



Общероссийский математический портал

С. В. Сай, А. В. Зинкевич, Е. С. Фомина, Сравнение дискретного косинус- и вейвлет-преобразований в системах сжатия RAW-изображений, *Компьютерная оптика*, 2022, том 46, выпуск 6, 929–938

<https://www.mathnet.ru/co1088>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<https://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.97.14.89

21 апреля 2025 г., 04:56:12



## Сравнение дискретного косинус- и вейвлет-преобразований в системах сжатия RAW-изображений

С.В. Сай<sup>1</sup>, А.В. Зинкевич<sup>1</sup>, Е.С. Фомина<sup>1</sup>

<sup>1</sup> Тихоокеанский государственный университет, Хабаровск, Россия

### Аннотация

В статье приводится описание особенностей цифровой обработки сигналов изображения в процессе кодирования на основе дискретного косинус- и вейвлет-преобразований, используемых в стандартах сжатия JPEG и JPEG2000. Для сравнения дискретного косинус-преобразования и дискретного вейвлет-преобразования разработана цифровая модель системы, в которой реализованы одинаковые этапы обработки сигналов, кроме этапов самих дискретных преобразований. Предложена методика анализа эффективности преобразований по объективным оценкам качества изображений в зависимости от коэффициента сжатия. К особенностям относится то, что, в отличие от популярных метрик PSNR и SSIM, для оценки качества предложено использовать коэффициент четкости изображения, вычисление которого связано с контрастной чувствительностью зрения. В результате исследований получены количественные оценки эффективности сжатия при заданных параметрах качества в зависимости от типа преобразования и детальности RAW-изображений. Предложены рекомендации по использованию результатов исследований в системах кодирования и передачи изображений с высокой четкостью.

**Ключевые слова:** анализ изображений, метрика искажений, дискретное косинус-преобразование, дискретное вейвлет-преобразование, эффективность сжатия.

**Цитирование:** Сай, С.В. Сравнение дискретного косинус- и вейвлет-преобразований в системах сжатия RAW-изображений / С.В. Сай, А.В. Зинкевич, Е.С. Фомина // Компьютерная оптика. – 2022. – Т. 46, № 6. – С. 929-938. – DOI: 10.18287/2412-6179-CO-1094.

**Citation:** Sai SV, Zinkevich AV, Fomina ES. Comparison of discrete cosine and wavelet transforms in RAW image compression systems. Computer Optics 2022; 46(6): 929-938. DOI: 10.18287/2412-6179-CO-1094.

### Введение

Цифровые изображения подвержены воздействию искажений, которые появляются в процессе формирования, цифровой обработки и передачи сигналов. Основные искажения качества возникают за счет использования сжатия с потерями, что обусловлено необходимостью передачи сигналов по каналам связи с ограниченной пропускной способностью, а также компактным хранением больших объемов видеоданных.

В сенсорах цифровых камер [1] на первой стадии обработки формируются изображения в исходном (RAW) формате, которые далее конвертируются и сжимаются обычно по стандарту JPEG [2].

В системах цифрового телевидения (H.264 и др.) наибольшее применение также имеют алгоритмы сжатия опорных (I) кадров видеоряда изображений по стандарту JPEG. В последних версиях (H.265) предлагается использовать сжатие опорных кадров по стандарту JPEG2000 [3–4].

Основные отличия стандартов JPEG и JPEG2000 заключаются в использовании разных типов преобразований сигналов. В JPEG используется блочное дискретное косинусное преобразование (ДКП), а в JPEG2000 – дискретное вейвлет-преобразование (ДВП). Искажения качества декодированного изображения возникают в процессе квантования и обнуления коэффициентов

трансформации ДКП или ДВП. Характеристики основных визуальных искажений практически одинаковы – это ухудшение четкости и резкости с увеличением коэффициентов сжатия, что приводит к размытию мелких структур и резких границ изображения. К специфическим искажениям JPEG относится появление «блокинг»-эффекта, обусловленного блочной структурой ДКП.

Сравнению эффективности JPEG и JPEG2000 посвящено достаточное количество публикаций [5–10]. В основном в этих работах приводятся результаты сравнения с помощью современных фоторедакторов и/или конверторов изображений. Методика сравнения основана на анализе качества эталонных изображений с заданными одинаковыми коэффициентами сжатия или битовой скоростью передачи. В качестве критериев качества [11] изображения традиционно используется пиковое отношение сигнала к шуму (PSNR) и/или индекс структурного подобия (SSIM).

Обзорные исследования показывают, что практически во всех работах предпочтение отдается стандарту JPEG2000, который по сравнению с JPEG обеспечивает лучшее качество передачи изображений при одинаковых битовых скоростях. Однако критерии PSNR и SSIM не всегда соответствуют зрительным оценкам качества фотореалистичных изображений. Также в стандартах используются разные алгоритмы

сжатия коэффициентов трансформации. В JPEG применяется модифицированный метод Хаффмана, а в JPEG2000 используется арифметическое сжатие. Следовательно, результаты сравнения не позволяют однозначно ответить на вопрос, какой количественный выигрыш по коэффициенту сжатия обеспечивает ДВП по сравнению с ДКП?

В статье предлагается единая цифровая модель кодека для обоих преобразований. Для анализа эффективности использована популярная метрика SSIM, а также способ сравнения эффективности ДКП и ДВП, основанный на анализе снижения четкости передаваемых изображений. В качестве тестовых изображений использованы оригинальные RAW-фотоизображения, представленные в обзорных галереях современных цифровых камер фирмы Sony [12].

Цель исследования – это разработка методики и получение наиболее объективных и точных оценок сравнения преобразований ДКП и ДВП в системах кодирования и передачи высококачественных изображений.

### 1. Цифровая модель кодека

Рассмотрим особенности построения цифровой модели кодека. Все алгоритмы обработки сигналов в цифровой модели и анализа качества изображений реализованы в виде программных модулей на языке C++.

В кодере выполняются следующие основные операции:

- а) преобразование цифровых RGB-сигналов тестового изображения в сигналы YVU;
- б) выполнение ДКП или ДВП с последующим квантованием коэффициентов трансформации;
- в) нормализация и преобразование коэффициентов трансформации в одномерный массив;
- г) энтропийное сжатие.

В декодере выполняются обратные преобразования.

Преобразование сигналов основных цветов в цифровые сигналы яркости (Y) и цветности (U, V) для вещественного режима выполняется следующим способом [13]:

$$\begin{aligned} Y &= 0,299R + 0,587G + 0,114B; \\ U &= -0,16875R - 0,33125G + 0,5B; \\ V &= 0,5R - 0,41869G - 0,08131B. \end{aligned} \tag{1}$$

Здесь выбрана цветовая модель 4:4:4, т.е. без прореживания сигналов цветности. В результате получим три матрицы Y, U и V с размером W × H, где W – ширина и H – высота изображения в пикселях. Далее каждая матрица сигналов преобразуется в матрицу коэффициентов трансформации ДКП или ДВП. Рассмотрим особенности этих преобразований.

Выполнение ДКП.

