УДК 004.932.4

НЕЙРОСЕТЕВАЯ ВИНЕРОВСКАЯ ФИЛЬТРАЦИЯ ИЗОБРАЖЕНИЙ В ОБЛАСТИ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Алимагадов К.А., *аспирант Национального исследовательского университета «МИЭТ», e-mail: alimagadovk@yandex.ru*
Умняшкин С.В., *д.ф-м.н., профессор Национального исследовательского университета «МИЭТ», e-mail: vrinf@miee.ru*

NEURAL NETWORK WIENER FILTERING OF IMAGES IN THE DOMAIN OF DISCRETE WAVELET TRANSFORM

Alimagadov K.A., Umnyashkin S.V.

The research is devoted to quality improvement of additive white Gaussian noise suppression in images. The architecture of neural network for image denoising based on Wiener filtering in the domain of discrete wavelet transform (DWT) has been developed. The key point of the processing scheme is the adaptation of image local filtering technique based on contour analysis in the domain of DWT. The number of trainable parameters of the proposed network is an order of magnitude smaller as compared with the modern machine learning models used to solve the same task.
The neural network was tested on «Berkeley segmentation data set 500». Estimation of filtering quality was based on analysis of peak signal to noise ratio and structure similarity index values. It was found that the developed approach provides higher quality of noise suppression than other considered algorithms based on Wiener filtering.

Key words: noise suppression, image processing, discrete wavelet transform, neural networks, Wiener filter.

Ключевые слова: подавление шумов, обработка изображений, дискретное вейвлет-преобразование, нейронные сети, фильтр Винера.

Исследование посвящено улучшению качества подавления аддитивного белого гауссового шума в изображениях. Разработана архитектура нейронной сети, выполняющей подавление шумов на основе винеровской фильтрации в области дискретного вейвлет-преобразования (ДВП). Ключевым моментом схемы обработки является адаптация способа локальной фильтрации изображения на основе контурного анализа в области ДВП. Предлагаемая архитектура сети имеет на порядок меньшее количество обучаемых параметров, чем у современных моделей машинного обучения, решающих аналогичную задачу.

Для сравнения качества обработки нейронная сеть была протестирована на наборе данных «Berkeley segmentation data set 500». Оценка качества фильтрации выполнялась на основе анализа значений мер пикового отношения сигнал/шум и индекса структурного сходства. Установлено, что разработанный подход обеспечивает более качественное подавление шумов, по сравнению с другими рассмотренными алгоритмами на основе винеровской фильтрации.

Предварительные сведения

Искажения и шумы, возникающие в процессах съёмки и передачи цифровых изображений, затрудняют восприятие фото- и видео данных человеком и могут приводить к некорректной работе систем компьютерного зрения. Подавление шумов способствует улучшению визуального качества изображений, а также снижению доли ошибок распознавания при последующем применении алгоритмов интеллектуального анализа данных [1].

На сегодняшний день наиболее качественные результаты подавления шумов демонстрируют методы обработки на основе нейронных сетей [2-4]. Однако следует отметить, что применение глубоких нейросетевых моделей, как правило, требует бóльших вычислительных ресурсов, по сравнению с классическими алгоритмами фильтрации изображений. Современные нейронные сети содержат сотни тысяч, а иногда и миллионы параметров [2-4]. Это приводит к необходимости сбора значительного объёма обучающих данных, искусственного расширения обучающей выборки за счёт аугментаций и использования регуляризации для того, чтобы избежать переобучения и сохранить способность модели к обобщению [5]. Помимо этого, глубокие нейронные сети являются плохо интерпретируемыми (работают по принципу «чёрного ящика») и подвержены влиянию спектрального сдвига в область низких частот, который приводит к деградации полезной высокочастотной информации при обработке [6, 7], выражающейся в потере резкости контуров на изображении.

Одним из направлений исследований в области нейросетевой фильтрации изображений является разработка алгоритмов, сочетающих преимущества машинного обучения и классических методов цифровой обработки, что позволяет снизить количество параметров моделей и повысить их интерпретируемость [3, 8].

Целью работы является улучшение качества подавления шумов в изображениях и развитие идей, описанных в статье [8].

Предлагаемый подход основан на нейросетевой винеровской фильтрации в базисе дискретного вейвлет-преобразования (ДВП) с учётом контурной информации

и позволяет обеспечить больший прирост метрик качества после обработки, по сравнению с большинством рассмотренных в работе аналогов.

Модель шума

В рамках исследования рассматривается модель белого аддитивного не коррелированного с изображением шума, распределённого по нормальному закону с нулевым математическим ожиданием μ = 0 и среднеквадратическим отклонением (СКО) σ.

Дискретные полутоновые изображения будем представлять в виде матриц, элементы которых определяют яркость пикселей. Искажение изображения в формализованном виде описывается выражением:

$$\mathbf{X} = \mathbf{Y} + \mathbf{N}, \quad (1)$$

где $\mathbf{X} = (x_{m,n})$, $\mathbf{Y} = (y_{m,n})$ и $\mathbf{N} = (\eta_{m,n})$ – матрицы зашумлённого изображения, исходного изображения и шума соответственно.

Дискретное вейвлет-преобразование

Фильтрацию изображений будем выполнять в базисе двумерного дискретного вейвлет-преобразования (ДВП). Вычисление ДВП-спектра входного изображения реализовано с применением биортогональных фильтров 9/7 [9], импульсные характеристики (ИХ) которых задают одномерные свёртки, применяющиеся последовательно вдоль строк, затем вдоль столбцов матрицы изображения.

В процессе обработки выполняется 3-х уровневое ДВП без прореживания откликов ВЧ и НЧ-фильтров, задающих преобразование [9]. При вычислении каждого i-го ($i \in \{1, 2, 3\}$) уровня ДВП получаются четыре набора вейвлет-коэффициентов (саббэндов): один саббэнд, соответствующий низкочастотной составляющей ДВП-спектра, $\mathbf{V}_i^A$, и три вейвлет-саббэнда $\mathbf{V}_i^H$, $\mathbf{V}_i^V$, $\mathbf{V}_i^D$, соответствующие горизонтальной, вертикальной и диагональной высокочастотным составляющим. Фильтрация применяется только к саббэндам $\mathbf{V}_i^H$, $\mathbf{V}_i^V$, $\mathbf{V}_i^D$, а саббэнд $\mathbf{V}_i^A$ используется для вычисления (i + 1)-го уровня ДВП.

