

**КОМПРЕССИЯ ВИДЕОСИГНАЛОВ ЦИФРОВОГО
ТЕЛЕВИДЕНИЯ ВЫСОКОЙ ЧЕТКОСТИ. ЧАСТЬ 2**

МАИН ХУССЕЙН, ГАССАН ХАЛИЛЬ
Университет Аль-Баас, г. Хомс, Сирийская Арабская Республика

**VIDEOSIGNALS COMPRESSIONS FOR
HIGH DEFINITION DIGITAL TELEVISION. PART 2**

MAIN HYSSEIN, GASSAN HALIL
Arabic Republic of Syria

Выполнен обзор методов компрессии (сжатия) сигналов цифрового телевизионного вещания, рассмотрены основные инструменты кодирования, используемые в стандарте H.264/MPEG-4/AVC, приведено описание методов и технологий кодирования, используемых в стандарте. Приведена оценка качества по сравнению с другими стандартами видеокодирования.

The review of encoding methods of signals for the digital television broadcasting is executed, encoding basic instruments are considered in-use in the standard of H.264/MPEG-4/AVC, types and levels of standard is resulted. Estimation of quality is resulted as compared to other standards of videoencoding.

В полном телевизионном изображении высокой четкости (Full HDTV) предусматривается 1920 элементов изображения (ЭИ) для каждого передаваемого кадра в горизонтальном направлении экрана и 1080 ЭИ (пикселей) – в вертикальном направлении. Произведение этих цифр дает более 2 млн. ЭИ, а с учетом того, что изображение должно быть цветным, т.е. состоящим из трех цветовых компонент (R, G, B), а также изменяться по яркости, общее число пикселей в каждом кадре превышает 6 млн. Поскольку скорость обновления информации в ТВ-вещании составляет 50 (60) передаваемых кадров в секунду, то передача такого значительного объема информации требует существенного частотного ресурса.

Сжатие (компрессия) телевизионного видеосигнала высокой четкости ставит своей целью сужение требуемой для передачи такого сигнала полосы радиочастот и увеличение числа транслируемых каналов в отведенной полосе частот. Для этой цели актуальным на сегодняшний день является применение стандарта H.264/MPEG-4/AVC. Он изначально ориентирован на использование, прежде всего, в телекоммуникационных системах, особенное внимание в нем уделяется увеличению помехоустойчивости и обеспечению гибкости представления информации, необходимой для обеспечения удобства передачи закодированной видеoinформации по разным сетям. Стандартом предусмотрены алгоритмы разделения закодированного потока на блоки, что обеспечивает необходимые для каждого случая характеристики информационного представления.

В первой части статьи отмечалось, что видеокадры в стандарте H.264/MPEG-4/AVC кодируются в виде одного или нескольких слоев, в каждом из которых содержится целое количество макроблоков - от 1 до полного количества макроблоков в кадре (в этом случае весь кадр содержится на одном слое). Количество макроблоков в кадре должно быть одинаковым в границах каждого кадра. Существует минимальная зависимость между закодированными слоями, которая помогает ограничивать распространение ошибок. Всего в стандарте обозначено пять типов закодированных слоев, кроме того, закодированные кадры могут содержать слои разных типов. Например, закодированные кадры базового профиля могут содержать слои I и P, а кадры основного и расширенного профилей могут содержать слои I, P и B.

Кроме того, предусмотрены SI- и SP-слои, которые представляют собой слои, закодированные специальным образом, что обеспечивает эффективное переключение между видеопотоками, и организывает быстрый произвольный доступ для видеокодеров. Общее требование к потоковому видео - необходимость переключения между одним или несколь-

кими кодируемыми потоками. Например, один и тот же видеопоток кодируют на нескольких скоростях цифрового потока для передачи Интернетом, а декодер пытается сначала декодировать самый высокоскоростной поток из тех, что он может получить, а впоследствии ему нужно будет переключаться на более низкую скорость, например из-за падения скорости соединения с сервером.