Дискретно-косинусное преобразование выполняется поблочно. Прямое ДКП для блоков N × N имеет следующий вид:

$$Y_{k,l}^T = \frac{C_{k,l}}{\sqrt{2N}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} Y_{m,n} \cos \frac{(2m+1)k\pi}{2N} \cos \frac{(2n+1)l\pi}{2N}, \tag{2}$$

где

$$C_{k,l} = \begin{cases} \frac{1}{\sqrt{2}} \text{if } k=0, l=0 \\ 1.0 \text{ if } k \neq 0, l \neq 0 \end{cases};$$

$Y_{k,l}^T$  – коэффициенты трансформации сигнала яркости; m, n – номер пикселя в блоке по вертикали и по горизонтали; k, l – номер коэффициента трансформации. Аналогично выполняются преобразования по сигналам цветности U и V.

Стандартом [2] рекомендован размер блока 8 × 8 пикселей, что реализовано в программном модуле.

Прямое вычисление ДКП (2) требует больших вычислительных затрат, поэтому на практике используется алгоритм быстрого преобразования Фурье. Также выполнение преобразований часто реализуется с помощью последовательности матричных перемножений. При этом прямое и обратное ДКП по яркости описываются следующими выражениями:

$$\begin{aligned} Y^T &= M_{\text{ДКП}} \cdot Y \cdot M_{\text{ДКП}}^T; \\ Y &= M_{\text{ДКП}}^T \cdot Y^T \cdot M_{\text{ДКП}}, \end{aligned} \tag{3}$$

где Y – блок пикселей яркости исходного изображения;  $Y^T$  – блок коэффициентов трансформации; M – прямая матрица преобразования;  $M^T$  – транспонированная матрица. В программном модуле реализовано матричное преобразование (3).

На следующем этапе выполняется квантование коэффициентов ДКП и обнуление малозначущих коэффициентов. Квантование коэффициентов можно выполнить с помощью матрицы квантования M:

$$\tilde{Y}_{k,l}^T = \frac{Y_{k,l}^T}{M_{k,l}^Y(Q)}, \tag{4}$$

где Q – регулируемый параметр качества.

Матрицы квантования рекомендованы стандартом [2] раздельно для блоков яркости и блоков цветности. С увеличением номера k и l значения коэффициентов в матрицах квантования увеличиваются, что обусловлено спадом контрастно-частотной характеристики зрения в области верхних пространственных частот. Для компонент цветности используется более грубое квантование, что обусловлено меньшей чувствительностью зрения к искажениям по цветности.

В цифровой модели значения коэффициентов матрицы квантования по яркости и по цветности вычислены по формулам:

$$\begin{aligned} M_{k,l}^Y &= 1 + (1 + k + l) \cdot (100 - Q) / 2; \\ M_{k,l}^{UV} &= 1 + C_{UV} (1 + k + l) \cdot (100 - Q) / 2, \end{aligned} \tag{5}$$

где коэффициент  $C_{U,V} > 1$  учитывает менее высокую чувствительность зрения к искажениям коэффициен-

тов трансформации блоков цветности в области высоких частот. Параметр  $Q$  изменяется от 1 до 100 с единичным шагом и образует шкалу квантования.

Как показали результаты экспериментальных исследований, использование матриц квантования (5) с коэффициентом  $C_{U,V}=8$  для блоков с размерами  $8 \times 8$  пикселей дает примерно одинаковые оценки качества восстановленных изображений по сравнению со стандартными матрицами при условии обеспечения одинаковых коэффициентов сжатия тестовых изображений. Также выражения (5) позволяют сформировать матрицы квантования с любым заданным размером  $N \times N$ .

Преобразования (2–5) являются обратимыми, т.е. при выполнении обратных преобразований и восстановлении изображения искажения в блоках яркости и цветности будут отсутствовать. Однако это условие выполняется при представлении результатов преобразования вещественными числами с плавающей запятой или с фиксированной точкой.

После выполнения операции квантования (4) вещественные коэффициенты ДКП преобразуются к целому типу с сохранением одного знака после запятой. В цифровой модели нормализация коэффициентов выполнена с помощью умножения вещественных значений коэффициентов на число 10 с последующим округлением до ближайшего целого

$$\bar{Y}_{k,l}^T = \lfloor \tilde{Y}_{k,l}^T \cdot 10 + 0,5 \rfloor, \quad (6)$$

где операция  $\lfloor \dots \rfloor$  означает округление вниз до ближайшего целого.

После выполнения нормализации (6) в блоках происходит обнуление малозначачих коэффициентов трансформации ДКП, что приводит к необратимости преобразований (2–5) и к появлению искажений в декодированном изображении. Очевидно, что уровень искажений будет зависеть от параметра качества  $Q$ , задаваемого в уравнении (5).

Таким образом, после выполнения операций ДКП, квантования и нормализации размер блока и, следовательно, всего файла увеличится в 2 раза. Здесь мы учитываем, что исходное изображение представлено 8-битным цифровым сигналом по каждому цветовому каналу и для каждого коэффициента трансформации в соответствии с (6) выделено 16-разрядное слово.

Однако ДКП выполняет декорреляцию блоков изображения, что после квантования приводит к появлению большого количества нулевых коэффициентов трансформации и является предпосылкой для использования эффективного сжатия. Для повышения эффективности сжатия использовано зигзаг-сканирование коэффициентов трансформации и преобразование блока в одномерный массив. Эта операция позволяет упорядочить коэффициенты от низкочастотных до высокочастотных составляющих с целью образования более длинных цепочек нулевых коэффициентов.

Выполнение ДВП.

Дискретное вейвлет-преобразование одномерного сигнала  $S_n$  длиной  $N$  можно выполнить по следующим формулам [9]:

$$S_{j,k}^L = \sum_{n=0}^{N-1} S_{j-1,n} h_{n-2k}; \quad S_{j,k}^H = \sum_{n=0}^{N-1} S_{j-1,n} g_{n-2k}, \quad (7)$$

где  $n$  – номера отсчетов исходного дискретного сигнала;  $k$  – номера коэффициентов ДВП преобразованного сигнала в диапазоне от 0 до  $(N-1)/2$ ;  $j$  – номер итерации, изменяющийся от 1 до заданного значения  $J$ ;  $h_n$  – низкочастотные коэффициенты и  $g_n$  – высокочастотные коэффициенты импульсной характеристики симметричного биортогонального фильтра. В стандарте JPEG2000 рекомендовано [4] использование симметричного биортогонального фильтра Коэна–Добеши–Фово (CDF 9/7).

Из преобразования (7) следует, что коэффициенты ДВП  $S^L$  представляют собой низкочастотные составляющие, а  $S^H$  – высокочастотные составляющие исходного дискретного сигнала. При этом за счет сдвига импульсной характеристики на  $2k$  в результате преобразования получаем две составляющие длиной  $N/2$ . Таким образом, в результате первой итерации ( $j=1$ ) имеем масштабированную (уменьшенную) в два раза копию ( $S^L$ ) исходного сигнала и его дополнение ( $S^H$ ) в виде высокочастотной составляющей. На следующей итерации ( $j=2$ ) процесс повторяется, где исходным сигналом будет низкочастотная составляющая, полученная после первой итерации. После выполнения всех итераций получим масштабированную (уменьшенную) копию исходного сигнала в  $2^J$  раза и  $J$  дополнительных высокочастотных составляющих.