Фильтр Винера

Для подавления шума предлагается использовать фильтр Винера, который минимизирует математическое ожидание квадрата ошибки для оценки полезного сигнала [10]. Двумерные ДВП-спектры (т.е. наборы коэффициентов ДВП) изображений $\mathbf{X}$, $\mathbf{Y}$ и шума $\mathbf{N}$ обозначим $\mathbf{V} = (v_{k,l})$, $\mathbf{U} = (u_{k,l})$ и $\mathbf{G} = (g_{k,l})$ соответственно. Из предположения о некоррелированности шума с полезным сигналом следует, что спектр мощности зашумлённого изображения представляет собой сумму спектров мощности исходного изображения и шумовой составляющей. Тогда коэффициенты фильтра Винера вычисляются по формуле [8]

$$w_{k,l} = \frac{\hat{s}_{k,l}}{\hat{s}_{k,l} + \hat{p}_{k,l}}, \quad (2)$$

где $\hat{s}_{k,l}$ и $\hat{p}_{k,l}$ – оценки коэффициентов ДВП-спектров мощности полезного сигнала $s_{k,l} = \mathrm{M}\left[u_{k,l}^2\right]$ и шума $p_{k,l} = \mathrm{M}\left[g_{k,l}^2\right]$ соответственно (обозначение M используется для математического ожидания).

Фильтрация осуществляется путём поэлементного перемножения коэффициентов фильтра Винера $w_{k,l}$ и коэффициентов ДВП $v_{k,l}$ зашумлённого изображения: $\hat{\mathbf{U}} = (\hat{u}_{k,l} = w_{k,l} v_{k,l})$. После этого для получения восстановленного изображения выполняется обратное ДВП (ОДВП): $\hat{\mathbf{Y}} = \text{ОДВП}(\hat{\mathbf{U}})$.

Общая схема метода подавления шума

Предлагаемый алгоритм подавления шумов перед фильтрацией выполняет нечёткую классификацию коэффициентов ДВП-спектра $\mathbf{V}$ на два класса в зависимости от локализации соответствующих базисных функций (двумерных вейвлетов) на изображении: 1) «контурные», соответствующие контурам на изображении, 2) «фоновые», соответствующие фоновым областям. Такая классификация позволяет более гибко подстраивать параметры $\mathbf{W} = (w_{k,l})$ фильтра Винера: необходимые в (2) оценки спектров мощности $\hat{u}_{k,l}^2$ формируются по-разному для каждого из двух классов вейвлет--коэффициентов.

Общая схема разработанного алгоритма подавления шума состоит из следующих шагов. Входными данными являются изображение $\mathbf{X}$ и заданная величина СКО σ шума $\mathbf{N}$.

Шаг 1. Вычисление ДВП-спектра входного изображения: $\mathbf{V} = \text{ДВП}(\mathbf{X})$.

Шаг 2. Оценка дисперсии шума в саббэндах $(\mathbf{V}_i^H, \mathbf{V}_i^V, \mathbf{V}_i^D)_{i=1}^3$ по величине σ.

Шаг 3. Определение степени принадлежности вейвлет-коэффициентов саббэндов $(\mathbf{V}_i^H, \mathbf{V}_i^V, \mathbf{V}_i^D)_{i=1}^3$ к классам «контурные» и «фоновые».

Шаг 4. Нахождение оценок спектров мощности полезного сигнала $\hat{\mathbf{S}}_{\mathbf{Y}} = (\hat{s}_{k,l})$ и шума $\hat{\mathbf{S}}_{\mathbf{N}} = (\hat{p}_{k,l})$ для каждого класса («контурных»/«фоновых») вейвлет-коэффициентов.

Шаг 5. Вычисление коэффициентов $\mathbf{W} = (w_{k,l})$ фильтра Винера для каждого класса вейвлет-коэффициентов по формуле (2).

Шаг 6. Фильтрация ДВП-спектра, состоящая в покомпонентном перемножении коэффициентов фильтра Винера и соответствующих им коэффициентов ДВП-спектра: $\hat{\mathbf{U}} = (\hat{u}_{k,l} = w_{k,l} \cdot v_{k,l})$.

Шаг 7. Восстановление изображения путём выполнения обратного ДВП: $\hat{\mathbf{Y}} = \text{ОДВП}(\hat{\mathbf{U}})$.

Шаг 8. Компенсация артефактов, возникших в результате винеровской фильтрации.

Перечисленные выше шаги были реализованы как операции, выполняемые слоями нейронной сети с архитектурой (см. рисунок 1), описание которой приведено в последующих разделах.

