

СПИСОК ЛИТЕРАТУРЫ

1. Ross N.W. A Painless Guide to CRC Error Detection Algorithms // Dr. Ross Williams. 1993. URL: http://www.ross.net/crc/download/crc_v3.txt (дата обращения: 23.10.2012).
2. Буркатовская Ю.Б., Мальчуков А.Н., Осокин А.Н. Быстродействующие алгоритмы деления полиномов в арифметике по модулю два // Известия Томского политехнического университета. – 2006. – Т. 309. – № 1. – С. 19–24.
3. Темников Ф.Е., Афонин В.А., Дмитриев В.И. Теоретические основы информационной техники. 2-е изд., испр. и доп. – М.: Энергия, 1979. – 512 с.
4. Koopman P. 32-Bit Cyclic Redundancy Codes for Internet Applications // Intern. Conf. on Dependable Systems and Networks (DSN). – Washington, July 2002. – P. 459–468.
5. Мальчуков А.Н., Осокин А.Н. Быстрое вычисление контрольной суммы CRC: таблица против матрицы // Прикладная информатика. – 2010. – № 2 (26). – С. 58–63.
6. Центр коллективного пользования «Суперкомпьютерный кластер «СКИФ-политех» // Сайт ЦКП СКК «СКИФ-политех» ТПУ. 2008. URL: <http://cluster.tpu.ru/> (дата обращения: 23.10.2012).
7. 32 bit Cyclic Redundancy Check Source Code for C++ // Create Window Website. 2011. URL: <http://www.createwindow.com/programming/crc32/> (дата обращения: 01.05.2011).

Поступила 25.12.2012 г.

УДК 004.62

ЭФФЕКТИВНОСТЬ РЕЖИМОВ ВНУТРЕННЕГО ПРЕДСКАЗАНИЯ БЛОКОВ В СОВРЕМЕННЫХ СТАНДАРТАХ СЖАТИЯ ВИДЕО

М.П. Шарабайко, Н.Г. Марков

Томский политехнический университет
E-mail: sme_box@tpu.ru

В 2013 г. ожидается финальная версия стандарта видеокompрессии H.265/HEVC. Предполагается, что с его внедрением начнется массовый переход на системы сверхвысокого разрешения. В статье приведены результаты анализа эффективности методов внутреннего предсказания блоков стандарта H.265/HEVC с точки зрения их влияния на степень сжатия опорных кадров цифрового видео. Кроме того, приводятся результаты сравнения эффективности сжатия опорных кадров в стандартах H.265/HEVC и H.264/AVC.

Ключевые слова:

Сжатие видео, стандарт H.265/HEVC, стандарт H.264/AVC, режимы внутреннего предсказания блоков, опорные кадры.

Key words:

Video compression, standard H.265/HEVC, standard H.264/AVC, intra prediction modes, key frames.

Актуальность задачи цифрового сжатия видео-последовательностей растет вместе с развитием сетей передачи данных, в том числе беспроводных, а также возможностей и мощностей персональных и мобильных компьютерных устройств. При этом требования к степени сжатия видео непрерывно возрастают, стимулируя, в свою очередь, развитие стандартов видеокodирования.

Первый общепринятый стандарт видеокompрессии MPEG2 [1] был окончательно принят в 1996 г., после чего началось быстрое развитие цифрового спутникового телевидения. Следующим стандартом стал MPEG4 part 10 (H.264/AVC) [2], принятый в 2003 г., что стало толчком к развитию систем DVB-T/C, Интернет ТВ и привело к появлению разнообразных сервисов видеобмена и видеосвязи.

В 2013 г. выходит стандарт цифрового сжатия видео-последовательностей H.265/HEVC [3], являющийся стандартом нового поколения и обеспечивающий повышение степени сжатия видео до двух раз по сравнению со степенью сжатия видео с помощью текущего индустриального стандарта H.264/AVC.

В статье исследуется эффективность методов, положенных в основу стандартов H.265/HEVC и H.264/AVC, касающихся предсказания и сжатия опорных видеокадров.

Задача сжатия видео-последовательностей

В цифровом представлении видео-последовательность является потоком видеоданных, состоящим из последовательности видеокадров. Каждый кадр видео-последовательности представляет определенный дискретный момент видеосигнала. Физически же кадр является матрицей пикселей изображения.

Большинство современных стандартов сжатия видео-последовательностей реализуют блочную модель кодирования, суть которой заключается в разбиении кадра видео-последовательности на множество непересекающихся блоков и выполнении алгоритма сжатия на каждом полученном блоке. В стандарте H.265/HEVC структура разбиения на блоки представляет собой квадро-дерево, подобное применяемому при фрактальном сжатии [4].

