

R-D ПРОБЛЕМА И ЭФФЕКТИВНОСТЬ СИСТЕМ СЖАТИЯ ИЗОБРАЖЕНИЙ

Мороз В.В.

Интенсивное развитие цифровых методов обработки информации оказало существенное влияние на все техники передачи и хранения информации. Это объясняется такими характеристиками цифровых систем, как помехозащищенность, возможность исправления ошибок, гибкости при коммутации сообщений, высокой надежности. Одновременно идет постоянное расширение использования изображений (видео) в различных областях науки и техники: медицине, астрономии, бесконтактной дефектоскопии, экспериментальной физике и т.д. Обычное изображение содержит много избыточной информации, которая приводит к экономическим потерям. Поэтому при передаче и хранении изображений предварительно осуществляют ее сжатие. Операция сжатия выполняется с целью уменьшения размера данных, описывающих изображение. Сжатие необходимо для скорейшей передачи информации по каналам связи и экономии памяти запоминающих устройств, т.к. современные объемы данных, описывающих изображения, достигают от нескольких сот килобайт до нескольких десятков гигабайт.

В зависимости от вида извлекаемой избыточности возможны два класса алгоритмов сжатия данных: с частичной потерей информации и без потери информации. Сжатие без потери информации, основанное на извлечении статистической избыточности, позволяет полностью восстановить исходные данные, т.е. реконструированные после сжатия данные будут идентичны данным до сжатия. Но, сжатие без потери информации позволяет достичь небольших степеней сжатия – в 3-5 раз. Для извлечения дополнительных типов избыточности используется психофизическое строение системы зрения человека. Установлено, что в системе зрения человека содержатся элементы гомоморфной обработки информации. А такие характеристики системы зрения, как способность восприятия яркости света и ее пространственно частотный отклик, позволяют извлекать из изображения цветовую и пространственную избыточность.

Сжатие с частичной потерей информации приводит к искажению восстанавливаемого изображения (видео). Предполагается такой уровень искажения, при котором оригинальное и восстановленное изображения неразличимы человеком. В зависимости от области приложения допускается разный уровень искажения. Поэтому важным элементом систем сжатия является контроль уровня искажения.

Классический подход к построению математической модели системы сжатия с потерей информации основывается на последовательном кодировании входного сигнала с использованием декоррелирующего преобразования, квантования и энтропийного кодирования. Типовая система сжатия изображений с подсистемой восстановления приведена на рис.1.

В задачах кодирования реалистических изображений с использованием дискретного вейвлетного преобразования (Discrete Wavelet Transform — DWT) [1] эффективность кодирования понимается как минимизация среднеквадратичной ошибки восстановленного после сжатия изображения для заданного битрейта. Битрейт есть количество бит, приходящееся на один пиксель. Эффективность кодирования изображений неразрывно связана с контролем уровня потерь информации и управлением степенью сжатия. Потери информации в процессе сжатия происходят на этапе квантования вейвлетных коэффициентов Q . Поэтому определение оптимальных, с точки зрения визуального качества, параметров квантования Q и энтропийного кодирования EC для заданного битрейта R является актуальной задачей. Кодер C используется для

подготовки контекстного моделирования квантованных вейвлетных коэффициентов и в дальнейшем рассматривается как часть энтропийного кодера.