Дисперсия шума в саббэндах ДВП-спектра

Помимо входного изображения, нейросетевая мо-

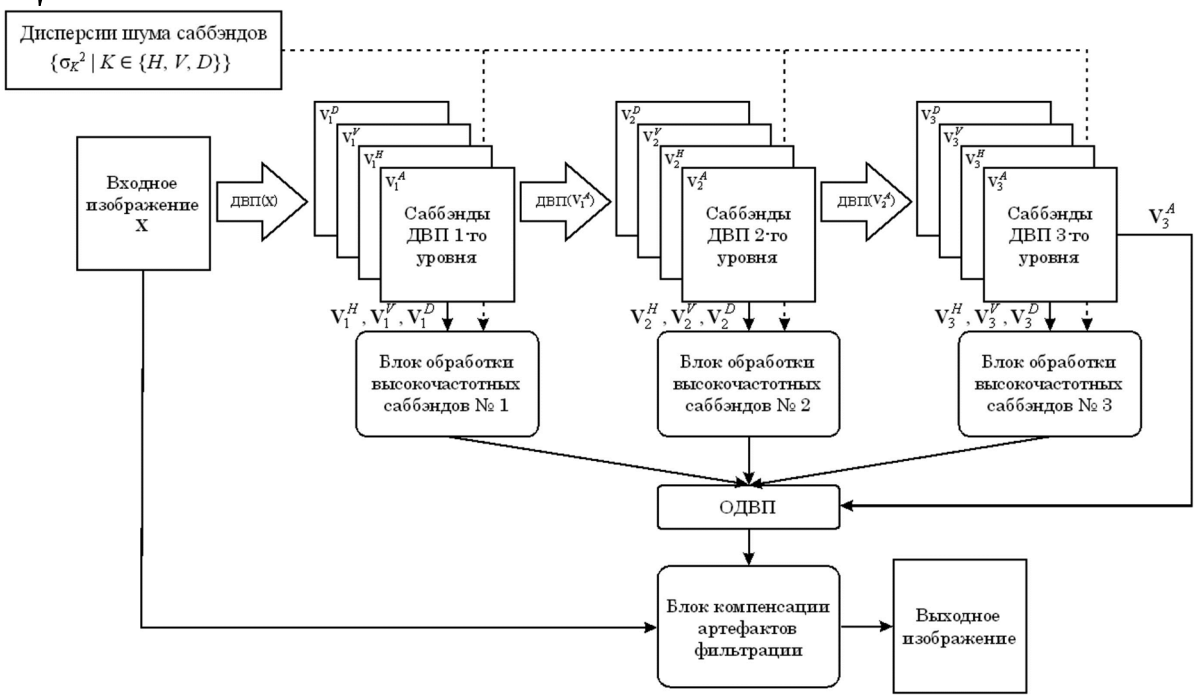


Рисунок 1 – Блок-схема архитектуры нейронной сети

дель также принимает значение СКО шума σ. Поскольку выбранный базис ДВП является биортогональным, в результате ДВП не выполняется сохранение евклидовой нормы сигнала, и дисперсия шумовой составляющей $\mathbf{G}$ саббэндов ДВП-спектра имеет значение, несколько отличающееся от дисперсии σ^2 исходного шума $\mathbf{N}$.

По заданному значению σ^2 дисперсии шумовой составляющей σ_K^2, $K \in \{A, H, V, D\}$, для саббэндов первого уровня ДВП $\mathbf{V}_1^K$ рассчитываются как [10]

$$\sigma_K^2 = \sigma^2 C_h C_f, \tag{3}$$

где $C_h = \sum_{i=1}^{L} h_i^2$, $C_f = \sum_{j=1}^{P} f_j^2$ – константы, $\{h_i\}_{i=1}^{L}$ и $\{f_j\}_{j=1}^{P}$ – КИХ фильтров, задающих ДВП и применяемых соответственно вдоль строк и столбцов изображения для вычисления коэффициентов саббэнда.

Для саббэндов второго и третьего уровней ДВП (i = 2,3) формула (3) применяется рекуррентно, где вместо σ^2 используется дисперсия σ_A^2, найденная для низкочастотного саббэнда предыдущего уровня преобразования $\mathbf{V}_{i-1}^A$.

Значения дисперсий σ_K^2 и соответствующие им саббэнды $\mathbf{V}^K$ передаются в блоки обработки, выполняющие обнаружение контуров и фильтрацию ДВП-спектра (см. рисунок 1).

Алгоритм обработки саббэндов ДВП-спектра

Блоки обработки высокочастотной информации, отвечающие за шаги 3-6 разработанного алгоритма, реализуют схему обработки, представленную на рисунке 2 и состоящую из нечёткой классификации вейвлет-коэффициентов, нахождения оценок ДВП-спектров мощности полезного сигнала и шума, формирования коэффициентов фильтра Винера $\mathbf{W}$ и собственно фильтрации ДВП-спектра $\mathbf{V}$. Далее поясним содержание схемы рисунка 2 более детально.

Нечёткая классификация коэффициентов ДВП

Учёт информации о положении контуров на изображении в процессе подавления шумов позволяет добиться более качественных результатов обработки [3, 8]. Поэтому перед фильтрацией ДВП-спектра изображения выполняется классификация его вейвлет-коэффициентов на «контурные» и «фоновые». Классификация применяется ко всем высокочастотным саббэндам всех уровней ДВП и включает в себя следующие этапы.

1. На основе значений вейвлет-коэффициентов саббэндов $\mathbf{V}^K$, $K \in \{H, V, D\}$, формируются массивы (карты признаков) $\mathbf{R}_1^K = \varphi_K(\mathbf{V}_K)$ с помощью функций $\varphi_K(x)$, имеющих вид

$$\varphi_K(x) = \left| \frac{x^3}{x^2 + (\sigma_K \lambda_K)^2} \right|, \tag{4}$$

где коэффициенты λ_K являются обучаемыми параметрами модели, σ_K находятся из (3). По аналогии с идеей, предложенной в работе [3], функция (4) используется для подавления шумовых вейвлет-коэффициентов с малыми абсолютными значениями, что можно интерпретировать как некоторую пороговую фильтрацию.

2. С целью морфологического сглаживания к полученным картам признаков $\mathbf{R}_1^K$, $K \in \{H, V, D\}$, последовательно применяются операции полутонового размыкания и замыкания [11] с примитивами:

$$\mathbf{Q}^H = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{Q}^V = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\mathbf{Q}^{D_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{Q}^{D_2} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

