

ОЦЕНКА КАЧЕСТВА ВНУТРИКАДРОВОГО СЖАТИЯ НОВОГО АЛГОРИТМА КОДИРОВАНИЯ ВЕЙВЛЕТ-КОЭФИЦИЕНТОВ

А.Л. Приоров, доцент Ярославского государственного университета им. П.Г. Демидова, д.т.н.; andcat@yandex.ru

И.С. Мочалов, инженер ООО «Эй-Ти Консалтинг», к.т.н.

А.С. Мочалов, студент Московского государственного университета экономики, статистики и информатики (МЭСИ, Ярославский филиал)

Ключевые слова: сжатие изображений, вейвлет-преобразование, JPEG2000, арифметическое кодирование.

Введение. В последние годы не прекращается процесс разработки новых алгоритмов сжатия. Современные методы используют адаптивные базисы для преобразования, различные направленные предсказания, сложные схемы квантования и контекстно-адаптивного энтропийного кодирования.

В плане сжатия визуальной информации большой значимостью обладают алгоритмы сжатия видео, особенно в рамках цифрового ТВ-вещания и видеоконференцсвязи. Важность задач цифрового ТВ-вещания в России сложно переоценить. Так, в соответствии с распоряжением правительства Российской Федерации № N287-р от 3 марта 2012 г. сроки начала перехода на новый стандарт вещания DVB-T2 были перенесены с 2015 на 2012 год. В последние два года телевидение по стандарту DVB-T2 в тестовом режиме заработало сразу в нескольких регионах России. Главное достоинство этого стандарта — возможность лучшей, по сравнению со старым стандартом DVB-T, адаптации к условиям вещания. За счет этого удастся поднять скорость передачи информации и одновременно снизить объем данных, требуемых для защиты от помех и шумов. По сути эффект от перехода на стандарт DVB-T2 равносителен замене кодека H.264 более совершенным.

Работы в данном направлении несколько лет ведутся в Международной организации стандартизации и Международном союзе электросвязи. В сентябре 2012 г. был представлен новый кодек H.265 [1], позволяющий сократить объем битового потока на 30—35% по сравнению с кодеком H.264 [2].

В видеоконференцсвязи задача стоит иначе, и потери пакетов, от которых в вещательных стандартах защищает помехоустойчивое кодирование с перемежением, периодически возникают. Кроме того, в видеоконференцсвязи разные получатели обладают разными каналами связи. В силу этого необходимо обеспечить высокое качество изображения как на низких скоростях передачи, так и на высоких. Важным фактором является также и вычислительная сложность. Несмотря на эффективность вейвлет-кодеков, например кодека Dirac [3], разработанного исследовательской группой BBC, они, как правило, не удовлетворяют условиям либо низкой вычислительной сложности, либо низкого объема выделяемой памяти, поэтому и в этой области применяются алгоритмы, основанные на блочных преобразованиях семейства H.26x.

Целью статьи является оценка качества полутоновых видеоизображений, закодированных и декодированных разработанным алгоритмом внутрикадрового сжатия.

Вейвлет-преобразование (ВП). С точки зрения декоррелирующих свойств ВП более эффективно, чем широко используемое в стандартах H.26x дискретное косинусное пре-

образование. Однако вейвлет-кодеки, из которых наиболее известны SPIHT и JPEG2000 [4], обладают слишком высокими требованиями к быстродействию и объему памяти. Поэтому возникает задача снижения вычислительной сложности алгоритмов сжатия, основанных на вейвлетах. Что касается самого ВП, то благодаря работам Свелдена и Добеши [5], удалось значительно повысить быстроту его выполнения.

Известно, что в основу ВП положено разложение произвольного сигнала по некоторому ортонормированному базису, построенному на основе вейвлет-функций (низкочастотных и высокочастотных), обладающих определенными свойствами, наиболее важными из которых являются частотно-временная локализация и масштабируемость. Как правило, для обработки изображений применяются двумерные разделимые вейвлеты (на практике обработка выполняется сначала по строкам, а затем по столбцам изображения). Известно, что ВП является мощным инструментом, позволяющим исследовать сигнал в различных частотных областях в требуемом масштабе. Так как низкочастотные вейвлет-коэффициенты фактически представляют собой прореженную копию исходного сигнала, к ним может быть повторно применено ВП.

