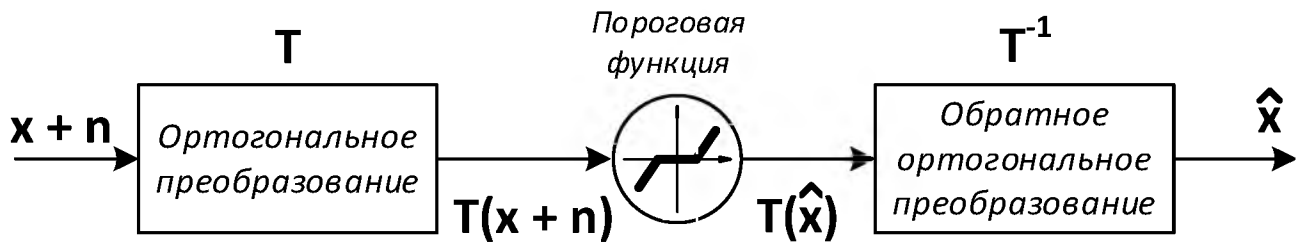


Рисунок 1. Схема порогового редактирования шумов

Классические «жесткая» и «мягкая» пороговые функции (рисунок 2), предложенные в [2], описываются следующими выражениями:

$$h(x, t) = \begin{cases} x, & |x| > t \\ 0, & |x| \leq t \end{cases} \quad (2)$$

$$s(x, t) = \begin{cases} x + t, & x < -t \\ x - t, & x > t \\ 0, & |x| \leq t \end{cases} \quad (3)$$

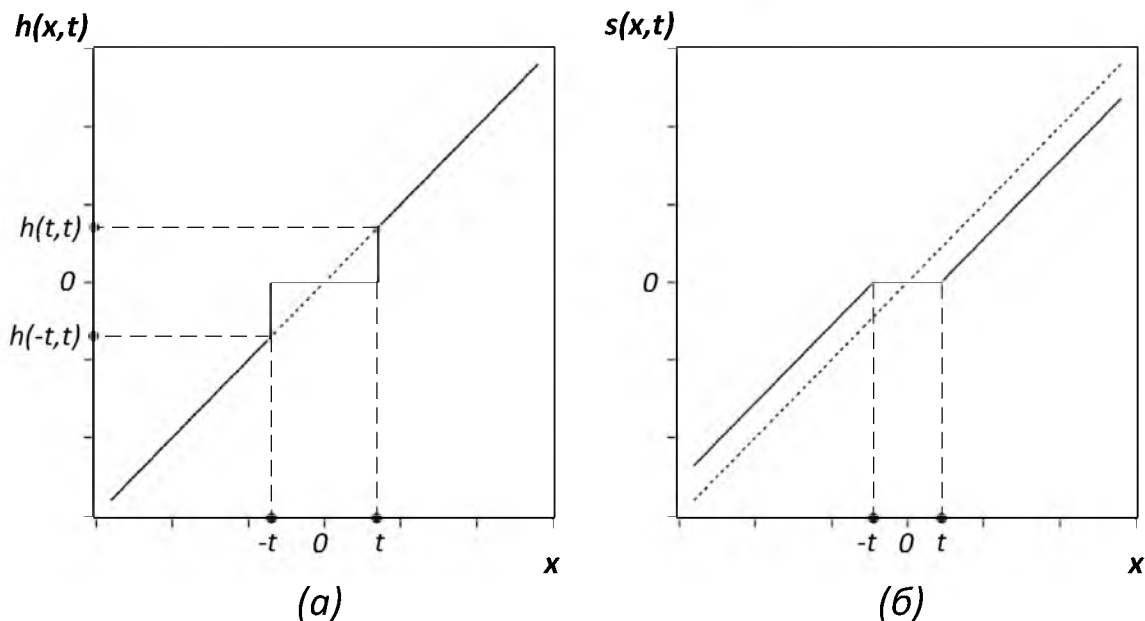


Рисунок 2. Пороговые функции: а) «жесткая»; б) «мягкая»

**2. Структура пороговой нейронной сети.** Концепция, лежащая в основе пороговой нейронной сети (thresholding neural network (TNN)) [3], существенно отличается от классических нейронных сетей. Обучение данного типа нейронных сетей, в отличие от классических нейронных сетей, где функции активации фиксированные, и обучение заключается в подстройке весовых коэффициентов, направлено на подстройку пороговой функции активации, с целью достижения оптимального порога усечения. Термин «нейронная сеть» здесь употребляется потому, что данная структура из базовых элементов классических нейронных сетей (связи, нелинейная функция активации) и является адаптируемой ко входному воздействию. Структура пороговой нейронной сети на основе вейвлет преобразования (WT-TNN) и процесса ее обучения представлена на рисунке 3.