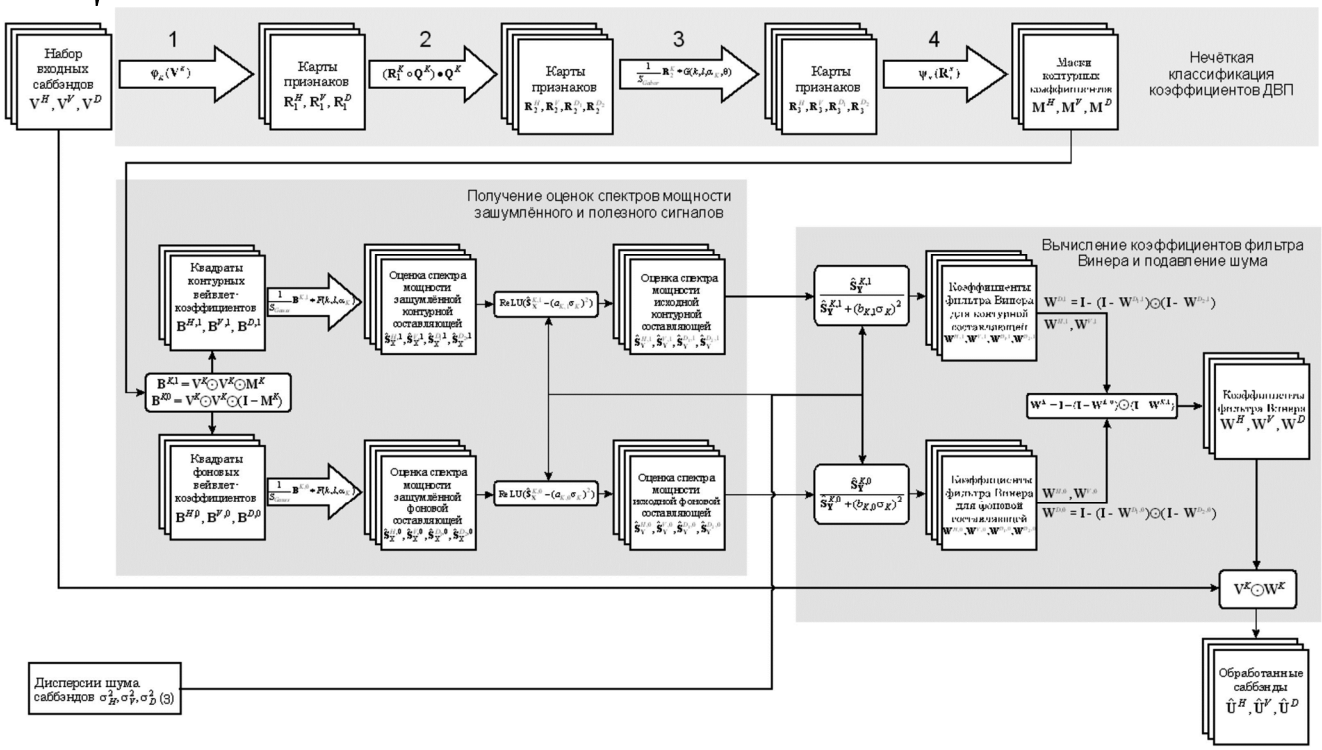


Рисунок 2 – Блок обработки высокочастотных саббэндов $\mathbf{V}^H$, $\mathbf{V}^V$, $\mathbf{V}^D$

При этом карты признаков $\mathbf{R}_1^H$ и $\mathbf{R}_1^V$ подвергаются обработке с примитивами $\mathbf{Q}^H$ и $\mathbf{Q}^V$ соответственно, а для обработки карты $\mathbf{R}_1^D$ используются два примитива $\mathbf{Q}^{D_1}$ и $\mathbf{Q}^{D_2}$ (по одному для каждого направления диагональных контуров). В результате выполнения данного этапа получаются четыре карты признаков: $\mathbf{R}_2^H$, $\mathbf{R}_2^V$, $\mathbf{R}_2^{D_1}$ и $\mathbf{R}_2^{D_2}$; в общем виде процедуру можно записать как

$$\mathbf{R}_2^K = (\mathbf{R}_1^K \circ \mathbf{Q}^K) \bullet \mathbf{Q}^K, \qquad (5)$$

где символы $\circ$ и $\bullet$ используются для обозначения морфологических операций полутонового размыкания и замыкания соответственно; $K \in \{H, V, D_1, D_2\}$.

3. Выполняется свёртка карт признаков, которые были получены на предыдущем этапе, с ИХ модифицированных фильтров Габора [12]:

$$G(k,l,\alpha,\theta) = \frac{1}{2}\exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right)\left(1+\cos\left(2\pi\theta x\right)\right), \qquad (6)$$

где $x = k\cos(\alpha) + l\sin(\alpha)$, $y = -k\sin(\alpha) + l\cos(\alpha)$, α – угол, задающий направление фильтра; σ_x, σ_y – параметры, отвечающие за затухание ИХ фильтра вдоль осей координат Ox и Oy соответственно; θ – пространственная частота гармонических колебаний. Параметры σ_x, σ_y, θ настраиваются в процессе обучения сети, а значение α выбирается из набора $\alpha_K \in \{0, \pm\pi/4, \pi/2\}$ на основе информации о типе саббэнда $\mathbf{V}^K$. $\alpha_H = 0$ – для горизонтальной составляющей $\mathbf{V}^H$, $\alpha_V = \pi/2$ – для вертикальной составляющей $\mathbf{V}^V$. Для диагональной составляющей $\mathbf{V}^D$ используются два значения угла фильтрации $\alpha_{D_1} = \pi/4$, $\alpha_{D_2} = -\pi/4$ и вычисляются две карты признаков (для каждого направления диагональных контуров). Размеры масок фильтров (6) выбраны равными 9×9, в соответствии с размерами ИХ используемых двумерных ДВП-фильтров. Результатом выполнения данного этапа являются карты признаков:

$$\mathbf{R}_3^K = \frac{1}{S_{Gabor}}\mathbf{R}_2^K * G(k,l,\alpha_K,\theta), \qquad (7)$$

где символ $*$ используется для обозначения операции двумерной дискретной свёртки, $K \in \{H, V, D_1, D_2\}$; нормировочный параметр $S_{Gabor} = \sum_{i=-4}^{4}\sum_{j=-4}^{4} G(i,j,\alpha_K,\theta)$.