Существуют, как минимум, три способа выполнения ВП: через банки фильтров, полифазное представление и лифтинг-схему, предложенную Свелденом. Последняя основана на применении обратимых шагов лифтинга, которые в общем случае могут быть нелинейными. В [5] доказано, что любое ВП может быть представлено в виде лифтинг-схемы. На этой базе построена лифтинг-схема для биортогональных вейвлетов 4/4, получившая название CDF9/7. В такой схеме сигнал разбивается на четные и нечетные отсчеты. Затем каждый четный отсчет обновляется взвешенной суммой двух соседних отсчетов (рис. 1).

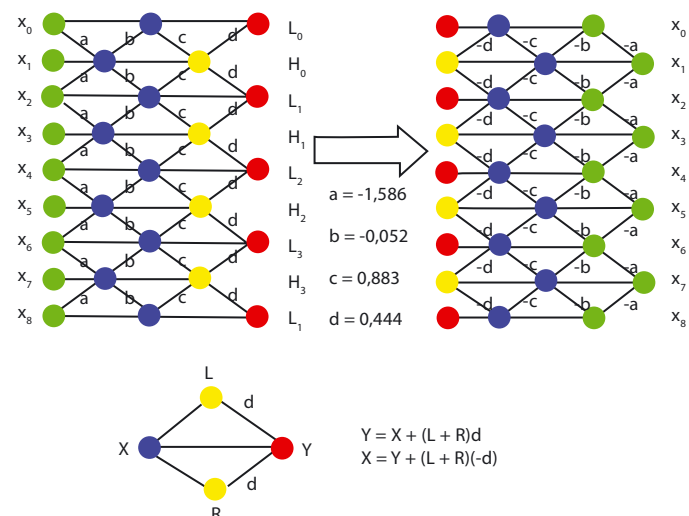


Рис. 1. Блок-схема лифтинга

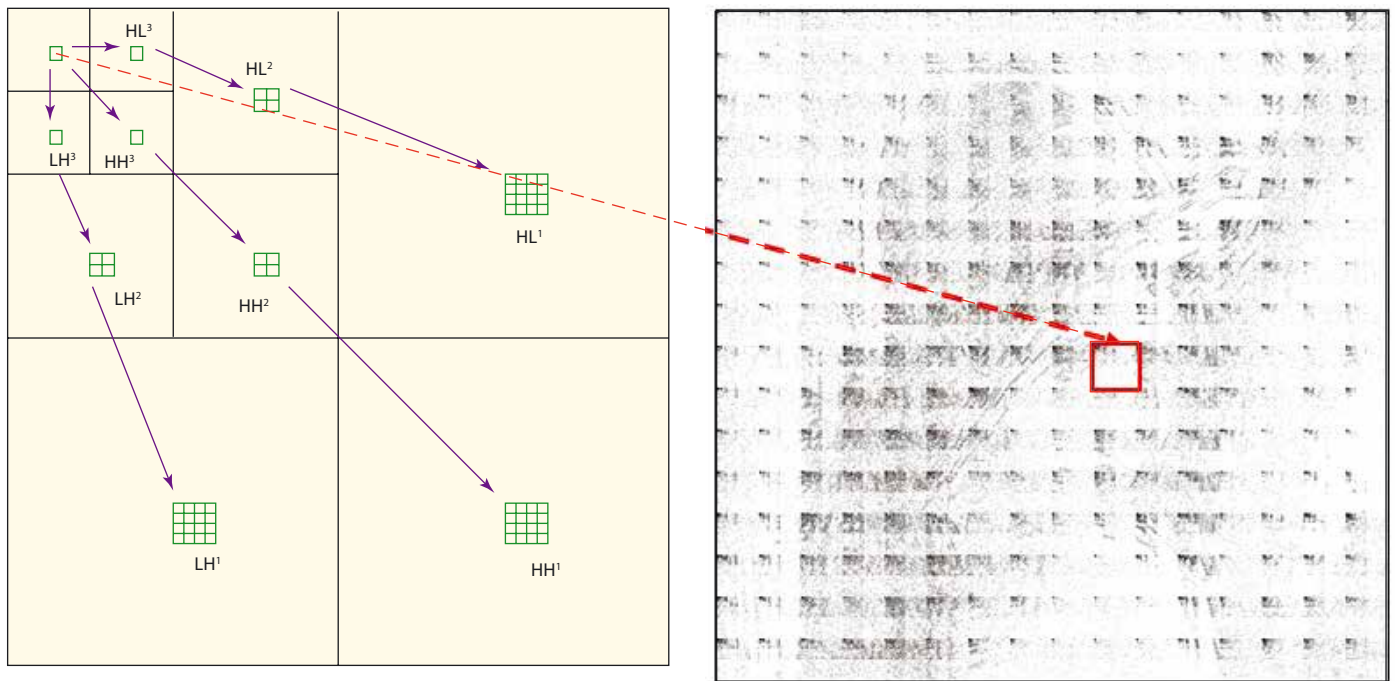


Рис. 2. Перегруппировка пространственно-ориентированных деревьев

Переупорядочивание. Значительно понизить требования к памяти можно, переупорядочив вейвлет-коэффициенты. Переупорядочивание 5-уровневого ВП позволяет отдельно обрабатывать блоки коэффициентов размером 32x32, что дает возможность выполнять параллельную обработку каждого блока. Переупорядочивание выполняется путем группировки пространственно-ориентированных деревьев, отражающих взаимосвязь коэффициентов на различных уровнях. Известно, что вейвлет-коэффициенты обладают подобием между уровнями, и поэтому каждому коэффициенту верхнего уровня может быть поставлено в соответствие дерево, в котором он будет корнем (рис. 2).

Квантование. В любом из вейвлет-кодеков используется тот или иной метод квантования, однако наибольший интерес представляет квантование с нулевой зоной, поскольку оно может быть выполнено в целочисленной арифметике за одну операцию на языке C: $Y = X/QS$.

Квантование с расширенной нулевой зоной, также как и равномерное квантование, применяется в вейвлет-сжатии. Его основное отличие от равномерного квантования состоит в том, что нулевая зона ($i = 0$) имеет больший интервал, чем другие зоны (у них один и тот же интервал QS). Квантование с расширенной нулевой зоной, как и равномерное квантование, симметрично относительно нуля и задается формулой:

$$i = \text{sign}(x) \text{floor} \left[\frac{|x|}{QS} + \lambda \right]_+, \quad (1)$$

где λ определяет размер нулевой зоны. При $\lambda = 0,5$ формула (1) переходит в формулу для равномерного квантования.

Кодовая книга такой системы задается следующим образом

$$\text{code}_i = \text{sign}(i)QP(|i| + 0,5 - \lambda)QS. \quad (2)$$

Квантованные целые числа Y в двоичном коде не являются однозначным кодом. Так, последовательность чисел $[3, 5] = [11101]_b$ может быть декодирована как $[1, 1101]_b = [1, 13]$ или $[110, 1]_b = [14, 1]$. Поэтому передача квантован-

ных коэффициентов требует отдельного подхода. В алгоритмах EZW [6], SPIHT и EBCOT (EBCOT — основа стандарта сжатия изображений JPEG2000) при передаче квантованных коэффициентов используется их представление в виде битовых плоскостей. Каждая битовая плоскость отвечает за тот или иной разряд для всех вейвлет-коэффициентов.

Однако может быть выбран и другой подход. Каждое целое число можно представить в виде: знак + число разрядов + уточняющие биты. Так как «1» — первый бит в любом числе, то его не надо передавать, если известно число разрядов. Поэтому уточняющие биты — это биты числа в двоичной записи без старшего бита. Для числа $13 = 1101_b$ уточняющими битами будут 101_b . Знак также может быть закодирован одним битом. Основная задача кодирования при таком представлении квантованных чисел заключается в передаче декодеру числа разрядов.

Арифметическое кодирование. Сегодня наилучшим решением для передачи числа разрядов является арифметическое кодирование [7]. Оно позволяет достичь результатов, всего на несколько бит отличающихся от оптимальных. Арифметическое кодирование широко используется в современных стандартах сжатия видеоданных и изображений. Это связано с его эффективностью при кодировании алфавитов из множества символов, со скоростью и низкими требованиями к хранению. Арифметическое кодирование позволяет использовать адаптивные модели вероятности символов в алфавите, которые могут меняться с течением времени, поэтому оно во многих отношениях превосходит метод кодирования Хаффмана.

Например, модель может определять вероятность каждого символа в алфавите. И тогда каждый символ будет сжиматься независимо от других без участия контекста. Вероятности каждого символа могут быть заранее подсчитаны на основе репрезентативных выборок. Такие модели с фиксированными вероятностями передаются заранее, чтобы модели кодера и декодера совпадали. После этого их используют для множества сообщений. Как альтернатива, вероятности могут меняться от символа к символу. Эти

вероятности могут пересчитываться на основе уже принятых символов, или же заранее может быть обговорено использование нескольких моделей, выбор между которыми формируется на основе уже переданных символов контекста. Такое кодирование называется контекстным.

Частным случаем кодирования алфавита из множества символов является кодирование бинарных сообщений. Для таких сообщений кодирование Хаффмана становится менее эффективным, если вероятность появления символа «0» много больше вероятности появления символа «1». Для этих случаев предложено несколько методов, в том числе кодирование длин серий.

Таким образом, основной проблемой при передаче разрядов является определение вероятностных моделей, в которых будут переданы разряды.

Алгоритм обхода. Разработанный однопроходный алгоритм обхода квантованных коэффициентов не требователен к ресурсам памяти и обладает низкой вычислительной сложностью. Алгоритм показывает результаты, сравнимые с результатами работы стандарта сжатия JPEG2000.

Исходно одно пространственно-ориентированное дерево представляет собой блок коэффициентов 32×32 . Соответствующая карта разрядов — это блок 32×32 неотрицательных целых чисел с четкой иерархией. Каждый коэффициент имеет четырех прямых потомков, кроме коэффициентов на первом уровне разложения (у них нет потомков).

Как известно из теории арифметического кодирования, эффективность кодирования зависит от правильности выбора модели распределения вероятностей. В разработанном алгоритме обхода вероятности разрядов описываются семью моделями, каждая из которых соответствует увеличению энергии сигнала. Вероятностные модели разрядов представлены на рис. 3.

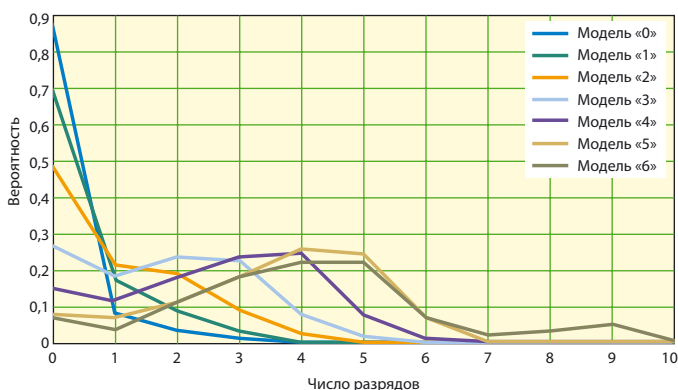


Рис. 3. Перегруппировка пространственно-ориентированных деревьев

Предлагаемый алгоритм обходит вершины дерева снизу вверх, справа налево. При кодировании коэффициента рассматриваются уже переданные значения его братьев и потомков. Используемая для кодирования модель совпадает со значением старшего разряда или старшего номера модели уже переданных коэффициентов из указанного контекста. На рис. 4 приведен контекст для разряда со значением 3.