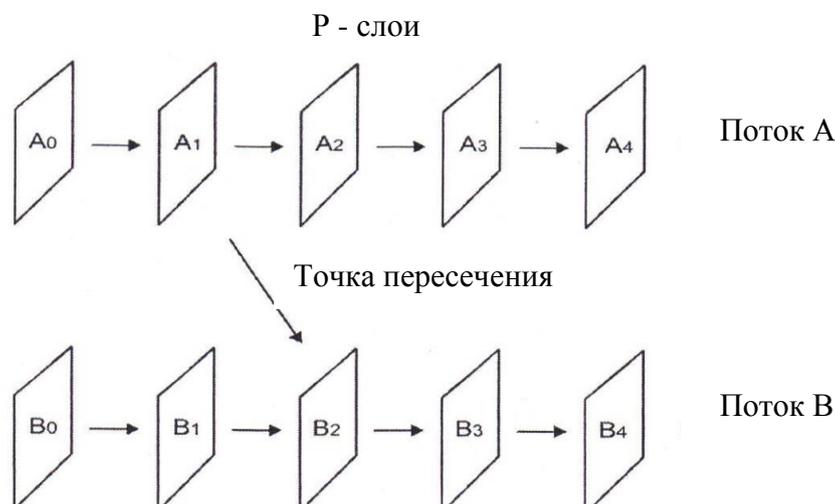


Рисунок 1 - Переключение при помощи I-слоев

Использование SP-слоев было предложено для обеспечения переключений между подобными кодируемыми видеопоследовательностями, например, для кодирования одного и того же видеоконтента с разными скоростями цифрового потока. При этом исключается такой недостаток, как увеличения скорости цифрового потока при использовании I-слоев. (рисунок 1). В точке переключения (кадр 2 в каждой последовательности) ставятся три SP-слоя, каждый из них кодируется с использованием предсказания компенсации движения (что делает их более эффективными, чем I-слои) (рисунок 2).

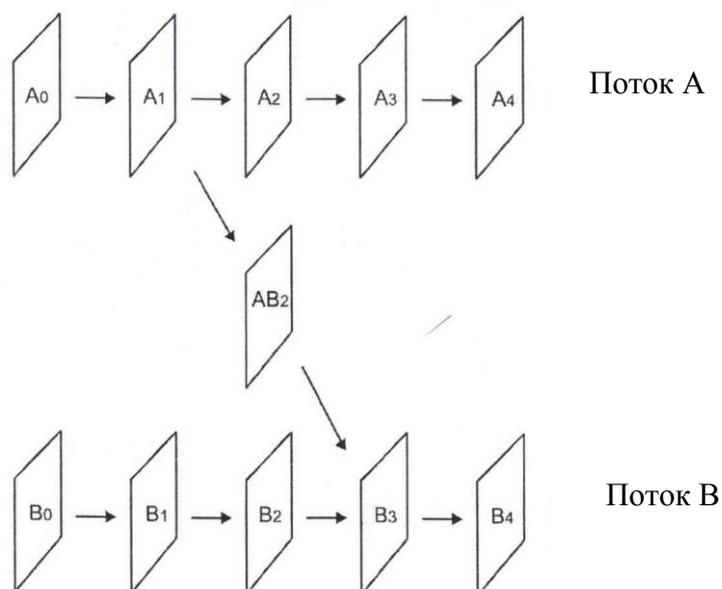


Рисунок 2 - Переключение с помощью SP- слоев

SP-слой A_2 можно декодировать с помощью опорного кадра A_1 , а слой B_2 - при помощи кадра B_1 . Ключевым элементом процесса переключения является SP-слой AB_2 (такие слои имеют название переключающих *SP-слоев*), который сделан таким образом, что его можно

декодировать с помощью опорного кадра A_1 для получения декодированного кадра B_2 (то есть выход декодера будет таким же, как если бы в качестве опорного использовали кадр B_1). В каждой точке переключения нужно расположить SP-слой (и еще один SP-слой BA_2 для переключения в обратном направлении), но это все равно даст меньший объем информации по сравнению с кодированием A_2 и B_2 как I-слоев. В таблице 1 содержится описание шагов, которые должен осуществить декодер при переключении с потока А на поток В.

Таблица 1 - Этапы переключения декодера

Вход декодера	Ссылочный кадр	Выход декодера
P-слой A_0	(предыдущий кадр)	Декодированный кадр A_0
P- слой A_1	Декодированный кадр A_0	Декодированный кадр A_1
SP-слой AB_2	Декодированный кадр A_1	Декодированный кадр B_2
P- слой B_3	Декодированный кадр B_2	Декодированный кадр B_3
...

На рисунке 3 показана упрощенная структурная схема процесса кодирования SP-слоя A_2 , при котором осуществляется вычитание версии с компенсацией движения A_1 (декодированного кадра A_1) из кадра A_2 , после которого идет кодирование полученного остатка. В отличие от обычного P-слоя вычитание осуществляется в области преобразования (после осуществления преобразования блока). SP-слой B_2 кодируют аналогично.