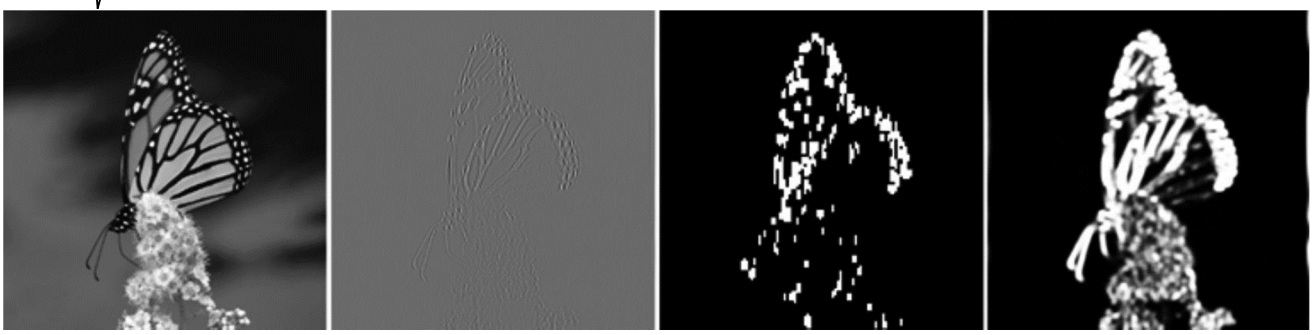
4. Найденные на третьем этапе карты признаков (7) проходят через нелинейную функцию активации $\psi_K(x)$, которая приводит входные значения к диапазону [0; 1) и имеет вид:

$$\psi_K(x) = (1 - \exp(-\mu_K^2 x^2))^2, \qquad (8)$$

где μ_K – обучаемый параметр, определяющий ширину полосы, при попадании в которую значение функции будет близко к нулю; $K \in \{H, V, D\}$. Для диагональных саббэндов полагаем $\psi_D(x) = \psi_{D_1}(x) = \psi_{D_2}(x)$. На основе откликов функции (8) по матрицам $\mathbf{R}_3^K = (r_{k,l}^K)$ формируются маски $\mathbf{M}^K = (m_{k,l} = \psi_K(r_{k,l}^K))$, задающие для каждого вейвлет-коэффициента $v_{k,l} \in \mathbf{V}^K$ степень его принадлежности к контурной составляющей ДВП-спектра изображения как непрерывную величину $m_{k,l} \in [0; 1)$:

$$\mathbf{M}^K = \psi_K(\mathbf{R}_3^K), \qquad (9)$$

где $K \in \{H, V, D_1, D_2\}$. Маски $\mathbf{M}^{D_1}$ и $\mathbf{M}^{D_2}$, соответствующие диагональным контурам различных направлений, объединяются в одну общую маску $\mathbf{M}^D = \mathbf{I} - -(\mathbf{I} - \mathbf{M}^{D_1}) \odot (\mathbf{I} - \mathbf{M}^{D_2})$, где $\mathbf{I}$ – матрица, состоящая из единиц, а символ $\odot$ обозначает операцию покомпонентного перемножения матриц.



а) *б)* *в)* *г)*

Рисунок 3 – Изображение Butterfly (а); его саббэнд $\mathbf{V}_2^V$ (б); помеченные «контурные» вейвлет-коэффициенты (алгоритм из статьи [8]) (в); помеченные «контурные» вейвлет-коэффициенты (разработанный алгоритм) (г)

Для каждого саббэнда $\mathbf{V}^K$ соответствующая ему маска $\mathbf{M}^K$ позволяет определить совокупность «контурных» вейвлет-коэффициентов как нечёткое множество [11], являющееся подмножеством $\mathbf{V}^K$: чем ближе значение $m_{k,l} = \psi_K(r_{k,l}^K)$ к единице, тем выше степень принадлежности $v_{k,l}$ к множеству контурных коэффициентов. Разработанный алгоритм «более мягко» выделяет структуру и положение контуров по высокочастотным саббэндам ДВП-спектра, в отличие от аналогичного детектора из работы [8], использующего бинарные признаки принадлежности для контурных точек {0; 1} (см. рисунок 3).

Получение оценок спектров мощности зашумлённого и полезного сигналов

Оценка ДВП-спектра мощности зашумлённого сигнала находится как свёртка квадратов значений коэффициентов ДВП-спектра с гауссовой ИХ вида:

$$F(k,l,\alpha) = \exp\left(-\frac{x^2}{\sigma_x^2} - \frac{y^2}{\sigma_y^2}\right), \quad (10)$$

где $x = k\cos(\alpha) + l\sin(\alpha)$, $y = -k\sin(\alpha) + l\cos(\alpha)$, α – угол, задающий поворот ИХ; параметры σ_x, σ_y отвечают за затухание ИХ фильтра вдоль осей координат Ox и Oy соответственно и настраиваются в процессе обучения сети. Как и ранее для фильтров (6), размеры масок фильтров (10) были выбраны равными 9×9.

Выражения для оценки спектров мощности контурной и фоновой составляющих зашумлённого изображения имеют вид

$$\hat{\mathbf{S}}_{\mathbf{X}}^{K,C} = \frac{1}{S_{Gauss}}\mathbf{B}^{K,C} * F(k,l,\alpha_K), \quad (11)$$

где $K \in \{H, V, D_1, D_2\}$; $C \in \{0,1\}$ – признак контурной (C = 1) либо фоновой (C = 0) составляющей; $\mathbf{B}^{K,1} = \mathbf{V}^K \odot \mathbf{V}^K \odot \mathbf{M}^K$, $\mathbf{B}^{K,0} = \mathbf{V}^K \odot \mathbf{V}^K \odot (\mathbf{I} - \mathbf{M}^K)$ – квадраты значений «контурных» и «фоновых» вейвлет-коэффициентов соответственно; $\alpha_K \in \{0, \pm\pi/4, \pi/2\}$ – параметр, значения которого выбирались, как и в (7), исходя из информации о типе саббэнда $\mathbf{V}^K$; $S_{Gauss} = \sum_{i=-4}^{4}\sum_{j=-4}^{4} F(i,j,\alpha)$. Для саббэнда $\mathbf{V}^D$ вычисляются две различных оценки $\hat{\mathbf{S}}_{\mathbf{X}}^{D_1,C}$ и $\hat{\mathbf{S}}_{\mathbf{X}}^{D_2,C}$ (при этом в (11) полагается, что $\mathbf{B}^{D_1,C} = \mathbf{B}^{D_2,C} = \mathbf{B}^{D,C}$).

