

КОДИРОВАНИЕ ОБЛАСТЕЙ ПОВЫШЕННОГО КАЧЕСТВА В СТАНДАРТЕ СЖАТИЯ ИЗОБРАЖЕНИЙ JPEG 2000

Власюк И.В., Романова Е.П., Сидорова А.И.

Аннотация. В настоящей работе рассматривается процесс кодирования изображений по стандарту JPEG2000, особое внимание уделено особенностям, появившимся впервые с введением JPEG2000 в системах сжатия изображений, в частности, возможность выделения областей повышенного качества (ROI) на изображении. Также проведён обзор методов распознавания образов, которые могут быть использованы для автоматизированного поиска и выделения ROI

Алгоритм JPEG2000 разработан той же группой экспертов в области фотографии, что и JPEG. Основные отличия алгоритма в JPEG2000 от алгоритма в JPEG заключаются в следующем:

1. Лучшее качество изображения при сильной степени сжатия. Или, что то же самое, большая степень сжатия при том же качестве для высоких степеней сжатия.
2. Поддержка кодирования отдельных областей с лучшим качеством. Отдельные области изображения критичны для восприятия человеком, в то время как качеством других можно пожертвовать.
3. Основной алгоритм сжатия заменен на вейвлет-преобразование. Помимо указанного повышения степени сжатия это позволило избавиться от 8-пиксельной «блочности», возникающей при повышении степени сжатия. Для повышения степени сжатия в алгоритме используется арифметическое сжатие.
4. Поддержка сжатия без потерь.

Базовая схема JPEG2000 (рис. 1) очень похожа на базовую схему JPEG.

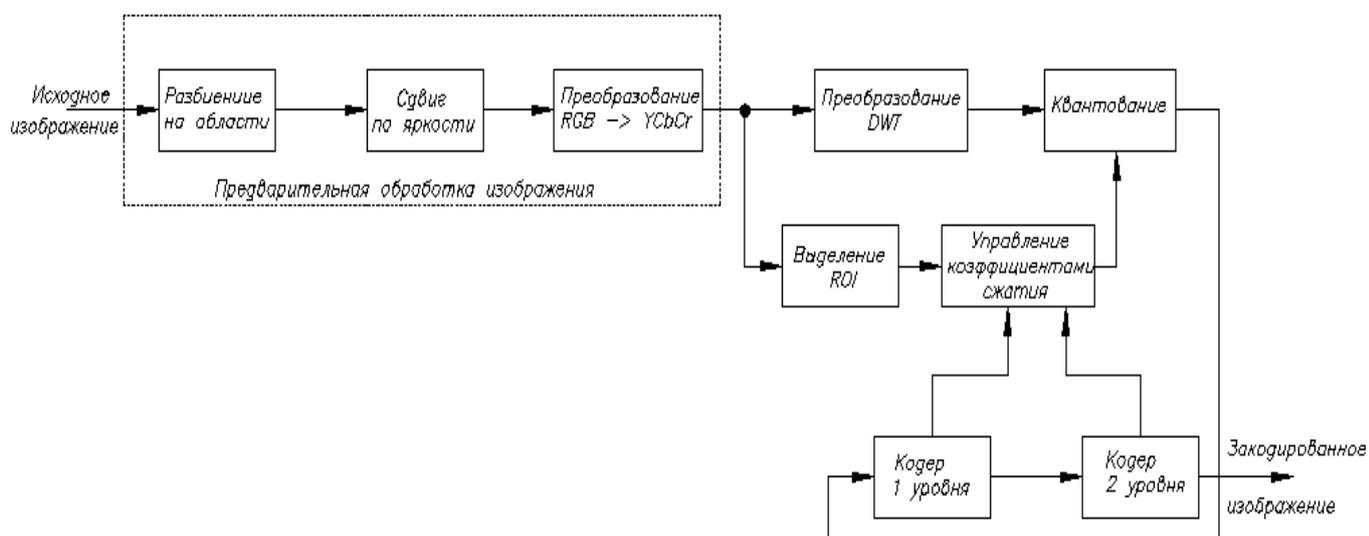


Рис. 1. Структурная схема сжатия изображения по стандарту JPEG2000

Рассмотрим алгоритм по шагам.

1. Разбиение на области. На этапе разбиения на области исходное изображение произвольно разбивается на множество неперекрывающихся прямоугольных блоков. Каждый из этих блоков называют областью (tile). Для выполнения на СБИС в качестве размеров областей выбирают 256 x 256 или 512 x 512.

2. Сдвиг яркости. В JPEG2000 предусмотрен сдвиг яркости (DC level shift) каждой компоненты (RGB) изображения перед преобразованием в YCbCr. Это делается для выравнивания динамического диапазона

3. Перевод изображения из цветового пространства RGB в YCbCr

4. Дискретное вейвлет-преобразование (DWT) Дискретное wavelet преобразование (DWT) также может быть двух видов — для случая сжатия с потерями и для сжатия без потерь.

