

JPEG XS – еще полшага в IP-будущее

Окончание. Начало в №№ 5, 6/2023

Часть 3

Предварительное и управляемое квантование

Константин Гласман

Квантование всегда связано с ошибками и искажениями при декодировании. Если бы квантованию был подвергнут исходный цифровой телевизионный сигнал, то отсечка даже одного-двух битов была бы заметной. Но квантование коэффициентов вейвлет-преобразования является визуально незаметным даже при степени компрессии до 10:1, то есть при отсечке большого числа битов. Это связано с тем, что все субматрицы, кроме одной, расположенной в верхнем левом углу матрицы коэффициентов вейвлет-преобразования (рис. 9-г), описывают высокочастотные составляющие изображения, а, как известно, шумы квантования, хорошо заметные на крупных участках изображения с плавно меняющейся яркостью, гораздо менее заметны на мелких деталях и границах. Поэтому большая часть коэффициентов вейвлет-преобразования может быть подвергнута сравнительно грубому квантованию с большим шагом при сохранении визуальной незаметности искажений и шумов квантования.

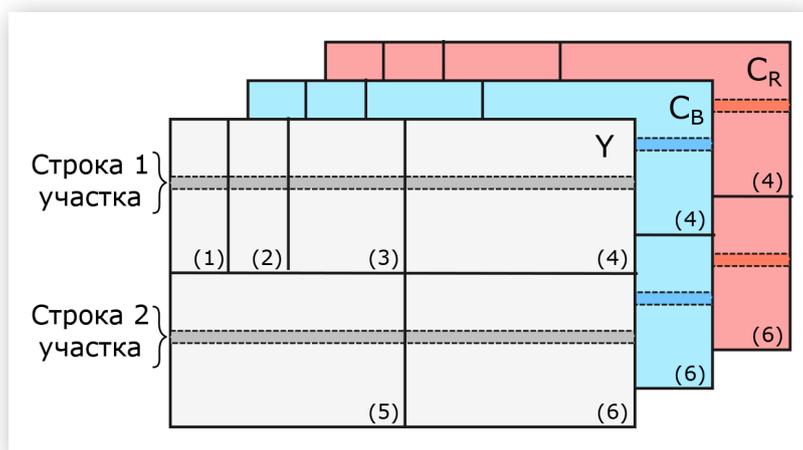


Рис. 12. Компоненты матрицы вейвлет-коэффициентов и участок изображения

Коэффициенты вейвлет-преобразования подвергаются квантованию последовательно. Изображение разбивается на участки. Участок – это наименьшая возможная совокупность строк всех субматриц, которые вносят вклад в одну и ту же пространственную область изображения (в стандарте JPEG XS такой участок называется precinct). Если выполняется декомпозиция в вертикальном направлении только первого уровня, то участок состоит из двух строк (рис. 12): строки 1, или верхней, и строки 2, или нижней. В пакет данных участка включаются последовательно блоки коэффициентов каждой из 6 субматриц из всех трех компонентов изображения (они обозначены на рис. 13 как 1Y, 1CB, 1CR, 2Y, ..., 6Y, 6CB, 6CR).

Данные участка сначала подвергаются предварительному квантованию. Напомним, что обратимое цветовое преобразование RCT и вейвлет-преобразование DWT выполняются как операции с числами длиной 20 битов. В предварительном квантователе коэффициенты сначала представляются в формате «знак-величина» (такой формат был описан выше), что дает числа в диапазоне $[-219-1, +219-1] = [-524287, +524287]$, а затем от двоичных кодов коэффициентов отсекаются 8 наименее значащих двоичных разрядов, или битовых плоскостей. Это эквивалентно квантованию с мертвой зоной с величиной шага квантования $\Delta_T = 2^8 = 256$ и размером мертвой зоны $2\Delta_T = 512$. Предварительное квантование, однако, не используется для целей управления скоростью потока кодированных данных и только гарантирует, что дальнейшая обработка будет осуществляться с числами, не выходящими за пределы диапазона $[-211-1, +211-1] = [-2047, +2047]$, что упрощает требования к аппаратным и программным средствам кодера.

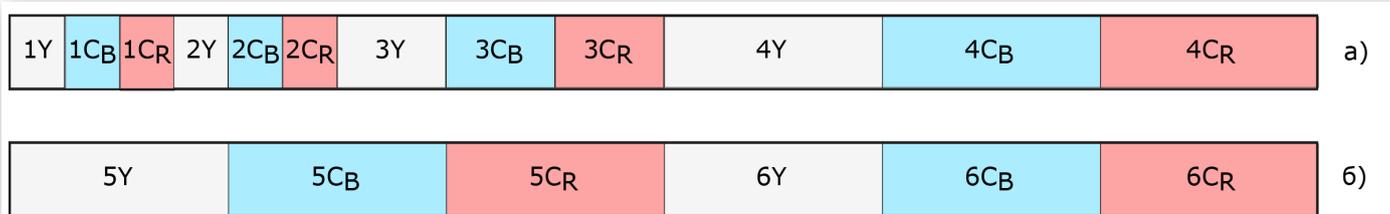


Рис. 13. Пакет данных участка изображения: а - строка 1 участка; б - строка 2 участка

Квантование в целях управления скоростью потока данных следует за этапом предварительного квантования. Значения предварительно квантованных вейвлет-коэффициентов сегментируются на группы, называемые «группами кодирования», по четыре коэффициента в каждой (рис. 10). Для каждой группы кодирования определяется параметр M_g , равный числу заполненных (ненулевых) битовых плоскостей группы кодирования и называемый в стандарте MSB (позиция MSB была описана выше). Стандарт допускает использование как квантования с мертвой зоной, так и равномерного квантования. Квантователь получает от блока управления скоростью потока параметр T – число двоичных разрядов, подлежащих отсечению в субматрице коэффициентов, к которой относится конкретная группа кодирования.

Если выбран режим квантования с мертвой зоной, то квантователь использует шаг квантования, равный $\Delta T = 2T$. Собственно квантование коэффициентов вейвлет-преобразования выполняется путем деления величины коэффициента на шаг квантования (такая процедура реализуется путем сдвига двоичного кода коэффициента вправо, то есть в сторону младших разрядов на число T). Для определения числа интервалов квантования полный интервал $(-2M_g, 2M_g)$ значений, возможных в группе кодирования, надо разделить на размер шага квантования. Это дает $(2M_g + 1 - T)$ интервалов (вычитание 1 связано с тем, что нулевой интервал (мертвая зона) вдвое больше обычного интервала и простирается от $-2T$ до $2T$).

Параметр T является управляющим воздействием и для равномерного квантователя. Если выбран режим равномерного квантования, то квантователь должен использовать тот же выходной диапазон и обеспечить то же количество интервалов квантования, что и в режиме квантования с мертвой зоной, но теперь потребуется применить другой шаг квантования:

$$\Delta_T = \frac{2^{M_g+1}}{2^{M_g+1-T} - 1} = 2^T \left(\frac{2^{M_g+1-T}}{2^{M_g+1-T} - 1} \right)$$

Как и следовало ожидать, шаг квантования оказался большим, чем шаг при квантовании с мертвой зоной. Важно отметить, что шаг квантования теперь не равен целой степени числа 2, и он зависит от данных, подлежащих квантованию. Рассчитав шаг квантования, равномерный квантователь выполняет квантование вейвлет-коэффициента, деля величину коэффициента на найденный шаг. В соответствии с приведенной выше формулой это можно сделать, умножив коэффициент на $(2M_g + 1 - T)$ и разделив на $(2M_g + 1)$. Это деление легко реализовать в соответствии с правилами двоичной арифметики – сначала делается сдвиг влево (в сторону старших разрядов) на $(M_g + 1 - T)$ разрядов, затем вычитается исходное значение коэффициента, потом выполняется сдвиг вправо (в сторону младших разрядов) на $(M_g + 1)$ разрядов. Но процедура в декодере, обратная квантованию, оказывается намного сложнее, так как она требует умножения на шаг квантования ΔT

и тем самым деления на величину, не являющуюся степенью числа 2. Для того, чтобы избежать такого деления и снизить аппаратную сложность декодера, разработана специальная процедура с использованием ряда Неймана [4].

