

# Система обучения сжатию изображений с управляемыми потерями

С.Н. Кочергин, А.Р. Мамлеева

*Астраханский государственный технический университет, Астрахань, Россия*

**Аннотация** – В работе представлена интеллектуальная система для обучения методам сжатия изображений с управляемыми потерями. Актуальность исследования обусловлена ростом объёмов визуальных данных и необходимостью эффективного их хранения и передачи. Разработанная система представляет собой интерактивное веб-приложение, которое визуализирует принципы работы алгоритмов сжатия, таких как JPEG, и позволяет пользователю экспериментировать с ключевыми параметрами, влияющими на соотношение «степень сжатия – качество изображения». В рамках системы реализованы модули для настройки коэффициента квантования, выбора цветового пространства и оценки метрик искажения (PSNR, SSIM). Проведенные эксперименты демонстрируют, что интерактивное обучение с использованием предложенной системы способствует более глубокому пониманию компромиссов, присущих сжатию с потерями по сравнению с традиционным изучением теоретического материала на лекции. Система предназначена для использования в образовательных целях студентами, изучающими компьютерную графику, обработку изображений и мультимедийные технологии.

**Ключевые слова**–сжатие изображений, сжатие с потерями, JPEG, обучение, интеллектуальная система, PSNR, SSIM.

## I. ВВЕДЕНИЕ

Рост разрешения и объёма цифровых изображений требует применения эффективных алгоритмов сжатия. В то время как сжатие без потерь (например, PNG) сохраняет все исходные данные, сжатие с управляемыми потерями (например, JPEG) достигает значительно более высоких коэффициентов сжатия за счёт необратимого отбрасывания части визуальной информации, малозаметной для человеческого восприятия [1, 2]. Понимание принципов и компромиссов такого сжатия является важной задачей в подготовке специалистов в области информационных технологий.

Существующие образовательные материалы зачастую носят теоретический характер и не предоставляют возможности для наглядного экспериментирования с параметрами сжатия [3]. Это затрудняет усвоение взаимосвязи между степенью сжатия, возникающими артефактами и объективными метриками качества. Целью данной работы является разработка интеллектуальной системы, которая

предоставляет интерактивную среду для обучения сжатию изображений с управляемыми потерями. Задачи исследования включают реализацию основных этапов алгоритма JPEG, создание веб-интерфейса для управления параметрами сжатия, интеграцию метрик оценки качества и педагогическую оценку эффективности системы.

## II. ПОСТАНОВКА ЗАДАЧИ

Разрабатываемая система должна решать следующие задачи:

1. Предоставить пользователю возможность загружать собственные изображения для обработки.
2. Реализовать визуализацию основных этапов алгоритма JPEG: преобразование цветового пространства, дискретное косинусное преобразование (ДКП) [4], квантование и энтропийное кодирование.
3. Обеспечить интерактивное управление ключевыми параметрами, такими как:
  - коэффициент качества, влияющий на матрицу квантования;
  - выбор между цветовыми пространствами RGB и YCbCr.
4. Рассчитывать и отображать объективные метрики качества реконструированного изображения: пиковое отношение сигнал-шум PSNR (Peak Signal-to-Noise Ratio) и индекс структурного сходства SSIM (Structural Similarity Index Measure).
5. Обеспечить сравнительный анализ исходного и сжатого изображений.

## III. ТЕОРИЯ

### *A. Алгоритмическая основа*

Основой системы является классический алгоритм JPEG [5]. Его этапы для серого изображения или отдельного канала цветного изображения могут быть описаны следующим образом.

1. Перевод в цветное пространство YCbCr: YCbCr – это цветное пространство, которое разделяет информацию на яркость (Y) и две цветоразностные компоненты (Cb и Cr). Оно широко используется в

системах обработки видео и сжатия изображений, так как позволяет эффективно уменьшать избыточность данных. Человеческий глаз менее чувствителен к изменениям цвета, чем к яркости, поэтому компоненты С<sub>b</sub> и С<sub>t</sub> часто подвергаются субдискретизации для экономии места без заметной потери качества [6].