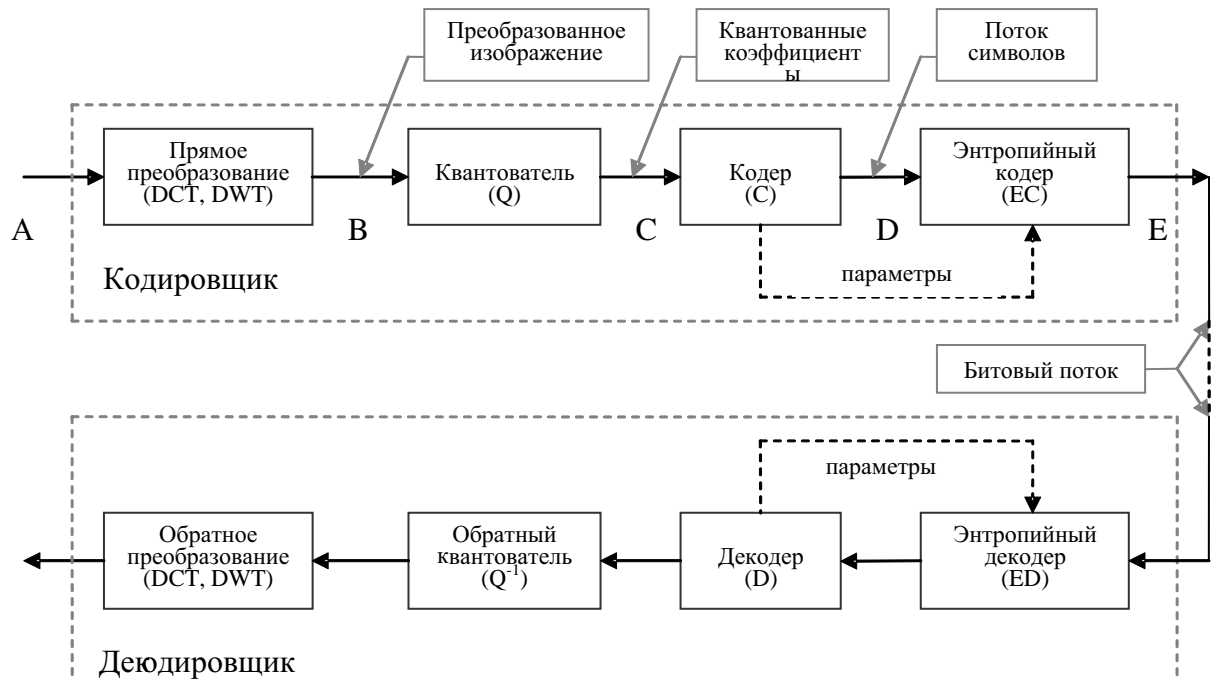


Рис. 1 Схема системы сжатия

Качество системы сжатия может быть измерено с помощью меры искажения $d_k(x, y)$. Данная мера определяет уровень искажения сигнала $y = (y_1, y_2, \dots, y_k)$ на выходе E системы сжатия по отношению к сигналу $x = (x_1, x_2, \dots, x_k)$ на входе A .

Наиболее известными алгоритмами решения задачи эффективного кодирования с использованием DWT являются EZW (Embedded Zerotree Wavelet) [2], SPIHT (Set Partitioning In Hierarchical Trees) [3] и EBCOT (Embedded Block Coding with Optimized Truncation) [4], в которых использована упрощенная математическая модель исходного изображения. Предполагается, что изображение является дискретным случайным процессом с независимыми и одинаково распределенными случайными величинами (как правило, с гауссовским распределением вероятностей). Кодирование рассматривается как последовательная обработка нескольких информационных потоков $k = 1, 2, \dots, K$, где K – количество частотных диапазонов, полученных в результате дискретного вейвлетного преобразования. Вейвлетные коэффициенты в каждом частотном диапазоне равномерно квантуются с заданным для каждого диапазона коэффициентом квантования. Энтропийное кодирование, как показано на рис. 2, выполняется за два шага: сначала выполняется контекстное моделирование (исследование окрестности уже квантованных коэффициентов) для выбора вероятностной модели с последующим адаптивным арифметическим кодированием, а затем — оптимальное распределение бит.

Для определения оптимального распределения бит между потоками используются кривые эффективности, которые характеризуют функциональную зависимость битрейта R от уровня искажения восстановленного изображения D . Минимальный уровень искажения при ограничении битрейта ищется путем минимизации функции $J_k = D_k + \lambda R_k$, где D_k — уровень искажения k -го потока, R_k — битрейт k -го потока, λ — множитель Лагранжа. Наибольшую сложность такого метода решения задачи

представляет определение кривых эффективности каждого потока, которые зависят от вероятностной модели изображения и используемой техники кодирования. В некоторых работах функциональная зависимость $R(D)$ строится для нескольких точек, а общая кривая получается в результате аппроксимации этих точек [5].