Квантователь в кодере работает под управляющим воздействием блока управления скоростью потока кодированных данных. Скорость потока в среднем поддерживается постоянной и равной заданному значению. Но динамичность телевизионного изображения приводит к тому, что возможны кратковременные колебания скорости потока кодированных данных. Для сглаживания таких колебаний скорости на выходе кодера используется сравнительно небольшая буферная память, которая так и называется – сглаживающая. Данные на ее вход поступают с переменной скоростью, а считываются со строго постоянной.

Энтропийное кодирование

В соответствии с целями проектирования кодека JPEG XS как системы низкой сложности, энтропийное кодирование выполнено чрезвычайно просто. Как уже было отмечено, вейвлет-коэффициенты объединяются в группы из четырех соседних по горизонтали коэффициентов, то есть в так называемые группы кодирования. Для каждой группы кодирования вычисляется самая верхняя заполненная битовая плоскость из всех четырех объединенных коэффициентов. Выше было показано использование этого параметра для реализации квантования. В разделе, посвященном энтропийному кодированию, уместно отметить, что этот счетчик битовых плоскостей, который называется в стандарте «позицией MSB», является единственной величиной, которая подвергается кодированию переменной длины, то есть собственно энтропийному кодированию.

Роль энтропийного кодирования в уменьшении скорости потока данных в системе JPEG XS невелика, поскольку требуемые общие коэффициенты компрессии относительно умеренные. Можно выделить четыре основные операции энтропийного кодирования:

- ◆ в кодере формируются группы значимости из восьми групп кодирования, то есть из 32 вейвлет-коэффициентов, и передается один битовый флаг для каждой группы значимости. Этот флаг указывает, несет ли вообще данные какая-либо из групп кодирования в группе значимости. Если флаг значимости не установлен, вся группа значимости пропускается, и декодер полагает все 32 коэффициента в группе значимости равными нулю;
- ◆ кодер передает позиции MSB всех оставшихся значимых групп кодирования. Стандарт JPEG-XS предоставляет несколько инструментов для кодирования позиций MSB: они передаются в необработанном виде (4 бита на группу кодирования) или предсказываются по их левому или верхнему соседу, а ошибка предсказания передается в унарном коде (унарное кодирование – это энтропийное кодирование, которое представляет число n в виде n единиц с замыкающим нулем);