Сжатие видеoinформации основано на двух важных принципах. Во-первых, каждый кадр видеопоследовательности обладает пространственной избыточностью вследствие корреляции значений пикселей внутри кадра. Во-вторых, большинство кадров видеопоследовательности похожи на предыдущие, а значит обладают временной избыточностью [5]. Алгоритмы сжатия видеопоследовательностей основаны на устранении этих двух видов избыточности.

В сжатой видеопоследовательности обычно присутствуют как кадры с устранением только пространственной избыточности, так и кадры с устранением обоих видов избыточности. Кадры первого типа называются ключевыми (опорными) и обозначаются символом *I* (сокр. от англ. *intra* – внутренний). Такие сжатые кадры позволяют начать их декодирование сразу, поскольку они не ссылаются на другие кадры данной видеопоследовательности.

Кадр, ссылающийся на один или несколько предыдущих, называется разностным и обозначается символом *P* (сокр. от англ. *predicted* – предсказанный). Кадр, для сжатия которого используются как предыдущие, так и последующие кадры видеопоследовательности называется двунаправленным и обозначается символом *B* (сокр. от англ. *bidirectional* – двунаправленный).

В статье исследования сосредоточены на сжатии опорных видео кадров. При их сжатии в стандарте *H.265/HEVC* используются только внутренние предсказания на основе уже сжатой части этого же кадра. В этом стандарте определено 35 режимов внутреннего предсказания: *DC*, плоский и 33 угловых предсказания (рис. 1).

В стандарте предыдущего поколения *H.264/AVC* определено всего 10 режимов внутреннего предсказания блоков: *DC*, плоский и 8 угловых (направления 6, 10, 14, 18, 22, 26, 30, 34 на рис. 1) [6].

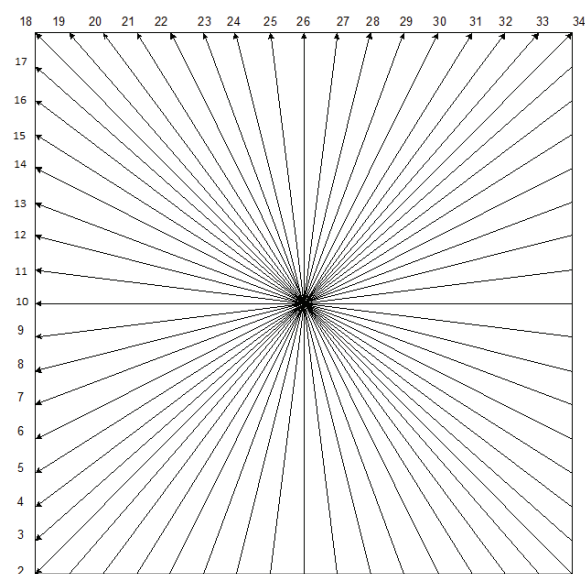


Рис. 1. Угловые внутренние предсказания в стандарте *H.265/HEVC*

При режиме *DC*-предсказания вычисляется среднее значение граничных пикселей, расположенных сверху и слева от предсказываемого блока (рис. 2), которым заполняется весь шаблон предсказания.



Рис. 2. *DC* предсказание в стандартах *H.265/HEVC* и *H.264/AVC*

При режиме плоского предсказания используется линейная функция от левых и верхних граничных пикселей предсказываемого блока (рис. 3).

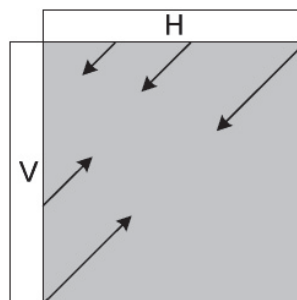


Рис. 3. Плоское предсказание в стандартах *H.265/HEVC* и *H.264/AVC*

В режимах угловых предсказаний шаблон предсказания формируется экстраполяцией значений граничных пикселей, попадающих под направление предсказания. Например, на рис. 4 схематично представлено вертикальное внутреннее предсказание, при котором используются пиксели *A–D*.

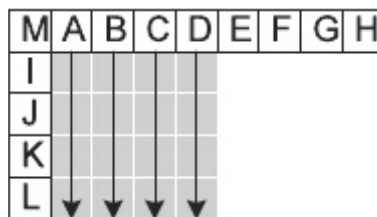


Рис. 4. Вертикальное предсказание в стандартах *H.265/HEVC* и *H.264/AVC*

Помимо введения дополнительных режимов внутреннего предсказания в стандарте *H.265/HEVC*, по сравнению со стандартом *H.264/AVC*, модифицировано энтропийное кодирование. Так, в стандарте *H.264/AVC* используется контекстно-адаптивное двоичное арифметическое кодирование *CABAC* (сокр. от англ. *Context Adaptive Binary Arithmetic Coding*), включающее 15 контекстных моделей для кодирования синтаксических элементов.