Преобразование в одномерном случае представляет собой скалярное произведение коэффициентов фильтра на строку преобразуемых значений (в нашем случае — на строку изображения). При этом четные выходящие значения (формула (1)) формируются с помощью низкочастотного преобразования, а нечетные (формула (2)) с помощью высокочастотного:

$$y_{output}(2n) = \sum_{j=0}^{N-1} x_{input}(j) \cdot h_H(j-2n) \quad (1)$$

$$y_{output}(2n+1) = \sum_{j=0}^{N-1} x_{input}(j) \cdot h_L(j-2n-1) \quad (2)$$

Далее к строке применяется чересстрочное преобразование, суть которого заключается в том, что все четные коэффициенты переписываются в начало строки, а все нечетные — в конец. В результате этого преобразования в начале строки формируется «уменьшенная копия» всей строки (низкочастотная составляющая), а в конце строки — информация о колебаниях значений промежуточных пикселей (высокочастотная составляющая).

Это преобразование применяется сначала ко всем строкам изображения, а затем ко всем столбцам изображения. В результате изображение делится на 4 квадранта (рис. 2). В первом квадранте будет сформирована уменьшенная копия изображения, а в остальных трех — высокочастотная информация. После чего преобразование повторно применяется уже только к первому квадранту изображения по тем же правилам (преобразование второго уровня).

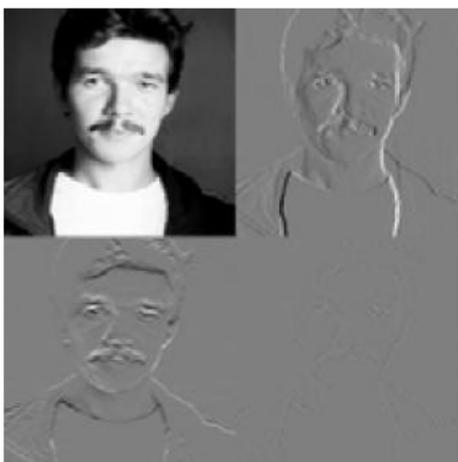


Рис. 2. Результат применения дискретного вейвлет-преобразования

Для корректного сохранения результатов под данные 2 и 3 квадрантов выделяется на один бит больше, а под данные 4-го квадранта — на 2 бита больше. Т.е. если исходные данные были 8-битные, то на 2 и 3 квадранты нужно 9 бит, а на 4-й — 10, независимо от уровня применения DWT.

5. Квантование. Так же, как и в алгоритме JPEG, после DWT применяется квантование. Вычисление величины шага квантования (Δ_b) для частотного диапазона (b) производится в соответствии с его динамическим диапазоном. Формула (3) однородного скалярного квантования с мертвой зоной вблизи нуля имеет вид:

$$q_b(i, j) = \text{sign}(y_b(i, j)) \left\lfloor \frac{|y_b(i, j)|}{\Delta_b} \right\rfloor \quad (3)$$

где $y_b(i, j)$ - коэффициент дискретного вейвлет-преобразования и Δ_b - величина шага квантования для частотного диапазона b.

6. Выделяемые регионы (ROI) Когда практически достигнут предел сжатия изображения в целом и различные методы дают очень небольшой выигрыш, мы можем существенно (в разы) увеличить степень сжатия за счет изменения качества разных участков изображения.

Проблемой этого подхода является то, что необходимо каким-то образом получать расположение наиболее важных для человека участков изображения.

Работы по автоматическому выделению таких областей активно ведутся. В частности, созданы алгоритмы автоматического выделения лиц на изображениях.

В JPEG2000 используется однобитное изображение-маска, задающее повышение качества в данной области изображения. Поскольку за качество областей у нас отвечают коэффициенты дискретного вейвлет-преобразования во 2, 3 и 4 квадрантах, то маска преобразуется таким

образом, чтобы указывать на все коэффициенты, соответствующие областям повышения качества (рис. 3):

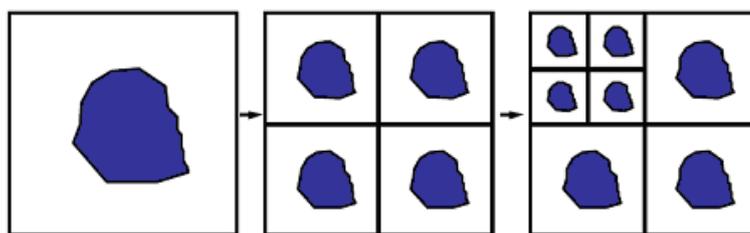


Рис. 3. Преобразование макси области повышения качества для обработки коэффициентов дискретного вейвлет-преобразования

Эти области обрабатываются далее другими алгоритмами (с меньшими потерями).

Методы распознавания образов делятся на две основные категории: методы, основанные на теории решений, и структурные методы. Первая категория имеет дело с образами, описанными с помощью количественных дескрипторов, таких как длина, площадь, текстура. Вторая категория методов ориентирована на образы, для описания которых лучше подходят качественные дескрипторы, например, реляционные.