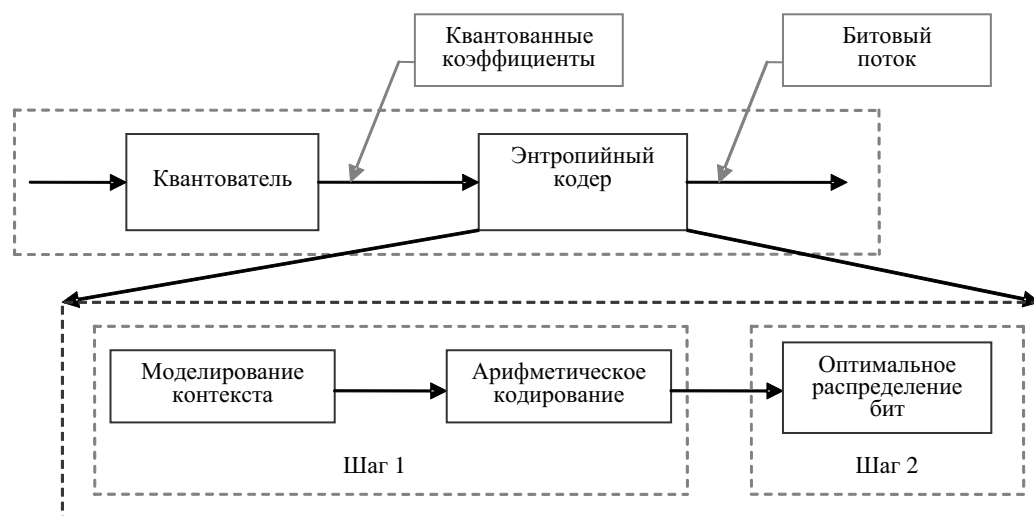


Рис. 2 Схема энтропийного кодера

Задача минимизации уровня искажения при заданном ограничении на битрейт также может быть решена с использованием алгоритма Ллойда-Макса [6], но он основан на обучении, в связи с чем возникает необходимость наличия информации об источнике. Построение квантователя на основе этой информации тоже требует больших вычислительных затрат. Следовательно решение задачи эффективного кодирования с оптимальным распределением бит должно быть направлено на поиск другой техники энтропийного кодирования. Так как кодированию подвергаются квантованные вейвлетные коэффициенты (индексы квантования), то было бы целесообразно совместить квантование, энтропийное кодирование и оптимальное распределение бит. При совместном рассмотрении процессов квантования и энтропийного кодирования, необходимо знать какой вклад приносит квантование и энтропийное кодирование в общую характеристику производительности кодека. Это позволило бы отслеживать квантователи, требующие больших вычислительных ресурсов, но без определения вычислительной сложности квантователя это сделать тяжело.

Поэтому предпочтительней использовать фиксированный равномерный квантователь для всех частотных диапазонов. Тогда энтропийный кодер может адаптироваться ко входной статистике и это не потребует перестройки квантователя. Для сбора статистических данных для энтропийного кодировщика используются индексы квантования, которые являются аппроксимацией вейвлетных коэффициентов. Это приводит к построению намного более простой системы сжатия и в вычислительном отношении, и концептуально.

Для построения оптимального распределения бит предлагается алгоритм кодирования вейвлетных коэффициентов, основанный на адаптивном квантовании с использованием контроля отношения битрейта к уровню искажения входного изображения. Предполагается, что коэффициенты вейвлетного преобразования представляют собой дискретный случайный процесс, распределение вейвлетных коэффициентов представляет собой смесь распределений. Каждому частотному диапазону присвоим свой весовой коэффициент. В пределах одного частотного диапазона вейвлетный коэффициент будет классифицироваться с учетом уже закодированных коэффициентов. Каждому классу соответствует своя модель распределения вероятностей.

На основании этой модели выбираются характеристики квантования и кодирования коэффициентов с учетом оценивания уровня искажения. После кодирования каждого вейвлетного коэффициента вероятностная модель энтропийного кодера, соответствующая данному классу, обновляется.

Пусть дискретная случайная величина x имеет плотность распределения вероятностей $f(x)$. Рассматривается равномерное квантование с размером шага квантования Δ и предполагается, что Δ настолько мал, что в пределах интервала квантования $R_i = (x_{i-1}, x_i)$, $x_i = i\Delta$ функция плотности вероятности $f(x)$ будет практически постоянной. С каждым интервалом квантования R_i связывается число \tilde{x}_i — уровень квантования, причем $i = 0, 1, \dots, N$, где N — количество уровней квантования. Входное значение x отображается в \tilde{x}_i , если, и только если $x \in R_i$. Мера уровня искажения, в общем случае, может быть введена как L_α норма [1]:

$$D_\alpha = \left(\sum_{i=0}^N \int_{R_i} f(x)(x - \tilde{x}_i)^\alpha dx \right)^{1/\alpha}. \quad (1)$$

Пусть максимальное и минимальное значения x необязательно конечные величины, и в процессе восстановления при обратном квантовании реконструированное значение \tilde{x}_i будет находиться точно посередине интервала $[i\Delta, (i+1) \cdot \Delta)$, т.е. будет равно $i\Delta + \Delta/2$.

Ошибка квантования для этого случая будет:

$$\begin{aligned} (L_\alpha\{x - \tilde{x}\})^\alpha &= \sum_{i=-\infty}^{\infty} \int_{i\Delta}^{(i+1)\Delta} (x - i\Delta - \Delta/2)^\alpha f(x) dx \approx \sum_{i=-\infty}^{\infty} \int_{i\Delta}^{(i+1)\Delta} (x - i\Delta - \Delta/2)^\alpha f(i\Delta) dx = \\ &= \sum_{i=-\infty}^{\infty} \int_{i\Delta}^{(i+1)\Delta} (x - i\Delta - \Delta/2)^\alpha f(i\Delta) dx = \sum_{i=-\infty}^{\infty} \frac{\Delta^{\alpha+1}}{2^\alpha(\alpha+1)} f(i\Delta) = \frac{\Delta^\alpha}{2^\alpha(\alpha+1)}. \end{aligned}$$

Тогда уровень искажения может быть определен как

$$D_\alpha = L_\alpha\{x - \tilde{x}\} = \frac{\Delta}{2(\alpha+1)^{1/\alpha}}. \quad (2)$$

Как видно из (2) уровень искажения является функцией шага квантования. Если используется равномерное квантование для всего изображения, то изменяя шаг квантования, можно изменять качество изображения. При достаточно малых значениях Δ с учетом выражения (2) среднеквадратичная ошибка в частотном домене будет $CKO = \Delta^2 / 12$. Если преобразование является ортонормированным, то CKO не будет изменяться в области изображения. Отношение сигнал/шум по мощности можно вычислить как:

$$PSNR \approx 10 \log_{10} \frac{12 \cdot 2^{2B}}{\Delta^2},$$

где: B — количество бит, приходящееся на один пиксель в исходном изображении.

Таким образом, можно контролировать качество восстанавливаемых изображений, изменяя шаг квантования. С другой стороны, перед началом сжатия можно, исходя из желаемого качества изображения, выбрать соответствующий размер шага квантования Δ . В соответствии с определением равномерного квантователя, индексы квантования являются аппроксимацией вещественной прямой, при которой некоторому целому i ставится в соответствие вещественное число $i\Delta$. Так как каждому вейвлетному коэффициенту x соответствует индекс квантования i , который является оптимальным с точки зрения отношения битрейта к уровню искажения при ограничении $R \leq R_0$ то, сводя задачу условной оптимизации к безусловной, получим функционал $J_\alpha = D_\alpha^\alpha + \lambda R$. Тогда

для некоторого λ можно получить одно из множества возможных i , которое будет минимизировать

$$J_{\alpha}(x, i\Delta) = (x - i\Delta)^{\alpha} + \lambda R(i). \quad (3)$$

Битрейт $R(i)$ можно получить из таблицы вероятностей арифметического кодировщика. Однако необходимо определить множитель Лагранжа λ , который является фиксированным для всех вейвлетных коэффициентов и зависит от заданного битрейта R_0 и данных для кодирования. Следовательно, до начала кодирования их необходимо предварительно оценить.

Для поиска λ возьмем математическое ожидание от функционала (3):

$$J_{\alpha}(\Delta) = M\{J_{\alpha}(x, i\Delta)\} = M\{|x - i\Delta|^{\alpha} + \lambda R(i)\}. \quad (4)$$

Предположим, что вейвлетные коэффициенты x имеют распределение Лапласа. Тогда оценка кодирования индексов квантования может быть определена как $R \approx h(x) - \log_2 \Delta$, где h – дифференциальная энтропия.

С учетом (2) и данного предположения, выражение (4) примет вид:

$$J_{\alpha}(\Delta) = \frac{\Delta^{\alpha}}{2^{\alpha}(\alpha + 1)} + \lambda(h(x) - \log_2 \Delta). \quad (5)$$

Тогда

$$\min_{R \leq R_0} J_{\alpha}(\Delta) = \min \left\{ \frac{\Delta^{\alpha}}{2^{\alpha}(\alpha + 1)} + \lambda(h(x) - \log_2 \Delta) \right\},$$

а ограничение $R \leq R_0$ может быть переписано как

$$\lambda(h(x) - \log_2 \Delta) \leq \lambda(h(x) - \log_2 \Delta_0) \Rightarrow \Delta \geq \Delta_0.$$

Дифференцируя правую часть в выражении (5) и приравнявая ее к нулю, получим

$$\lambda = \frac{\alpha \Delta_0^{\alpha}}{2^{\alpha}(\alpha + 1)} \ln 2,$$

или

$$\lambda = \lambda_{\alpha} \cdot \Delta_0^{\alpha}, \quad (6)$$

где: $\lambda_{\alpha} = \frac{\alpha \ln 2}{2^{\alpha}(\alpha + 1)}$.