В стандарте *H.265/HEVC* применяется синтаксически адаптивное двоичное арифметическое кодирование *SBAC* (сокр. от англ. *Syntax-Based Context-Adaptive Binary Arithmetic Coding*), использующее 31 контекстную модель.

Исследование эффективности методов внутреннего предсказания блоков

Ставилась задача исследования эффективности режимов внутреннего предсказания блоков в стандартах *H.264/AVC* и *H.265/HEVC* в целом на видеопоследовательности и на участках видеокадра, а также изучения выгоды от использования 24 новых режимов внутреннего предсказания блоков в стандарте *H.265/HEVC*.

Исследования проводятся на известном наборе видеопоследовательностей университета Ганновера [7], представленном в таблице.

Оценка эффективности сжатия видеопоследовательности осуществляется на основе общепринятого в области сжатия видеоинформации отношения объективно оцениваемого искажения сжатой видеопоследовательности к скорости потока данных (размеру) сжатой видеопоследовательности. Для возможности соотнесения полученных в данной работе результатов с результатами других исследователей в качестве оценки искажения используется метрика пикового соотношения сигнала к шуму *PSNR* (сокр. от англ. *Peak Signal to Noise Ratio*), вычисляемого по формуле:

$$PSNR = 10 \cdot \lg \left(\frac{255^2}{MSE} \right), \quad (1)$$

при этом

$$MSE = \frac{1}{m \cdot n} \cdot \sum_{i=0}^m \sum_{j=0}^n |S(i, j) - D(i, j)|^2,$$

где $S(i, j)$ и $D(i, j)$ – i и j пиксели исходного и сжатого видео кадров размером $m \times n$ соответственно. Для оценки искажений при сжатии видеопоследовательностей применяется метрика *OPSNR* (сокр. от англ. *Overall Peak Signal to Noise Ratio*), в которой значения *MSE* (сокр. от англ. *Mean Squared Error* – среднеквадратическая ошибка) для каждого сжатого кадра видеопоследовательности суммируются и подставляются в формулу (1).

Выбор режима внутреннего предсказания и процесс сжатия видеопоследовательности опорными кадрами осуществляется с помощью авторской реализации видеоэнкодера, совместимого с проектом стандарта *H.265/HEVC* версии № 9 [1]. В данной реализации выбор режима внутреннего предсказания выполняется по минимальному значению суммы абсолютных значений матрицы коэффициентов, получаемых преобразованием Адамара–Уолша матрицы остаточных коэффициентов для каждого режима предсказания. Преобразование Адамара–Уолша, именуемое как *SATD* (сокр. от англ. *Sum of Absolute Transformed Differences* – сумма абсолютных значений преобразованных разностей), является наиболее часто используемым

методом принятия решения при сжатии видео [4], так как является достаточно хорошим приближением к дискретно-косинусному преобразованию (ДКП), выполняемому при непосредственном сжатии остаточных коэффициентов блоков видео кадра. Преобразование Адамара–Уолша применяется при переборе режимов предсказания для ускорения процесса перебора, т. к. его вычисление на ПЭВМ происходит намного быстрее вычисления ДКП.

Таблица. Тестовый набор видеопоследовательностей

Название	<i>BasketballDrill</i>	<i>PeopleOnStreet</i>
Разрешение	832×480	2560×1600
Частота (кадр/с)	50	30
Всего кадров	501	150
Битрейт (Мбит/с)	228	1406,25

На рис. 5 приведен первый кадр тестовой видеопоследовательности *BasketballDrill*, разбитый сеткой на блоки размером 16×16 пикселей. Это максимальный размер макроблока в стандарте *H.264/AVC*, и один из допустимых размеров блока кодирования (*CU – coding unit*) в стандарте *H.265/HEVC*. *CU* может иметь размер 64×64, 32×32, 16×16 и 8×8 [1]. На таком разбиении интересно сравнить режимы внутреннего предсказания, т. к. чем ближе к исходному будет предсказан блок большего размера, тем меньше необходимость дальнейшего разбиения, и меньше информации требуется закодировать.

При любом виде кодирования чем больше предсказание подобно текущему кодируемому блоку, тем меньше разность (остаток) этих двух блоков (кодируемого и предсказанного). Чем меньше разность, тем меньше высоких частот, а значит, в результате дискретно-косинусного преобразования получится меньше коэффициентов, которые необходимо закодировать и передать в декодер. Таким образом, чем лучше предсказание, тем выше эффективность сжатия.

На рис. 5 приведена исследуемая область видеокадра. Область подобрана специально из-за наличия угловых шаблонов (контуры номера игрока, футболки, руки, линии разметки идут под разными углами). На таком участке интересно оценить эффективность и пользу от новых режимов внутреннего предсказания.

