Национальный исследовательский университет

ВИДЕОТЕХНОЛОГИИ

КОДИРОВАНИЕ ВИДЕО (ТЕОРИЯ)

Денис Королев

Природа информации (в видео)

- Аналоговая
- Требуется оцифровка
- Аналого-цифровое преобразование – что это?
- Какие два основных процесса происходят при АЦП?
- **Дискретизация (Sampling)**
- **Квантование (Quantization) цветовые пространства и модели**

+ Дискретизация (Sampling)

- О том, как превратить свет и звук в цифры
- Могут встречаться формулы!
- Надо будет вспомнить физические основы из КГ
- Все это по материалам Параграфа 2 из **Книги***

- * Рекомендованная вам книга **Compression for Great Video and Audio** – B. Waggoner, Elsevier, 2010, p. 15

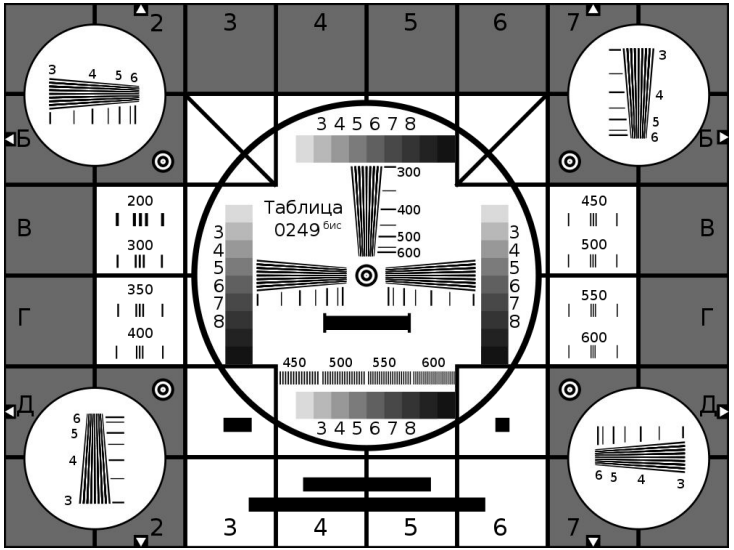
Аналоговая природа и цифровая форма сигнала

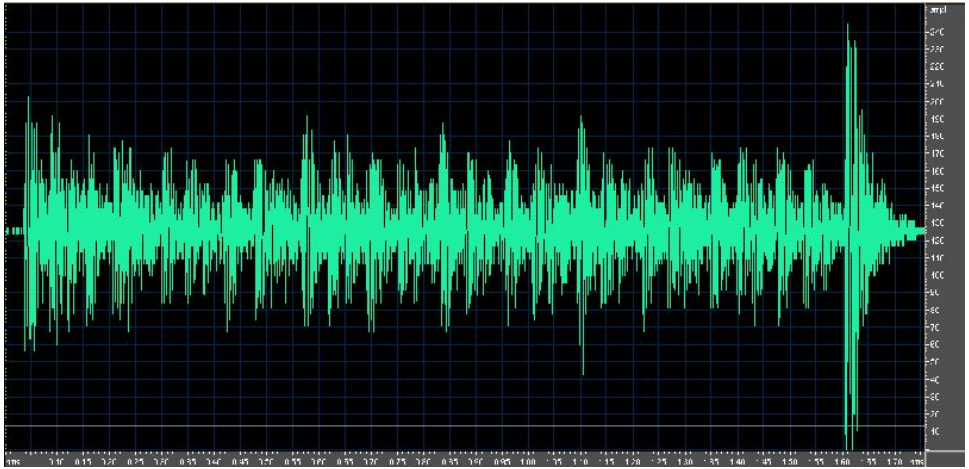
- Свет и звук – это непрерывный **аналоговый** сигнал.
- Сигнал для датчиков (глаза, уха, камеры, микрофона)
- Непрерывный – значит бесконечно подробный
- Мы не можем записать в дискретной форме **всё ->**
- **Как часто мы будем замерять сигнал?**
- **Это частота дискретизации**

Что будем дискретизировать?