2. Разбиение на блоки: изображение разделяется на блоки размером 8×8 пикселей.

3. Дискретное косинусное преобразование (ДКП): для каждого блока вычисляется ДКП, которое переводит сигнал из пространственной области в частотную и вычисляется по формуле:

$$F(u, v) = \frac{1}{4} C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \times \cos \left[ \frac{(2x+1)u\pi}{16} \right] \times \cos \left[ \frac{(2y+1)v\pi}{16} \right], \quad (1)$$

где  $f(x, y)$  – значение яркости (или цвета) пикселя в позиции  $(x, y)$  внутри блока 8×8. Перед применением ДКП значения пикселей сдвигаются из диапазона [0, 255] в [-128, 127].

$F(u, v)$  – результирующий коэффициент ДКП в частотной области на позиции  $(u, v)$ . Это коэффициенты, которые затем подвергаются квантованию и сжатию.

$x$  и  $y$  – координаты в пространственной области (координаты пикселя),  $0 \leq x, y \leq 7$ .

$u$  и  $v$  – координаты в частотной области (координаты коэффициента),  $0 \leq u, v \leq 7$ .

**Квантование:**

Коэффициенты ДКП делятся на элементы матрицы квантования и округляются до целых чисел:

$$F_q(u, v) = \text{round} \left( \frac{F(u, v)}{Q(u, v)} \right), \quad (2)$$

где  $F(u, v)$  – коэффициент ДКП на позиции  $(u, v)$ ;  $Q(u, v)$  – значение из таблицы квантования на позиции  $(u, v)$ .

Матрица квантования является основным инструментом управления степенью сжатия и потерями.

**Энтропийное кодирование:**

Квантованные коэффициенты кодируются с использованием методов RLE и Хаффмана для получения итогового сжатого битового потока [7].

**В. Метрики качества**

Для количественной оценки искажений используются две метрики  $PSNR$  (дБ) и  $SSIM$  [8, 9].  $PSNR$  даёт соотношение между максимально возможной мощностью сигнала и мощностью шума (ошибки), внесенной сжатием:

$$PSNR = 10 \times \lg \left( \frac{MAX_i^2}{MSE} \right) \quad (3)$$

где  $MAX$  – максимально возможное значение пикселя (например, 255),  $MSE$  – среднеквадратическая ошибка между исходным и сжатым изображениями.

$SSIM$  – это метрика оценки качества изображений, которая лучше соответствует человеческому восприятию, чем  $PSNR$ .

$SSIM$  анализирует три аспекта:

- Яркость (Luminance) – сравнение средней яркости
- Контраст (Contrast) – сравнение дисперсии (контрастности)
- Структуру (Structure) – сравнение паттернов и текстур

$$SSIM(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (4)$$

где  $\mu_x$  – среднее значение пикселей изображения  $x$  (оценка яркости),  $C_1$  – константа для стабилизации (чтобы избежать деления на ноль).  $C_2 = (k^2 L)^2$ ,  $k^2$  обычно равно 0,03,  $\sigma_x$  – стандартное отклонение пикселей изображения  $x$  (оценка контраста).

**IV. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ**

Пользовательский интерфейс разделен на две основные колонки (Рис. 1). В левой части отображается оригинал изображения с метаданными (размер, формат), в правой – результат сжатия. Загрузка осуществляется через боковую панель.