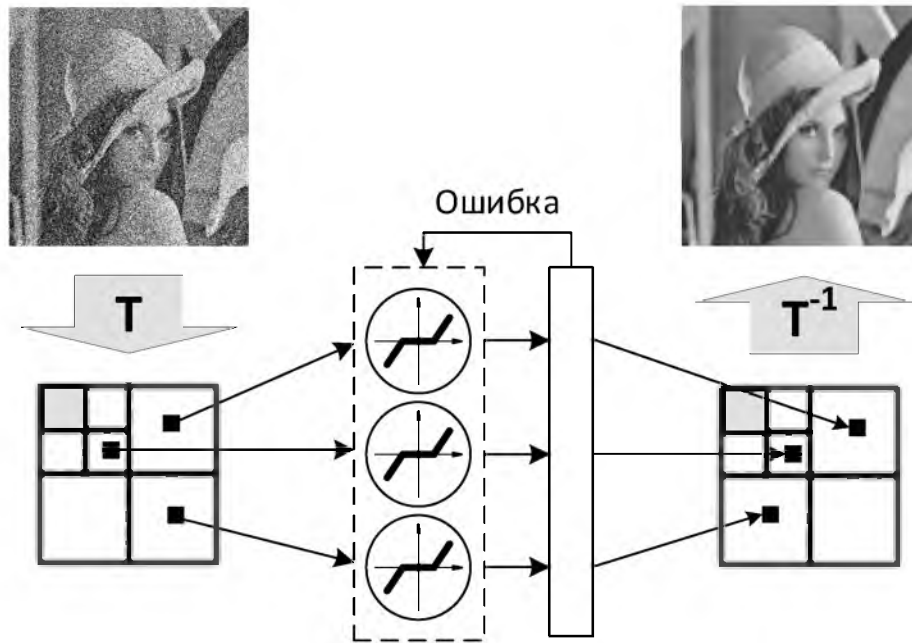


Рисунок 3. Структура пороговой нейронной сети (WT-TNN)

Согласно структуре, показанной на рисунке 3, процесс обучения пороговой нейронной сети заключается в адаптации отдельной пороговой функции под распределение коэффициентов преобразования. Пороговые функции активации используются в качестве нелинейных функций активации нейронной сети. Обратное преобразование применяется для восстановления сигнала из коэффициентов с уменьшенным в пространстве преобразования уровнем шума.

Поскольку градиентные методы оптимизации требуют дифференцируемости функций, необходимым условием которого множество является непрерывность функции на заданном множестве аргументов. Классические пороговые функции, сформулированные в выражениях 1 и 2, не могут быть применены. Поэтому в [3,4,5] были предложены разновидности дифференцируемых версий пороговых функций. Например, функция, показанная в [4], а также применяемая в [5], помимо значения порога  $t$  также имеет 2 параметра  $m$  и  $k$  отвечающих за форму функции и, в зависимости от их значений, приблизить ее к «жесткой» либо «мягкой» пороговым функциям, описанным в выражениях (2), (3) и показанных на рисунке 2:

$$f(x, t, m, k) = \begin{cases} x + (k - 1)t - \frac{kt^m}{2x^{m-1}}, & x > t \\ \frac{k|x|^{m+(2-k)/k}}{2t^{m+2(1-k)/k}} \operatorname{sign}(x), & |x| \leq t \\ x - (k - 1)t + \frac{k(-t)^m}{2x^{m-1}}, & x < -t \end{cases} \quad (4)$$

Таким образом, обучение пороговой нейронной сети с активационной функцией, описанной выражением 4, заключается в поиске оптимальных значений  $t, m, k$ , обеспечивающих наилучшее решение задачи

