

H.264

H.264, MPEG-4 Part 10 или **AVC** (*Advanced Video Coding*) — лицензируемый стандарт сжатия видео, предназначенный для достижения высокой степени сжатия видеопотока при сохранении высокого качества.

О стандарте

Он был создан ITU-T Video Coding Experts Group (VCEG) совместно с ISO / IEC Moving Picture Experts Group (MPEG) в рамках совместной программы Joint Video Team (JVT).

Стандарты ITU-T H.264 и ISO/IEC MPEG-4 Part 10 (формальное название — ISO/IEC 14496-10) технически полностью идентичны. Финальный черновой вариант первой версии стандарта был закончен в мае 2003 года.

Используется в цифровом телевидении высокого разрешения (HDTV) и во многих других областях цифрового видео.

Возможности

Стандарт H.264 / AVC / MPEG-4 Part 10 содержит ряд новых возможностей, позволяющих значительно повысить эффективность сжатия видео по сравнению с предыдущими (такими, как ASP) стандартами, обеспечивая также большую гибкость применения в разнообразных сетевых средах. Основные из них:

- Многокадровое предсказание:
 - Использование сжатых ранее кадров в качестве опорных (то есть с заимствованием части материала из них) куда более гибко, чем в предыдущих стандартах. Позволяется использование до 32 ссылок на другие кадры, тогда как в ASP и более ранних число ссылок ограничено одним или, в случае В-кадров, двумя кадрами. Это поднимает эффективность кодирования, так как позволяет кодеру выбирать для компенсации движения между большим количеством изображений. В большинстве сцен данная функция обеспечивает не очень большое улучшение в качестве и не даёт заметного понижения битрейта. Однако, для некоторых сцен, например, с частыми повторяющимися участками, возвратно-поступательным движением и т. п. данный подход при сохранении качества позволяет очень сильно снизить затраты битрейта.
 - Независимость порядка воспроизведения изображений и порядка опорных изображений. В предшествующих стандартах устанавливалась жёсткая зависимость между порядком следования изображений для использования при компенсации движения и порядком следования изображений при воспроизведении. В новом стандарте эти ограничения в значительной мере устранены, что позволяет кодеру выбирать порядок изображений для компенсации движения и для воспроизведения с высокой степенью гибкости, которая ограничена только объёмом памяти, который гарантирует возможность декодирования. Устранение ограничения также позволяет в ряде случаев устранить дополнительную задержку, ранее связанную с двунаправленным предсказанием.
 - Независимость методов обработки изображений и возможности их использования для предсказания движения. В предшествующих стандартах изображения, закодированные с использованием некоторых методов (например, двунаправленного предсказания), не могли использоваться в качестве опорных для предсказания движения других изображений видеопоследовательности. Устраняя это ограничение, новый стандарт обеспечивает кодеру большую гибкость и, во многих случаях, возможность использовать для предсказания движения изображение, более близкое по содержанию к кодируемому.
 - Компенсация движения с переменным размером блока (от 16x16 до 4x4 пикселя) позволяет крайне точно выделять области движения.
 - Векторы движения, выходящие за границы изображения. В MPEG-2 и предшествовавших ему стандартах векторы движения могли указывать только на пиксели, находящиеся в границах декодированного опорного изображения. Методика экстраполяции за границы изображения, появившаяся как опция в H.263, включена в новый стандарт.
 - Шеститочечная фильтрация компонента яркости для полупиксельного предсказания с целью уменьшения зубчатости краев и, в конечном счёте, обеспечения большей чёткости изображения.
 - Точность до четверти пикселя (Qpel) при компенсации движения обеспечивает очень высокую точность описания движущихся областей (что особенно актуально для медленного движения). Цветность, как правило, хранится с разрешением, уменьшенным вдвое по вертикали и горизонтали (прореживание цвета), поэтому компенсация движения для компонента цветности использует точность в одну восьмую пикселя цветности.
 - Взвешенное предсказание, позволяющее использовать масштабирование и сдвиг после компенсации движения на величины, указанные кодером. Такая методика может чрезвычайно сильно поднять эффективность кодирования для сцен с изменением освещённости, например при эффектах затемнения, постепенного появления изображения.
- Пространственное предсказание от краёв соседних блоков для I-кадров (в отличие от предсказания только коэффициента трансформации в H.263+ и MPEG-4 Part 2, и дискретно-косинусного коэффициента в MPEG-2 Part 2). Новая методика экстраполяции краёв ранее декодированных частей текущего изображения повышает качество сигнала, используемого для предсказания.
- Сжатие макроблоков без потерь:
 - Метод представления макроблоков без потерь в РСМ, при котором видеоданные представлены непосредственно, позволяющий точно описывать определённые области и допускающий строгое ограничение на количество закодированных данных для каждого макроблока.
 - Улучшенный метод представления макроблоков без потерь, позволяющий точно описывать определённые области, при этом обычно затрачивая существенно меньше битов, чем РСМ (поддерживается не во всех профилях).

- Гибкие функции чересстрочного сжатия (поддерживается не во всех профилях):
 - Адаптивное к изображению кодирование полей (PAFF), позволяющее кодировать каждый кадр как кадр или как пару полей (полукадров) — в зависимости от отсутствия/наличия движения.
 - Адаптивное к макроблокам кодирование полей (MBAFF), позволяющее независимо кодировать каждую вертикальную пару макроблоков (блок 16×32) как прогрессивные или чересстрочные. Позволяет использовать макроблоки 16×16 в режиме разбиения на поля (сравните с 16×8 полумакроблоками в MPEG-2). Почти всегда эффективнее PAFF.
- Новые функции преобразования:
 - Точное целочисленное преобразование пространственных блоков 4×4 (концептуально подобное широко известному DCT, но упрощенное и способное обеспечить точное декодирование^[1]), позволяющее точное размещение разностных сигналов с минимумом шума, часто возникающего в предыдущих кодеках.
 - Точное целочисленное преобразование пространственных блоков 8×8 (концептуально подобное широко известному DCT, но упрощенное и способное обеспечить точное декодирование; поддерживается не во всех профилях), обеспечивающее большую эффективность сжатия схожих областей, чем 4×4.
 - Адаптивный выбор кода между размерами блока 4×4 и 8×8 (поддерживается не во всех профилях).
 - Дополнительное преобразование Адамара, применяемое к дискретно-косинусным коэффициентам основного пространственного преобразования (к коэффициентам яркости, и, в особом случае, цветности) для достижения большей степени сжатия в однородных областях.
- Квантование:
 - Логарифмическое управление длиной шага для упрощения распределения битрейта кодером и упрощенного вычисления обратной длины квантования.
 - Частотно-оптимизированные матрицы масштабирования квантования, выбираемые кодером для оптимизации квантования на основе человеческих особенностей восприятия (поддерживается не во всех профилях).
- Внутренний фильтр деблокинга в цикле кодирования, устраняющий артефакты блочности, часто возникающие при использовании основанных на DCT техниках сжатия изображений.
- Энтропийное кодирование квантованных коэффициентов трансформации:
 - Context-adaptive binary arithmetic coding (CABAC, контекстнозависимое адаптивное бинарное арифметическое кодирование) — алгоритм сжатия без потерь для синтаксических элементов видеопотока на основе вероятности их появления. Поддерживается только в Main Profile и выше. Обеспечивает более эффективное сжатие, чем CAVLC, но требует значительно больше времени на декодирование.
 - Context-adaptive variable-length coding (CAVLC, контекстнозависимое адаптивное кодирование с переменной длиной кодового слова) — альтернатива CABAC меньшей сложности. Тем не менее, оно сложнее и эффективнее, чем алгоритмы, применяемые для тех же целей в более ранних технологиях сжатия видео (как правило это алгоритм Хаффмана).
 - Часто используемое, простое и высоко структурированное кодирование словами переменной длины многих элементов синтаксиса, не закодированных CABAC или CAVLC, известное как коды Голомба (экспоненциальное кодирование Голомба).
- Функции устойчивости к ошибкам:
 - Определение уровня сетевой абстракции (NAL), позволяющее использовать один и тот же синтаксис видео в различных сетевых окружениях, включая наборы параметров последовательности (sequence parameter sets, SPSS) и наборы параметров изображения (picture parameter sets, PPSs), которые обеспечивают большую надёжность и гибкость, чем предыдущие технологии.
 - Гибкое упорядочивание макроблоков (FMO), также известное как группы частей (поддерживается не во всех профилях) и произвольное упорядочивание частей (ASO) — методы реструктурирования порядка представления фундаментальных областей (макроблоков) в изображениях. При эффективном использовании гибкое упорядочивание макроблоков может существенно повысить устойчивость к потере данных.