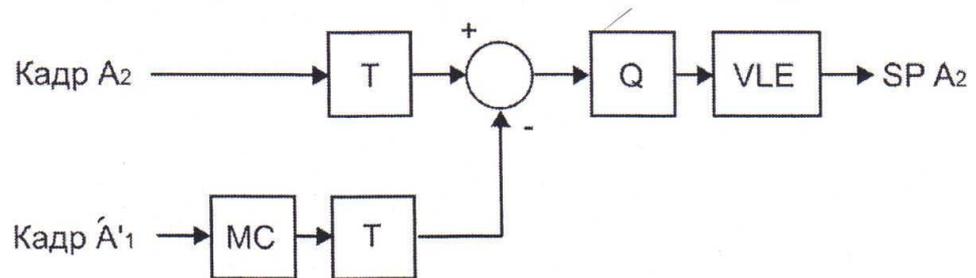


Рисунок 3 - Кодирование SP-слоя A_2

Декодер, который имеет информацию о предыдущем кадре A_1 , может декодировать SP-слой A_2 , так, как это изображено на рисунке 4. Следует заметить, что приведенные блок-схемы являются упрощенными - на практике существуют еще шаги квантования и деквантования для согласования кодера и декодера.

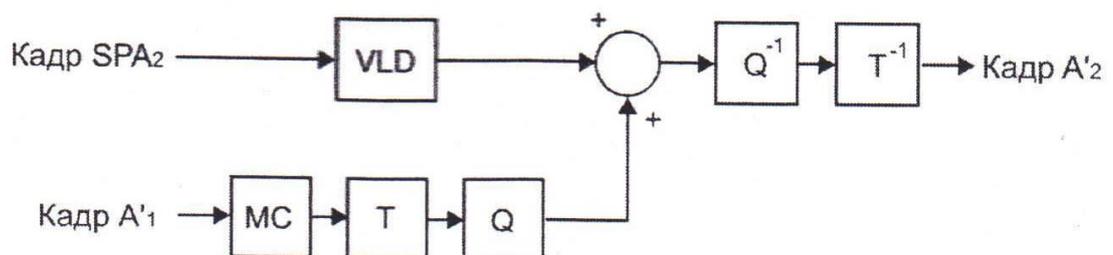


Рисунок 4 - Декодирование SP-слоя A_2

Упрощенный процесс кодирования SP-слоя AB_2 показан на рисунке 5. Кадр B_2 (кадр потока, с которого идет переключение) преобразовывают, и предсказания компенсации движения строят по A'_1 (кадр потока, с которого идет переключение). Блок MC на этой схеме пытается найти наилучшее совпадение для каждого макроблока кадра B_2 , при этом он использует в качестве опорного кадр A_1 . Предсказание компенсации движения преобразуют, потом вычитают из преобразованного B_2 (то есть в случае слоя SP вычитание выполняют в области преобразования). Остаток (после вычитания) квантуют, кодируют и посылают в исходный поток.

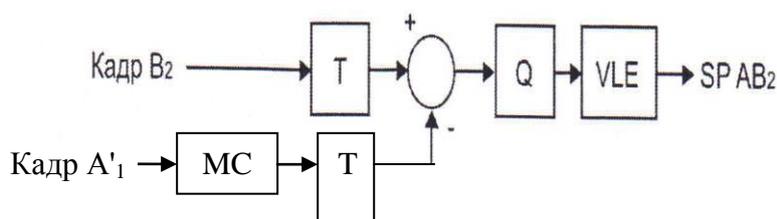


Рисунок 5 - Кодирование SP-слоя AB_2

SI-слои используют таким же образом, как SP-слои, за исключением того, что предсказания формируют в режимах предсказаний *intra*-блоков 4×4 по ранее декодированным ЭИ кадра, который его возобновляет. Этот тип слоев можно использовать при переключении от одной видеопоследовательности к другой.

Взвешенное предсказание

Взвешенное предсказание представляет собой метод модификации (масштабирования) ЭИ предсказания компенсации движения макроблоков слоев P или B. В стандарте H.264/MPEG-4/AVC существует три типа взвешенного предвидения:

- макроблок P-слоя: "явное взвешенное предсказание";
- макроблок B-слоя: "явное взвешенное предсказание";
- макроблок B-слоя: "неявное взвешенное предсказание".

Каждый ЭИ предсказания $pred_0(i, j)$ или $pred_1(i, j)$ умножают на весовой множитель w_0 или w_1 до предсказания компенсации движения. В "явном виде" множители определяют кодером и передают в заголовке слоя. При использовании "неявного типа" предсказания, множители w_0 и w_1 рассчитывают по относительным временным позициям опорных кадров списков 0 и 1. Чем более близок по времени опорный кадр к текущему, тем большим является соответствующий весовой множитель.

Один из возможных вариантов использования взвешенного предсказания заключается в контроле относительной степени участия опорных кадров в процессе компенсации движения. Например, взвешенное предсказание может быть эффективным при кодировании плавных переходов от одной сцены к другой.

Видеоформат

Стандарт H.264/MPEG-4/AVC поддерживает кодирование и декодирование видеопотока с прогрессивной и чересстрочной разверткой в форматах $4 : 2 : 0$, $4 : 2 : 2$ и $4 : 4 : 4$.

Принятым по умолчанию является формат $4 : 2 : 0$ (рисунок 6), в котором цветные ЭИ (Cb и Cr) выровнены по горизонтали с каждым вторым ЭИ яркости и расположены по вертикали между двумя соседними ЭИ яркости. Кадр чересстрочной развертки состоит из двух полукадров (верхнего и нижнего), разделенных по времени. Их формат показан на рисунке 7.