Для двумерного сигнала изображения ДВП выполняется следующим образом. На первой итерации преобразование (7) выполняется последовательно по строкам и столбцам, после чего формируются четыре матрицы (LL, LH, HL, HH) коэффициентов ДВП с размерами в два раза меньшими ширины и высоты изображения. Принятые обозначения матриц следующие: LL – низкочастотная фильтрация по строкам и столбцам; LH – низкочастотная фильтрация по строкам и высокочастотная по столбцам; HL – высокочастотная фильтрация по строкам и низкочастотная по столбцам; HH – высокочастотная фильтрация по строкам и столбцам. На второй итерации выполняется аналогичное преобразование для LL-матрицы и т.д. Обычно достаточно 4–6 итераций. Такой процесс преобразований эквивалентен операции фильтрации сигналов изображения по частотным субдиапазонам для каждой итерации  $j$ .

После выполнения преобразования коэффициенты ДВП квантуются, где для каждого субдиапазона итерации  $j$  выбирается свой постоянный коэффициент  $\Delta$  (шаг квантования), на который делятся все элементы матрицы.

Процесс квантования коэффициентов ДВП для каждой итерации ( $j$ ) можно описать следующими формулами:

$$\widehat{S}_j^{HH} = \frac{S_j^{HH}}{\Delta_j^{HH}}; \quad \widehat{S}_j^{LH} = \frac{S_j^{LH}}{\Delta_j^{LH}}; \quad \widehat{S}_j^{HL} = \frac{S_j^{HL}}{\Delta_j^{HL}}, \quad (8)$$

где  $S_j$  – исходное значение коэффициента;  $\widehat{S}_j$  – квантованное значение. Исключение составляет отрезок «мертвой зоны» около нуля, в которой квантованные вещественные коэффициенты обнуляются с помощью условия:

$$if \quad |S / \Delta| < 1; \quad S = 0; \quad else \quad S = S.$$

Для самого низкочастотного субдиапазона шаг квантования выбирается равным единице ( $\Delta_j^{LH} = 1$ ). Значения коэффициентов  $\Delta_j^{HH}$ ,  $\Delta_j^{LH}$  и  $\Delta_j^{HL}$  образуют шкалу квантования. Коэффициенты квантования рекомендованы стандартом и имеют более высокие значения для матриц НН, что обусловлено спадом пространственно-частотной характеристики зрения в области верхних частот.

В стандарте [4] рекомендовано оптимальное соотношение параметров квантования для разных субдиапазонов:

$$\frac{\Delta_b}{\Delta_0} = \sqrt{\frac{\gamma_0}{\gamma_b}}, \quad (9)$$

где  $\gamma_b$  оценивается среднеквадратичной величиной искажения, которое вносится в изображение при единичном искажении коэффициента из субдиапазона  $b$ . Эта величина может быть определена на основании анализа вейвлет-преобразования «CDF 9/7». Из таблицы значений соотношения шагов квантования (9), представленной в работе [15], следует, что соотношение шагов квантования для соседних субдиапазонов приблизительно равно двум.

В табл. 1 в качестве примера приведена начальная шкала квантования для пяти итераций при заданном шаге для высокочастотного диапазона:  $\Delta_1^{HH} = 32$ .

Табл. 1. Шкала квантования коэффициентов ДВП

| $j$ | $\Delta_j^{HH}$ | $\Delta_j^{LH}$ | $\Delta_j^{HL}$ |
|-----|-----------------|-----------------|-----------------|
| 1   | 32              | 16              | 16              |
| 2   | 16              | 8               | 8               |
| 3   | 8               | 4               | 4               |
| 4   | 4               | 2               | 2               |
| 5   | 2               | 1               | 1               |

В программном модуле цифровой модели кодека предусмотрено изменение шкалы квантования с помощью соотношения  $\Delta_{1,Y}^{HH} = 1 + 2 \cdot (100 - Q)$  для канала яркости и  $\Delta_{1,U,V}^{HH} = 2 \cdot \Delta_{1,Y}^{HH}$  для каналов цветности, где параметр  $Q$  изменяется от 1 до 100 с единичным шагом.

Аналогично выполнению ДКП, квантованные вещественные коэффициенты ДВП нормализуются (6)

и преобразуются к целому типу с сохранением одного знака после запятой.

В процессе квантования происходит обнуление малозначачих коэффициентов трансформации ДВП, что позволяет выполнить эффективное сжатие на следующем этапе. В то же время этот процесс вносит основные потери качества в декодированное изображение. Изменение коэффициентов квантования с помощью параметра  $\Delta_1^{HH}$  позволяет регулировать качество изображения и, следовательно, скорость цифрового потока сжатых визуальных данных.

Оценка эффективности сжатия.

В цифровой модели тестировались различные современные архиваторы на основе алгоритмов семейства LZ (7z, zip, rar, bzip2, gzip и др.), на вход которых подавались файлы коэффициентов трансформации ДКП или ДВП. В результате экспериментов получено, что лучшее сжатие обеспечивает архиватор bzip2 [14]. При этом архиваторы давали разные оценки как по степени сжатия, так и по сравнению эффективности ДКП и ДВП.

Таким образом, сравнение эффективности ДКП и ДВП по результатам сжатия тестового изображения с помощью алгоритмов семейства LZ или других алгоритмов сжатия без потерь не дает однозначных оценок количественного выигрыша того или иного преобразования при заданном качестве. Для решения этого вопроса было предложено использовать классический подход оценки потенциальной возможности сжатия источника на основе вычисления его энтропии

$$H = -\sum_i P(\widehat{S}_i) \log_2 P(\widehat{S}_i), \quad (10)$$

где  $P(\widehat{S}_i)$  – вероятность появления коэффициента трансформации с величиной  $\widehat{S}_i$ . Значение энтропии определяется количеством бит/элемент и, следовательно, позволяет оценить потенциальный коэффициент сжатия изображения с помощью простого выражения  $K_{сж} = 24/H$ . В цифровой модели реализовано вычисление (10), где в качестве источника использованы матрицы коэффициентов трансформации по яркости ( $Y$ ) и по цветности ( $U$  и  $V$ ) с размерами, соответствующими размеру изображения.