Рассмотрим работу разработанного алгоритма. Каждый обрабатываемый разряд имеет предков, находящихся на более высоком уровне ВП. Кроме того, каждый обрабатываемый разряд кодируется с помощью арифметического кодера с определенной вероятностной моделью. Номер модели для данного коэффициента определя-

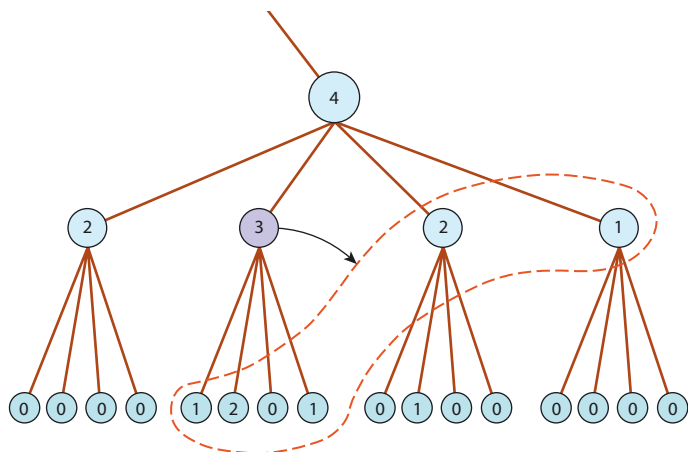


Рис. 4. Контекст, определяющий модель для разряда со 2-го уровня ВП

ется заранее при обработке предыдущих коэффициентов. Моделью по умолчанию является модель с номером «0».

Для всех разрядов с 1-го уровня ВП до 5-го выполняются следующие действия:

- если разряд не превосходит номера модели, в которой он кодируется, то передается «0» и значение разряда в текущей модели;
- если разряд превосходит номер модели, в которой он кодируется, то передается «1» и значение разряда в текущей модели;
- модели предков данного разряда становятся равными значению обрабатываемого разряда;
- модели разрядов не переданных братьев становятся равными значению обрабатываемого разряда.

Значения бит, сигнализирующие о том, что разряд превосходит номер модели, кодируются арифметическим кодером в одной из семи моделей с разными вероятностями «0» и «1».

Результаты тестирования. Такая относительно простая схема энтропийного кодирования позволяет добиться высоких результатов в качестве изображений. Для тестирования были выбраны изображения стандартной базы изображений университета Южной Калифорнии (<http://sipi.usc.edu/database/>), базы изображений Kodak (<http://r0k.us/graphics/kodak/>) и опорные кадры видеопоследовательностей базы данных университета Фраунгофера, используемых при тестировании нового видеокодека H.265 (<ftp://hvc:US88Hula@ftp.tnt.uni-hannover.de/testsequences>).

Для сравнения визуального качества изображений использовались широко распространенная метрика пик отношения сигнал/шум (Пик ОСШ, в дБ) и метрика коэффициентов структурного подобия (КСП), коэффициент корреляции которой с оценками экспертов составляет более 90% [8]. Для того чтобы оценить не только качество изображения в целом, но и отдельных его частей, рассчитывался локальный КСП в окне размером 11×11 .

Для тестирования использованы алгоритм SPIHT [4] (автор W.A. Pearlman) и алгоритм JPEG2000, реализованный в программе FastStone Image Viewer. Результаты Пик ОСШ для изображения «Барбара», двух тестовых изображений из базы Kodak и двух опорных кадров видеопоследовательностей представлены в таблице.

С точки зрения локальных оценок КСП, в 62% случаев результаты протестированного алгоритма лучше, чем у алгоритма сжатия изображений JPEG2000. Пример кадра тестовой последовательности «Баскетбол» приведен

Барбара (Barbara)			
Бит на пиксель	Разработанный алгоритм	SPiNT	JPEG2000
0,25	27,06	26,39	26,68
0,5	30,88	30,34	30,30
0,75	33,70	33,07	33,07
1	36,00	35,56	35,30
Дом (Img11)			
Бит на пиксель	Разработанный алгоритм	SPiNT	JPEG2000
0,25	31,20	30,93	30,96
0,5	34,15	33,42	33,36
0,75	36,30	35,55	35,58
1	38,14	37,28	37,39
Маяк (Img8)			
Бит на пиксель	Разработанный алгоритм	SPiNT	JPEG2000
0,25	30,70	30,33	30,93
0,5	33,84	33,30	33,67
0,75	36,28	35,62	35,91
1	38,18	37,47	37,98
Баскетбол, 1-й кадр (BasketballPass_416x240)			
Бит на пиксель	Разработанный алгоритм	SPiNT	JPEG2000
0,25	32,17	31,52	30,82
0,5	35,90	34,97	34,77
0,75	38,62	37,57	37,46
1	40,65	39,87	39,24
Четыре человека, 20-й кадр (FourPeople_1280x720)			
Бит на пиксель	Разработанный алгоритм	SPiNT	JPEG2000
0,2	35,13	34,87	35,03
0,3	37,78	37,31	37,48
0,4	39,91	39,43	39,42
0,5	41,60	40,91	40,86