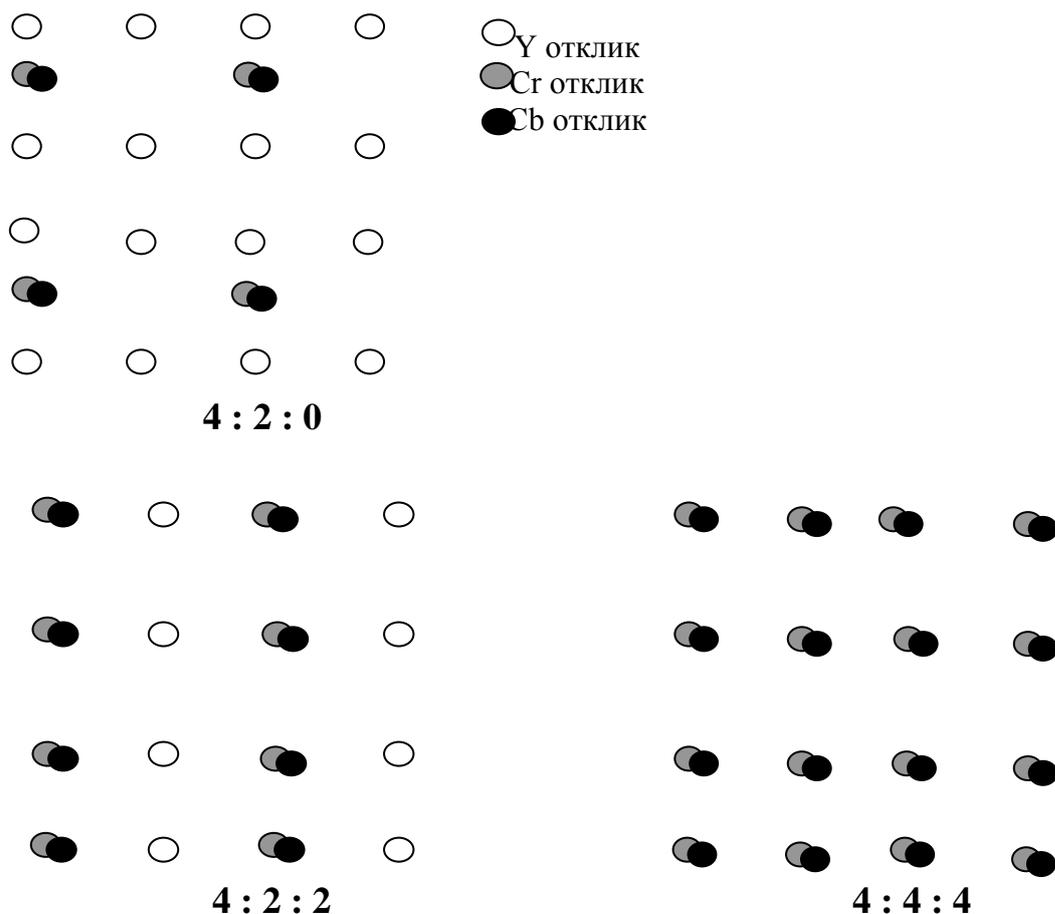


Рисунок 6 – Форматы цветности

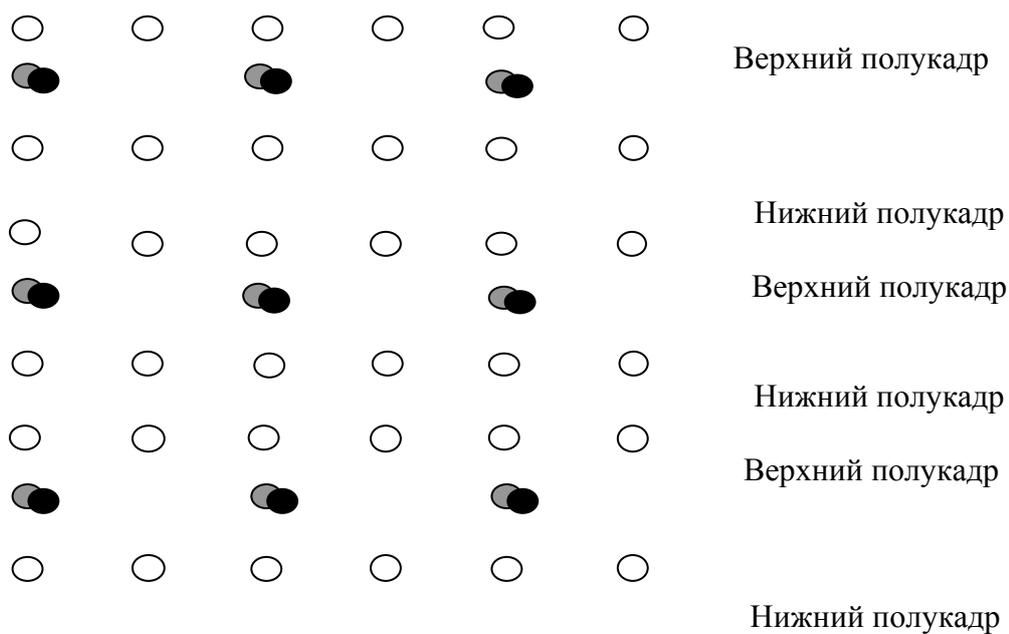


Рисунок 7 – Распределение ЭИ формата 4 : 2 : 0 между верхним и нижним полукадром

Чересстрочная развертка

При кодировании видеоинформации с чересстрочной разверткой требуются методы и инструменты, которые оптимизированы для кодирования макроблоков, состоящих из полукадров. При использовании полукадров в заголовке каждого слоя записывают тип используемых кадров (кадровый или полукадровый). В режиме адаптивного кодирования макроблоков кадр/полукадр (MB-AFF, MacroBlock-Adaptive Frame/Field) выбор кадра или полукадра можно определить на уровне макроблоков. В этом режиме текущий слой обрабатывают единицами, которые имеют ширину 16 ЭИ, и высоту 32 ЭИ яркости, каждый из которых кодируется как "пара макроблоков" (рисунок 8). Кодер может выбрать способ кодирования каждого макроблока в виде двух кадровых макроблоков или двух полукадровых макроблоков. При этом можно избрать оптимальный режим кодирования для каждой области на кадре.