Поскольку в рассматриваемой модели шум является белым и не коррелированным с полезным сигналом, его спектр мощности является константой. С учётом этого оценку спектра мощности полезного сигнала $\hat{\mathbf{S}}_{\mathbf{Y}}^{K,C} = (\hat{s}_{k,l}^{K,C})$ для контурной и фоновой составляющих предлагается формировать как

$$\hat{\mathbf{S}}_{\mathbf{Y}}^{K,C} = \mathrm{ReLU}\,(\hat{\mathbf{S}}_{\mathbf{X}}^{K,C} - (a_{K,C}\sigma_K)^2), \quad (12)$$

где $\mathrm{ReLU}(x) = \max(0, x)$ – функция активации, $K \in \{H, V, D_1, D_2\}$, $C \in \{0,1\}$; $a_{K,C}$ – обучаемый параметр, причём $a_{D_1,C} = a_{D_2,C}$.

Вычисление коэффициентов фильтра Винера и подавление шума

Значения коэффициентов фильтра Винера $\mathbf{W}^{K,C}$ для контурной и фоновой составляющих ДВП-спектра рассчитываются в соответствии с формулой (2) следующим образом:

$$w_{k,l}^{K,C} = \frac{\hat{s}_{k,l}^{K,C}}{\hat{s}_{k,l}^{K,C} + (b_{K,C}\sigma_K)^2}, \quad (13)$$

где $K \in \{H, V, D_1, D_2\}$, $C \in \{0,1\}$, $\hat{s}_{k,l}^{K,C}$ – оценки коэффициентов ДВП-спектра мощности $\hat{\mathbf{S}}_{\mathbf{Y}}^{K,C}$, $b_{K,C}$ – обучаемый параметр. Набор коэффициентов фильтра $\mathbf{W}^{D,C} = (w_{k,l}^{D,C})$ для обработки саббэнда $\mathbf{V}^D$ вычисляется как $w_{k,l}^{D,C} = 1 - (1 - w_{k,l}^{D_2,C})(1 - w_{k,l}^{D_1,C})$.

На основе полученных значений $w_{k,l}^{K,C}$ (13) для каждого саббэнда $\mathbf{V}^K$, $K \in \{H, V, D\}$, вычисляются коэффициенты фильтра $\mathbf{W}^K = (w_{k,l}^K)$, учитывающие степень принадлежности ДВП-коэффициентов к контурам и фону:

$$w_{k,l}^K = 1 - (1 - w_{k,l}^{K,0})(1 - w_{k,l}^{K,1}). \quad (14)$$

Фильтрация ДВП-спектра производится путём покомпонентного перемножения матриц коэффициентов фильтра (14) и вейвлет-коэффициентов саббэндов: $\hat{\mathbf{U}}^K = \mathbf{W}^K \odot \mathbf{V}^K$. Так преобразованные саббэнды ДВП-спектра используются для получения восстановленного изображения с помощью ОДВП.

Компенсация артефактов фильтрации

В процессе винеровской фильтрации вместе с шумовыми коэффициентами ДВП-спектра также могут быть подавлены вейвлет-коэффициенты, соответству-

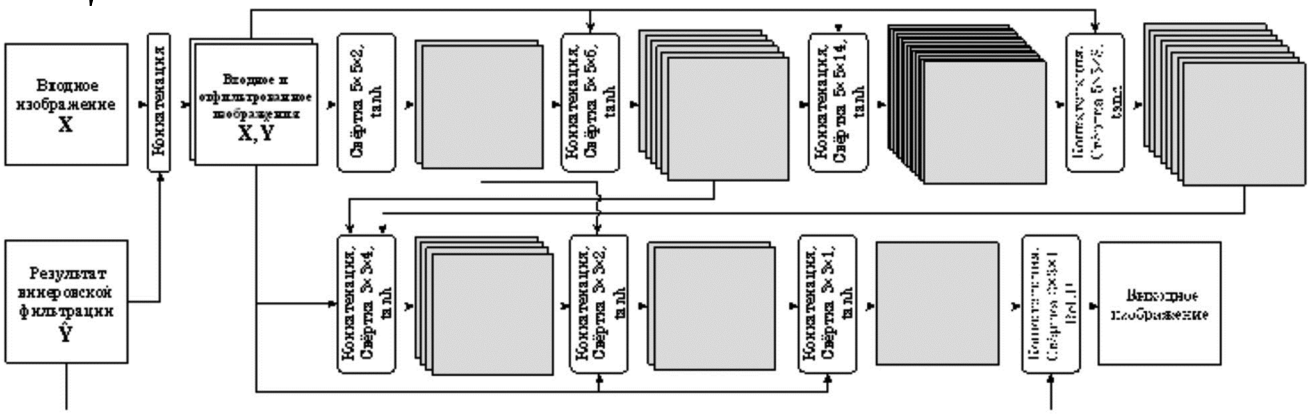


Рис. 4. Блок компенсации артефактов фильтрации

ющие полезной высокочастотной информации. Для того чтобы снизить влияние возникающих вследствие этого искажений и артефактов, в архитектуру модели дополнительно был добавлен блок (см. рисунок 4), разработанный на основе идей, изложенных авторами нейронной сети U-Net в работе [13]. Серым цветом на рисунке выделены карты признаков нейронной сети, а выполняемые операции представлены блоками, где под конкатенацией подразумевается объединение матриц карт признаков и изображений в 3-х мерный тензор (матрицы конкатенируются как каналы тензора); рядом с каждой свёрткой подписан размер её ядра в формате (высота×ширина×количество выходных каналов); обозначения tanh (гиперболический тангенс) и ReLU соответствуют использованию нелинейных функций активации после вычисления свёрток.

