# MPEG-4 Visual (группа Part 2 стандартов MPEG-4)

**1999й год**

MPEG-4 Visual (Part 2 ISO/IEC 14496, «кодирование аудиовизуальных объектов») представляет собой объемный документ, описывающий довольно широкий класс функций, относящихся к кодированию и представлению визуальной информации. Стандарт имеет дело со следующими типами данных:

* движущиеся изображения (прямоугольные кадры):
* видеообъекты (области произвольной формы с движущимися предметами);
* двухмерные и трехмерные сеточные объекты (представляющие деформируемые объекты);
* анимированные лица и фигуры людей;
* статические текстуры (неподвижные изображения).

КРОМЕ ПРЯМОУГОЛЬНЫХ КАДРОВ, ОСТАЛЬНОЕ ПРАКТИЧЕСКИ НЕ ИСПОЛЬЗУЕТСЯ

Несмотря на многообразие инструментов, задаваемых стандартом, в сердцевине MPEG-4 Visual находится довольно простой механизм видеокодирования, использующий кодек на основе блоков с компенсацией движения и с последующим преобразованием DCT, квантованием и энтропийным кодированием. Синтаксис этого стержневого кодека (при некоторых ограничениях) идентичен ядру Н.263.

# Стандарт H.264 / AVC / MPEG-4 Part 10

**2003й год**

Группы экспертов ITU-Т и MPEG объединились в JVT (Joint Video Team) и разработали новый стандарт, который превосходит по степени сжатия видеоизображений опубликованные ранее стандарты MPEG-4 и Н.263. Этот стандарт получил название AVC (Advanced Video Coding, усовершенствованное кодирование видео) и был опубликован одновременно как стандарт MPEG-4 Part 10 и как рекомендация ITU-T H.264.

Стандарт Н.264 был разработан для эффективного кодирования видео прямоугольного формата. Его первоначальной задачей было обеспечение функциональности, схожей с Н.263, но с лучшим сжатием и более надежным механизмом передачи данных. Большинство функциональных блоков кодера (прогноз, преобразование, квантование, энтропийное кодирование) - такие же, как и в Н.261, Н.263, - сохранились, но подверглись существенной переработке.

Для энтропийного кодирования выбраны Exp-Golomb коды (экспоненциальные коды Golomb) - коды переменной длины с регулярной структурой, кодирование переменной длины с адаптацией на основе контекста (CAVLC), адаптивное двоичное арифметическое кодирование на основе контекста (CABAC).

Компенсация движения выбрана с точностью в четверть пикселя, при этом размер блока компенсации движения может изменяться от 16х16 до 4х4 пикселя.

Дополнительно для повышения сжатия и качества используются (ВИКИПЕДИЯ):

* внутрикадровое кодирование. Эта техника сжимает похожие цвета лучше, чем Н.263, особенно в сложных для кодирования сценах с быстрым движением. Благодаря этой возможности подобные фрагменты кодируются с меньшей битовой скоростью (битрейтом);
* адаптивный фильтр для уменьшения артефактов блочности (Adaptive Deblocking Filter). В рекомендации H.264 используется фильтр удаления блочности для блоков размером 4x4.
* 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 более ранних число ссылок ограничено одним или, в случае [B-кадров](https://ru.wikipedia.org/wiki/%D0%A2%D0%B8%D0%BF%D1%8B_%D0%BA%D0%B0%D0%B4%D1%80%D0%BE%D0%B2), 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](https://ru.wikipedia.org/wiki/%D0%91%D0%B8%D1%82%D1%80%D0%B5%D0%B9%D1%82). 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.
  + [Компенсация движения](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BC%D0%BF%D0%B5%D0%BD%D1%81%D0%B0%D1%86%D0%B8%D1%8F_%D0%B4%D0%B2%D0%B8%D0%B6%D0%B5%D0%BD%D0%B8%D1%8F) с переменным размером блока (от 16x16 до 4x4 пикселя) позволяет крайне точно выделять области движения.
  + Векторы движения, выводящие за границы изображения. В [MPEG-2](https://ru.wikipedia.org/wiki/MPEG-2) и предшествовавших ему стандартах векторы движения могли указывать только на пикселы, находящиеся в границах декодированного опорного изображения. Методика экстраполяции за границы изображения, появившаяся как опция в [H.263](https://ru.wikipedia.org/wiki/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 пиксела (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](https://ru.wikipedia.org/w/index.php?title=%D0%9F%D1%80%D0%BE%D1%80%D0%B5%D0%B6%D0%B8%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5_%D1%86%D0%B2%D0%B5%D1%82%D0%B0&action=edit&redlink=1)), поэтому компенсация движения для компонента цветности использует точность в одну восьмую пиксела цветности.
* Сжатие макроблоков без потерь:
  + Метод представления макроблоков без потерь в [PCM](https://ru.wikipedia.org/wiki/PCM), при котором видеоданные представлены непосредственно, позволяющий точно описывать определённые области и допускающий строгое ограничение на количество закодированных данных для каждого макроблока.
* Новые функции преобразования:
  + Точное целочисленное преобразование пространственных блоков 4×4 (концептуально подобное широко известному [DCT](https://ru.wikipedia.org/wiki/DCT), но упрощенное и способное обеспечить точное декодирование[[1]](https://ru.wikipedia.org/wiki/H.264#cite_note-1)), позволяющее точное размещение разностных сигналов с минимумом шума, часто возникающего в предыдущих кодеках.
  + Точное целочисленное преобразование пространственных блоков 8x8 (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 кодеком между размерами блока 4×4 и 8×8 (поддерживается не во всех профилях).**
  + Дополнительное [преобразование Адамара](https://ru.wikipedia.org/w/index.php?title=%D0%9F%D1%80%D0%B5%D0%BE%D0%B1%D1%80%D0%B0%D0%B7%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5_%D0%90%D0%B4%D0%B0%D0%BC%D0%B0%D1%80%D0%B0&action=edit&redlink=1), применяемое к дискретно-косинусным коэффициентам основного пространственного преобразования (к коэффициентов 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](https://ru.wikipedia.org/w/index.php?title=%D0%94%D0%B5%D0%B1%D0%BB%D0%BE%D0%BA%D0%B8%D0%BD%D0%B3&action=edit&redlink=1) в цикле кодирования, устраняющий [артефакты](https://ru.wikipedia.org/wiki/%D0%90%D1%80%D1%82%D0%B5%D1%84%D0%B0%D0%BA%D1%82%D1%8B_%D1%81%D0%B6%D0%B0%D1%82%D0%B8%D1%8F) блочности, часто возникающие при использовании основанных на [DCT](https://ru.wikipedia.org/wiki/DCT) техниках сжатия изображений.
* [Энтропийное кодирование](https://ru.wikipedia.org/wiki/%D0%AD%D0%BD%D1%82%D1%80%D0%BE%D0%BF%D0%B8%D0%B9%D0%BD%D0%BE%D0%B5_%D0%BA%D0%BE%D0%B4%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5) квантованных коэффициентов трансформации:
  + **Context-adaptive binary arithmetic coding** ([CABAC](https://ru.wikipedia.org/wiki/CABAC), контекстнозависимое адаптивное бинарное [арифметическое кодирование](https://ru.wikipedia.org/wiki/%D0%90%D1%80%D0%B8%D1%84%D0%BC%D0%B5%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B5_%D0%BA%D0%BE%D0%B4%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5)) — 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](https://ru.wikipedia.org/w/index.php?title=CAVLC&action=edit&redlink=1), 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](https://ru.wikipedia.org/wiki/%D0%90%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC_%D0%A5%D0%B0%D1%84%D1%84%D0%BC%D0%B0%D0%BD%D0%B0)).
  + Часто используемое, простое и высоко структурированное кодирование словами переменной длины многих элементов синтаксиса, не закодированных CABAC или CAVLC, известное как [коды Голомба](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%B4%D1%8B_%D0%93%D0%BE%D0%BB%D0%BE%D0%BC%D0%B1%D0%B0) (экспоненциальное кодирование Голомба).
* Функции устойчивости к ошибкам:
  + Определение уровня сетевой абстракции ([NAL](https://ru.wikipedia.org/w/index.php?title=Network_Abstraction_Layer&action=edit&redlink=1)), 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**](https://ru.wikipedia.org/w/index.php?title=Flexible_Macroblock_Ordering&action=edit&redlink=1)**),** также известное как группы частей (поддерживается не во всех профилях) и произвольное упорядочивание частей ([ASO](https://ru.wikipedia.org/w/index.php?title=Arbitrary_Slice_Ordering&action=edit&redlink=1)) — 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.