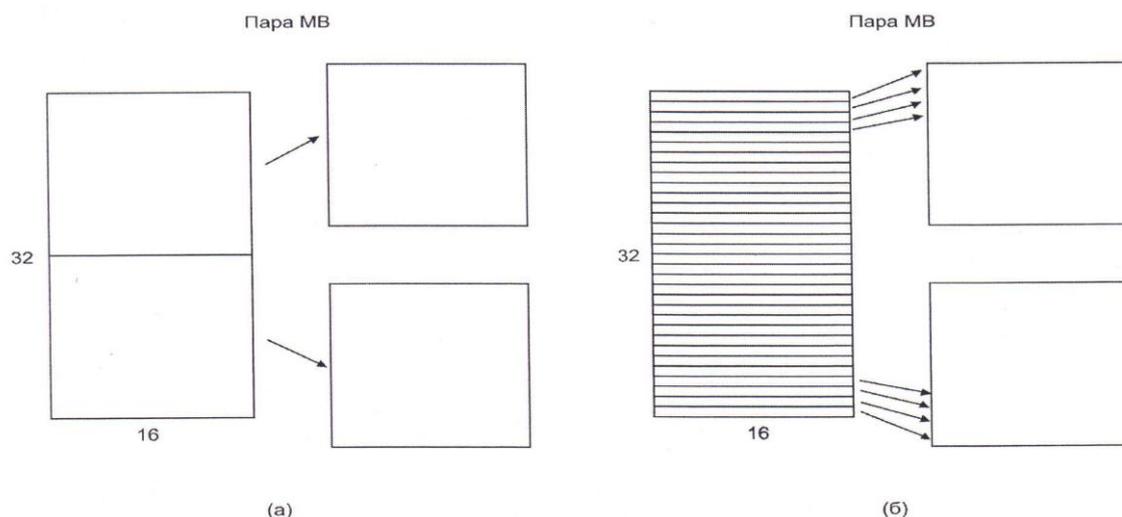


Рисунок 8 - Пары макроблоков при прогрессивном (а) и чересстрочном (б) формате изображения

Кодирование слоя или пары макроблоков в полукадровом режиме требует изменения некоторых этапов кодирования. Например, каждый кодируемый полукадр используют как отдельный опорный кадр с целью предсказания по типу P или B. Предсказание режимов кодирования макроблоков intra и векторов движения макроблоков inter также требуют изменений, в зависимости от того, как закодированы прилегающие макроблоки (по кадрам или полукадрам). В частности, обычный порядок сканирования зигзагом заменяют схемой сканирования, которая показана на рисунке 9.

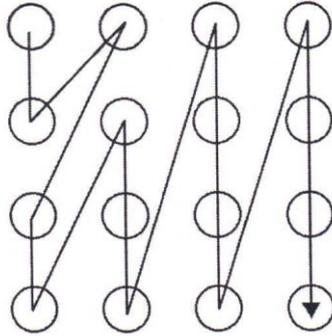


Рисунок 9 - Сканирование блоков яркости 4 x 4 в полукадровом режиме.

Преобразование и квантование

В стандарте H.264/MPEG-4/AVC применяют три разных типа преобразования, в зависимости от типа требующих кодирования данных:

1. преобразования Адамара для матриц 4 x 4 коэффициентов DC яркости макроблоков предсказаний intra в режиме 16 x 16;
2. преобразования Адамара для матриц 2 x 2 коэффициентов DC цветности (любых макроблоков);
3. преобразования на основе DCT для всех других блоков 4 x 4 данных.

Данные внутри макроблока передаются в последовательности, которая приведена на рисунке 10.

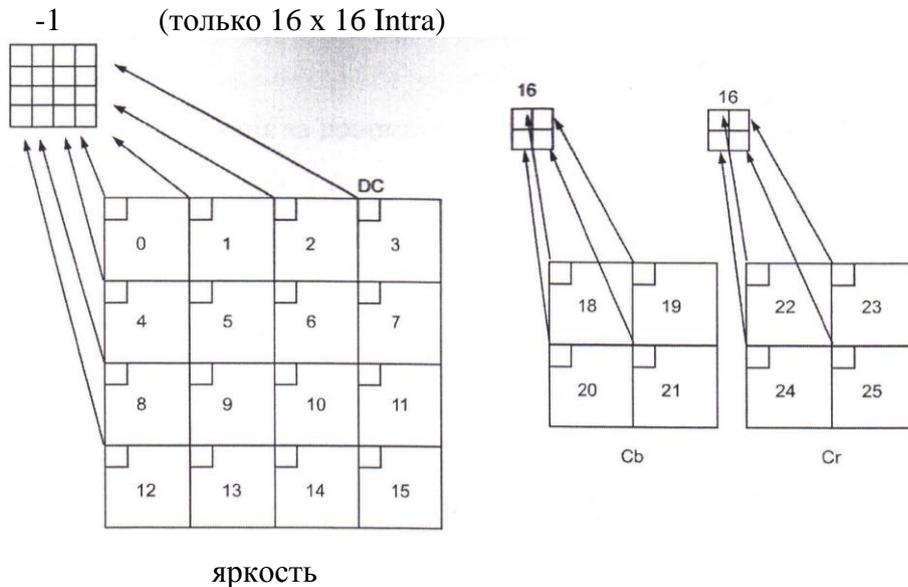


Рисунок 10 - Порядок сканирования блоков внутри макроблока

Если макроблок кодируют в режиме intra 16 x 16, то блок, обозначенный индексом -1 и содержащий коэффициенты DC преобразования каждого блока яркости 4 x 4, передают в первую очередь. После этого передают блоки яркости 0-15 в обозначенном порядке (коэффициенты DC макроблоков, закодированные в режиме 16 x 16 intra не передают). Далее передаются информация о блоках 16 и 17, содержащих матрицу 2 x 2 коэффициентов DC цветных компонентов Cb и Cr, и, наконец, посылают блоки цветности 18-25 (без коэффициентов DC).

