

Компьютерная графика

Алгоритмы сжатия
видео

MPEG

Сжатие последовательности изображений (видео) и звука:

- Несимметричный – сжатие гораздо сложнее, чем разжатие.
- Сжимает примерно в 100 раз (реально 30).
- Использует похожесть соседних кадров друг на друга.

Принципы кодирования видео

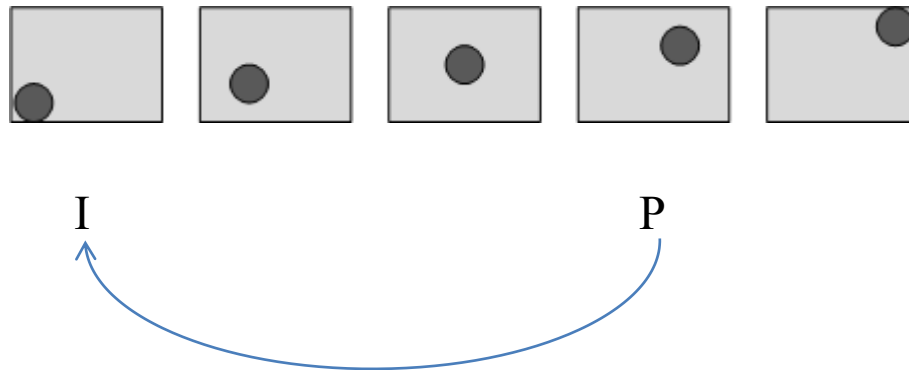
- I-frame (intra) – аналогичен JPEG



I

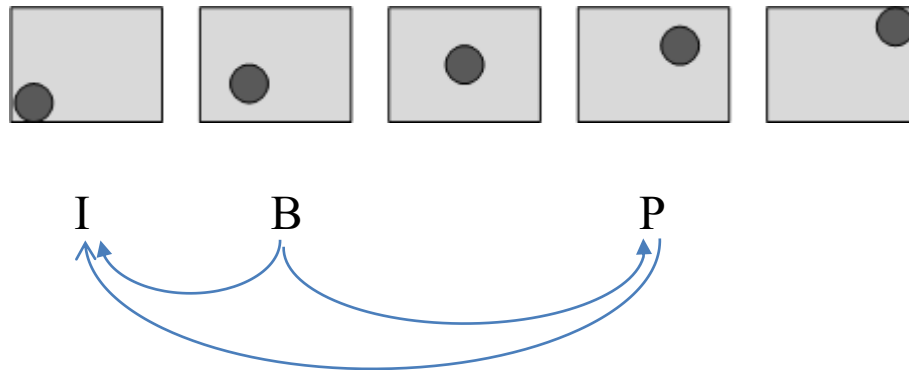
Принципы кодирования видео

- I-frame (intra) – аналогичен JPEG
- P-frame (predicted) – один опорный кадр (из прошлого)

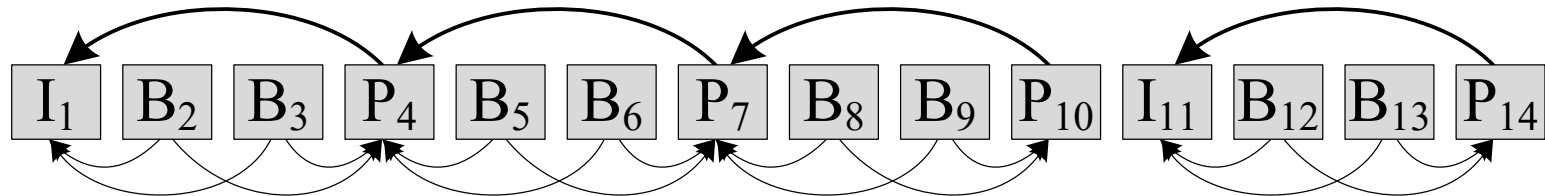


Принципы кодирования видео

- I-frame (intra) – аналогичен JPEG
- P-frame (predicted) – один опорный кадр (из прошлого)
- B-frame (bidirectional) – два опорных кадра (из прошлого и из будущего)