Пространство Время

DPI/LPI, lines, pixels fps, Hz





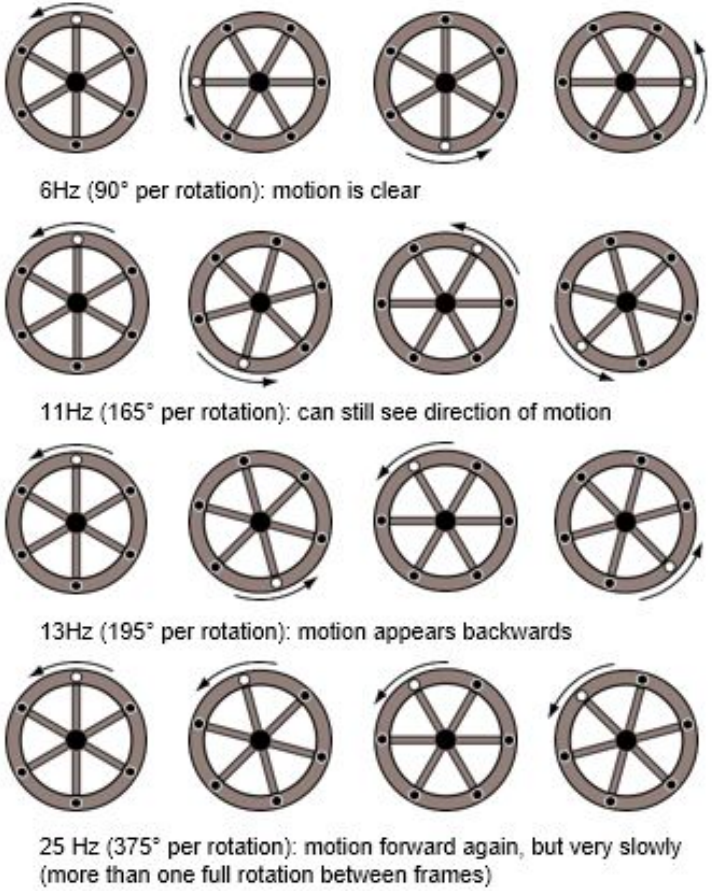
Теорема Найквеста-Шеннона, она же – Котельникова

- Теорема Найквеста-Шеннона
- В России называется теоремой Котельникова

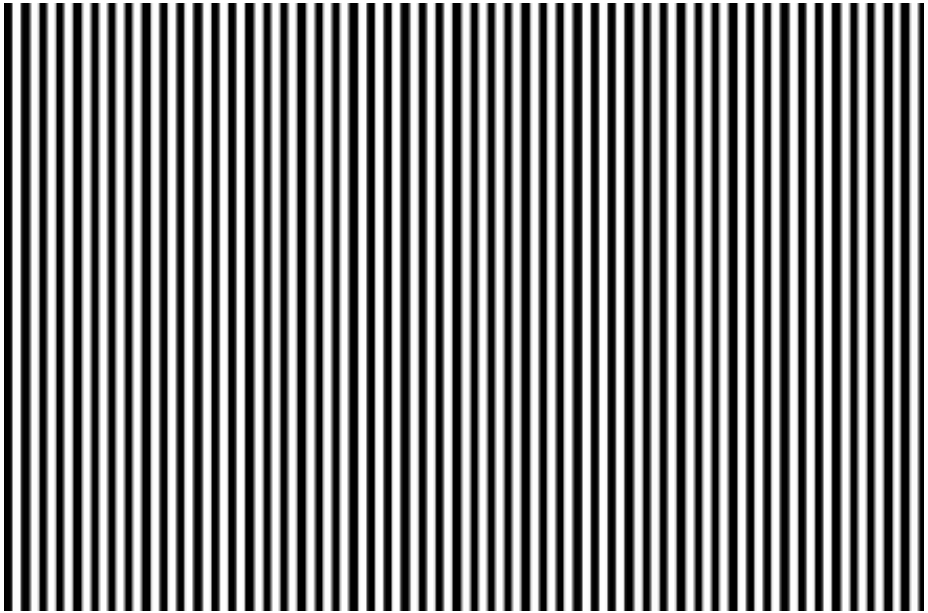
Если аналоговый сигнал имеет конечный (ограниченный по ширине) спектр,
то он может быть восстановлен однозначно и без потерь по своим отсчётам, взятым с частотой,
большей или равной удвоенной верхней частоте