Контекстно-адаптивное кодирование с переменной длиной (CAVLC)

Контекстно-адаптивное кодирование с переменной длиной (Context-Adaptive Variable Length Coding, CAVLC) является стандартным видом энтропийного кодирования, используемого в стандарте H.264/MPEG-4/AVC. В методе CAVLC число коэффициентов, не равных 0, кодируется отдельно от их значения и положения в векторе. После зигзаг-сканирования коэффициентов преобразования, большие значения коэффициентов, соответствующие низкочастотной составляющей, находятся в начале вектора, а меньшие (соответствующие высокочастотной составляющей) - в конце вектора.

Рассмотрим кодирование с помощью CAVLC на следующем примере:

Пусть вектор равен 7; 6; -2; 0; -1; 0; 0; 1; 0; 0; 0; 0; 0; 0; 0. Для передачи информации об этом векторе используют следующие данные:

1. Количество ненулевых коэффициентов (N) и количество коэффициентов, равных единице, которые находятся после последнего ненулевого элемента (T1). Для данного примера $T1 = 2$ (коэффициенты -1 и 1), а $N = 5$ (коэффициенты 7; 6; -2; -1; 1).

2. Закодированные значения коэффициентов. При этом для ненулевых элементов, равных единице, необходимо сохранять только знак. Коэффициенты кодируют в обратном порядке, то есть первым закодированным коэффициентом в данном случае будет -2. Для кодирования используют 6 таблиц кодов переменной длины (VLC) Голомба. Для кодирования первого коэффициента используют стартовую таблицу. Контекстная адаптивность алгоритма предоставляет возможность изменения таблицы в зависимости от предыдущих коэффициентов.

3. Информация о знаках. Для кодирования знаков используют один бит. Для кодирования единичных коэффициентов не нужна никакая информация кроме этого бита, а для других коэффициентов знаковый бит добавляют в код Голомба. Местонахождения каждого ненулевого коэффициента кодируют путем обозначения позиций нулей перед последним ненулевым коэффициентом. Эта информация разбивается на две части:

3.1. Общее количество нулей. Это число определяет количество нулей между последним ненулевым коэффициентом вектора и его начальным элементом. Для данного примера это число равно 3. Так как уже известно, что количество ненулевых коэффициентов равняется 5, то это число должно быть в пределах [0; 11]. Для N в пределах от 1 до 15 доступно 15 таблиц. Число N, равное 16, будет означать, что в векторе нет нулевых коэффициентов.

3.2. Размещение нулей в векторе. В данном примере нужно определить положение 3 нулей. Сначала кодируют количество нулей перед последним ненулевым коэффициентом (для данного примера -2). Это число должно находиться в пределах [0; 3], поэтому используется соответствующая таблица VLC. Осталось закодировать положение последнего нуля. Количество нулей перед предпоследним ненулевым коэффициентом должно находиться в пределах [0; 1]. В данном примере это число равно 1. Больше нулей не осталось, поэтому на данной операции кодирование заканчивается.

Контекстно-адаптивное арифметическое кодирование (CABAC)

Контекстно-адаптивное арифметическое кодирование CABAC (Context-based Adaptive Binary Arithmetic Coding) используют для сжатия синтаксических элементов H.264/MPEG-4/AVC. При этом методе кодирования можно получить хорошие результаты по сжатию данных, если выполняются следующие условия: вероятностная модель для каждого синтаксического элемента выбрана в соответствии с его контекстом; адаптивные оценки вероятностей основаны на локальных статистиках; используют арифметическое кодирование вместо VLC. Система кодирования CABAC обозначена флажком `entropy_coding_mode = 1` во множителе параметров кадра.

Этот метод кодирования предусматривает следующие стадии:

1. Переход к двоичной системе: САВАС использует только двоичные коды, то есть кодируется лишь двоичный набор (1 или 0). Символы, принимающие недвоичные значения (например, коэффициенты преобразования или векторы движения, и все другие символы, которые принимают более, чем два значения), приводятся к двоичному виду до применения кодирования.