## Помехозащита

Низкая устойчивость к равновероятно распределенным случайным ошибкам при передаче видеоданных по каналам связи. Это обусловлено прежде всего явлением размножения ошибок при восстановлении подвижных изображений. Проведенные эксперименты по декодированию изображений по видеоданным, имеющим искаженные двоичные разряды, позволили установить, что даже одиночные ошибки в видеоданных приводят к потере основного содержания определенного фрагмента изображений.



## Профили

**Constrained Baseline Profile (Ограниченный базовый профиль)**

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

**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 отошёл на второй план, так как все потоки Constrained Baseline Profile соответствуют Baseline Profile, и оба этих профиля имеют общий код идентификатора.

**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 ещё один вариант кодирования цветности и работу с 14-битной глубиной кодирования.

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 |
| **B slices** | 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 |
| **Monochrome (4:0:0)** | Нет | Нет | Нет | Нет | Да | Да | Да | Да |

# H.265 или HEVC ( High Efficiency Video Coding). MPEG-H Part 2

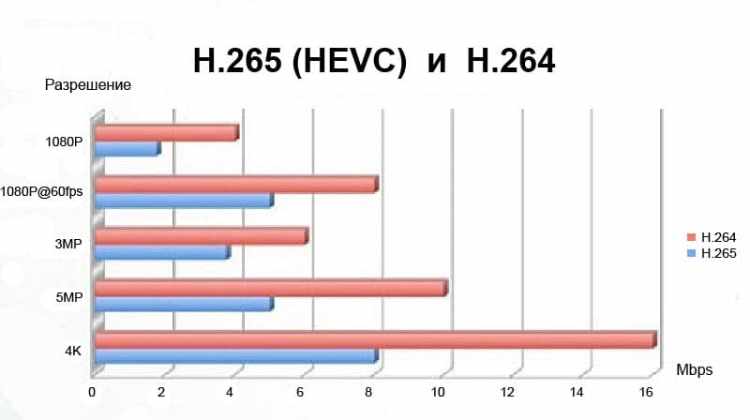
2010й год

Вместо применяющихся в H.264 макроблоков в HEVC используются блоки с древовидной структурой кодирования. Выигрыш кодера HEVC — в применении блоков большего размера.



Другие дополнения:

* [Компенсация движения](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BC%D0%BF%D0%B5%D0%BD%D1%81%D0%B0%D1%86%D0%B8%D1%8F_%D0%B4%D0%B2%D0%B8%D0%B6%D0%B5%D0%BD%D0%B8%D1%8F) с точностью до 1/8 пикселя (Qpel)
* Адаптивное предсказание ошибок кодирования (APEC) в пространственной и частотной областях
* Адаптивный выбор матрицы квантования (AQMS)

 ЯКОБЫ!!!

А вот реальность:

