



# Современные системы цифрового телевидения

Старт 2-клик  
Стоп - 1 клик

## Лекция 4

### Стандарты сжатия видеоинформации

ФИО преподавателя: Смирнов

Александр Витальевич

e-mail: [av\\_smirnov@mirea.ru](mailto:av_smirnov@mirea.ru)





# Стандарты сжатия видеоинформации

Motion Picture Expert Group - Группа экспертов по движущимся изображениям

Международные стандарты:

MPEG-1 – 1992 год – компьютерные видео-диски: ISO/IEC 11172.

MPEG-2 – 1994 год – телевизионное вещание: ISO/IEC 13818.

MPEG-4 - 1999 год.

MPEG-4 Part 10 AVC (H.264) - 2004 год.

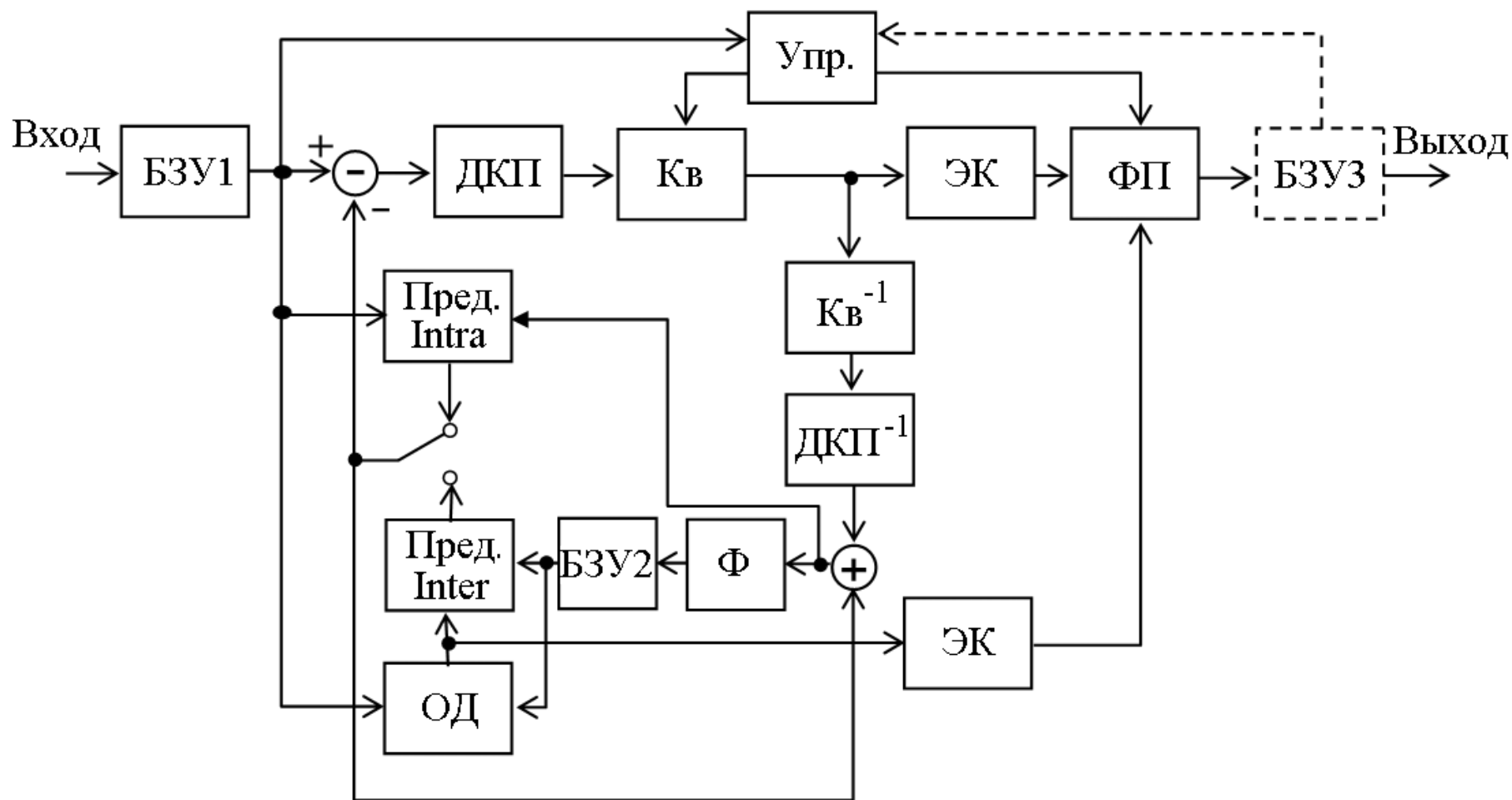
HEVC (H.265) - 2013 год.

VVC (H.266) – 2020 год.

Помимо этого, используются видеокодеки, не являющиеся международными стандартами. Среди них VP9, VC1, AV1 и др.



## Обобщенная структура видеокодера





## Типы изображений в MPEG-2

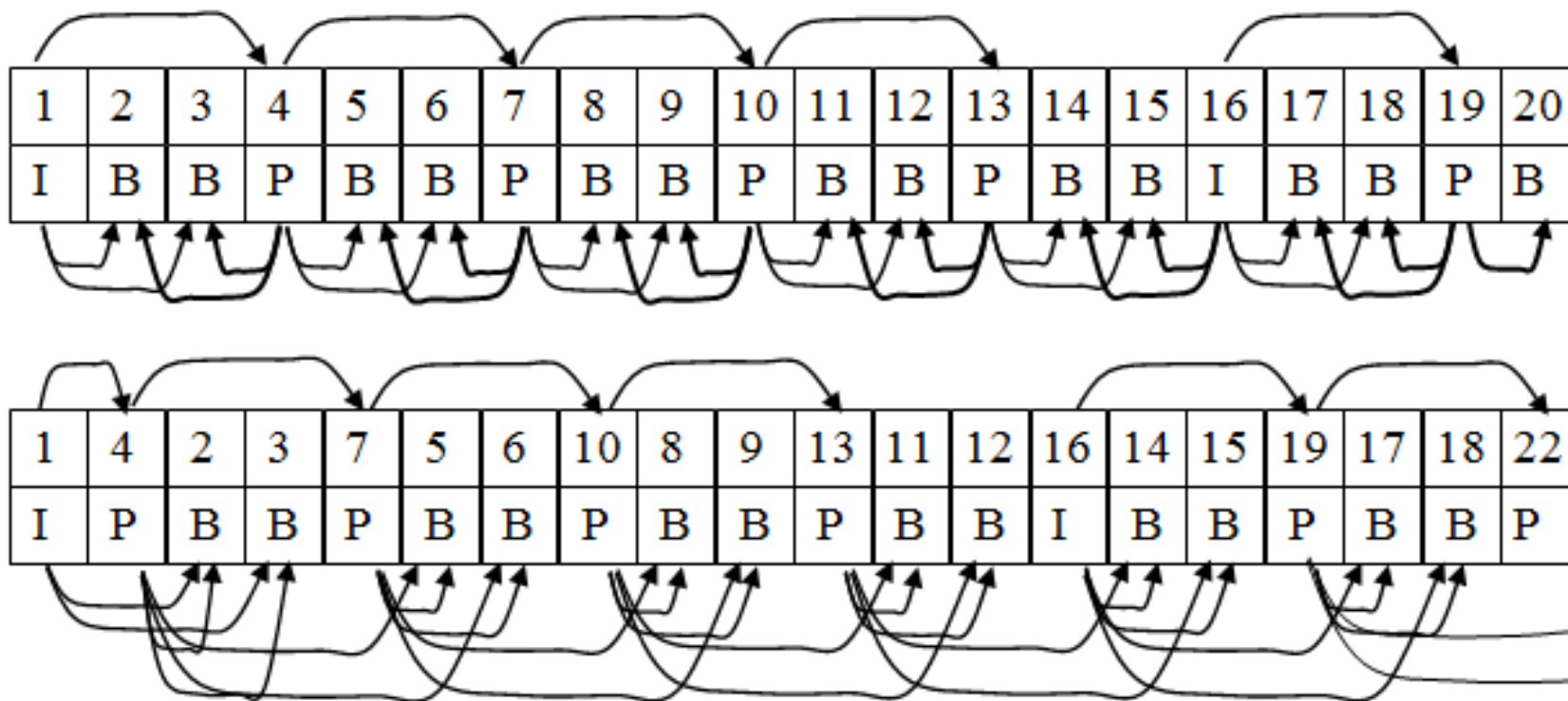
**I-изображения** – внутрикадровые (Intraframe), все макроблоки которых кодируются внутрикадровым методом.