Критерий Найквеста во времени

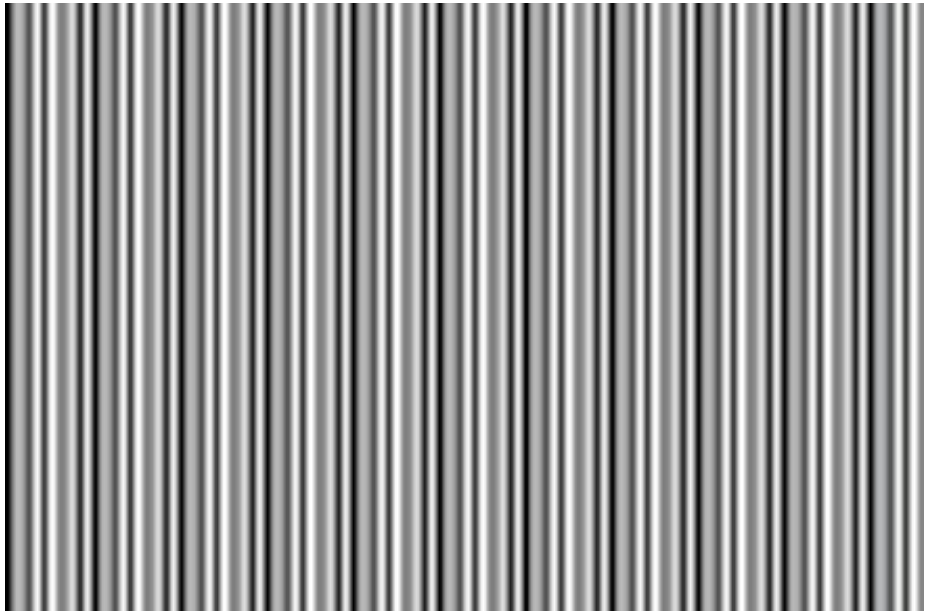
Почему колеса в кино крутятся «назад»?



Критерий Найквеста в пространстве



3 px



2 px

256/320*3=**2,4**

256/320*2=**1,6**

320x240 -> 256x192

Квантование

- Сколько значений может принимать каждый дискретный элемент (sample) 2^x bit

- Помним допустимые пределы:
  8bit RGB: 0 255 (256)
  8bit YCbCr: 16 235 (219) – «так принято»

Динамический диапазон

- Помните, что это?
- Для монитора отношение яркости около 4000:1
- В кинозале шагов может быть чуть больше 100.
- Человек видит разницу, а не абсолютное значение
- Для яркости могут выделять больше бит (10-16)
- Цветность обычно урезают (субдискретизация)

Работа с видеокодеками

- Профиль
- Уровень
- Поток (data rate)

Кодек	Профиль	Цветовая субдискретизация	Уровень
H.264	**High**	**4:2:0**	**@ 2.1**

Эффективность сжатия

A

800kbps **160x120** Apple Video codec

B

800kbps **640x480** H.264
High profile 3-pass Quality optimized

Терминология объёма памяти (вдруг забыли?)

Computer industry numbers

- “K” = Kib = 2^{10} = 1024
- “M” = Mib = 2^{20} = 1,048,576
- “G” = Gib = 2^{30} = 1,073,741,824