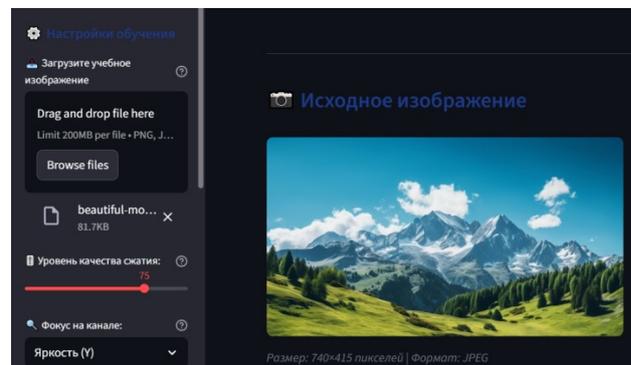


Рис. 1. Главное меню

После загрузки RGB-изображения программа автоматически преобразует его в цветовое пространство YCbCr. Программа визуально демонстрирует, как выглядит каждый канал в отдельности (Рис. 2).

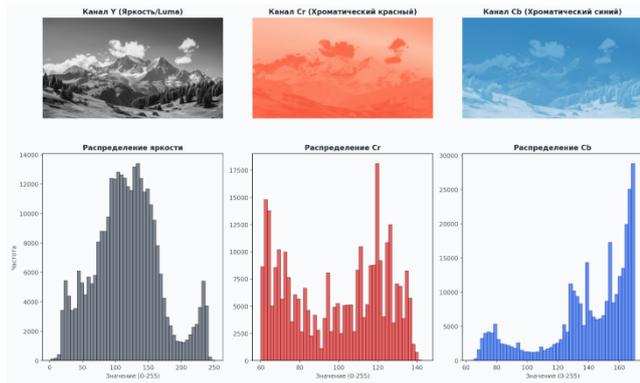


Рис. 2. Разделение на каналы

Для каналов Cb и Cr применяется понижающая дискретизация 4:2:0 (Рис. 3).

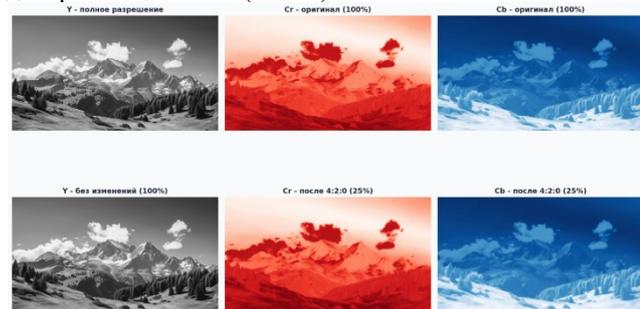


Рис. 3. Понижающая дискретизация

Размер каналов цветности уменьшается в два раза по горизонтали и вертикали. На каждый блок  $2 \times 2$  пикселя приходится одно значение Cb и Cr. Канал Y остается без изменений.

Этот этап дает первое значительное сокращение данных (примерно 50 %) без заметной потери качества.

Каждый канал разбивается на непересекающиеся блоки размером  $8 \times 8$  пикселей (Рис. 4).



Рис. 4. Разбиение на блоки

Для изображения  $512 \times 512$  пикселей получается  $64 \times 64 = 4096$  блоков в канале Y.

В каналах Cb и Cr после прореживания –  $32 \times 32 = 1024$  блока каждого.

Программа позволяет выбрать любой блок для детального изучения и показывает его числовые значения.

К каждому блоку  $8 \times 8$  применяется ДКП (Рис. 5). ДКП преобразует сигнал из пространственной области в частотную. Вместо хранения значений пикселей хранятся коэффициенты косинусных волн разной частоты.

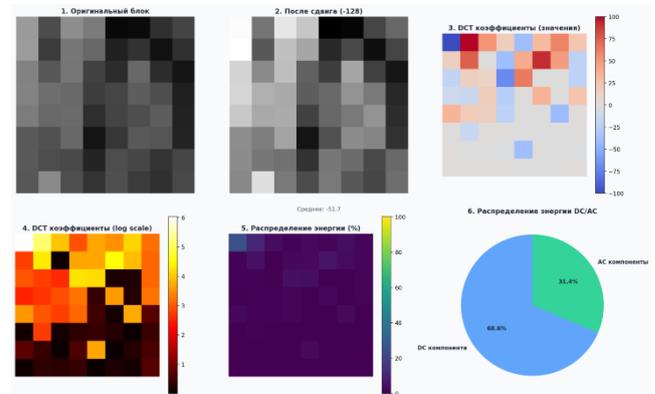


Рис. 5. ДКП для одного блока

Можно выделить высокие частоты, к которым человек менее восприимчив и удалить их без особого ущерба для качества изображения с помощью умножения на матрицу квантования (Рис. 6).