На рис. 6–10 показана эффективность сжатия блоков при различных режимах внутреннего предсказания. Эффективность предсказания определяется получаемым размером сжатого блока, так как это непосредственная характеристика степени сжатия, не отражающая, тем не менее, напрямую качество предсказания и потери в декодированном блоке.

Рис. 6 показывает эффективность сжатия при использовании только *DC* режима предсказания. Каждый предсказанный блок однороден, так как полностью заполняется вычисленным средним

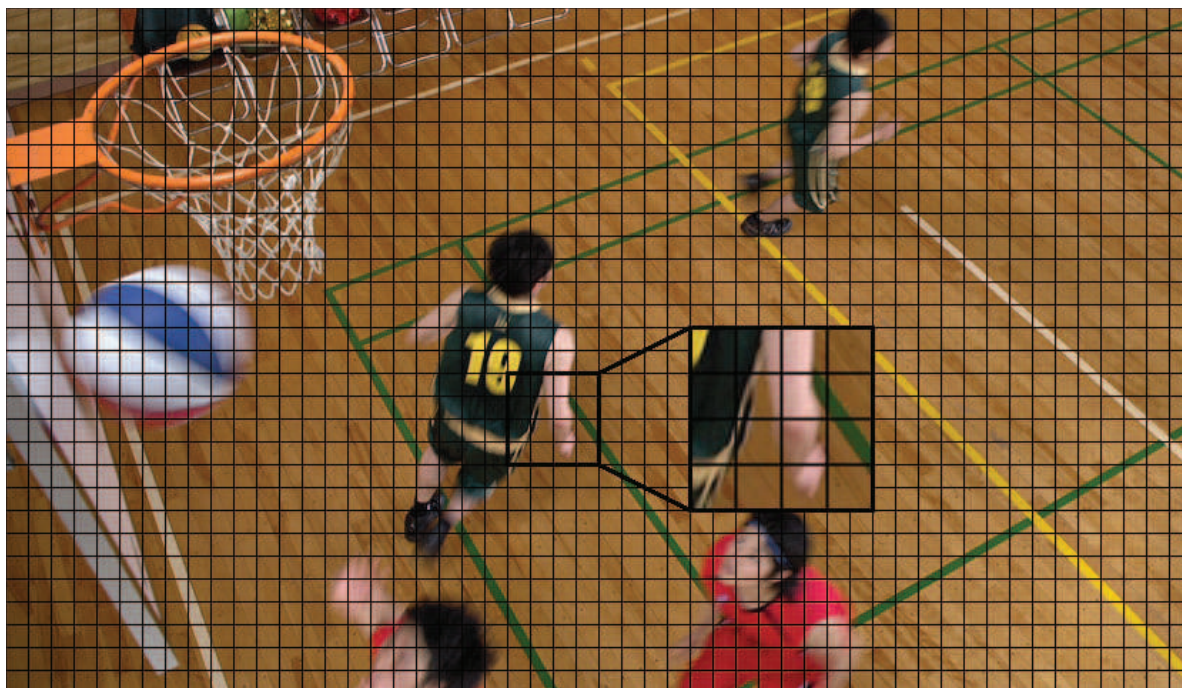


Рис. 5. Разбиение кадра видео последовательности BasketballDrill на блоки размером 16×16 с выделением исследуемого участка

значением соседних пикселей. Поскольку сам участок кадра неоднороден, такое предсказание является крайне неточным и достаточно близко предскажет лишь однородные области: блоки (3;0), (0;1), (3;3).

Рис. 7 показывает эффективность сжатия при использовании только плоского режима предсказания. Каждый блок экстраполируется линейной функцией от левой и верхней границ блока. В результате предсказание получается не такое однородное, как при DC режиме, и некоторые блоки предсказываются лучше, например, блоки (0;0) и (0;1) имеют заметно меньший размер сжатия.

При использовании всего доступного набора режимов внутреннего предсказания стандарта H.265/HEVC (рис. 8) удается повысить эффективность сжатия блоков (1;0), (3;1), (3;2). Для последних двух блоков выбран режим внутреннего предсказания 20, соответствующий примерному углу интерполяции 56° . Линия разметки баскетбольной

площадки, попадающая в этот участок области, очень хорошо предсказывается с помощью такого режима углового предсказания (рис. 6, б, в), угол направления которого максимально близок к углу разметки на данном участке. Размеры сжатых коэффициентов для этого участка на 48 и на 25 % меньше, чем при предсказании от среднего и плоском предсказании. Режимы внутреннего предсказания 20 нет в стандарте H.264/AVC. При использовании только режимов внутреннего предсказания, доступных в этом стандарте, наиболее эффективными угловыми режимами сжатия блоков (3;1), (3;2) становятся 22 и 18 (рис. 9). При этом их эффективность заметно уступает режимам DC и плоскому режиму (рис. 1, 2). Видно, что предсказание разметки баскетбольной площадки режимом внутреннего предсказания 18 с углом в 45° (блок (3;2) на рис. 9) дает большую разность, более высокую ошибку предсказания (рис. 9, в) и, как следствие, больший размер сжатия (рис. 9, г).



а



б

1231	947	690	486
378	1056	843	785
1210	1013	712	896
1218	353	730	702

в

Рис. 6. Иллюстрация эффективности сжатия блоков при DC режиме внутреннего предсказания: а) блоки, предсказанные режимом DC; б) блоки после вычитания предсказываемого и предсказанного блоков (остатки); в) соответствующие размеры сжатых блоков, бит

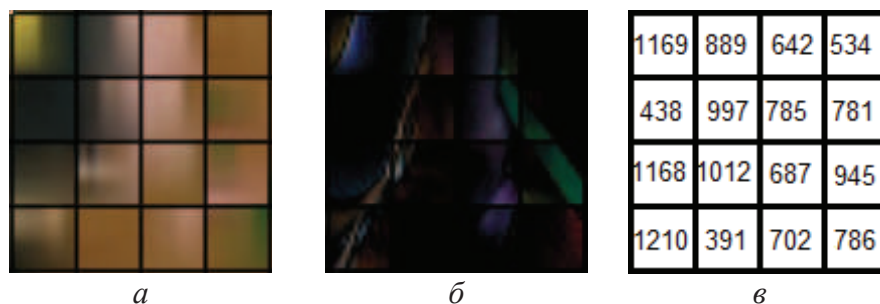


Рис. 7. Иллюстрация эффективности сжатия блоков при плоском режиме внутреннего предсказания: а) блоки, предсказанные плоским режимом; б) блоки после вычитания предсказываемого и предсказанного блоков (остатки); в) соответствующие размеры сжатых блоков, бит

Наконец, если оставить только режимы, доступные в стандарте H.264/AVC для блоков размером 16×16 , то можно заметить снижение эффективности сжатия блоков, например, блоки (1;2) и (1;3) теперь имеют размер 845 и 783 бит соответственно (рис. 10), а эффективность снизилась на 9 и на 48 % соответственно по сравнению с тем, когда были доступны все режимы предсказания.

В стандарте H.264/AVC существуют 3 размера блоков предсказания: 16×16 , 8×8 и 4×4 . Для блоков размером 8×8 и 4×4 допустимо использовать режимы предсказания 1, 6, 10, 14, 18, 22, 26, 30, 34. Таким образом, плоский режим предсказания для блоков такого размера не может использоваться. Для блоков размером 16×16 доступно всего 4 режима внутреннего предсказания: 0 (плоский), 1 (от среднего), 10 (горизонтальный) и 26 (вертикальный).

Добавленные режимы внутреннего предсказания эффективны не на всех блоках предсказания, а скорее на участках с отчетливым шаблоном, идущим под соответствующим углом, как, например, линии разметки площадки в рассмотренном примере. Однако эффективность и частота выбора новых режимов внутреннего предсказания зависит от характеристик самого видео: чем больше угловых шаблонов внутри кадра, тем более эффективны новые режимы предсказания, а значит, тем более эффективно сжатие внутренних блоков в стандарте H.265/HEVC по сравнению со стандартом H.264/AVC, в котором этих предсказаний нет.

Следующей задачей является оценка эффективности дополнительных режимов предсказания при сжатии видео последовательностей. Для этого используется тестовая модель HM 9.0 – общедоступная программная реализация стандарта

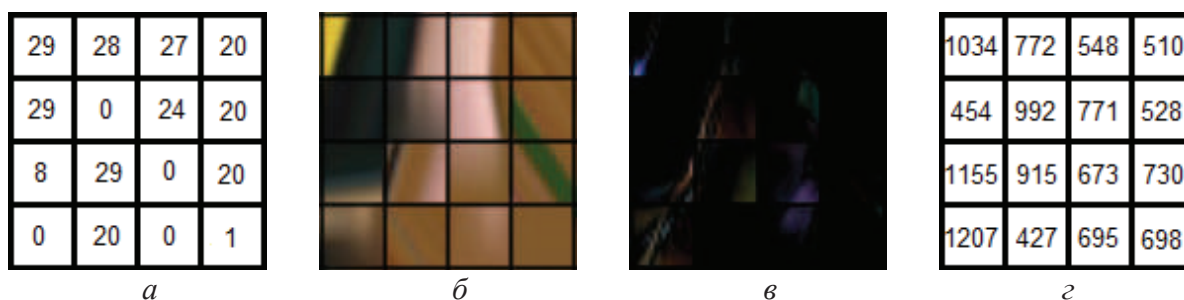


Рис. 8. Иллюстрация эффективности сжатия блоков при выборе из всех режимов внутреннего предсказания: а) режим внутреннего предсказания для данного блока; б) предсказанные блоки; в) блоки после вычитания предсказываемого и предсказанного блоков (остатки); г) соответствующие размеры сжатых блоков, бит

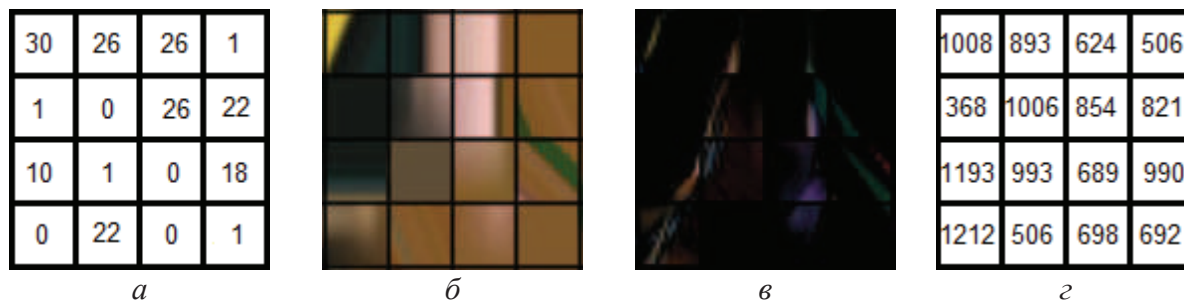


Рис. 9. Иллюстрация эффективности сжатия блоков при выборе из всех режимов внутреннего предсказания, которые были в H.265/AVC: а) режим внутреннего предсказания для данного блока; б) предсказанные блоки; в) блоки после вычитания предсказываемого и предсказанного блоков (остатки); г) соответствующие размеры сжатых блоков, бит

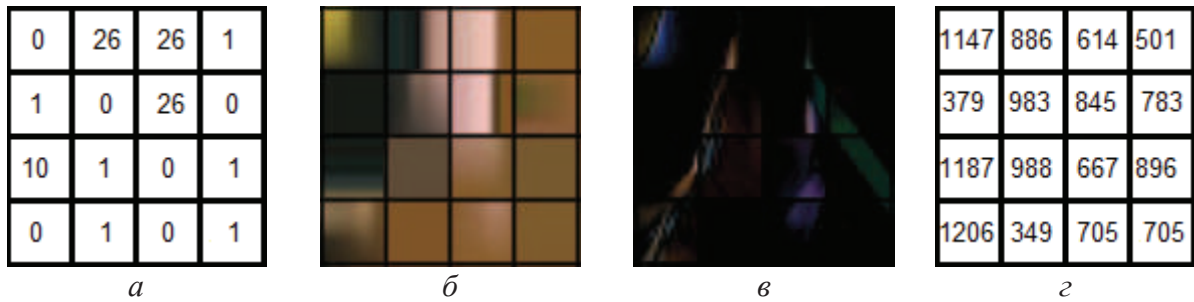


Рис. 10. Иллюстрация эффективности сжатия блоков при выборе из всех режимов внутреннего предсказания, доступных в H.265/AVC для блоков размером 16×16: а) режим внутреннего предсказания для данного блока; б) предсказанные блоки; в) блоки после вычитания предсказываемого и предсказанного блоков (остатки); г) соответствующие размеры сжатых блоков, бит

H.265/HEVC, созданная разработчиками стандарта JCT-VC. Эта реализация выполняет выбор режима предсказания на основе полного перебора всех вариантов предсказания с вычислением SATD и модели RDO (от англ. *rate-distortion optimization* – оптимизация соотношения степени сжатия к шуму). При использовании полной модели RDO выполняется кодирование нескольких наиболее удачных вариантов предсказаний, вычисляются размеры сжатия и потери в декодированных окнах. Полученные метрики используются при окончательном выборе режима предсказания, что позволяет определить наиболее эффективный вариант.

Используется три конфигурации сжатия по стандарту H.265/HEVC: первая с максимальным размером блока кодирования 64×64 пикселей, глубина разбиения 4 (обозн. HM 64-4 на рис. 11, 12); вторая с максимальным размером блока кодирования 32×32 пикселей, глубина разбиения 3; третья с максимальным размером блока кодирования 16×16 пикселей, глубина разбиения 2, подобно разбиению кадра на блоки в стандарте H.264/AVC (обозн. HM 16-2 на рис. 1, 12). В первом случае

блок предсказания может быть выбран размером 64×64, 32×32, 16×16, 8×8 или 4×4; во втором – размером 32×32, 16×16, 8×8 или 4×4, в третьем – 16×16, 8×8 или 4×4. Дополнительно используется конфигурация с отключенными внутренними предсказаниями, отсутствующими в стандарте H.264/AVC (HM 16-2-AVC). Для сравнения используется тестовая модель JM 18.3 – общедоступная программная реализация стандарта H.264/AVC комитетом JCT-VC. Исследования проводятся на наборе видеопоследовательностей, приведенном в таблице.