Рис. 5. Кадр тестовой видеопоследовательности

определяются размером буфера для хранения изображений. Также требуются семь дополнительных буферов по 1024 байта для хранения разрядов, уточняющих бит модели и сигнализирующих бит. Символы и вероятностные модели определяются во время обхода, и арифметический кодер получает последовательно пару «символ–вероятность». Процесс арифметического кодирования может быть ускорен за счет перехода к бинарному арифметическому кодированию.

Заключение. Сравнение качества декодированных изображений, полученных на выходе разработанного алгоритма, алгоритма SPiNT и стандарта JPEG2000, показывает превосходство первого. Однако, поскольку кодеры JPEG2000 различных производителей отличаются друг от друга по эффективности сжатия, то полученные результаты работы разработанного кодера нельзя интерпретировать, как однозначно превосходящие JPEG2000. Тем не менее, очевидно, что предложенный алгоритм имеет ряд серьезных преимуществ, благодаря своей простоте. Данный алгоритм может быть улучшен путем введения адаптивного шага квантования для каждого пространственно-ориентированного дерева и использования квантования, оптимизированного по критерию число бит/искажение.

ЛИТЕРАТУРА

1. High Efficiency Video Coding (HEVC) [Электронный ресурс]. Режим доступа: <https://hevc.hhi.fraunhofer.de/> (дата обращения: 05.12.12).
2. ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4) AVC. Advanced Video Coding for Generic Audiovisual Services.— Version 3: 2005.
3. Dirac specification by BBC research Group [Электронный ресурс]. Режим доступа: <http://diracvideo.org/> (дата обращения: 05.12.12).
4. Приоров А. Л., Апальков И. В., Хряшев В. В. Цифровая обработка изображений: Уч. пособ.— Ярославль: ЯрГУ, 2007.— 235 с.
5. Daubechies I., Sweldens W. Factoring wavelet transforms into lifting steps // J. Fourier Anal.— Appl. 1998.— Vol. 4, № 3.— P. 245–267.
6. Shapiro J. M. Embedded image coding using zero-trees of wavelets coefficients // IEEE Trans. Signal Processing.— 1993.— Vol. 41, № 12.— P. 3445–3462.
7. Moffat A., Neal R., Witten I. H. Arithmetic coding revisited // ACM Transactions on Information Systems.— July 1998.— Vol. 16, № 3.— P. 256–294.
8. Wang Z., Simoncelli E. P., Bovik A. C. Multi-scale structural similarity for image quality assessment // IEEE 37th Asilomar Conf. on Signals, Systems and Computers.— 2003.— Vol. 2.— P. 1398–1402.

на рис. 5. На этом кадре сравниваются значения локальных оценок КСП предложенного алгоритма и стандарта JPEG2000. Здесь затемненные участки показывают области, где JPEG2000 показывает лучшие результаты.

Анализируя рис. 5, можно сказать, что на ряде текстур разработанный алгоритм стабильно превосходит JPEG2000. Кроме того, он показывает сопоставимые результаты при сжатии изображений цифр и букв. Общий итог анализа внутрикадрового кодирования: несмотря на простоту реализации, алгоритм позволяет достичь результатов, не уступающих результатам работы популярного стандарта JPEG2000.

С точки зрения числа операций, требуемых для работы данного алгоритма, можно отметить, что основная вычислительная сложность связана с арифметическим кодированием. Сложность ВП не превышает 20 сложений и 10 умножений на отсчет. Необходимые объемы памяти

Получено 11.03.13