2. Выбор контекстной модели. Под контекстной моделью понимают вероятностную модель для одного или нескольких битов двоичных символов, которые содержат вероятности каждого бита 0 или 1. Эту модель выбирают из набора доступных моделей, в зависимости от статистики кодирования символов в недавно прошедшей видеoinформации.

3. Арифметическое кодирование: арифметический кодер кодирует каждый бит в зависимости от выбранной вероятностной модели. Следует отметить, что при этом существует всего два субинтервала для каждого бита (соответственно для 0 и 1).

4. Обновление вероятностей: избранная контекстная модель корректируется на основе фактически закодированных значений (изменяются состояния счетчиков соответственно значению 0 или 1).

В стандарте H.264/MPEG-4/AVC определены контекстные модели и схемы двоичного определения всех синтаксических элементов. В целом существует около 400 различных контекстных моделей для разных синтаксических элементов. В начале кодирования каждого слоя происходит инициализация контекстных моделей в зависимости от начальных значений параметра квантования QP (поскольку этот параметр имеет большое влияние на вероятность появления разных синтаксических символов). Дополнительно при кодировании P-, SP- и B-слоев, кодер может выбрать одно из трех множеств параметров инициализации контекстных моделей (см. табл. 2), что позволяет лучше адаптироваться к разным типам видеоконтента.

Таблица 2 - Контекстные модели для первого бита

e_k	Контекстная модель для первого бита
$0 < e_k < 3$	Модель 0
$3 < e_k < 33$	Модель 1
$33 < e_k$	Модель 2

Арифметический кодер имеет три специфических свойства:

1. Оценку вероятностей осуществляют переходным процессом между 64 отдельными состояниями вероятностей для «символа с наименьшей вероятностью» LPS (он является менее всего вероятным из двух возможных состояний 0 или 1)

2. Область R представления текущего состояния арифметического кодера перед каждым шагом квантуют на маленькие области значений (заданных предварительно), при этом имеется возможность табличного определения новых областей (то есть без использования операции умножения).

3. Для символов с распределением вероятностей, близким к равномерному, используют упрощенные процессы кодирования и декодирования (без контекстного моделирования).

Процесс декодирования строится таким образом, чтобы облегчить программную реализацию схемы арифметического кодирования и декодирования. В среднем алгоритм САВАС обеспечивает лучшую степень сжатия по сравнению с алгоритмом кодирования VLC.

ВЫВОДЫ

Осуществлен анализ кодирования видеоинформации по стандарту H.264|MPEG4/AVC, сделано предложение относительно использования стандарта H.264|MPEG-4/AVC в системах цифрового телевизионного вещания, созданных по стандарту DVB-T.

Во второй части статьи представлено описание кодирования в стандарте H.264|MPEG-4/AVC для чересстрочной развертки, контекстно-адаптивное кодирование с переменной длиной (CAVLC), контекстно-адаптивное арифметическое кодирование (CA-VAC).

Выполнен обзор процесса кодирования по стандарту H.264|MPEG-4/AVC, рассмотрены основные инструменты кодирования, которые введены в этом стандарте, приведено описание профилей и уровней стандарта. Приведена оценка качества кодирования по сравнению с другими стандартами видеокодирования.

Стандарт H.264|MPEG-4/AVC является значительным шагом вперед по сравнению со стандартом MPEG-4 и позволяет достичь эффективность сжатия, которая значительно (до 30%) повышает эффективность MPEG-4.

ЛИТЕРАТУРА

1. S.Wenger, T.Stockhamter. An overview on H.264 NAL concept. - JVT-B028, Feb 2002.
2. S.Kumar, L.Xu, G.K.Mandal, S.Panchanathan. Error Resiliency Schemes in H.264/AVC Standard. - Visual Communication & Image Representation (Special issue on Emerging H.264/ AVC Video Coding Standard), Vol. 17(2), pp. 425-450, April 2006.
3. Ричардсон Ян. Видеокодирование H.264 и MPEG-4 - стандарт нового поколения. - М: Техносфера, 2005.
4. Семенюк В.В. Современные методы и стандарты экономного кодирования видеоинформации. – Санкт-Петербург, 2002.
5. T.Wedi, B.Kashiwagi. Subjective quality evaluation H.264/AVC FRExt for HD movie content. - Joint Video Team document JVT-L033, July 2004.