**P-изображения** – предсказанные (Predictive), макроблоки которых, кроме активных и пропущенных, кодируются с применением предсказания по ближайшему предшествующему I- или P-изображению. Для таких макроблоков передаются вектор движения и кодированная ошибка предсказания. *Активными* признаются макроблоки, для которых не удается найти соответствующий фрагмент в опорном изображении. Они кодируются внутрикадровым методом. *Пропущенными* (Skipped) объявляются макроблоки, не изменившиеся по сравнению с опорным изображением. Их видеоданные не передаются, что позволяет уменьшить поток информации.

**B-изображения** – двунаправленные (Bidirectional), макроблоки которых, кроме активных и пропущенных, кодируются с применением предсказания по ближайшим к ним предшествующему и последующему I- и P-изображениям, а сами не могут использоваться для предсказания других изображений.



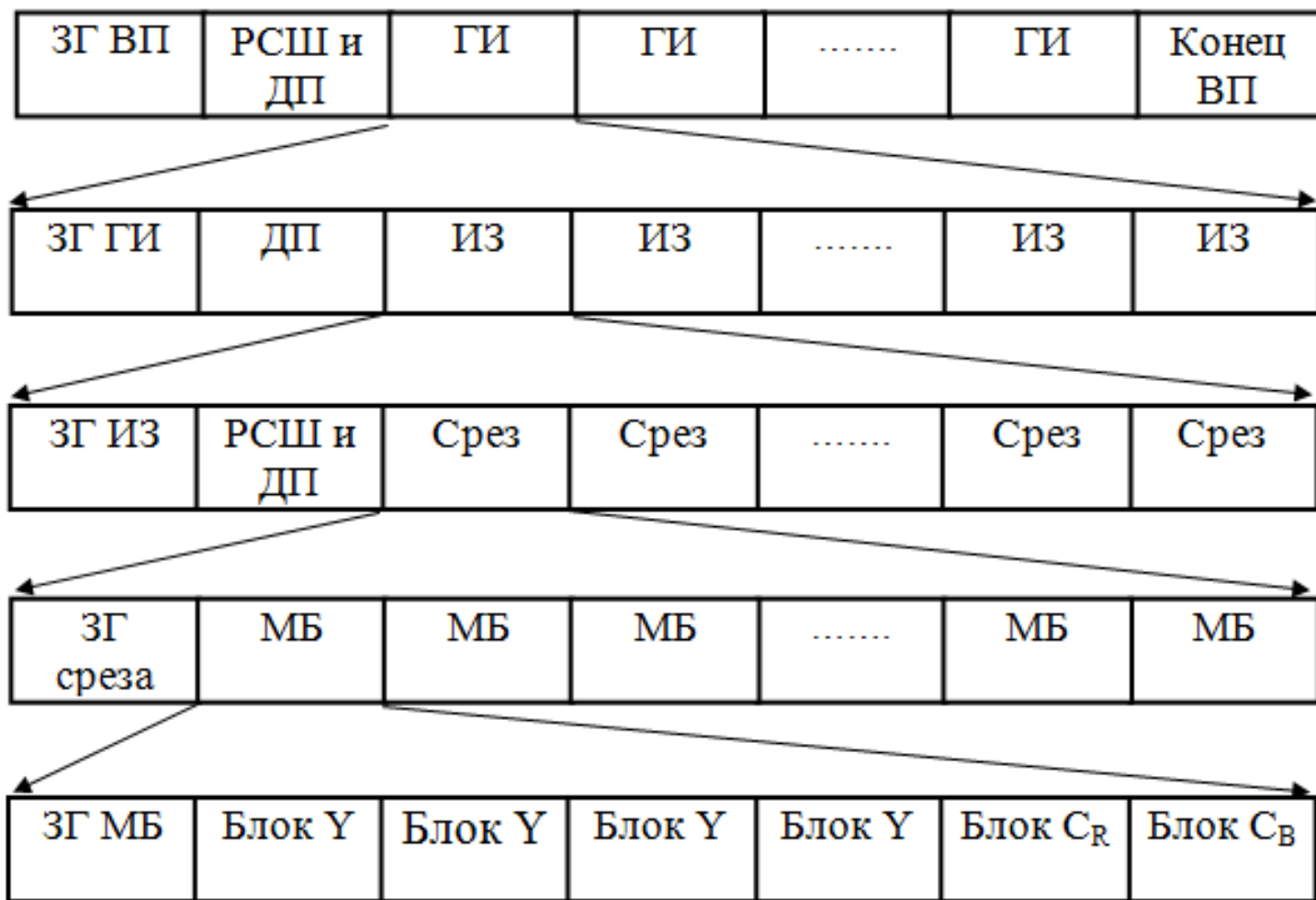
## Предсказание в MPEG-2



Видеопоследовательность разделяется на группы изображений GOP (Group of Pictures). Каждая группа начинается с изображения типа I. Затем следуют, чередуясь, изображения типов P и B.