## 2. Сравнение эффективности сжатия на основе метрики SSIM

Метрика SSIM, или индекс структурного подобия [16], в настоящее время является одним из наиболее популярных критериев оценки качества изображений. Метрика SSIM определяет степень сходства соответствующих участков (окон) сравниваемых изображений. Разница между двумя окнами  $x$  и  $y$  с размером  $N \times N$  определяется как

$$SSIM_{x,y} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (11)$$

где  $x, y$  – окна оригинального и искаженного изображений;  $\mu_x, \mu_y$  – средние значения цветовых координат пикселей в окнах  $x$  и  $y$ ;  $\sigma_x, \sigma_y$  – среднеквадратичные отклонения;  $\sigma_{x,y}$  – значение ковариации;  $c_1, c_2$  – постоянные коэффициенты. Для цветных изображений параметры  $\mu_x, \sigma_x, \mu_y, \sigma_y$  и  $\sigma_{x,y}$  вычисляются отдельно по сигналам основных ( $R, G, B$ ) цветов оригинального изображения и по сигналам ( $\hat{R}, \hat{G}, \hat{B}$ ) искаженного изображения. Далее параметры усредняются. Как правило, метрика (11) рассчитана на окно с размером  $8 \times 8$  пикселей. Для анализа всего изображения окно смещается на заданное число пикселей последовательно по горизонтали и по вертикали. Интегральная оценка определяется как

$$SSIM = \frac{1}{L} \sum_{l=1}^L SSIM_l, \tag{12}$$

где  $l$  – номер окна;  $L$  – количество окон.

Значения (11) и (12) равны единице только при полном совпадении параметров. Отметим, что существуют различные модификации [16] метрики SSIM, например, MS-SSIM, CW-SSIM и др. В цифровой модели кодека реализован программный модуль – анализатор качества изображений на основе вычислений SSIM.

Анализ эффективности сжатия выполнен по следующей методике. На первом этапе было выбрано приложение Adobe Photoshop CS6, в котором реализованы стандартные алгоритмы JPEG и JPEG2000 с регулируемыми параметрами качества.

К особенностям конверторов относится следующее. Шкала качества JPEG 12-бальная, в которой пользователь может установить значение параметра  $Q$  от 1 до 12 с единичным шагом. При этом считается, что в диапазоне параметров 8, 9 обеспечивается высокое качество, а в диапазоне 10, 11, 12 – максимальное.

Шкала качества JPEG2000 изменяется от 1 до 100 баллов с единичным шагом, при этом максимальное качество обеспечивается в диапазоне 70...100.

Для анализа выбрана шкала качества JPEG с диапазоном параметра  $Q=8,9 \dots 12$ . В процессе исследований тестовое RAW-изображение конвертировалось в формат JPEG и с помощью программного анализатора качества вычислялась зависимость

$$SSIM = f(Q). \tag{13}$$

Далее полученные значения SSIM использовались для настройки шкалы качества JPEG2000. Для каждого значения SSIM оценивался коэффициент сжатия  $K_{сж}$ .

Для примера в табл. 2 показаны зависимости эффективности сжатия JPEG и JPEG2000 от параметра качества  $Q$  и значения SSIM тестовых изображений с различной детальностью: «Лес» (рис. 1а), «Велосипедист» (рис. 1б) и «Собор» (рис. 1в).

В процессе экспериментов были использованы тестовые RAW-изображения с различной детальностью из коллекции фотоизображений фирмы Sony. Для ана-

лиза использовались изображения с исходными форматами 24 или 12 мегапикселей с последующей конвертацией в формат \*.bmp. Визуальная оценка качества декодированных изображений выполнялась с помощью 2К монитора с разрешением экрана  $2560 \times 1440$ . При этом все тестовые изображения масштабировались до разрешения  $1424 \times 2128$  или  $2128 \times 1424$  в зависимости от исходного формата сторон кадра.

Табл. 2. Эффективность JPEG и JPEG2000 в приложении Adobe Photoshop CS6

| «Лес»          |       |       |       |       |       |        |
|----------------|-------|-------|-------|-------|-------|--------|
| $Q$            | 12    | 11    | 10    | 9     | 8     | JPEG   |
| SSIM           | 0,995 | 0,987 | 0,979 | 0,970 | 0,959 |        |
| $K_{сж}$       | 2,43  | 3,98  | 5,73  | 7,83  | 9,94  |        |
| $Q$            | 85    | 69    | 60    | 51    | 46    | JP2000 |
| $K_{сж}$       | 3,00  | 4,53  | 6,04  | 8,02  | 10,10 |        |
| «Велосипедист» |       |       |       |       |       |        |
| $Q$            | 12    | 11    | 10    | 9     | 8     | JPEG   |
| SSIM           | 0,996 | 0,990 | 0,983 | 0,977 | 0,970 |        |
| $K_{сж}$       | 3,65  | 5,87  | 8,45  | 11,26 | 14,07 |        |
| $Q$            | 91    | 75    | 65    | 58    | 50    | JP2000 |
| $K_{сж}$       | 4,38  | 7,05  | 9,71  | 12,06 | 16,92 |        |
| «Собор»        |       |       |       |       |       |        |
| $Q$            | 12    | 11    | 10    | 9     | 8     | JPEG   |
| SSIM           | 0,992 | 0,986 | 0,982 | 0,977 | 0,972 |        |
| $K_{сж}$       | 4,62  | 9,52  | 14,90 | 20,93 | 26,48 |        |
| $Q$            | 82    | 69    | 63    | 57    | 51    | JP2000 |
| $K_{сж}$       | 7,23  | 12,56 | 16,90 | 22,71 | 30,00 |        |



Рис. 1. Примеры тестовых RAW-изображений: а) лес, б) велосипедист, в) собор

Анализ полученных результатов при использовании стандартных алгоритмов Adobe Photoshop CS6 (табл. 2) показывает, что JPEG2000 обеспечивает более высокий коэффициент сжатия и при этом выигрыш

увеличивается с уменьшением детальности. Например, при обеспечении максимального ( $Q=10$ ) качества сжатых изображений выигрыш JPEG2000 составляет: «Лес» – 5,4%, «Велосипедист» – 14,9%, «Собор» – 13,4%. Также наблюдается разброс значений SSIM при фиксированных параметрах квантования.

На втором этапе сравнение эффективности ДКП и ДВП выполнялось по аналогичной методике, где для конвертации сигналов изображения использовалась разработанная цифровая модель кодека. В отличие от Adobe Photoshop CS6 в цифровой модели реализована более точная шкала квантования JPEG, что позволило выполнять настройку качества тестового изображения на любое заданное значение SSIM.

В табл. 3 приведены результаты сравнения ДКП и ДВП по значению энтропии  $H$  для заданных параметров SSIM.

Табл. 3. Эффективность ДКП и ДВП в цифровой модели

| SSIM           | 0,99 | 0,98 | 0,97 | 0,96 |     |
|----------------|------|------|------|------|-----|
| «Лес»          |      |      |      |      |     |
| Q              | 97   | 91   | 80   | 65   | ДКП |
| H              | 8,21 | 5,32 | 3,75 | 2,87 |     |
| Q              | 90   | 78   | 65   | 49   | ДВП |
| H              | 7,08 | 4,52 | 3,38 | 2,62 |     |
| «Велосипедист» |      |      |      |      |     |
| Q              | 91   | 78   | 60   | 35   | ДКП |
| H              | 3,93 | 2,58 | 1,88 | 1,45 |     |
| Q              | 84   | 66   | 42   | 13   | ДВП |
| H              | 3,59 | 2,30 | 1,61 | 1,22 |     |
| «Собор»        |      |      |      |      |     |
| Q              | 96   | 81   | 56   | 16   | ДКП |
| H              | 3,65 | 1,61 | 0,99 | 0,65 |     |
| Q              | 91   | 72   | 41   | 3    | ДВП |
| H              | 3,22 | 1,45 | 0,83 | 0,54 |     |

Выигрыш в коэффициенте сжатия можно оценить отношением значений энтропии при заданном параметре качества SSIM. Например, при высоком качестве ( $SSIM=0,98$ ) выигрыш ДВП составляет: «Лес» – 17,7%, «Велосипедист» – 11,2%, «Собор» – 11,0%.

В случае применения цифровой модели кодека из данных табл. 3 следует, что разброс в оценках эффективности сжатия с уменьшением качества и детальности изображений становится меньше. Для сравнения приведем результаты оценки эффективности при использовании архиватора bzip2. Для заданного параметра  $SSIM=0,98$  получены следующие коэффициенты сжатия: ДКП – «Лес» (4,8), «Велосипедист» (10,7), «Собор» (18,4) и ДВП – «Лес» (5,5), «Велосипедист» (12,2), «Собор» (20,7). Соответственно, выигрыш ДВП составляет: «Лес» – 14,5%, «Велосипедист» – 14,0%, «Собор» – 12,5%.

Таким образом, ДВП обеспечивает лучшую эффективность сжатия во всех случаях. Отметим, что результаты анализа эффективности преобразований ДКП и ДВП в цифровой модели на основе энтропийной оценки являются более объективными, так как

здесь исключено влияние особенностей алгоритмов сжатия без потерь.

К недостаткам метрики SSIM относится следующее. Искажения структурного подобия (11) определяются отклонениями сигналов основных цветов ( $R, G, B$ ) и, следовательно, не учитывают свойства контрастной чувствительности зрения. Также метрика SSIM дает интегральную оценку и не позволяет оценивать локальные характеристики качества изображения, такие как четкость, резкость, точность цветопередачи, зашумленность. Поэтому из данных табл. 2–3 мы не можем однозначно определить, при каком значении SSIM искажения четкости или резкости становятся незаметными или малозаметными глазом.

В следующем параграфе предлагается оценка качества изображения на основе измерений коэффициента четкости.

### 3. Сравнение эффективности сжатия на основе коэффициента четкости

Для измерения четкости по рекомендациям стандарта ISO-12233:2017 [17] используется пространственная частотная характеристика (SFR), которая определяется как функция спада контраста мелких деталей от пространственной частоты  $f_s$ . В дополнение к SFR используется характеристика спада модуляционной функции (MTF) синусоидального сигнала с увеличением пространственной частоты.

Как правило, контраст уменьшается с увеличением пространственной частоты до уровня, при котором мелкие детали больше не различаются глазом. Это предельное значение частоты и определяет реальное разрешение (четкость) изображения, полученное с камеры или после его сжатия и реставрации. Таким образом, разрешение оценивается самой высокой пространственной частотой, на которой контраст мелких деталей ещё различается глазом.

Для измерения SFR используется два метрологических метода – на основе выделения границ и на основе синусоидальных сигналов [18]. Для обоих методов используются специальные тестовые таблицы [19], в изображениях которых присутствуют необходимые объекты для измерения.

В предыдущей авторской работе [20] предложен способ оценки четкости без использования тестовых таблиц, основанный на анализе искажений активных пикселей реальных фотореалистичных изображений. В настоящей работе этот способ с небольшими модификациями использован для сравнения эффективности ДКП и ДВП. Рассмотрим особенности предлагаемого метода измерения четкости.

В [20] введено определение активного пикселя: пиксель с пространственными координатами  $(i, j)$  будет считаться активным, если его цветовой контраст относительно двух соседних пикселей в горизонтальном, вертикальном или в диагональных направлениях будет превышать нормированный пороговый контраст зрения.

Вычисление цветового контраста выполняется в нормированной колориметрической системе N-CIELAB на основе выражения [21]:

$$K_{j,i} = \sqrt{(\Delta\bar{L}_{j,i})^2 + (\Delta\bar{a}_{j,i})^2 + (\Delta\bar{b}_{j,i})^2} > 1, \quad (14)$$

где  $\Delta\bar{L}_{i,j} = \Delta L_{i,j}^* / L_\delta$ ,  $\Delta\bar{a}_{i,j} = \Delta a_{i,j}^* / a_\delta$  и  $\Delta\bar{b}_{i,j} = \Delta b_{i,j}^* / b_\delta$  – нормированные разностные значения цветовых координат в системе N-CIELAB;  $L_\delta$ ,  $a_\delta$  и  $b_\delta$  – весовые коэффициенты по светлоте и цветности, заданные количеством минимальных цветовых порогов МЦП в зависимости от размера ( $\delta$ ) детали. Система N-CIELAB позволяет оценивать цветовой контраст деталей в зависимости от их размера с учетом контрастной чувствительности зрения. Экспериментально доказано [21], что при значении цветового контраста  $K_{i,j} \approx 1$  мелкие детали начинают различаться глазом, и этот порог мало зависит от цвета детали и яркости окружающего фона. В (14) использованы следующие значения весовых коэффициентов:  $L_1 \approx 6$ ,  $a_1 \approx 40$  и  $b_1 \approx 55$ , что соответствует самым мелким деталям с размером в один пиксель ( $\delta = 1$ ).

Для идентификации активного пикселя используется скользящее окно в виде блока с размером  $3 \times 3$  пикселя, в центре которого идентифицируется активный пиксель по следующему алгоритму.

На первом этапе последовательно вычисляется контраст центрального пикселя относительно двух соседних: слева-справа по горизонтали ( $p1$ ); сверху-снизу по вертикали ( $p2$ ); снизу-сверху по диагонали ( $p3$ ); сверху-снизу по диагонали ( $p4$ ).

Параметр  $p$  устанавливается в единичное значение, если выполняется условие (14), и при этом значения градиента центрального пикселя относительно двух соседних будут одинаковыми. В противном случае параметр  $p$  устанавливается в нулевое значение.

На втором этапе выполняется анализ блока по значению параметров  $P = (p1 p2 p3 p4)$ , которое позволяет идентифицировать активный пиксель и оценить структуру блока. Если все значения  $p$  равны нулю  $P = (0 0 0 0)$ , то это будет означать, что такой пиксель не будет различаться глазом на фоне соседних пикселей и ему присваивается статус «пассивный». Пассивные пиксели относятся к фоновым или мало изменяющимся областям изображения, а также к фрагментам их границ.

Если любой из параметров  $p$  не равен нулю, т.е. четырехбитный код  $P > 0$ , пикселю присваивается статус «активный». При этом в изображении маркируются все пиксели в окне ( $3 \times 3$ ) с координатами центрального пикселя ( $i, j$ ).

После анализа всего изображения выполняется подсчет активных пикселей  $N_a$ .

На рис. 2а показан увеличенный фрагмент тестового изображения «Лес» и на рис. 2б этот же фрагмент с маркированными пикселями.

Комбинация бит в кодовом слове позволяет идентифицировать тип фрагмента мелкой структуры. Например, если  $P = (1 1 1 1)$ , то это будет означать, что обнаружен точечный объект. Если  $P = (0 1 1 1)$ , то это будет фрагмент тонкой горизонтальной линии. Код  $P = (0 0 1 1)$  означает фрагмент «шахматное поле».

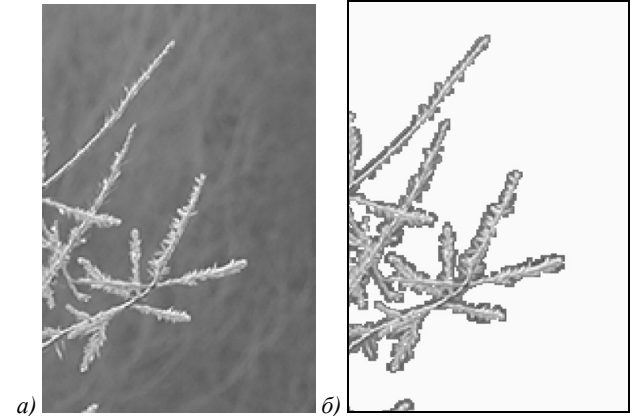


Рис. 2. Фрагмент тестового изображения

Оценка четкости основана на сравнении кодов активных блоков оригинального ( $P$ ) и искаженного ( $\tilde{P}$ ) изображений. Если выполняется условие

$$\tilde{P} = P, \quad (15)$$

то принимается решение о том, что структура блока искаженного изображения подобна структуре оригинального блока.

Если условие (15) не выполняется, то принимается решение о том, что в изображении блока произошли структурные изменения. В частности, если для оригинального блока  $P > 0$ , а для искаженного  $\tilde{P} = 0$ , то это будет означать, что произошло полное размытие изображения мелких деталей. Для оценки четкости предложено использовать следующий коэффициент:

$$R_d = \tilde{N}_a / N_a, \quad (16)$$

где  $\tilde{N}_a$  – количество активных пикселей в искаженном изображении, вычисленное на основании выполнения условия (15).

Значение коэффициента  $R_d$  позволяет оценивать ухудшение визуальной четкости оригинального изображения, возникающее в процессе его цифрового сжатия с потерями. Очевидно, что чем ближе значение  $R_d$  к единице, тем ближе качество восстановленного изображения к оригинальному.

Квантование и обнуление малозначимых коэффициентов трансформации ДКП или ДВП приводит к искажениям мелких структур изображения, что будет снижать коэффициент четкости (16) за счет частичного или полного размытия мелких структур блока.

Здесь возникает вопрос, при каком значении  $R_d$  искажения мелких структур становятся незаметными для глаза?



Для решения этого вопроса использована авторская метрика [21] оценки искажений мелких структур изображений (MFSD):

$$dE_a = \frac{1}{N_{mark}} \sum_{m=1}^{N_{mark}} dE_m, \tag{17}$$

$$dE_m = \max_{i,j} |K_{m,j,i} - \tilde{K}_{m,j,i}|,$$

где значения  $K_{m,j,i}$  вычисляются по формуле (14) для идентифицированного блока  $m$  оригинального изображения и  $\tilde{K}_{m,j,i}$  для блока искаженного изображения;  $N_{mark}$  – количество маркированных блоков с мелкими структурами. Выражение  $dE_m$  оценивает максимальное отклонение контраста цветовых переходов оригинального и искаженного изображений в микроблоке с номером  $m$ .

Из результатов исследований [21] следует, что при значении  $dE_a \approx 0,5$  искажения становятся незаметными (малозаметными) для глаза.

Идентификация мелких структур выполняется на основе разбиения изображения на блоки с размером  $3 \times 3$  пикселя и анализа контраста (14) каждого блока. Если выполняется условие  $K_{m,j,i} > 1$ , то принимается решение о том, что в блоке присутствуют мелкие структуры, различимые глазом.

Относительное количество маркированных блоков оригинального изображения позволяет оценить его коэффициент детальности:

$$FDL = \frac{9 \cdot N_{mark}}{W \cdot H}. \tag{18}$$

Значение коэффициента детальности (18) фотореалистичных RAW-изображений зависит от их структуры, а также от характеристик камеры, в первую очередь, от точности фокусировки. В частности, коэффициент детальности тестового изображения «Лес» равен 51 %, «Велосипедист» – 29 % и «Собор» – 12 %.

Обнуление коэффициентов трансформации ДКП или ДВП также приводит к искажениям остальных фоновых или мало изменяющихся участков изображения, к которым относятся пассивные пиксели. Для оценки таких искажений использован дополнительный параметр:

$$dE_F = \frac{1}{N_F} \sum_{k=1}^{N_F} dE_k, \tag{19}$$

где значение

$$dE_k = \frac{1}{9} \sum_{i=1}^9 \sqrt{(L_{k,i}^* - \tilde{L}_{k,i}^*)^2 + (a_{k,i}^* - \tilde{a}_{k,i}^*)^2 + (b_{k,i}^* - \tilde{b}_{k,i}^*)^2}$$

оценивает среднее отклонение цветовых координат оригинального и искаженного изображения в системе CIELAB [13] внутри блока ( $k$ ), в котором отсутствуют различимые глазом мелкие структуры;  $i$  – номер пикселей в блоке ( $3 \times 3$ ). Выражение (19) дает среднюю оценку по всем «фоновым» блокам  $N_F$ .

К особенностям измерения искажений в системе CIELAB относится следующее: если  $dE \leq 2 \dots 3$  МЦП, то считается, что цветовые различия незаметны глазом. Отметим, что в отличие от оценки цветового контраста (14) в системе N-CIELAB, в системе CIELAB в формуле (19) не используются весовые коэффициенты по светлоте и цветности.

Для анализа эффективности ДКП и ДВП в цифровую модель кодека добавлен программный модуль, реализующий вычисления: коэффициента четкости (16); коэффициента детальности (18); искажений (17) и (19).

В процессе экспериментов параметры квантования ДКП или ДВП настраивались таким образом, чтобы после обратного преобразования и восстановления изображения получить заданное значение  $R_d$ . Далее для выбранных параметров квантования оценивалось значение энтропии (10).

В табл. 4 показаны зависимости энтропии и искажений тестовых изображений от коэффициента четкости  $R_d$  (15). Эффективность  $E$  оценивалась отношением значений энтропии при заданном  $R_d$ . Для сравнения в таблице приведены параметры оценки качества SSIM.

Табл. 4. Эффективность ДКП и ДВП в зависимости от коэффициента четкости

| $R_d$                       | 0,90  | 0,85  | 0,80  | 0,75  | 0,70  |     |
|-----------------------------|-------|-------|-------|-------|-------|-----|
| «Лес» $FDL = 51\%$          |       |       |       |       |       |     |
| Q                           | 95    | 90    | 85    | 79    | 73    | ДКП |
| H                           | 6,77  | 5,09  | 4,27  | 3,66  | 3,26  |     |
| dEa                         | 0,15  | 0,26  | 0,35  | 0,46  | 0,56  |     |
| dEf                         | 1,99  | 2,41  | 2,65  | 2,94  | 3,23  |     |
| SSIM                        | 0,986 | 0,979 | 0,974 | 0,969 | 0,965 |     |
| Q                           | 89    | 82    | 74    | 66    | 58    | ДВП |
| H                           | 6,76  | 5,01  | 4,11  | 3,45  | 2,99  |     |
| dEa                         | 0,17  | 0,26  | 0,37  | 0,47  | 0,55  |     |
| dEf                         | 1,67  | 2,11  | 2,44  | 2,6   | 2,81  |     |
| SSIM                        | 0,990 | 0,982 | 0,977 | 0,971 | 0,966 |     |
| E%                          | 0,1   | 1,6   | 3,9   | 6,1   | 9,0   |     |
| «Велосипедист» $FDL = 29\%$ |       |       |       |       |       |     |
| Q                           | 93    | 88    | 82    | 76    | 69    | ДКП |
| H                           | 4,37  | 3,42  | 2,78  | 2,44  | 2,14  |     |
| dEa                         | 0,17  | 0,27  | 0,38  | 0,47  | 0,57  |     |
| dEf                         | 1,43  | 1,69  | 1,98  | 2,29  | 2,47  |     |
| SSIM                        | 0,992 | 0,987 | 0,983 | 0,978 | 0,974 |     |
| Q                           | 88    | 81    | 72    | 64    | 53    | ДВП |
| H                           | 4,18  | 3,27  | 2,60  | 2,22  | 1,86  |     |
| dEa                         | 0,19  | 0,28  | 0,38  | 0,46  | 0,57  |     |
| dEf                         | 0,89  | 1,08  | 1,27  | 1,42  | 1,59  |     |
| SSIM                        | 0,993 | 0,988 | 0,983 | 0,979 | 0,974 |     |
| E%                          | 4,5   | 4,6   | 6,9   | 9,9   | 15,1  |     |
| «Собор» $FDL = 12\%$        |       |       |       |       |       |     |
| Q                           | 96    | 93    | 89    | 85    | 78    | ДКП |
| H                           | 3,65  | 2,75  | 2,17  | 1,84  | 1,49  |     |
| dEa                         | 0,09  | 0,16  | 0,23  | 0,27  | 0,35  |     |
| dEf                         | 0,93  | 1,31  | 1,72  | 1,95  | 2,27  |     |
| SSIM                        | 0,991 | 0,986 | 0,984 | 0,982 | 0,978 |     |
| Q                           | 92    | 88    | 83    | 76    | 68    | ДВП |
| H                           | 3,51  | 2,61  | 2,04  | 1,69  | 1,32  |     |
| dEa                         | 0,10  | 0,17  | 0,28  | 0,27  | 0,34  |     |
| dEf                         | 0,79  | 1,03  | 1,12  | 1,21  | 1,29  |     |
| SSIM                        | 0,993 | 0,988 | 0,985 | 0,982 | 0,978 |     |
| E%                          | 3,9   | 5,3   | 6,3   | 8,8   | 12,9  |     |

Анализ полученных результатов позволяет сделать следующие выводы. Высокое качество изображений на основе метрики MFSD ( $dE_a < 0,5$  и  $dE_f < 3,0$ ) обеспечивается при коэффициенте четкости  $R_d > 0,75$ . При этом, как показывают визуальные оценки, искажения практически незаметны глазом. Лучшую эффективность обеспечивает ДВП, которая увеличивается с ухудшением четкости. Для всех значений  $R_d$ , по результатам оценки параметра  $dE_f$ , лучшее качество передачи фоновых или мало изменяющихся участков изображения обеспечивает ДВП.

При настройке кодера на фиксированный параметр четкости  $R_d$  наблюдается разброс оценок по метрике SSIM в зависимости от детальности RAW-изображений. Из данных табл. 4 следует, что при  $R_d = 0,80$  значение SSIM равно 0,974 для изображения «Лес», 0,983 для «Велосипедиста» и 0,985 для «Собора». При одинаковом коэффициенте четкости лучшее качество по метрике SSIM обеспечивается при сжатии изображений с более низкой детальностью.

Таким образом, оценка качества по коэффициенту четкости позволяет более объективно оценивать количественный выигрыш ДВП и настраивать систему сжатия RAW-изображений на высокую четкость передачи мелких структур изображений.

### Заключение

В результате исследования эффективности дискретного косинус- и вейвлет-преобразований в цифровой модели сжатия RAW-изображений получены объективные количественные оценки, доказывающие преимущество использования ДВП. Из результатов экспериментов следует, что в целом сжатие коэффициентов ДВП обеспечивает выигрыш от 4 % до 30 % в зависимости от детальности RAW-изображений, при условии обеспечения высокой четкости  $R_d > 0,75$  с минимальными искажениями.

У исследователей в области обработки изображений может возникнуть вопрос, зачем нужны такие точные оценки, когда и так известно, что, во-первых, JPEG2000 лучше и, во-вторых, если искажения незаметны, зачем проводить сравнение?

Цель результатов данного исследования – это использование точных оценок в перспективных системах кодирования и передачи высококачественных изображений. Во-первых, точные оценки позволяют определить объективные границы «качество – сжатие», при которых искажения ещё не заметны, и настроить шкалу квантования в зависимости от детальности изображения. Во-вторых, шкала квантования может быть адаптивна к различным классам изображения, что позволит повысить эффективность сжатия при передаче больших объемов данных.

К перспективным направлениям развития данного исследования относится следующее:

- применение нейросетевых методов сегментации и классификации участков изображения по признакам их детальности;
- повышение качества с помощью нейросетевых методов [22] коррекции искажений мелких структур и резких границ изображения;
- развитие нейросетевых методов SR CNN [23] и алгоритмов повышения качества масштабируемых изображений в формат с более высоким разрешением.

### Благодарности

Исследование выполнено при финансовой поддержке Российского научного фонда в рамках научного проекта № 22-21-00394 «Развитие нейросетевых методов повышения качества передачи цифровых изображений в интеллектуальных видеосистемах».

### References

- [1] Digital camera sensors. Source: <https://www.cambridgeincolour.com/tutorials/camera-sensors.htm>.
- [2] Pennebaker WB, Mitchell JL. JPEG still image data compression standard. New York, USA: Springer; 1992.
- [3] Drozdov SN, Zhiglaty AA, Kravchenko PP, Lutai VN, Skorokhod SN, Khusainov NS. JPEG2000 standard: basic algorithms, implementation examples, and application prospects [In Russian]. Rostov-on-Don: Publishing House of SFU; 2014.
- [4] Taubman D, Marcellin MD. JPEG2000: Image compression fundamentals, standard and practice. Kluwer Academic Publishers; 2002.
- [5] Kapustin VV, Kamenskiy AV. Estimation of intra-frame compression quality of a high-resolution image for JPEG and JPEG2000 Standards. Proceedings of TUSUR University 2016; 19(3): 27-31.
- [6] Elamaran E, Praveen P. Comparison of DCT and wavelets in image coding. 2012 Int Conf on Computer Communication and Informatics 2012: 1-4.
- [7] Jain K, Agrawal V. A image comparative study using DCT, fast Fourier, wavelet transforms and Huffman algorithm. Int J Eng Res Generic Sci 2015; 3(4): 594-599.
- [8] Jeengar V, Omkar S, Singh A, Yadav M, Keshri S. A review comparison of wavelet and cosine image transforms. Int J Image Graph Signal Process 2012; 11: 16-25.
- [9] Xiong Z, Ramchandran K, Orchard M, Zhang Y-Q. A comparative study of DCT- and wavelet-based image coding. IEEE Trans Circuits Syst Video Technol 1999; 9(5): 692-695.
- [10] Mehala R. Comparison of DCT and DWT in image compression techniques. Int J Adv Res Trends Eng Technol 2016; 3(20): 291-295.
- [11] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. IEEE Trans Image Process 2004; 13(4): 600-612.
- [12] Sony Alpha ILCE-A6000 Gallery. Source: <https://www.imaging-resource.com/cameras/sony/reviews/>.
- [13] Fairchild MD. Color appearance models. John Wiley and Sons; 2005.
- [14] Bzip2 and libzip2. Source: <http://bzip2.org/>.
- [15] Balster EJ, Fortener BT, Turri WF. Post-compression rate-distortion development for embedded block coding with optimal truncation in JPEG2000 imagery. Int J Image Graph 2011; 11(4): 611-627.
- [16] Lin W, Kuo C-CJ. Perceptual visual quality metrics: A survey. J Vis Commun Image Represent 2011; 22(4): 297-312.

- [17] ISO 12233:2017 Photography – Electronic still picture imaging – Resolution and spatial frequency responses. Source: (<https://www.iso.org/standard/71696.html>).
- [18] Imatest. Source: (<https://www.imatest.com/>).
- [19] High resolution test patterns. Source: (<http://www.bealecorner.org/red/test-patterns/>).
- [20] Sai SV. A method for assessing photorealistic image quality with high resolution. *Computer Optics* 2022; 46(1): 121-129. DOI: 10.18287/2412-6179-CO-899.
- [21] Sai SV. Metric of fine structures distortions of compressed images. *Computer Optics* 2018; 42(5): 829-837. DOI: 10.18287/2412-6179-2018-42-5-829-837.
- [22] Svoboda P, Hradis M, Barina D, Zemcik P. Compression artifacts removal using convolutional neural networks. 2016. arXiv Preprint. Source: (<https://arxiv.org/abs/1605.00366>).
- [23] Wang Z, Chen J, Steven CHH. Deep learning for image super-resolution: A survey. 2020. arXiv Preprint. Source: (<https://arxiv.org/abs/1902.06068>).

#### **Сведения об авторах**

**Сай Сергей Владимирович**, 1960 года рождения, в 1983 году окончил Томский институт автоматизированных систем управления и радиоэлектроники (ТИАСУР) по специальности «Радиоэлектронные устройства». Доктор технических наук, профессор, заведующий кафедрой вычислительной техники Тихоокеанского государственного университета (ТОГУ). Область научных интересов: объективные методы анализа изображений, распознавание образов. E-mail: [000493@pnu.edu.ru](mailto:000493@pnu.edu.ru).

**Зинкевич Алексей Владимирович**, 1985 года рождения, в 2007 году окончил Хабаровский институт инфокоммуникаций (ХабИИК) по специальности «Радиосвязь, радиовещание и телевидение». Кандидат технических наук, доцент кафедры вычислительной техники Тихоокеанского государственного университета (ТОГУ). Область научных интересов: обработка изображений, видеосистемы на кристалле. E-mail: [006526@pnu.edu.ru](mailto:006526@pnu.edu.ru).

**Фомина Екатерина Сергеевна**, 1993 года рождения, в 2015 году окончила Тихоокеанский государственный университет (ТОГУ) по специальности «Многоканальные телекоммуникационные системы». Преподаватель кафедры вычислительной техники Тихоокеанского государственного университета (ТОГУ). Область научных интересов: цифровая обработка сигналов, кодирование и передача информации. E-mail: [009373@pnu.edu.ru](mailto:009373@pnu.edu.ru).

ГРНТИ: 28.23.15

Поступила в редакцию 11 января 2022 г. Окончательный вариант – 15 апреля 2022 г.

---

# Comparison of discrete cosine and wavelet transforms in RAW image compression systems

S.V. Sai<sup>1</sup>, A.V. Zinkevich<sup>1</sup>, E.S. Fomina<sup>1</sup>  
<sup>1</sup> Pacific National University, Khabarovsk, Russia

## Abstract

The article describes features of digital processing of image signals in the process of coding based on discrete cosine (DCT) and wavelet transforms (DWT) that are used in the JPEG and JPEG2000 compression standards. To compare DCT and DWT, a digital model of the system has been developed that implements the same stages of signal processing, except for the stages of the proper discrete transforms. A method for analyzing the efficiency of the transformations based on objective assessments of image quality depending on the compression ratio is proposed. The peculiarities include the fact that, in contrast to the popular PSNR and SSIM metrics, it is proposed that the quality be assessed using the reduction factor for the RAW image format, the calculation of which is associated with the contrast sensitivity of vision. As a result of the research, quantitative estimates of the compression efficiency are obtained for the given quality parameters, depending on the type of conversion and the detail of the RAW images. Recommendations are made regarding the development of methods for increasing the efficiency of image compression based on DWT or DCT.

**Keywords:** image analysis, distortion metric, discrete cosine transform, discrete wavelet transform, compression efficiency.

**Citation:** Sai SV, Zinkevich AV, Fomina ES. Comparison of discrete cosine and wavelet transforms in RAW image compression systems. *Computer Optics* 2022; 46(6): 929-938. DOI: 10.18287/2412-6179-CO-1094.

**Acknowledgements:** This work was supported by the Russian Science Foundation (Project No. 22-21-00394) "Development of neural network methods for improving the quality of digital image transmission in intelligent video systems".

---

## Authors' information

**Sergey Vladimirovich Sai**, born in 1960, in 1983 graduated from the Tomsk Institute of Automated Control Systems and Radioelectronics (TIASUR), specializing in Radioelectronics Devices. Doctor of Technical Sciences, Professor, Head of Computing Technology department, Pacific National University (PNU). Area of scientific interests: objective methods of image analysis, pattern recognition. E-mail: [000493@pnu.edu.ru](mailto:000493@pnu.edu.ru).

**Alexey Vladimirovich Zinkevich**, born in 1985, in 2007 graduated from the Khabarovsk Institute of Infocommunications (KhabIIK), specializing in Radio Communication, Radio Broadcasting and Television. Candidate of Technical Sciences, Associate Professor of Computer Engineering department, Pacific National University (PNU). Research interests: image processing, video systems on a chip. E-mail: [006526@pnu.edu.ru](mailto:006526@pnu.edu.ru).

**Ekaterina Sergeevna Fomina**, born in 1993, in 2015 graduated from Pacific National University (PNU), specializing in Multichannel Telecommunication Systems. Lecturer at Computer Science department, Pacific National University (PNU). Research interests: digital signal processing, coding and transmission of information. E-mail: [009373@pnu.edu.ru](mailto:009373@pnu.edu.ru).

---

Received January 11, 2022. The final version – April 15, 2022.

---