Графики отношения скорости видеопотока (битрейта, кбит/с) к качеству сжатия (PSNR, дБ) приведены на рис. 11, 12. На используемых тестовых видеопоследовательностях применение блоков размером 64×64 пикселей с глубиной разбиения 4 дает практически идентичный результат сжатия, что и при использовании блоков размером 32×32 пикселей с глубиной 3 и 16×16 пикселей с глубиной 2. Графики HM 64-4, HM 16-2 на рис. 11, 12 совпадают. Для удобства восприятия график HM 32-3 не отображен на рис. 11, 12, т. к. он совпадает с графиками HM 64-4 и HM 16-2.

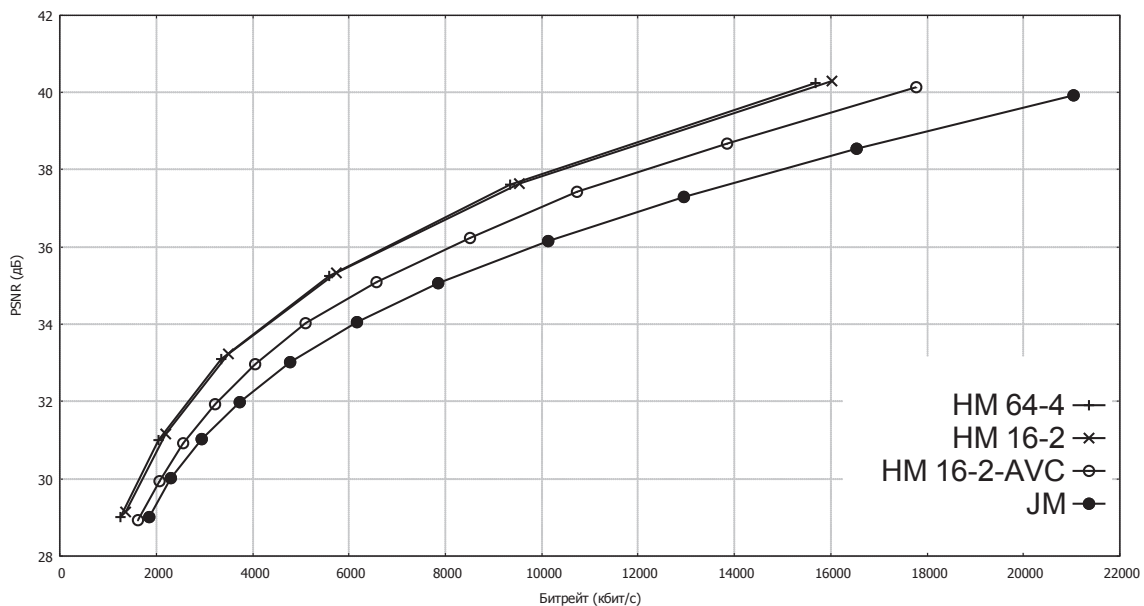


Рис. 11. Сравнение эффективности сжатия, последовательность BasketballDrill

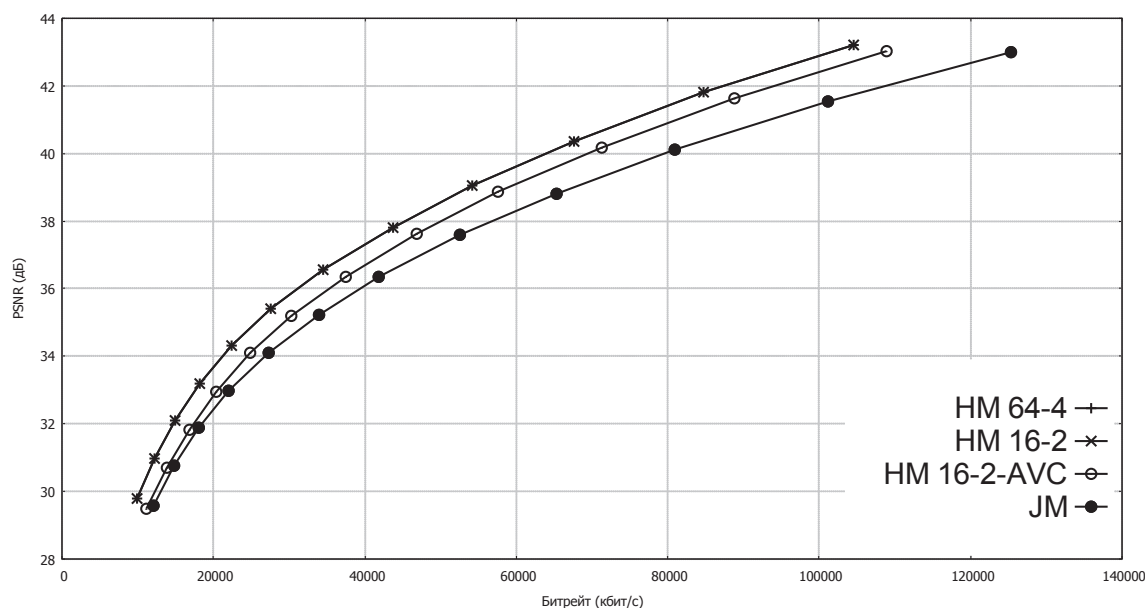


Рис. 12. Сравнение эффективности сжатия, последовательность *PeopleOnStreet*

Последовательности *BasketballDrill* (рис. 11) и *PeopleOnStreet* (рис. 12) имеют разное разрешение, при этом исследуемые конфигурации сжатия дают практически одинаковые результаты относительно разных конфигураций.

Скорости потока несжатых данных у исследуемых последовательностей отличаются в 6 раз. Степень сжатия при одинаковой степени потерь в декодированном видео отличается в 4,5 раза. При качестве 36 дБ эффективность сжатия последовательности *BasketballDrill* стандартом *H.265/HEVC* на 42 % выше, чем при использовании стандарта *H.264/AVC*. Аналогичная разница для последовательности *PeopleOnStreet* составляет 28 %. При этом для обеих последовательностей эффективность от введения дополнительных режимов внутреннего предсказания достигает всего 18 и 14 % соответственно, т. е. 43 и 50 % от общего повышения эффективности.

Видно, что новый стандарт *H.265/HEVC* позволяет на 30...40 % эффективнее стандарта *H.264/AVC* сжимать видеопоследовательности уже с помощью опорных видеокадров (см. графики *HM 16-2* и *JM* на рис. 11, 12). При этом 40...50 % эффективности достигается за счет увеличенного набора внутренних предсказаний (см. графики *HM 16-2* и *HM*

16-2-AVC на рис. 11, 12). Остальное повышение степени сжатия достигается за счет иных улучшений стандарта, в том числе измененной модели энтропийного кодирования.

Выводы