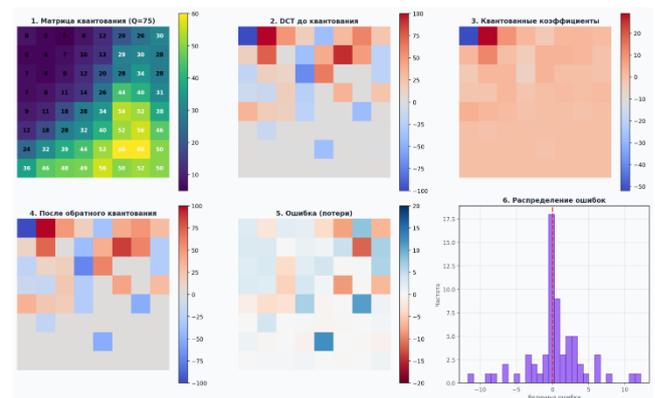


Рис. 6. Применение матрицы квантования

После квантования матрица коэффициентов ДКП содержит множество нулей, особенно в области высоких частот. Чтобы эффективно упаковать эти данные, применяется цепочка алгоритмов без потерь.

1. Зигзаг-сканирование переупорядочивает коэффициенты в одномерный вектор, концентрируя нули в его конце.

2. Кодирование длин серий (RLE) заменяет цепочки нулей компактными кодами вида (количество\_нулей, следующее\_значение).

3. Кодирование Хаффмана – это финальный этап, на котором полученные символы (пары RLE и разницы DC) заменяются на двоичные коды переменной длины. Частые символы кодируются короткими битовыми последовательностями, редкие – длинными. Именно это и обеспечивает финальное сжатие потока данных.

Таким образом, этап энтропийного кодирования «упаковывает» уже отфильтрованную и

упорядоченную информацию, удаляя избыточность без изменения самих числовых значений.

В конце можно просмотреть метрики получившегося сжатого изображения (Рис. 7).

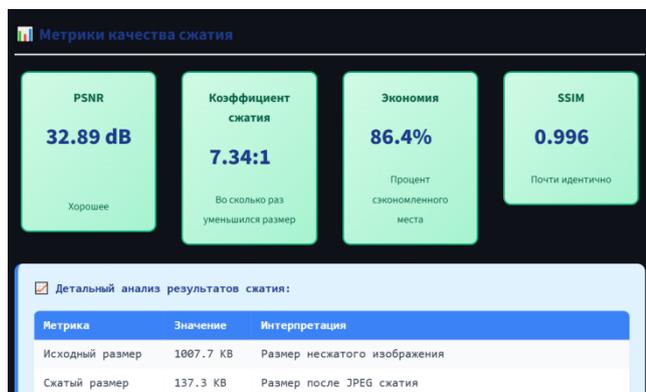


Рис. 7. Метрики качества

## V. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

### *Влияние коэффициента качества*

Было проведено сжатие изображения размером 2200 Кбайт с коэффициентами качества 10 %, 50 % и 90 %. Результаты представлены в Табл. I.

ТАБЛИЦА I  
ВЛИЯНИЕ КОЭФФИЦИЕНТА КАЧЕСТВА НА ПАРАМЕТРЫ СЖАТИЯ

Коэффициент качества	Размер файла, Кбайт	PSNR, дБ	SSIM
10 %	46,7	27,5	0,81
50 %	154,3	33,1	0,95
90 %	356,3	36,92	0,99

Таблица показывает ожидаемую зависимость: с ростом качества увеличивается размер файла. Параметр SSIM показывает наибольшее визуальное сходство с оригиналом при наибольшем коэффициенте качества.

На Рис. 8 представлено исходное изображение.



Рис. 8. Исходное изображение

При сжатии до 90 % качества (Рис. 9) отсутствуют заметные визуальные изменения, что подтверждает SSIM, но размер изображения уменьшился в шесть раз.



Рис. 9. Качество сжатия 90 %

При сжатии до 50 % (Рис. 10) визуальное качество также остается на хорошем уровне, а размер изображения уменьшается в 14 раз.



Рис. 10. Качество сжатия 50%

При сжатии до 10 % (Рис.11) можно заметить типичные артефакты сжатия алгоритма JPEG:

- Блочная пикселизация – на картинке становятся заметны мелкие или крупные квадратные блоки, залитые различными оттенками или градиентами. Они появляются, когда в силу экстремального сжатия пиксели похожих цветов «склеиваются» друг с другом.
- Шумовые ореолы вокруг краёв возникают из-за попыток сжать резкие контрасты, что приводит к размытию или эффекту «кольца».
- Отдельные полосы цвета в областях, которые должны иметь плавный градиент появляются, когда количество цветов в градиенте слишком уменьшено, что приводит к заметным переходам [10].



Рис. 11. Качество сжатия 10 %

## VI. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Разработанная система наглядно демонстрирует фундаментальный компромисс при сжатии с потерями: невозможно одновременно достичь максимальной степени сжатия и высочайшего визуального качества. Эксперименты показывают, что даже при относительно высоких значениях PSNR (свыше 35 дБ) визуальный анализ может выявить специфические артефакты, такие как «блокинг» и «размытие» контуров, что подчеркивает ограниченность любой единственной метрики и важность субъективной оценки.

## VII. ВЫВОДЫ И ЗАКЛЮЧЕНИЕ

1. Разработана и реализована интеллектуальная интерактивная система для обучения основам сжатия изображений с управляемыми потерями.

2. Система предоставляет инструменты для экспериментирования с ключевыми параметрами алгоритма JPEG и визуализирует как промежуточные этапы обработки, так и итоговый результат.

3. Интеграция объективных метрик качества (PSNR, SSIM) и средств визуального сравнения позволяет проводить всесторонний анализ эффективности сжатия.

Перспективы развития системы включают внедрение поддержки современных форматов, таких как WebP и HEIF, а также добавление модулей, объясняющих основы сжатия без потерь.

## ЛИТЕРАТУРА

- [1] Wallace G.K. The JPEG still picture compression standard // Communications of the ACM. – 1991. – Vol. 34. – № 4. – P. 30–44.
- [2] Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2012. – 1104 с.
- [3] Иванов Д.В., Кузнецова О.П. Мультимедийные технологии в образовании: учеб. пособ. – СПб.: Лань, 2019. – 320 с.
- [4] Ahmed N., Natarajan T., Rao K.R. Discrete Cosine Transform // IEEE Transactions on Computers. – 1974. – Vol. C-23. – № 1. – P. 90–93.
- [5] Козлов С.А., Чернявский А.В. Методы сжатия изображений: учебное пособие. – М.: Горячая линия – Телеком, 2018. – 256 с.

- [6] В.В. Старовойтов, Д.С. Шпилев. Цифровая обработка сигналов и изображений. – Минск: БГУИР, 2015. – 367 с.
- [7] Иванов Д.В. Основы мультимедийных технологий: сжатие данных. – СПб.: БХВ-Петербург, 2017. – 320 с.
- [8] Прокопов А.Н., Федотов М.А. Алгоритмы и методы сжатия изображений. – Новосибирск: НГТУ, 2020. – 180 с.
- [9] Kwon M.J., Nam S.H., Yu I.J. et al. Learning JPEG Compression Artifacts for Image Manipulation Detection and Localization // International Journal of Computer Vision, 2022, v. 130, pp. 1875–1895, doi: 10.1007/s11263-022-01617-5.
- [10] Сэйдж К., Мерсер М. JPEG: всё о формате сжатия изображений / пер. с англ. – М.: ДМК Пресс, 2019. – 144 с.