оптимизации установленной в выражении (1). Тогда процесс подстройки параметров функции на основе градиентного спуска может быть выражен следующими итеративными формулами:

$$\begin{aligned}
 t(i+1) &= t(i) - \alpha \frac{\partial E(x, \hat{x})}{\partial t}, \\
 m(i+1) &= m(i) - \beta \frac{\partial E(x, \hat{x})}{\partial m}, \\
 k(i+1) &= k(i) - \gamma \frac{\partial E(x, \hat{x})}{\partial k},
 \end{aligned}$$

где  $\alpha, \beta, \gamma$  – параметры контролирующие скорость обучения параметров  $t, m, k$  соответственно;  $E(x, \hat{x})$  – функция потерь (выражение 1), определенная в пространстве преобразования.

**3. Редактирование шума с использованием пороговой нейронной сети на базе пакета вейвлет преобразования.** Для конкретного класса сигналов могут быть выбраны соответствующие ортогональные преобразования, позволяющие локализовать энергию сигнала по отношению к шуму. В случае нестационарного сигнала, имеющего сложную структуру и взаимосвязи, в частности, в частотной области, возникает необходимость адаптивного выбора подходящего ортогонального преобразования, позволяющего наилучшим образом разделить полезный сигнал и шум для последующей пороговой обработки с целью снижения уровня последнего.

Вейвлет-преобразование обладает рядом уникальных свойств, одним из которых является разреженность в пространстве преобразования. Это означает, что большая часть энергии исходного сигнала в вейвлет-домене сконцентрирована в некотором ограниченном наборе коэффициентов. Данное свойство преобразования предоставляет широкий спектр возможностей для компрессии изображений, путем отбора коэффициентов с наибольшей энергией, а также для снижения уровня шумов. Поскольку вейвлет-преобразование зашумленного сигнала представляет собой линейную комбинацию полезного сигнала и шума, энергия шума в вейвлет-домене равномерно распределяется среди всех коэффициентов преобразования. Следовательно, уровень шумов может быть существенно снижен путем обработки пороговой функцией с подходящим порогом, оставляя при этом некоторый набор наиболее энергетически значимых коэффициентов, и удаляя энергетически незначимые коэффициенты, несущие большую часть шума в пространстве преобразования. Значение универсального порога может выбираться для всех коэффициентов вейвлет-преобразования зашумленного изображения, как сформулировано в методе VisuShrink [6], согласно выражению

$$t = \sigma \sqrt{2 \log(N)/N} \tag{5}$$

где  $\sigma$  – среднеквадратическое отклонение коэффициентов;  $N$  – количество коэффициентов.

Адаптивный выбор порога для каждой из субполос вейвлет-преобразования [7] позволяет более точно восстанавливать полезный сигнал из его искаженной версии, поскольку энергия сигнала и шума, как правило, распределяется неравномерно среди субполос. Лучших результатов, в сравнении с адаптивными субполосными, позволяют достичь методы, основанные на подборе порога для каждого отдельно взятого вейвлет-коэффициента [8]. Кроме того, моделирование зависимостей между вейвлет-коэффициентами в субполосах [9], позволяет более точно подстроить значение порога, и, тем самым, повысить производительность упомянутых ранее подходов.

Для оценки эффективности описанных ранее пороговых методов редактирования шумов, в качестве источника искажения был выбран гауссовский шум с различными уровнями мощности (дисперсии)  $\sigma$ , а в качестве полезного сигнала стандартное тестовое изображение «Лена» размером  $512 \times 512$  пикселей. Полученные в [5] результаты оценки качества на основе метрики PSNR сведены в таблицу 1

Таблица 1

Результаты оценки PSNR, дБ				
$\sigma$	Зашумленное изображение	Мягкая пороговая функция	Жесткая пороговая функция	WT-TNN
10	28,16	28,10	30,49	33,82
20	22,14	25,75	27,49	30,56
30	18,62	24,61	25,93	28,77

Согласно результатам, представленным в таблице 1, подход с применением пороговой нейронной сети WT-TNN на основе функции приведенной в (4) позволяет достичь лучших результатов в сравнении с классическими «жесткой» и «мягкой» пороговыми функциями, сформулированных в выражениях (2) и (3) соответственно, для которых значение порога выбиралось в соответствии с (5).