По мнению разработчиков, целью создания стандарта *H.265/HEVC* являлось повышение степени сжатия видеопоследовательностей до двух раз по сравнению с возможностями стандарта *H.264/AVC* при том же качестве. Только на сжатии опорных кадров стандарт *H.265/HEVC* дает до 40 % увеличения степени сжатия при том же качестве. При этом до 50 % повышения эффективности достигается за счет введения дополнительных вариантов внутреннего предсказания, оставшиеся, по-видимому, достигаются за счет модификаций энтропийного кодирования и других усовершенствований стандарта. Такое существенное повышение эффективности сжатия уже на ключевых кадрах позволяет говорить о допустимости двукратного увеличения эффективности сжатия стандарта *H.265/HEVC* по сравнению со стандартом *H.264/AVC*. Для изучения этого вопроса необходимо исследовать возможности стандарта *H.265/HEVC* по сжатию разностных кадров.

СПИСОК ЛИТЕРАТУРЫ

1. Generic coding of moving pictures and associated audio information: Video // MPEG-2 ISO/IEC 13818-2. – 1998. – P. 202.
2. Advanced video coding for generic audiovisual services // Recommendation ITU-T H.264. – 2011. – P. 686.
3. Bross B. et al. High-Efficiency Video Coding (HEVC) text specification draft 9 // JCT-VC: 11th meeting. – Shanghai, CN, 10–19 October, 2012. – 390 p. URL: http://phenix.int-evry.fr/jct/doc_end_user/documents/11_Shanghai/wg11/JCTVC-K1003-v13.zip (дата обращения: 16.12.2012).
4. Шарабайко М.П., Осокин А.Н. Быстродействующий алгоритм фрактального сжатия изображений // Известия Томского по-

литехнического университета. – 2011. – Т. 318. – № 5. – С. 52–57.

5. Сэлмон Д. Сжатие данных, изображений и звука. – М.: Техносфера, 2004. – 368 с.
6. Richardson I. White Paper: H.264/AVC Intra Prediction // Vcodex. – 2011. – P. 1–7. URL: http://www.vcodex.com/files/H264_intraped_wp.pdf (дата обращения: 16.12.2012).
7. Bossen F. Common test conditions and software reference configurations // JCT-VC: 6th Meeting. – Torino, IT, 14–22 July 2011. – 5 p. URL: http://phenix.it-sudparis.eu/jct/doc_end_user/documents/6_Torino/wg11/JCTVC-F900-v1.zip (дата обращения: 16.12.2012).

Поступила 16.12.2012 г.