Благодаря ASO, так как каждая часть изображения может быть декодирована независимо от других (при определённых ограничениях кодирования), новый стандарт позволяет посылать и получать их в произвольном порядке друг относительно друга. Это может снизить задержку в приложениях реального времени, особенно при использовании на сетях, имеющих режим работы **доставка вне очереди**. Эти функции могут также использоваться для множества других целей помимо восстановления ошибок.

- - Разбиение данных — функция, обеспечивающая разделение данных разной важности (например, векторы движения и другая информация предсказания имеет большую значимость для представления видеоконтента) по разным пакетам данных с разными уровнями защиты от ошибок (поддерживается не во всех профилях).
 - Избыточные части. Возможность посылки кодером избыточного представления областей изображений, позволяя воспроизвести области изображений (обычно с некоторой потерей качества), данные о которых были потеряны в процессе передачи (поддерживается не во всех профилях).
 - Нумерация кадров, позволяющая создание «подпоследовательностей» (включая временное масштабирование включением дополнительных кадров между другими) а также обнаружение (и скрытие) потерь целых кадров при сбоях канала или пропаже пакетов.

Профили

Стандарт определяет комплекты возможностей, которые называются профили, ориентированные на конкретные классы приложений.

Baseline Profile (Базовый профиль)

Применяется в недорогих продуктах, требующих дополнительной устойчивости к потерям. Используется для видеоконференций и в мобильных продуктах. Включает все возможности Constrained Baseline Profile и, дополнительно, возможности для большей устойчивости к потерям при передаче. С появлением Constrained Baseline Profile отошел на второй план, т.к. все потоки Constrained Baseline Profile соответствуют Baseline Profile, и оба этих профиля имеют общий код идентификатора.

Constrained Baseline Profile (Ограниченный базовый профиль)

Рассчитан на применение в недорогих продуктах. Включает набор возможностей, общих для профилей Baseline, Main, и High профилей.

Main Profile (Основной профиль)

Применяется для цифрового телевидения стандартной четкости в трансляциях, использующих сжатие MPEG-4 в соответствии со стандартом DVB.

Extended Profile (Расширенный профиль)

Предназначен для потокового видео, имеет относительно высокую степень сжатия и дополнительные возможности для повышения устойчивости к потере данных.

High Profile (Высокий профиль)

Является основным для цифрового вещания и видео на оптических носителях, особенно для телевидения высокой четкости. Используется для Blu-Ray видеодисков и DVB HDTV вещания.

High 10 Profile (Высокий профиль 10)

Дополнительно поддерживает 10-битовую глубину кодирования изображения.

High 4:2:2 Profile (Hi422P)

В основном нацелен на профессиональное использование при работе с чересстрочным видеопотоком. Поддерживает дополнительный вариант кодирования цветности.

High 4:4:4 Predictive Profile (Hi444PP)

Базируясь на Hi422P, включает еще один вариант кодирования цветности и работу с 14-битной глубиной кодирования.

Для профессионального применения стандарт содержит четыре дополнительных all-Intra («всё внутри») профиля, которые характеризуются отсутствием межкадрового сжатия. То есть, при кодировании одного кадра информация о соседних не используется:

High 10 Intra Profile

High 4:2:2 Intra Profile

High 4:4:4 Intra Profile

CAVLC 4:4:4 Intra Profile

С принятием расширения Scalable Video Coding (SVC) к стандарту были добавлены три профиля, соответствующие базовым, с добавлением возможности включать потоки более низкого разрешения.

Scalable Baseline Profile

Scalable High Profile

Scalable High Intra Profile

Добавление расширения Multiview Video Coding (MVC) принесло еще два дополнительных профиля:

Stereo High Profile

Этот профиль рассчитан на стереоскопическое 3D видео (два изображения).

Multiview High Profile

Этот профиль поддерживает два или несколько изображений (каналов) в потоке с использованием как межкадрового, так и межканального сжатия, но не поддерживает некоторые возможности MVC.

Функции поддержки в отдельных профилях

Функции	CBP	BP	XP	MP	HiP	Hi10P	Hi422P	Hi444PP
Chroma formats	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0/4:2:2	4:2:0/4:2:2/4:4:4
Sample depths (bits)	8	8	8	8	8	8 to 10	8 to 10	8 to 14
Flexible macroblock ordering (FMO)	Нет	Да	Да	Нет	Нет	Нет	Нет	Нет
Arbitrary slice ordering (ASO)	Нет	Да	Да	Нет	Нет	Нет	Нет	Нет
Redundant slices (RS)	Нет	Да	Да	Нет	Нет	Нет	Нет	Нет
Data partitioning	Нет	Нет	Да	Нет	Нет	Нет	Нет	Нет
SI and SP slices	Нет	Нет	Да	Нет	Нет	Нет	Нет	Нет
B slices	Нет	Нет	Да	Да	Да	Да	Да	Да
Interlaced coding (PicAFF, MBAFF)	Нет	Нет	Да	Да	Да	Да	Да	Да
CABAC entropy coding	Нет	Нет	Нет	Да	Да	Да	Да	Да
8×8 vs. 4×4 transform adaptivity	Нет	Нет	Нет	Нет	Да	Да	Да	Да
Quantization scaling matrices	Нет	Нет	Нет	Нет	Да	Да	Да	Да
Separate C _b and C _r QP control	Нет	Нет	Нет	Нет	Да	Да	Да	Да
Monochrome (4:0:0)	Нет	Нет	Нет	Нет	Да	Да	Да	Да
Separate color plane coding	Нет	Нет	Нет	Нет	Нет	Нет	Нет	Да
Predictive lossless coding	Нет	Нет	Нет	Нет	Нет	Нет	Нет	Да

Уровни

Согласно определению стандарта, «уровень» является определенным набором ограничений, указывающих степень требуемой производительности декодера для профиля. Например, поддержка уровня в профиле будет указывать максимальное разрешение изображения, частоту кадров и битрейт так, что декодер можно будет использовать. Декодер, который соответствует данному уровню, обязан декодировать все потоки битов, которые кодируются для этого уровня и для всех более низких уровней.

Уровни с максимальными параметрами

Уровень	Макс. кол-во макроблоков		Макс. скорость видеопотока (VCL) кбит/с				Примеры максимального разрешения@частоты кадров (макс. кол-во сохраненных кадров)
	в секунду	в кадре	BP, XP, MP	HiP	Hi10P	Hi422P, Hi444PP	
1	1,485	99	64	80	192	256	128×96@30.9 (8) 176×144@15.0 (4)
1b	1,485	99	128	160	384	512	128×96@30.9 (8) 176×144@15.0 (4)
1.1	3,000	396	192	240	576	768	176×144@30.3 (9) 320×240@10.0 (3) 352×288@7.5 (2)
1.2	6,000	396	384	480	1,152	1,536	320×240@20.0 (7) 352×288@15.2 (6)
1.3	11,880	396	768	960	2,304	3,072	320×240@36.0 (7) 352×288@30.0 (6)
2	11,880	396	2,000	2,500	6,000	8,000	320×240@36.0 (7) 352×288@30.0 (6)
2.1	19,800	792	4,000	5,000	12,000	16,000	352×480@30.0 (7) 352×576@25.0 (6)
2.2	20,250	1,620	4,000	5,000	12,000	16,000	352×480@30.7(10) 352×576@25.6 (7) 720×480@15.0 (6) 720×576@12.5 (5)
3	40,500	1,620	10,000	12,500	30,000	40,000	352×480@61.4 (12) 352×576@51.1 (10) 720×480@30.0 (6) 720×576@25.0 (5)
3.1	108,000	3,600	14,000	17,500	42,000	56,000	720×480@80.0 (13) 720×576@66.7 (11) 1280×720@30.0 (5)
3.2	216,000	5,120	20,000	25,000	60,000	80,000	1,280×720@60.0 (5) 1,280×1,024@42.2 (4)
4	245,760	8,192	20,000	25,000	60,000	80,000	1,280×720@68.3 (9) 1,920×1,080@30.1 (4) 2,048×1,024@30.0 (4)
4.1	245,760	8,192	50,000	62,500	150,000	200,000	1,280×720@68.3 (9) 1,920×1,080@30.1 (4) 2,048×1,024@30.0 (4)
4.2	522,240	8,704	50,000	62,500	150,000	200,000	1,920×1,080@64.0 (4) 2,048×1,080@60.0 (4)
5	589,824	22,080	135,000	168,750	405,000	540,000	1,920×1,080@72.3 (13) 2,048×1,024@72.0 (13) 2,048×1,080@67.8 (12) 2,560×1,920@30.7 (5) 3,680×1,536@26.7 (5)
5.1	983,040	36,864	240,000	300,000	720,000	960,000	1,920×1,080@120.5 (16) 4,096×2,048@30.0 (5) 4,096×2,304@26.7 (5)

Патенты

В странах, где действуют патенты на программное обеспечение, разработчики программного обеспечения, использующего алгоритмы H.264/AVC, обязаны платить лицензионные отчисления держателям патентов. Держателями таковых, в частности, являются Microsoft, Fujitsu, Philips, Apple, Samsung, Cisco, Toshiba, Panasonic ^{[2][3]}. Также существует организация MPEG LA, которая является администратором консолидированного пула патентов ^{[4][5]}. Всего существует более сотни патентов, так или иначе затрагивающих или описывающих алгоритмы H.264. Сроки действия части из них уже истекли, однако некоторые будут продолжать действовать в США вплоть до 2028 года ^{[6][2]}.

В марте 2011 г. Министерство юстиции США начало расследование против MPEG LA по подозрению в использовании патентного права с целью устранения конкурента - WebM от Google. Поводом к началу расследования стали обвинения в нарушении патентов третьих разработчиков.^[7]

Недостатки

Кодеки для MPEG-4 AVC более требовательны к ресурсам, нежели кодеки на основе MPEG-4 ASP (такие, как DivX и XviD)^[8], однако это компенсируется другими достоинствами^[9].

Формат запатентован, и создатели кодеков обязаны платить за их распространение путём покупки лицензий. С 2011 года MPEG LA могла бы начать взимать плату и с тех, кто участвует в кодировании и/или бесплатном предоставлении пользователям видеопотока в AVC.^{[10][11]} Однако позже этот срок был изменён на 2015 год, а 26 августа 2010 года компания MPEG LA объявила, что за бесплатное предоставление пользователям видеопотока в H.264 плата взиматься не будет.^[12]

Примечания

- ↑ В предшествующих стандартах преобразование оговаривалось только в рамках допустимых ошибок из-за практической нереализуемости точного обратного преобразования. В итоге, каждая реализация декодера могла давать слегка различное декодированное видео (что обуславливалось несовпадением представления декодированного видео на кодере и декодере), что приводит к снижению эффективного качества видео.
- ↑^{1 2} <http://www.mpegla.com/main/programs/avc/Documents/avc-att1.pdf>
- ↑ MPEG LA — The Standard for Standards — AVC Patent List
- ↑ MPEG LA License for MPEG-4 Video FAQ.
- ↑ MPEG LA — The Standard for Standards — AVC Introduction
- ↑ [whatwg] Codecs for <audio> and <video>
- ↑ В США MPEG LA подозревают в попытках нечестно конкурировать с Google. *ИТС.иа* (7 марта 2011). Архивировано из первоисточника 15 февраля 2012. Проверено 7 марта 2011.
- ↑ *Филипп Казаков*. h264. Год спустя. Кодирование видео в MPEG-4 AVC. Часть первая // *Компьютерра-онлайн*. — 16 октября 2006.
- ↑ *И. В. Олейник*. H.264. Некоторая правда о заблуждениях // *Системы безопасности* : журнал. — 2009. — № 2.
- ↑ *Jan Ozer* H.264 Royalties: what you need to know (англ.). Streaming Learning Center (22 июня 2009). Архивировано из первоисточника 15 февраля 2012. Проверено 7 июля 2009.
- ↑ *Tim Siglin* The H.264 Licensing Labyrinth (англ.). Streaming Media Inc. / Information Today Inc. (12 февраля 2009).^(недоступная ссылка — история) Проверено 7 июля 2009.
- ↑ MPEG LA’s AVC License Will Not Charge Royalties for Internet Video that is Free to End Users through Life of License (англ.)

Источник: ru.wikipedia.org