После подстановки (6) в (3) получим

$$J_{\alpha}(x, \tilde{x}) = |x - i\Delta_0|^{\alpha} + \Delta_0^{\alpha} \lambda_{\alpha} R(i)$$

или, после нормализации,

$$J'_{\alpha}(x, i) = \frac{J_{\alpha}(x, i\Delta)}{\Delta_0^{\alpha}} = \left| \frac{x - i\Delta_0}{\Delta_0} \right|^{\alpha} + \lambda_{\alpha} R(i) = |\tilde{x} - i|^{\alpha} + \lambda_{\alpha} R(i). \quad (7)$$

Следовательно, после вычисления $x = x/\Delta$ можно найти такое i , которое минимизирует $J'_{\alpha}(x, i)$.

В качестве тестовых изображений использовались Lena (размером 512x512 пикселей), Barbara и Goldhill (размером 720x256 пикселей) с глубиной цвета 24 бита. Изображения сначала переводились из цветового пространства RGB в YUV. Затем к ним применялось DWT с использованием фильтра 9/7. Последующее кодирование производилось с использованием предложенного алгоритма.

Как показывают экспериментальные исследования, данный подход позволяет значительно ускорить поиск оптимального индекса квантования и снизить

вычислительные затраты не менее чем в два раза. При этом качество восстанавливаемых после сжатия тестовых изображений не уступает, а для битрейтов ниже 0,6 бит/пиксель и превосходит на 0,1 – 0,3 децибела качество изображений, полученным с помощью верификационной модели кодека JPEG2000.

Исключение составляют классы изображений, содержащих высокочастотные области. В качестве таких тестовых изображений были выбраны Cha1520 (размером 1520x1008 пикселей) и House (размером 512x512 пикселей). Несмотря на высокую производительность кодирования, для таких изображений наблюдался эффект «размывания» высокочастотных областей. Для устранения данного артефакта необходимо производить дополнительный анализ контекста текущего пикселя, но это может привести к повышению вычислительной сложности алгоритма.

Таким образом, предложенный алгоритм эффективного кодирования может быть использован в системах сжатия с целью уменьшения вычислительных затрат. Время кодирования/декодирования кодека, разработанного на базе данного алгоритма, меньше по сравнению с верификационной моделью JPEG2000 Jasper 7.1. Но судить о скорости работы практических реализаций стандарта пока рано, количество таких реализаций на сегодняшний день очень невелико. Имеющиеся plug-in JPEG2000, например в графическом браузере ACDSsee 8.0, не позволяют провести детальный анализ их эффективности. Тем не менее, можно сказать, что доступные реализации на некоторых битрейтах не только уступают по качеству, предложенному алгоритму, но и медленнее его. Так для тестового изображения CHA1520 у кодека, реализующего предложенный алгоритм, время кодирования для степени сжатия 80:1 (0,3 бит/пиксель) ниже на 18%, чем у кодека Jasper 7.1.

Достоинством кодека, реализующего данный алгоритм, является также его совместимость со стандартом сжатия JPEG2000.

The rate-distortion relation is a key problem in systems of images compression. For increase of efficiency of compression systems search of an optimum rate-distortion parity offered. The embedded encoding with change of scanning wavelet coefficients is using for truncation of the least significant bits.

1. Добеши И. Десять лекций по вейвлетам. — Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001, 464 с.
2. Shapiro J. Embedded image coding using zerotree of wavelet coefficients // IEEE Trans. Signal Processing. — 1993 — Vol. 41. — Pp. 3445 — 3462.
3. Said A., Pearlman W. A new, fast and efficient image codec based on set partitioning in hierarchical trees // IEEE Trans. Circuits Syst. Video Technol. — 1996 — Vol. 6. — Pp. 243 — 250.
4. Taubman D. High performance scalable image compression with EBCOT // IEEE Trans. Image Proc. — 2000 — Vol. 9. — Pp. 1158 — 1170.
5. Jin Li, Shawmin Lei. An embedded still image coder with rate-distortion optimization // IEEE Trans. on image processing. — 1999 — Vol. 8, №. 7. — P. 913 — 924.
6. Gersho A., Gray R. M., Vector Quantization and Signal Compression. Norwell: Kluwer Academic Publishers, 1992. — 732 p.
7. Adams M. D. The JPEG-2000 Still Image Compression Standard // ISO/IEC JTC 1/SC 29/WG1 N 2412. — 2001.