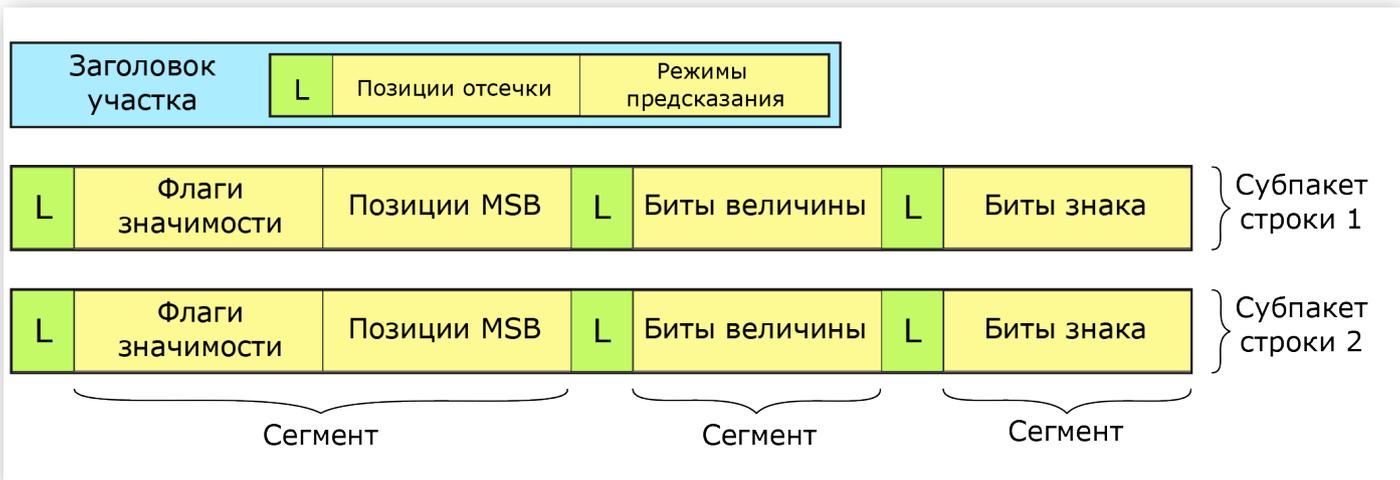


Рис. 14. Структура пакета с кодированными данными участка изображения

- ♦ для всех значимых групп кодирования передаются абсолютные значения квантованных вейвлет-коэффициентов. Они включают в себя все биты от позиции MSB до битовой плоскости отсечки, выбранной устройством управления потоком данных;
- ♦ знаковые биты передаются в соответствии с одним из двух режимов. Если кодер выбирает использование упаковки знаков, то передаются знаковые биты только тех коэффициентов, которые имеют ненулевую величину после усечения. В этом режиме биты знака будут расположены в сегменте кодового потока отдельно от битов абсолютных значений. Альтернативным вариантом является чередование битов знака с битами данных. Упаковка знаков связана с немного более высокой вычислительной сложностью, но обеспечивает более высокую эффективность кодирования, поскольку знаковые биты коэффициентов с нулевой величиной опускаются.

Структура пакета с кодированными данными одного участка показана на рис. 14. Пакет состоит из двух субпакетов, каждый из которых содержит данные одной строки участка. В трех сегментах субпакета передаются: закодированные позиции MSB и флаги значимости, биты абсолютных величин коэффициентов и биты знаков, чередующиеся в соответствии с порядком следования субматриц и компонентов телевизионного сигнала, как показано на рис. 13. Каждому сегменту предшествует поле длины. В заголовке пакета участка содержится информация, важная для декодера: позиции отсечки при квантовании и режимы предсказания, используемые при энтропийном кодировании позиций MSB.

Поток закодированных данных на выходе кодера имеет иерархическую структуру. Участки группируются в срезы, группа срезов образует изображение. На каждом иерархическом уровне имеются заголовки с различной служебной информацией, необходимой при декодировании.

Заключение

Система видеокompрессии JPEG XS сыграет важную роль в развитии телевизионного вещания, упрощая и сглаживая переход между последовательными поколениями

форматов, протоколов и инфраструктур на пути в IP-будущее. Но сфера применения JPEG XS как системы видеокompрессии с малой задержкой и невысокой сложностью, обеспечивающей при этом качество изображения на уровне «без визуальных потерь», много шире. Автомобильные информационные системы, системы промышленного зрения, медицинской визуализации, видеонаблюдения и безопасности, профессиональные аудиовидеосистемы, видеосенсоры Интернета вещей – это далеко не полный список областей, в которых тормозом развития становится недостаточная пропускная способность используемых сегодня для передачи видеосигналов каналов связи и шин. JPEG XS – элегантное и конкурентоспособное решение с широчайшим спектром применения!

Литература

1. International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) JTC1 SC29 WG1, Information Technology – JPEG XS low-latency lightweight image coding system – Part 1: Core coding system, vol. 2019. May 2019.
2. A. Descampe, J. Keinert, Th. Richter, S. Foessel, G. Rouvroy: «JPEG XS: a new standard for visually lossless low-latency lightweight image compression», Appl. of Digital Image Proc. XL (Proc. of SPIE Optics+Photonics), (2017).
3. V. Bruns, Th. Richter, B. Ahmed, J. Keinert, S. Foessel. «Decoding JPEG XS on a GPU». Proc. SPIE 9971, Applications of Digital Image Processing XXXIX, Sept. 2018.
4. Th. Richter, J. Keinert, A. Descampe, G. Rouvroy: «Entropy Coding and Entropy Coding Improvements of JPEG XS», Proc. of Data Compression Conf. 2018 (DCC), 2018.
5. Th. Richter, J. Keinert, S. Foessel, A. Descampe, G. Rouvroy, J.B. Lorent: «JPEG-XS, A high quality mezzanine image codec for video over IP», SMPTE Motion Imaging Journal, (Oct. 2018).
6. A. Descampe et al., «JPEG XS—A New Standard for Visually Lossless Low-Latency Lightweight Image Coding», Proc. of the IEEE, vol. 109, no. 9, pp. 1559–1577, Sep. 2021.
7. S. Foessel and T. Richter. Proposed Extensions to the Third Edition of JPEG XS (ISO/IEC 21122) Standard. SMPTE Motion Imaging Journal, May 2023, pp. 26-32. ■