Блок компенсации артефактов фильтрации состоит из 8 свёрточных слоёв и имеет «короткие соединения» на основе конкатенации карт признаков. Количество слоёв и размеры свёрток в блоке, а также карты признаков, между которыми выполняется конкатенация, были определены в ходе экспериментов таким образом, чтобы результат обработки обеспечивал наибольший прирост пикового отношения сигнал/шум (PSNR) [11]. Двумерные свёрточные слои были реализованы как разделимые свёртки [2]. Перед вычислением свёрток выполняется операция дополнения отсчётами, симметрично отражёнными относительно границ изображений и карт признаков, чтобы размеры (высота и ширина) откликов всех слоёв оставались равными размерам изображения, поступающего на вход.

Как показали проведённые нами эксперименты, на тестовых данных при среднем ($\sigma = 15$) и высоком ($\sigma = 30$) уровнях шума добавление блока компенсации даёт прирост PSNR в среднем на 0,46 дБ.

Результаты экспериментов

Обучение нейронной сети

Для обучения описанной нейросетевой модели шумоподавления использовался набор данных ILSVRC [14], состоящий из 1 281 167 изображений с разрядностью 8 бит/пиксель и диапазоном значений яркости пикселей {0, …, 255}. Исходные цветные изображения были преобразованы к полутоновому виду. Для каждой эпохи из тренировочного набора случайным образом отбирались 500 изображений, затем из случайно выбранной области каждого изображения вырезался фрагмент размера 128×128 пикселей. Данные фрагменты использовались в качестве эталонных меток в последующем обучении. При формировании входного изображения фрагмент-метка подвергался искажению белым аддитивным нормальным шумом с нулевым математическим ожиданием и СКО σ, значение которого выбиралось как реализация случайной величины, имеющей равномерное распределение U(0, 30). Зашумлённый таким способом фрагмент изображения использовался затем в качестве образа обучающей выборки.

В качестве функции потерь использовался квадратичный штраф. Обучение модели выполнялось с помощью библиотеки PyTorch версии 2.4.1 [15]. Для минимизации выборочного штрафа использовался алгоритм Adam [16] со значениями параметров: $\alpha = 10^{-2}$; $\beta_1 = 0{,}9$; $\beta_2 = 0{,}999$; $\varepsilon = 10^{-8}$, количество эпох было выбрано равным 600 (при дальнейшем обучении рост метрик PSNR и SSIM не наблюдался), размер батча выбран равным 8. Для всех параметров алгоритма Adam, за исключением скорости обучения α, использовались значения по умолчанию. Параметр α, размеры фрагментов-образов обучающей выборки, а также количество обучающих примеров в батче были подобраны эмпирически как показавшие наибольшие значения метрик качества фильтрации.

Тестирование методов шумоподавления

Тестирование сети выполнялось на наборе данных BSDS500 [17], состоящем из 500 полутоновых изображений. Оценка качества обработки выполнялась на основе сравнения значений мер PSNR и индекса структурного сходства (SSIM) [18] до и после обработки.

Проведено сравнение характеристик разработанной нейронной сети с существующими алгоритмами подавления шумов: винеровской фильтрацией в области ДВП [19], винеровской фильтрацией с применением ансамбля нейросетевых моделей [8], нейросетевой моделью DCT2net [3]. Рассмотрены три фиксированных уровня шума, соответствующие различным значениям СКО: низкий уровень шума ($\sigma = 7$), средний ($\sigma = 15$) и высокий ($\sigma = 30$). Значения метрик, рассчитанные по результатам экспериментов и усреднённые по тестовому набору, представлены в таблице 1.

Таблица 1 – Средние значения PSNR и SSIM до и после фильтрации при разных значениях СКО шума, вычисленные по набору BSDS500

Метрика	PSNR (средн.) [дБ]	SSIM (средн.)	PSNR (средн.) [дБ]	SSIM (средн.)	PSNR (средн.) [дБ]	SSIM (средн.)
СКО шума	σ = 7		σ = 15		σ = 30	
Значение до обработки	31,26	0,80	24,75	0,56	18,94	0,32
Фильтр Винера [19]	34,59	*0,93*	29,84	0,83	25,93	0,67
Фильтр Винера + нейросетевой ансамбль [8]	34,75	*0,93*	30,10	0,85	26,43	0,72
Разработанный алгоритм	*35,20*	**0,94**	*31,17*	*0,87*	*27,49*	*0,76*
DCT2net [3]	**35,38**	**0,94**	**31,22**	**0,88**	**27,72**	**0,77**

Примечание. Наибольшие значения в каждом столбце выделены жирным шрифтом, курсивом – следующие по убыванию после наибольших.

Таблица 2 – Значения PSNR (дБ) до и после фильтрации при разных значениях СКО шума

Изображение	Barbara	Boat	Lena	Goldhill	Mandrill
СКО шума	σ = 7				
PSNR до обработки	31,22	31,22	31,26	31,22	31,22
Фильтр Винера [19]	34,87	34,49	36,19	34,95	*35,64*
Фильтр Винера + нейросетевой ансамбль [8]	35,00	34,64	36,49	35,11	**35,77**
Разработанный алгоритм	*35,32*	*34,93*	*36,60*	*35,38*	34,62
DCT2net [3]	**35,78**	**35,22**	**36,97**	**35,79**	34,71
СКО шума	σ = 15				
PSNR до обработки	24,61	24,64	24,77	24,62	24,60
Фильтр Винера [19]	30,21	30,35	31,77	30,61	30,20
Фильтр Винера + нейросетевой ансамбль [8]	30,47	30,76	32,26	31,02	*30,41*
Разработанный алгоритм	*31,62*	*31,75*	**33,39**	*31,81*	**30,52**
DCT2net [3]	**31,95**	**31,92**	*33,23*	**32,07**	30,40
СКО шума	σ = 30				
PSNR до обработки	18,76	18,74	19,00	18,76	18,64
Фильтр Винера [19]	26,29	26,66	27,67	27,14	25,76
Фильтр Винера + нейросетевой ансамбль [8]	26,64	27,43	28,34	28,02	*26,16*
Разработанный алгоритм	*27,70*	*28,53*	**29,72**	*28,76*	**26,70**
DCT2net [3]	**28,30**	**28,89**	*29,66*	**29,13**	**26,70**



а) *б)* *в)* *г)*

Рисунок 5 – Фрагменты изображений, искажённых шумом при различных значения СКО, до (сверху) и после фильтрации (снизу): σ = 0 (а), σ = 7 (б), σ = 15 (в), σ = 30 (г)

Полученные значения метрик позволяют сделать вывод, что предложенная нейронная сеть обеспечивает более качественное шумоподавление по сравнению с другими алгоритмами винеровской фильтрации, уступая лишь сети DCT2net. При сравнении с последней разница в приросте PSNR в среднем не превышает 0,18 дБ для низкого, 0,05 дБ для среднего и 0,23 дБ для высокого уровня шума, а разница в приросте SSIM в среднем не превышает 0,01.

Рассмотренные алгоритмы также были протестированы на полутоновых изображениях Barbara, Boat, Lena, Goldhill и Mandrill, широко применяющихся для проверки алгоритмов цифровой обработки (см. таблицу 2).

Найденные значения PSNR согласуются с результатами из таблицы 1, при этом можно видеть, что в отдельных случаях при среднем и высоком уровнях шума

предложенный алгоритм демонстрирует лучшее качество обработки по сравнению с нейросетевой моделью DCT2net. Следует отметить, что DCT2net содержит 28 561 обучаемый параметр [3], нейросетевой ансамбль из работы [8] содержит 2772 параметра, в то время как у предложенной в настоящем исследовании модели их всего 1491. Примеры зашумлённых и отфильтрованных изображений для различных значений СКО шума представлены на рисунке 5.

Заключение

Значения метрик качества, полученные в экспериментах, свидетельствуют о том, что предложенный алгоритм нейросетевой фильтрации обеспечивает лучшее подавление шумов в сравнении с рассмотренными альтернативными реализациями фильтра Винера. Разработанная для подавления белого шума в изображениях нейросеть в среднем несколько уступает DCT2net по метрике PSNR, однако в некоторых случаях демонстрирует сопоставимые, а иногда и превосходящие результаты обработки.

Количество параметров предлагаемой нами сети шумоподавления примерно в 19 раз меньше чем DCT2net и в 1,8 раз меньше, чем у ансамбля нейронных сетей из нашей предыдущей работы [8]. Таким образом, описанный в настоящей работе алгоритм нейросетевой винеровской фильтрации в области ДВП, представляющий собой развитие подходов из [8], является более предпочтительным решением с точки зрения баланса сложности и качества обработки.

Литература

1. Алимагадов К.А., Умняшкин С.В. Аугментация данных на основе вейвлет-фильтрации при обучении нейронных сетей. Труды Международной конференции по компьютерной графике и зрению «Графикон». М.: Институт прикладной математики им. М.В. Келдыша РАН, 2023. С. 437-442. DOI 10.20948/graphicon-2023-437-442.

2. Huang J.-J., Dragotti P.L. WINNet: Wavelet-inspired invertible network for image denoising. IEEE Transactions on Image Processing. IEEE, 2022, pp. 4377-4392.

3. Herbreteau S., Kervrann C. DCT2net: An interpretable shallow CNN for image denoising. IEEE Trans. Image Process. IEEE, 2022, pp. 4292-4305.

4. Tu Z., Talebi H., Zhang H., Yang F., Milanfar P., Bovik A., Li Y. MAXIM: Multi-Axis MLP for image processing. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022, pp. 5759-5770.

5. Haykin S. Neural networks and learning machines. 3rd ed. New York: Pearson, 2009. 936 p.

6. Rahaman N., Baratin A., Arpit D., Draxler F., Lin M., Hamprecht F., Bengio Y., Courville A. On the spectral bias of neural networks. In International Conference on Machine Learning. PMLR. 2019, pp. 5301-5310.

7. Gal R., Cohen D., Bermano A., Cohen-Or D. SWAGAN: A style based wavelet-driven generative model. CoRR. 2021, pp. 1-11.

8. Алимагадов К.А., Умняшкин С.В. Подавление белого шума на изображениях на основе винеровской фильтрации в области дискретного вейвлет-преобразования с применением нейросетевых технологий. // Известия высших учебных заведений. Электроника. 2022. Т. 27, № 6. С. 807-818. DOI 10.24151/1561-5405-2022-27-6-807-818.

9. Guo T. et al. A review of wavelet analysis and its applications: Challenges and opportunities. IEEE Access. 2022, no. 10, pp. 58869-58903. DOI 10.1109/ACCESS.2022.3179517.

10. Умняшкин С.В. Основы теории цифровой обработки сигналов. М.: Техносфера, 2025. 550 с.

11. Gonzalez R.C., Woods R.E. Digital image processing. 4th ed. New York: Pearson Education, 2018. 1168 p.

12. Feichtinger H.G., Strohmer T. Gabor analysis and algorithms: Theory and applications. Springer Science & Business Media, 2012.

13. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention. Munich: Springer International Publishing, 2015, pp. 234-241.

14. Russakovsky O. et al. Imagenet large scale visual recognition challenge // International journal of computer vision. - 2015. - №115. - P. 211-252.

15. PyTorch. GitHub URL: https://github.com/pytorch/pytorch (дата обращения: 21.09.2024).

16. Kingma D.P., Ba J.L. Adam: a method for stochastic optimization. arXiv URL: https://arxiv.org/abs/1412.6980v1 (дата обращения: 21.09.2024).

17. Berkeley segmentation data set and benchmarks 500 (BSDS500). Berkeley Computer Vision Group URL: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html (дата обращения: 15.09.2024).

18. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing. 2004, no. 13, pp. 600-612.

19. Alimagadov K.A., Umnyashkin S.V. Application of Wiener filter to suppress white noise in images: wavelet vs Fourier basis. 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). St. Petersburg, Moscow: IEEE, 2021, pp. 2059-2063. DOI 10.1109/ ElConRus51938.2021.9396470.