В методах распознавания, основанных на сопоставлении, каждый класс представляется вектором признаков образа, являющегося прототипом этого класса. Незнакомый образ приписывается к тому классу, прототип которого оказывается ближайшим в смысле заранее заданной метрики. Простейший подход состоит в использовании классификатора, основанного на минимальном расстоянии, который, как ясно из названия, вычисляет евклидовы расстояния между вектором признаков неизвестного объекта и каждым вектором прототипа. Решение о принадлежности объекта к определенному классу принимается по наименьшему из таких расстояний.

Статистически оптимальные классификаторы основываются на вероятностном подходе. Как и в большинстве областей, связанных с измерением и интерпретацией физических явлений, вероятностные подходы оказываются важными в задаче распознавания образов из-за случайностей, влияющих на порождение классов образов. Можно выработать такой метод классификации, который будет оптимальным в том смысле, что при его использовании будет достигаться наименьшая (в среднем) вероятность появления ошибок классификации.

Среди классификаторов данного типа выделим байесовский классификатор с нормальным распределением. Одним из самых успешных применений байесовского классификатора является его использование в задаче классификации данных дистанционного зондирования, регистрируемых с помощью мультиспектральных сканеров, установленных на борту самолета, спутника или орбитальной станции. Учитывая большой объем изображений, получаемых с

помощью такого оборудования, задача автоматического анализа и классификации изображений вызывает значительный интерес.

При построении байесовского классификатора образов с гауссовым распределением требуется получить оценки вектора математического ожидания и ковариационной матрицы для каждого класса. В прикладных задачах дистанционного зондирования эти оценки вычисляются путем сбора мультиспектральных данных для каждого интересующего вида областей.

7. Регулирование скорости. Регулирование скорости - процесс, посредством которого скорости битового потока (иногда называемые скоростями кодирования) распределяются в каждом блоке коэффициентов частотного диапазона для достижения запланированной скорости передачи кодированной информации всего изображения. При этом минимизируются искажения (ошибки), введенные в восстановленное изображение квантованием и усечением кодов.

Кодер JPEG 2000 при кодировании блоков генерирует много независимых потоков битов. Соответственно, алгоритм оптимизации соотношения искажений и скорости (PCDR - algorithm) производит точки усечения для этих потоков битов оптимальным способом, чтобы минимизировать искажение и добиться запланированной скорости передачи. После того, как изображение полностью сжато, алгоритм оптимизации соотношения искажений и скорости (PCDR - algorithm) применяется один раз в конце, используя всю скорость и информацию градиента искажения скорости каждого модуля кодирования.

8. Кодирование. Энтропийное кодирование и создание сжатого потока битов в JPEG2000 разделяются на два шага: 1 уровень и 2 уровень кодирования.

На 1 уровне блоки квантованных вейвлет-коэффициентов кодируются независимо. В JPEG2000 для кодирования битовой плоскости был принят алгоритм вложенного блочного кодирования с оптимизированным усечением (EBCOT). EBCOT кодирует каждую битовую плоскость за три прохода вместе с частью битовой плоскости, закодированной в каждом проходе без перекрывания с двумя другими проходами кодирования. Во время первого прохода осуществляется обработка информации о значимости разрядов, во время второго - уточнение разрядов величины, и наконец - обработка данных, не обработанных во время предыдущих проходов.

Значения двоичного выбора, сгенерированные EBCOT, кодируются с использованием двоичного арифметического кодирования, названного MQ-кодером.

После генерирования первым уровнем сжатых битов второй уровень механизма кодирования эффективно представляет слой суммарного кода и итоговую блочную информацию для каждого блока коэффициентов.

Итоговая блочная информация состоит из последовательности сжатых кодовых слов, наиболее значительная среди других битовая плоскость та, в которой любая выборка в блоке кода является отличной от нуля, а также точка усечения между битовым потоком слоя суммарного кода. Декодер получает эту информацию в закодированном виде в форме двух тэг-деревьев. Это кодирование помогает представлять информацию в очень компактной форме без образования слишком большого количества заголовков в конечном сжатом файле.

Заключение. В докладе были рассмотрены этапы кодирования изображения по стандарту JPEG2000. Когда практически достигнут предел сжатия изображения в целом и различные методы дают очень небольшой выигрыш, мы можем существенно (в разы) увеличить степень сжатия за счет изменения качества разных участков изображения. Дальнейшую работу будем вести в направлении разработки алгоритма по автоматическому выделению областей повышенного качества.

Список литературы

1. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. - М.: ДИАЛОГ-МИФИ, 2003. - 384 с.
2. Tinku Acharya, Ping-Sing Tsai. JPEG2000 standard for image compression: concepts, algorithms and VLSI architectures. Published by John Wiley & Sons, Inc., Hoboken, New Jersey, 2004. -296 page.
3. Р. Гонсалес, Р. Вудс Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.