Correct numbers

- K = 10^3 = 1000
- M = 10^6 = 1,000,000
- G = 10^9 = 1,000,000,000

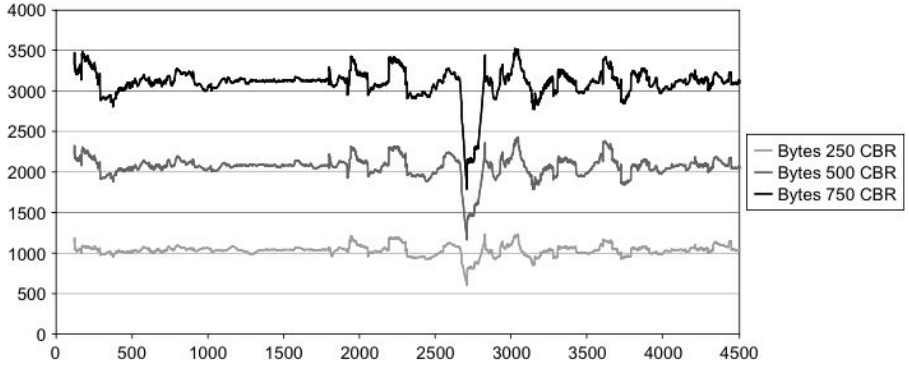
Difference between values

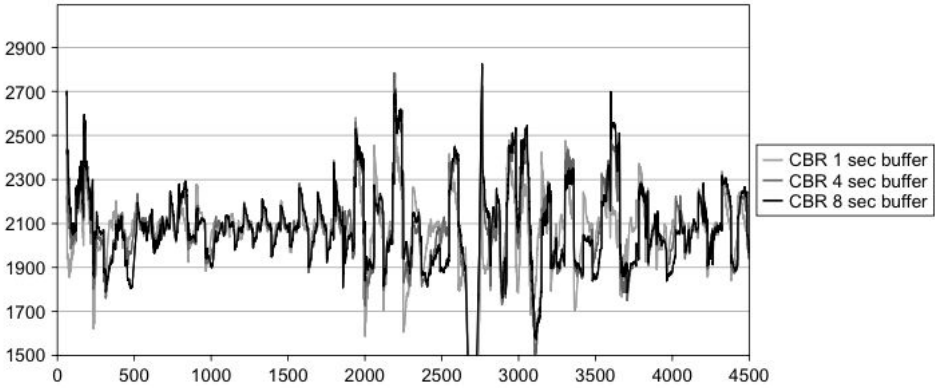
- K v. Ki = 2.4 percent
- M v. Mi = 4.8 percent
- G v. Gi = 7.37 percent

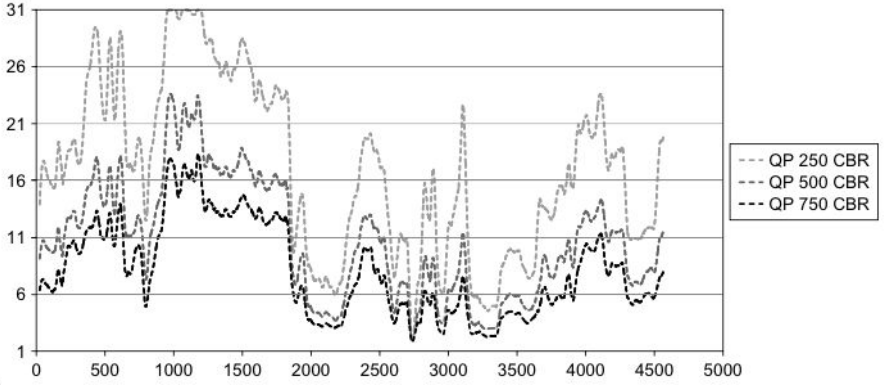
Контроль потока (VBR and CBR)

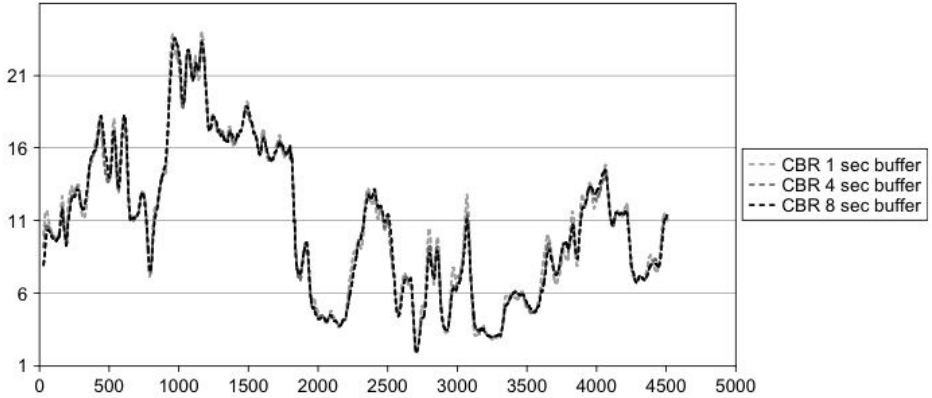
- VBV (Video buffering verifier): under/over-run
- CBR for streaming
- VBR for download
- Quality-limited VBR

Поток и квантование

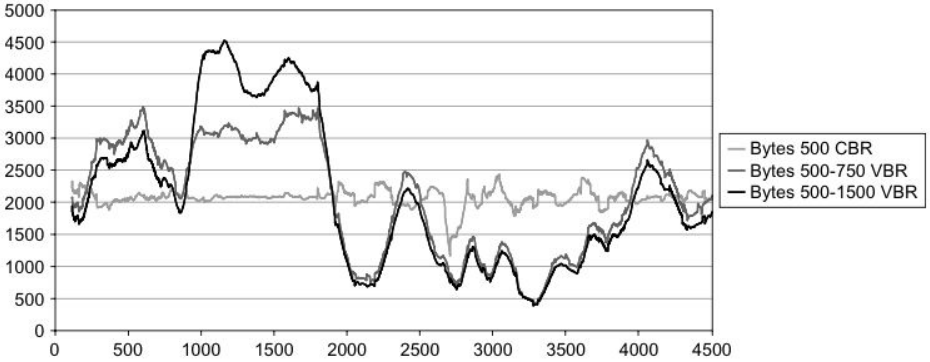


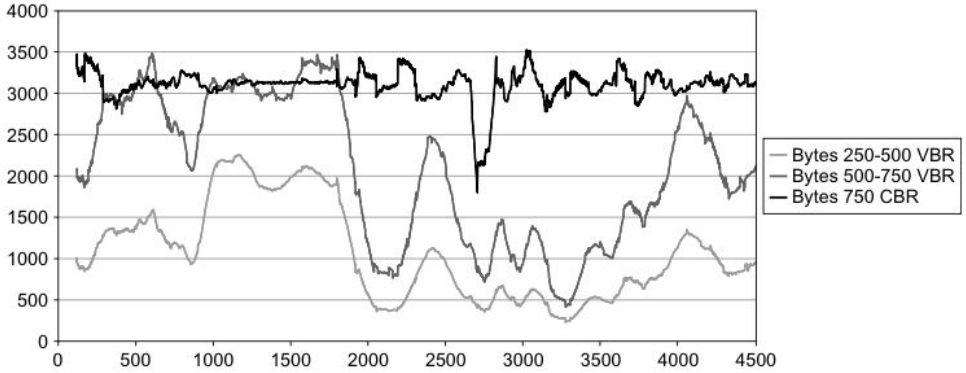


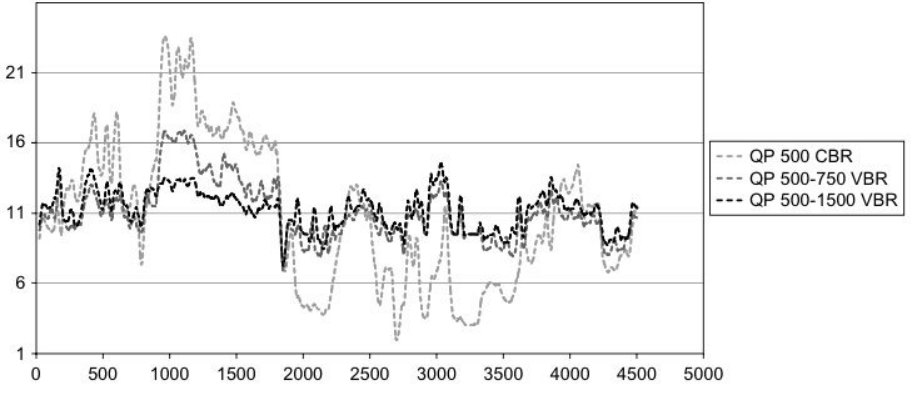


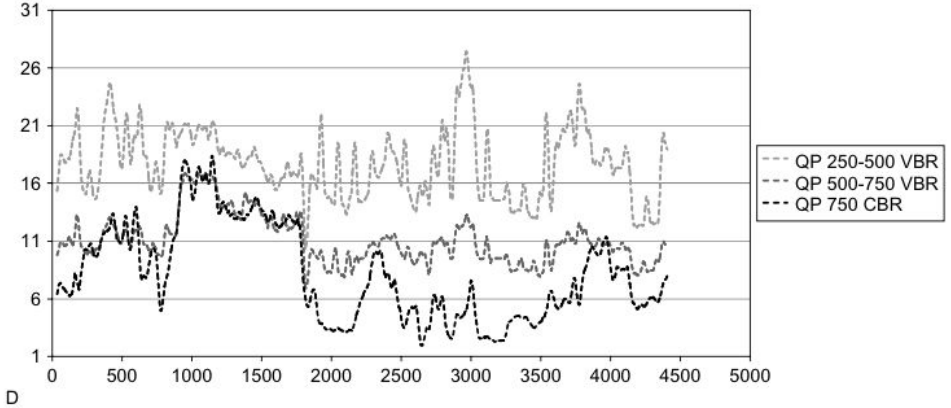


Поток и квантование









Кодирование в несколько проходов

- 1-pass
- 1-Pass lookahead
- 2-pass
- Segment re-encoding

Размер кадра

Rule of “^0,75”

640 x 360 @ 1000 Kbps

1280 x 720 @ ?

(1280 x 720)/(640 x 360) = 4

$4^{0.75}$ = 2.828

1000 Kbps x 2.828 = 2828 Kbps

Неквадратный пиксель

- Когда это требуется выходным форматом (DVD).
- Когда это определено исходным форматом.
- При активном движении по одной оси.

- Компьютерное видео имеет квадратный пиксель.

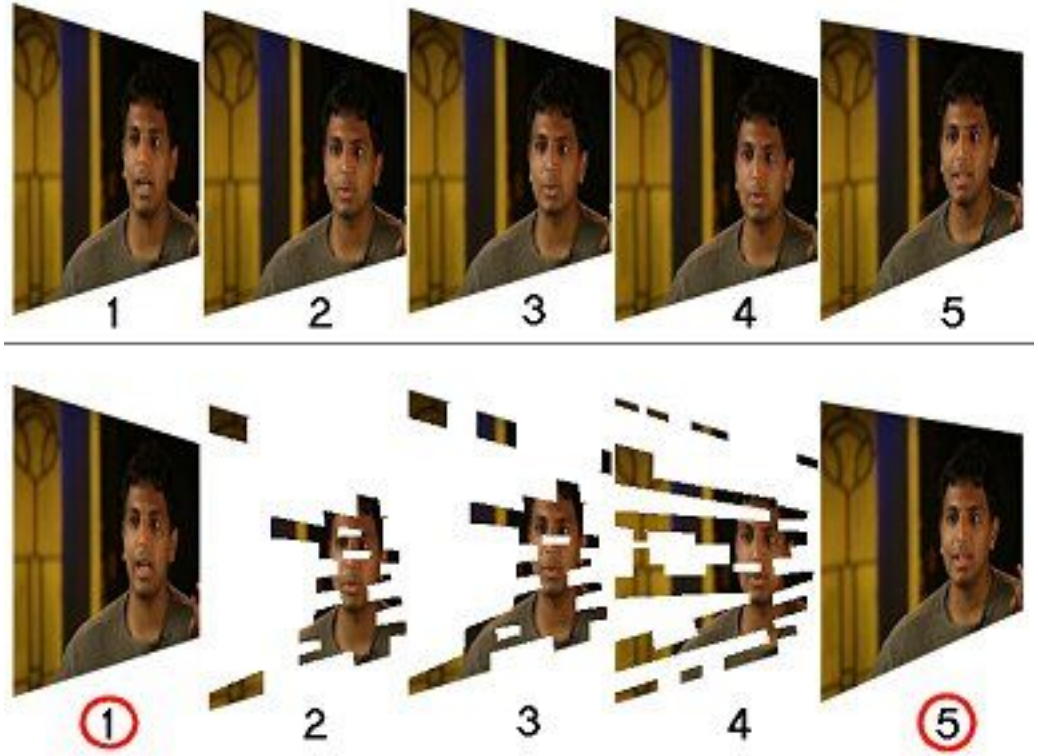
Глубина цвета

- Обычно: 8 бит/канал
- Профессиональные кодеки: 10, 12, 16 бит/канал
- Для съемки и промежуточного сохранения
- Просмотр – только 8 бит/канал.

Частота кадров

Frame Rate	Full	Half	Third	Quarter	Fifth
23.976	23.976	11.988	7.992	5.994	4.795
24.000	24.000	12.000	8.000	6.000	4.800
25.000	25.000	12.500	8.333	6.250	5.000
29.970	29.970	14.985	9.990	7.493	5.994
30.000	30.000	15.000	10.000	7.500	6.000
50.000	50.000	25.000	16.667	12.500	10.000
59.940	59.940	29.970	19.980	14.985	11.988
60.000	60.000	30.000	20.000	15.000	12.000
120.000	120.000	60.000	40.000	30.000	24.000