## Структура потока видеоданных





## Уровни и профили MPEG-2

Профиль Уровень	Простой (Simple)	Основной (Main)	Масштаб. по С/Ш	Простран ственный	Высокий (High)
High 1920x 1152	–	80 Мбит/с	–	–	100 Мбит/с
High 1440x 1152	–	60 Мбит/с	–	60 Мбит/с	80 Мбит/с
Main 720x576	15 Мбит/с	15 Мбит/с	15 Мбит/с	–	20 Мбит/с
Low 352x280	–	4 Мбит/с	4 Мбит/с	–	–

Еще есть профиль 4:2:2 для студийной аппаратуры



## Сжатие в MPEG-2

Входной поток 4:2:0 – 162 Мбит/с или 202,5 Мбит/с

Максим. битрейт Main@Main – 15 Мбит/с

Отличное качество – 9 Мбит/с

Хорошее качество – 6 Мбит/с

Удовлетворительное качество – 3 Мбит/с





## Стандарт MPEG-4 AVC (H.264)

AVC - Advanced Video Coding, наиболее распространенный на сегодня стандарт сжатия видео. Обеспечивает в 1,5 - 2 раза большее сжатие чем MPEG-2 при том же качестве изображения.

Это достигается за счет:

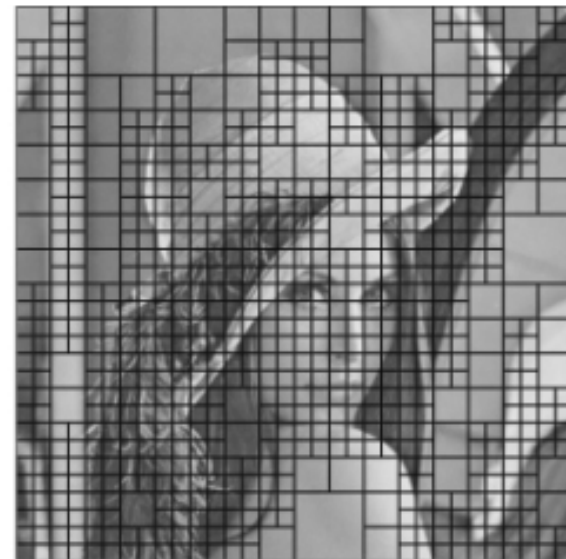
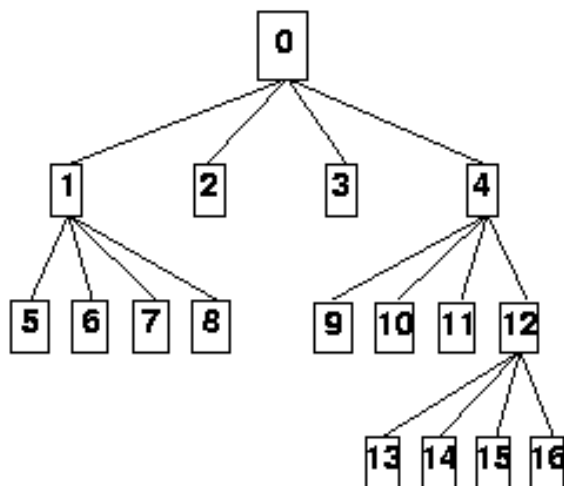
- разделения макроблоков на подмакроблоки;
- более эффективного и гибкого предсказания;
- более эффективного кодирования квантованных коэффициентов ДКП;
- фильтрации для уменьшения заметности блочной структуры.

Типы I, P, B присваиваются не целым изображениям, а слайсам. Слайс может включать неск. рядов макроблоков, даже целое изображение. Для предсказания макроблока в реж. Inter выбирается одно или два изображения из хранящихся в БЗУ1. Разные макроблоки слайса можно предсказывать по разным изображениям. В режиме Intra выполняется внутрикадровое предсказание.



## Структура макроблоков в H.264

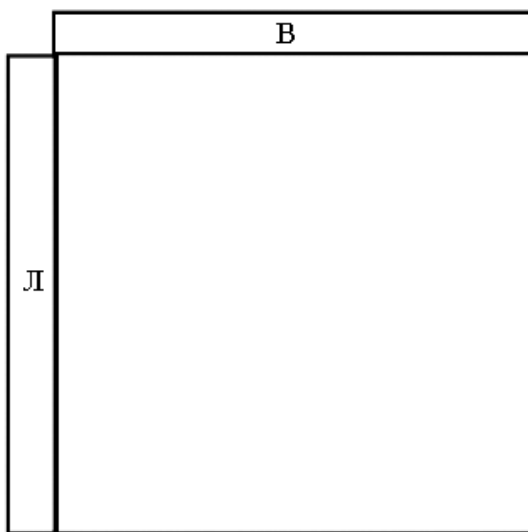
16	15	11	3
13	14		
9		10	
8		7	2
5		6	



Для макроблока 16x16 пикселей выбирается 1 (Р-слайс) или 2 (В-слайс) опорных изображения, положения которых в видеопоследовательности произвольны. Части макроблока размером 8x16, 16x8, 8x8, 8x4, 4x8 и 4x4 пикселя могут предсказываться разными фрагментами этих опорных изображений, то есть для каждой части может быть свой вектор движения. Разбиение изображения адаптивно к размерам деталей.



## Внутрикадровое предсказание в H.264



	A	B	C	D	E	F	G	H
a								
b								
c								
d								

Внутрикадровое предсказание может осуществляться для макроблоков 16x16 пикселей или подмакроблоков 4x4 пикселя. В первом случае для предсказания используются пограничные пиксели левого (Л) и верхнего (В) соседних макроблоков. Возможны 4 варианта.

Для предсказания отсчетов сигнала яркости в подмакроблоках 4x4 пикселя используются пиксели, отмеченные буквами на рисунке справа. Возможны 9 вариантов предсказания.



## Внутрикадровое кодирование

1. ДКП выполняется в блоках  $4 \times 4$  пикселей. Используются приближенные значения функции  $\cos x$ , так что необходимо выполнять только целочисленные вычисления. Это уменьшает нагрузку процессора.
2. Квантованные коэфф-ты ДКП на простых профилях кодируются контекстно-адаптивными кодами переменной длины (CAVLC), а на более сложных применяется контекстно-адаптивное двоичное арифметическое кодирование (CABAC).
3. После деквантования и обратного ДКП производится деблочная фильтрация для уменьшения заметности границ блоков. Параметры фильтра выбираются в зависимости от вида границы (между макроблоками или внутри макроблока) и скорости изменения сигнала у этой границы.



## Видеопоток H.264

Поток состоит из модулей (unit) NAL - Network Abstract Level. Модуль NAL состоит из индикатора типа содержимого модуля и нагрузки, называемой RBSP - Raw Byte Sequence Payload. Типы RBSP: наборы параметров; кодированные слайсы; вспомогательные и/или необязательные RBSP. Модули NAL передаются в порядке выполнения декодирования. Наборы параметров должны передаваться в начале.

При передаче модулей NAL в пакетизированном потоке каждый модуль упаковывается в отдельный пакет. При объединении модулей NAL в непрерывный поток перед ними вставляются префиксы, сходные с таковыми в видеопотоке MPEG-2.



## Профили и уровни H.264

Базовый **профиль** (Baseline - BP).

Основной (Main - MP) (используется в ТВ вещании с чс разв.).

Расширенный (Extended - XP).

Ограниченный базовый профиль (Constrained baseline - CBP)  
(имеет минимальный набор функций).

Высокий профиль (High - HP) для вещания ТВЧ.

Профили для студийной работы High 10, High 4:2:2, High 4:4:4,  
High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra.

Масштабируемые профили Scalable Baseline, Scalable High,  
Scalable High Intra.

Профиль для стереоскопического ТВ Stereo High.

Профиль для многоракурсного ТВ Multiview High.

16 **уровней** 1; 1b; 1.1; 1.2; ... 5; 5.1 с размером изображения от 128x96 до 4096x2304 пикселей. Битрейт зависит и от профиля.



## Стандарт HEVC (H.265) 2013 г.

**HEVC - High Efficient Video Coding.** Предназначен для видео с разрешением - до 8192x4320.

Усовершенствованы алгоритмы предсказания, энтропийное кодирование и антиблоковая фильтрация.

Уменьшение битрейта при том же PSNR:

- по сравнению с MPEG-2 на 70%;
- по сравнению с H.264 AVC на 35%.

Помимо деления на слайсы, аналогичному H.264, введено деление на тайлы (tile) для организации параллельных вычислений на неск. процессорах.



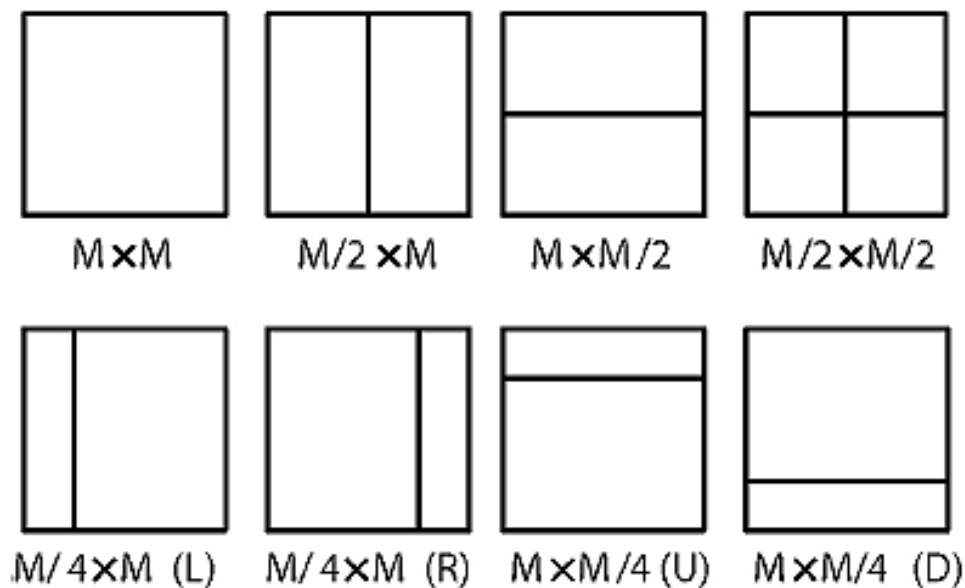
## Иерархия блоков H.265

1. Единицы кодирования - Coded Tree Unit – CTU,  $L \times L$  пикселей, где  $L$  может быть равно 16, 32 или 64. CTU включает Coded Tree Block – CTV яркости и цветности. Размер CTV яркости совпадает с размером CTU, а размеры CTV цветоразностных сигналов  $L/2 \times L/2$  пикселей.
2. CTV может разделяться на квадратные CB (Coded Block). Минимальный размер CB яркости  $8 \times 8$  пикселей. CB яркости и связанные с ним CB цветоразностных сигналов образуют модуль кодирования CU (Coded Unit). Для CU задается режим предсказания inter или intra.
3. CU разделяются на модули предсказания PU (Prediction Unit) и модули преобразования TU (Transform Unit). CB яркости и цветности делятся на блоки предсказания PB (Prediction Block) и блоки преобразования TB (Transform Block).

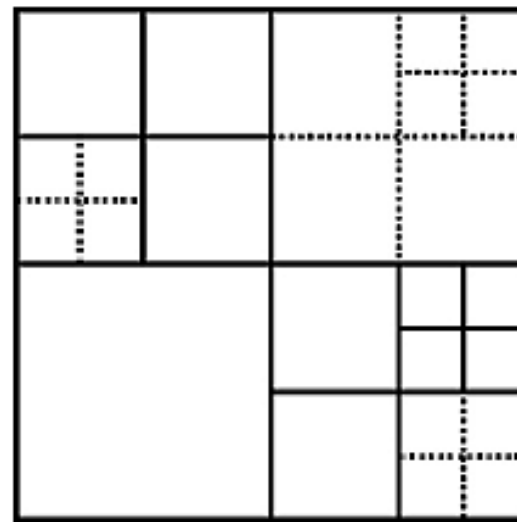




## Варианты структуры СВ



Деление СВ на РВ



Деление СВ на СВ  
(сплошные линии)  
и на ТВ (пунктиры)  
Размеры ТВ 16x16;  
8x8 и 4x4.



## Другие улучшения в Н.265

1. Введен объединенный (merge) режим предсказания, в котором информация о движении извлекается из соседних по пространственным координатам или по времени блоков.
2. Во внутрикадровом предсказании число вариантов увеличено до 35.
3. Деблочная фильтрация выполняется на границах блоков 8x8 п.
4. После деблочной в случае межкадр. предсказания выполняется SAO - Sample Adaptive Offset (адаптивное смещение отсчетов). Эта нелинейная обработка заключается в изменении значений декодированных отсчетов на величины, выбираемые из таблицы, формируемой в кодере по гистограммам значений сигнала.



## Профили H.265

**Основной (Main):** 4:2:0, 8 бит/отсчет, БЗУ декодера до 6 кадров.

**Main 10** - 10 бит/отсчет.

**Main Still Picture** для кодирования отдельных изображений.

**Main 4:2:2** с числом бит на отсчет 10 и 12.

Группы профилей **Main 4:4:4** и **High Throughput**.

В этих группах есть профили **Intra**.

Профили **Monochrome** с числом бит на отсчет 8, 12 и 16.

Масштабируемые версии профилей группы **Main**.

Профили **3D Main** и **Multiview Main**.

Профили **Screen-Extended**, каждый из которых объединяет свойства нескольких ранее введенных профилей.

Профили **High Throughput 4:4:4** с числом бит на отсчет до 14 для каждого сигнала.



## VVC (H.266) 2020 г.

VVC – Versatile Video Coding – Универсальное видеокодирование.

Цель – получить на 30-50 % большую степень сжатия чем в HEVC при том же визуально воспринимаемом качестве.

Среди других требований наличие режимов сжатия без потерь и сжатия субъективно без потерь, форматы изображений от 4K до 16K, работа с панорамным видео, форматы дискретизации 4:4:4, 4:2:2 и 4:2:0, квантование 8-16 битов на канал, наличие режимов high dynamic range (HDR) с квантованием более 16 битов на канал.

При этом сложность кодирования оценивается до 10 раз выше, чем у HEVC, а сложность декодирования – в 2 раза выше.



## Изменения в VVC по сравнению с HEVC

Принципы структурирования изображения и терминология в основном такие же, как в HEVC.

Число режимов внутрикадрового предсказания увеличено с 33 до 65, что соответствует 65 возможным направлениям предсказания.

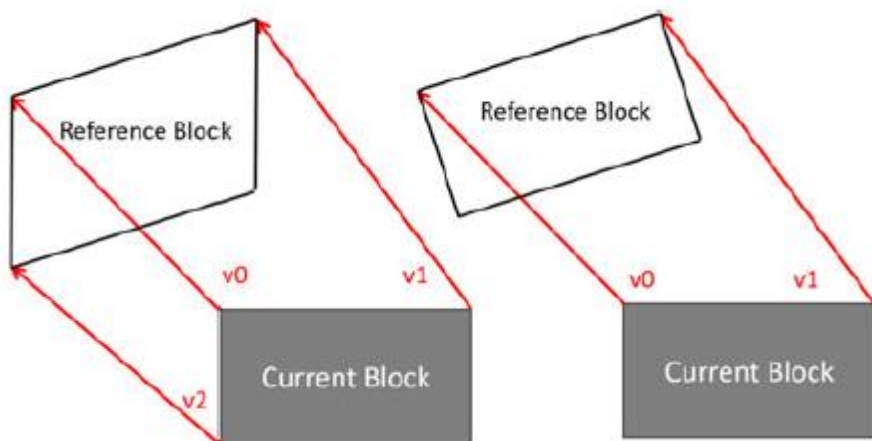
Новые виды внутрикадрового предсказания: комбинация по нескольким опорным отсчетам и межкомпонентное.

При межкадровом предсказании по двум кадрам возможна интерполяция с разными весовыми коэффициентами, а не только как среднее арифметическое.



## Аффинное преобразование блоков

Наиболее существенное усовершенствование – предсказание по прямоугольным блокам. Предсказываемый блок получается путем аффинного преобразования опорного блока.



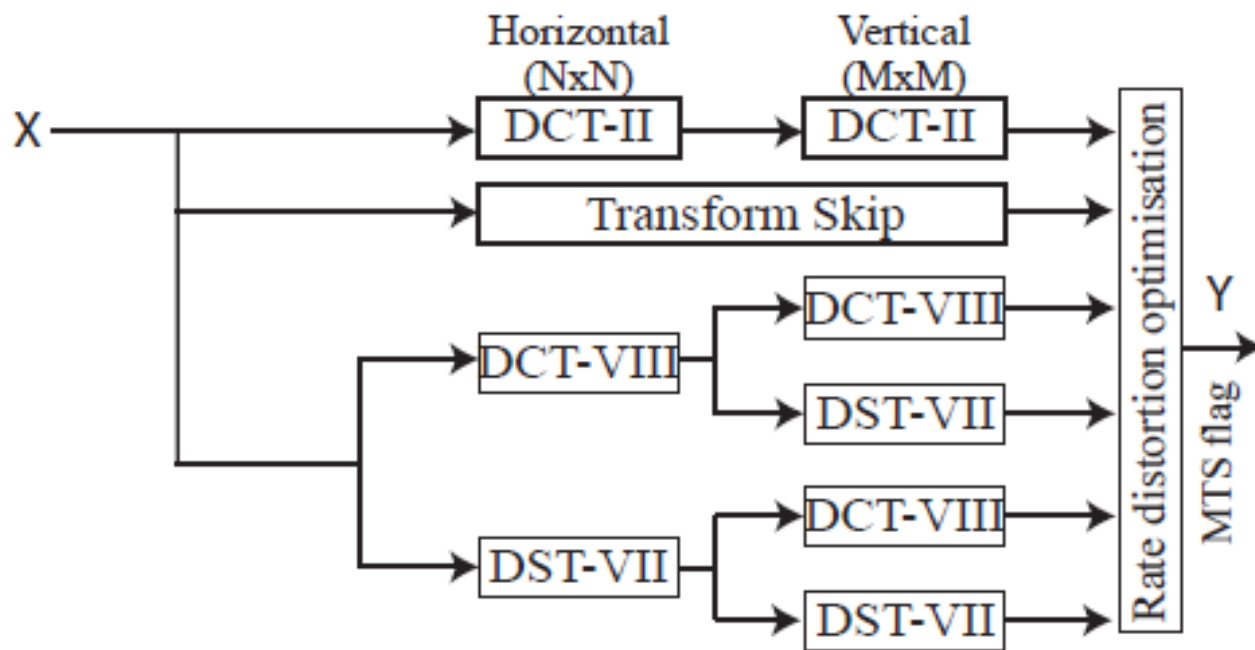
$$\begin{cases} v_x = \frac{v_{1x} - v_{0x}}{w} x - \frac{v_{1y} - v_{0y}}{w} y + v_{0x} \\ v_y = \frac{v_{1y} - v_{0y}}{w} x - \frac{v_{1x} - v_{0x}}{w} y + v_{0y} \end{cases}$$

Эти равенства для 4-параметрической модели (справа). Здесь  $w$  – ширина блока,  $x$  и  $y$  – координаты элемента изображения, отсчитываемые от верхнего левого угла. В случае 6-параметрической модели (слева) соотношения значительно усложняются.



## Поблочное преобразование

Максимальный размер блоков преобразования увеличен до 64x64. Кодер в новом стандарте дополнен блоком выбора преобразования Multiple Transform Selection (MTS).





## Настройка параметров VVC

Таким образом, для каждой операции в стандарте VVC/H.266 предусмотрено несколько вариантов выполнения и возможности настройки параметров. В стандарте нет правил выбора вариантов и параметров, а только описано, как результаты этого выбора должны фиксироваться в выходном потоке и использоваться в декодере. Построение алгоритмов, реализующих указанный выбор, предоставляется разработчикам средств кодирования. От эффективности этих алгоритмов во многом зависит соотношение достигаемого сжатия и возникающих искажений изображений.

Одним из перспективных направлений в этой области может быть применение машинного обучения, в частности, искусственных нейронных сетей, и других интеллектуальных технологий.





## Другие видеокодеки

1. VC-1 (Microsoft) - стандарт SMPTE. Применен в программах Windows Media Player, Windows Media Encoder и других. Основной формат при записи Blu-Ray Disks и HD DVD.

2. VP8, VP9, AV1 (Google) – открытые и бесплатные. Основная область применения - Интернет, в частности, сервис You Tube и браузеры. Последний кодек AV1 немного лучше HEVC, но уступает VVC.

Полное название AOMedia Video 1, где АО происходит от Alliance for Open Media – альянс, в состав которого входят компании Amazon, Cisco, Google, Intel, Microsoft, Mozilla и Netflix.

3. Dirac или VC-2 (BBC). Используется вейвлет-преобразование. Кодек может свободно распространяться и применяться.



Спасибо за внимание!