## Информация об авторах

Кочергин Сергей Николаевич – бакалавр направления “Информатика и вычислительная техника” Астраханского государственного технического университета, г. Астрахань, Россия, [sergey.cochergin2016@gmail.com](mailto:sergey.cochergin2016@gmail.com)

Мамлеева Аделя Рифкатовна, старший преподаватель кафедры АСОИУ, Астраханского государственного технического университета, г. Астрахань, Россия, [amamleeva@ilabsltd.com](mailto:amamleeva@ilabsltd.com)

## Intelligent System for Teaching Lossy Image Compression

S. N. Kochergin, A. R. Mamleeva

*Astrakhan State University, Astrakhan, Russia*

**Abstract** – The paper presents an intelligent system for teaching of lossy image compression methods. The relevance of the research is driven by the growing volume of visual data and the need for its efficient storage and transmission. The designed system is an interactive web application visualizes the principles of compression algorithms, such as JPEG, and allows experiments with key parameters as "compression ratio – image quality". The system implements modules for the quantization parameter adjusting, selecting color space, and assessing distortion metrics (PSNR, SSIM). Experiments demonstrate that interactive learning supported the proposed system leads to a deeper understanding of the compromises inherent in lossy compression compared to traditional lecture-based methods. The system is intended for students studying computer graphics, image processing, and multimedia technologies.

**Keywords** – image compression, lossy compression, JPEG, learning, intelligent system, PSNR, SSIM.

## References

- [1] Wallace G.K. The JPEG still picture compression standard // Communications of the ACM. – 1991. – Vol. 34. – № 4. – P. 30–44.
- [2] Gonzalez R., Woods R. Digital Image Processing. Moscow, Tekhnosfera, 2012. – 1104 p.
- [3] Ivanov D.V., Kuznetsova O.P. Multimedia Technologies in Education: textbook. – Saint-Petersburg, Lan', 2019. – 320 p.

- [4] Ahmed N., Natarajan T., Rao K.R. Discrete Cosine Transform // IEEE Transactions on Computers. – 1974. – Vol. C-23. – № 1. – P. 90–93.
- [5] Kozlov S.A., Chernyavsky A.V. Image Compression Methods: Textbook. Moscow: Goryachaya Liniya – Telekom, 2018.
- [6] Starovoitov V.V., Shpilev D.S. Digital Signal and Image Processing. Minsk, BGUIR, 2015.
- [7] Ivanov D.V. Fundamentals of Multimedia Technologies: Data Compression. St. Petersburg, BHV-Petersburg, 2017.
- [8] Prokopov A.N., Fedotov M.A. Algorithms and Methods of Image Compression. Novosibirsk, NSTU, 2020.
- [9] Kwon M.J., Nam S.H., Yu I.J. et al. Learning JPEG Compression Artifacts for Image Manipulation Detection and Localization // International Journal of Computer Vision, 2022, v. 130, pp. 1875–1895, doi: 10.1007/s11263-022-01617-5.
- [10] Sage K., Mercer M. JPEG: All about the Image Compression Format, Moscow, DMK Press, 2019.

#### Information about the authors

Sergey N. Kochergin, bachelor of Science in Computer Science and Engineering, Astrakhan State Technical University, Astrakhan, Russia, [sergey.cochergin2016@gmail.com](mailto:sergey.cochergin2016@gmail.com)

Adelya R Mamleeva, senior lecturer of the AIPCS Department, Astrakhan State Technical University, Astrakhan, Russia, [amamleeva@ilabsltd.com](mailto:amamleeva@ilabsltd.com)