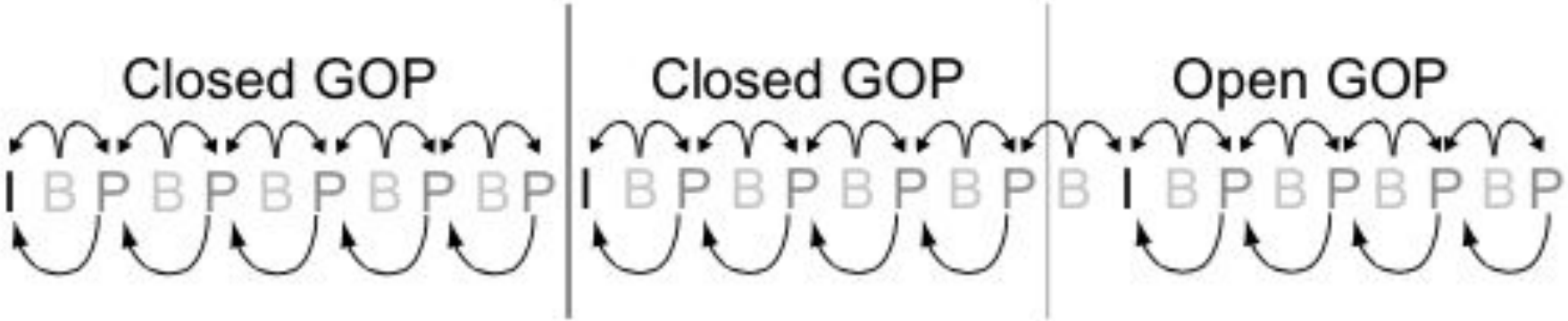
Частота ключевых кадров



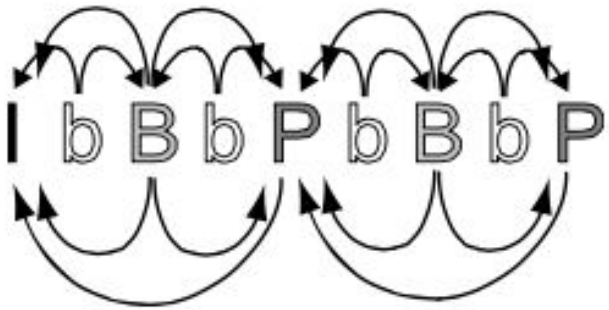
B-кадры

- Минимальный размер,
- Могут быть хуже, на них никто не ссылается.
- Сэкономленное место уходит в I, P кадры, на которых строятся те же B кадры.
- Их можно пропускать, от них другие не зависят
- Упрощают навигацию – меньше P кадров декодировать

Open/Closed GOP



Pyramid B-frames



Они все-таки ссылаются на обычные b-кадры.
Но и на I и P тоже.

Некоторые особенности H.264

- CABAC Entropy Coding
- Context Adaptive Binary Arithmetic Coding

- +40% нагрузки на декодер*
- +10-20% эффективности сжатия (при сильном сжатии)
- Не используется в Baseline профиле.

* По сравнению с ранее применявшимся CAVLC.

Профили и уровни H.264

	Baseline	Extended	Main	High	High 10	High 4:2:2	High 4:4:4 Predictive
I and P Slices	Yes	Yes	Yes	Yes	Yes	Yes	Yes
B Slices	No	Yes	Yes	Yes	Yes	Yes	Yes
Multiple Reference Frames	Yes	Yes	Yes	Yes	Yes	Yes	Yes
In-Loop Deblocking Filter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CAVLC Entropy Coding	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CABAC Entropy Coding	No	No	Yes	Yes	Yes	Yes	Yes
Interlaced Coding (PicAFF, MBAFF)	No	Yes	Yes	Yes	Yes	Yes	Yes
8x8 vs. 4x4 Transform Adaptivity	No	No	No	Yes	Yes	Yes	Yes
Quantization Scaling Matrices	No	No	No	Yes	Yes	Yes	Yes
Separate Cb and Cr QP control	No	No	No	Yes	Yes	Yes	Yes
Separate Color Plane Coding	No	No	No	No	No	No	Yes
Predictive Lossless Coding	No	No	No	No	No	No	Yes

Профили и уровни H.264

Итого: Критерии выбора кодека

- Эффективность сжатия
- Производительность воспроизведения
- Доступность на стороне зрителя