#### 4. Неразделимые многомерные банки фильтров в пороговых нейронных сетях.

Многомерный сигнал – изображение является неразделимым, поэтому многомерное преобразование в, например, WT-TNN, представлялось в виде прямого произведения одномерных функций координат, что не соответствует природе обрабатываемого сигнала (изображения). В [10] был предложен 2D неразделимый параунитарный банк фильтров (2D Q-ПУБФ) на основе умножителя кватернионов для числа каналов 4 (структура 16in-16out) и 8 (структура 64in-64out), характеризующийся высокими показателями эффективности кодирования, возможностью работать с замещением входных данных и высокой скоростью обработки. Двумерная декомпозиция позволяет получить более достоверное частотное отображение двумерного сигнала, и, тем самым, предоставляет широкий спектр возможностей для улучшения характеристик систем редактирования шумов на изображениях.

В таблице 2 представлены первые экспериментальные результаты применения пороговой нейронной сети на основе преобразования 2D Q-ПУБФ. Условия проведения экспериментов соответствуют таковым для полученных результатов в [5] и приведенных в таблице 1. Искаженное гауссовским шумом ( $\sigma = 30$ ) тестовое изображение, а также его очищенная копия представлены на рисунках 4а и 4б соответственно. Результаты оценки PSNR, представленные в таблице 2, ниже таковых представленных в таблице 1. В дальнейшем предполагается работа по совершенствованию пороговой функции и согласованию ее с процессом обработки изображения банком фильтров 2D-Q-ПУБФ.

Таблица 2

Результаты оценки PSNR и SNR, дБ

$\sigma$	PSNR		SNR	
	Зашумленное изображение	2D-Q-ПУБФ-TNN	Зашумленное изображение	2D-Q-ПУБФ-TNN
10	28,16	32,30	22,46	25,96
20	22,14	29,82	16,48	23,45
30	18,62	28,25	13,03	21,77



Рисунок 4. Тестовое изображение «Лена»: а) зашумленное ( $\sigma = 30$ ); б) очищенное

#### Литература

1. Krim H., Tucker D., Mallat S., Donoho D. On Denoising and Best Signal Representation // IEEE Transactions on Information Theory. – 1999. – V. 45. – P. 2225–2238.
2. D.L. Donoho, I.M. Johnstone. Ideal spatial adaptation by wavelet shrinkage // Biometrika. — 1994. – V. 81. – P. 425–455.
3. Zhang X.-P., Desai M. Nonlinear Adaptive Noise Suppression Based On Wavelet Transform // Proceedings of the ICASSP. – 1998. – V. 3. – P. 1589–1592.
4. Nasri M., Nezamabadi-pour H. Image denoising in the wavelet domain using a new adaptive thresholding function // Elsevier Journal of Neurocomputing. – 2009. – V. 72. – P. 1012–1025.
5. Bhutada G.G., Anand R.S., Saxena S.C. Image enhancement by wavelet-based thresholding neural network with adaptive learning rate // IET Image Processing. – 2011. – V. 5, No. 7. – P. 573–582.

6. Donoho D. L. De-noising by soft-thresholding // IEEE Transactions on Information Theory. – 1995. – V. 41. – P. 613–627.
7. Donoho D.L., Johnstone I.M. Adapting to unknown smoothness via wavelet shrinkage // Journal of the American Statistical Association. – 1995. – V. 90, No. 432. – P. 1200–1224.
8. Mihcak M.K., Kozintsev I., Ramchandran K., Moulin P. Low-complexity image denoising based on statistical modeling of wavelet coefficients // IEEE Signal Processing Letters. – 1999. – V. 6. – P. 300–303.
9. Chen G.Y., Bui T.D., Krzyzak A. Image denoising with neighbour dependency and customized wavelet and threshold // Pattern Recognition. – 2005. – V. 38. – P. 115–124.
10. Nick A. Petrovsky, Eugene V. Rybenkov, Alexander. A. Petrovsky. Two-dimensional non-separable quaternionic paraunitary filter banks // IEEE Int. conf. on Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Sep 2018, Poznan, Poland. – P. 120–125.