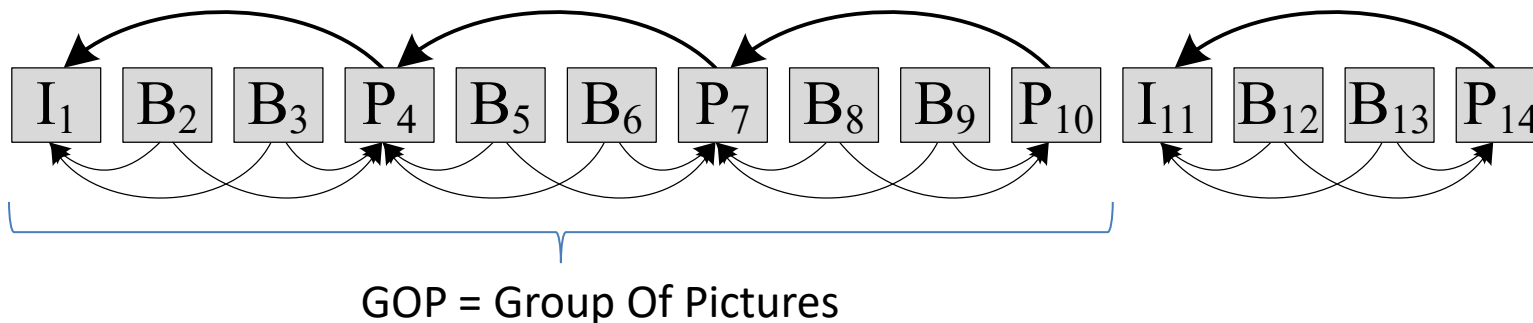


Принципы кодирования видео



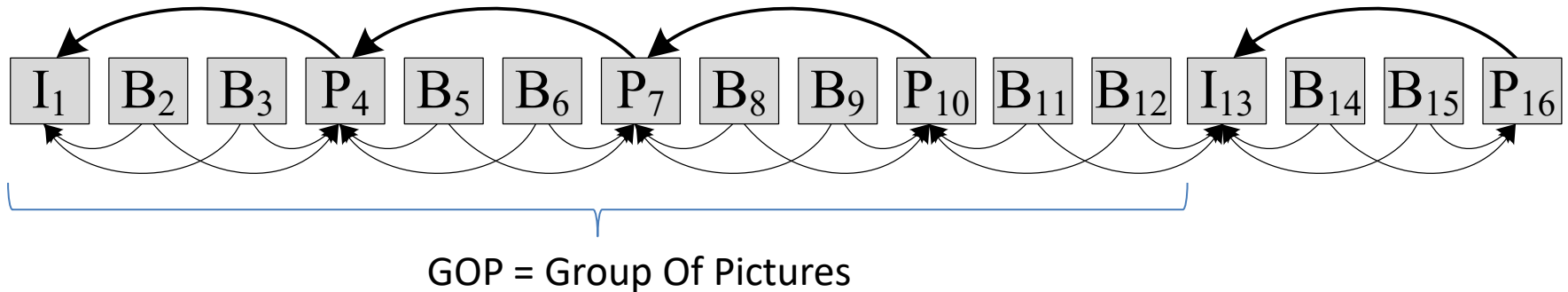
- Порядок на входе:
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, ...
- Порядок на выходе:
1, 4, 2, 3, 7, 5, 6, 10, 8, 9, 11, 14, 12, 13, ...

Принципы кодирования видео



- Порядок на входе:
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, ...
- Порядок на выходе:
1, 4, 2, 3, 7, 5, 6, 10, 8, 9, 11, 14, 12, 13, ...

Принципы кодирования видео



«Открытый» GOP – B-кадры в конце группы ссылаются на первый кадр следующей группы.

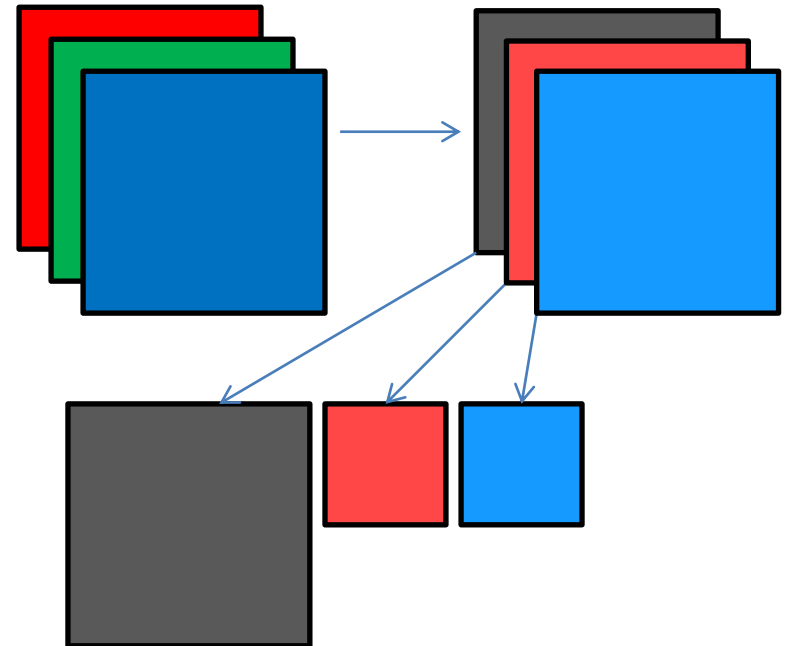
Кодирование кадра

Шаг 1. Матрицирование (попиксельно)

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

$$Cr = R - Y$$

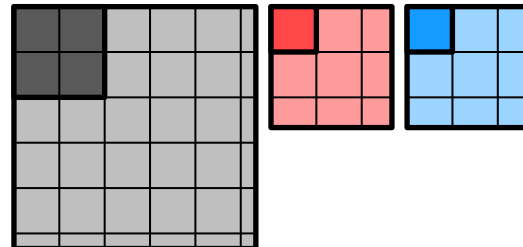
$$Cb = B - Y$$



Шаг 2. Децимация (прореживание) компонент Cr и Cb (в 4 раза)

Далее данные разбиваются на макроблоки 16x16:

- четыре блока 8x8 Y
- одним блок 8x8 Cr
- одним блок 8x8 Cb

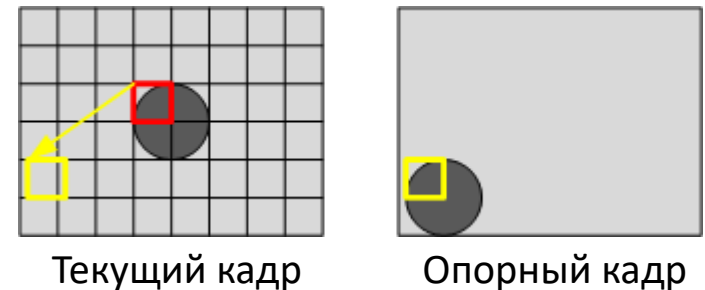


Кодирование с опорой



Шаг 3. Поиск вектора смещения

Для каждого макроблока сжимаемого кадра в опорном кадре ищется максимально похожий на него квадрат 16x16 (с точностью до долей пиксела)



Шаг 4. Кодировается попиксельное отличие

Поскольку абсолютная разность меньше, то данные сжимаются сильнее.

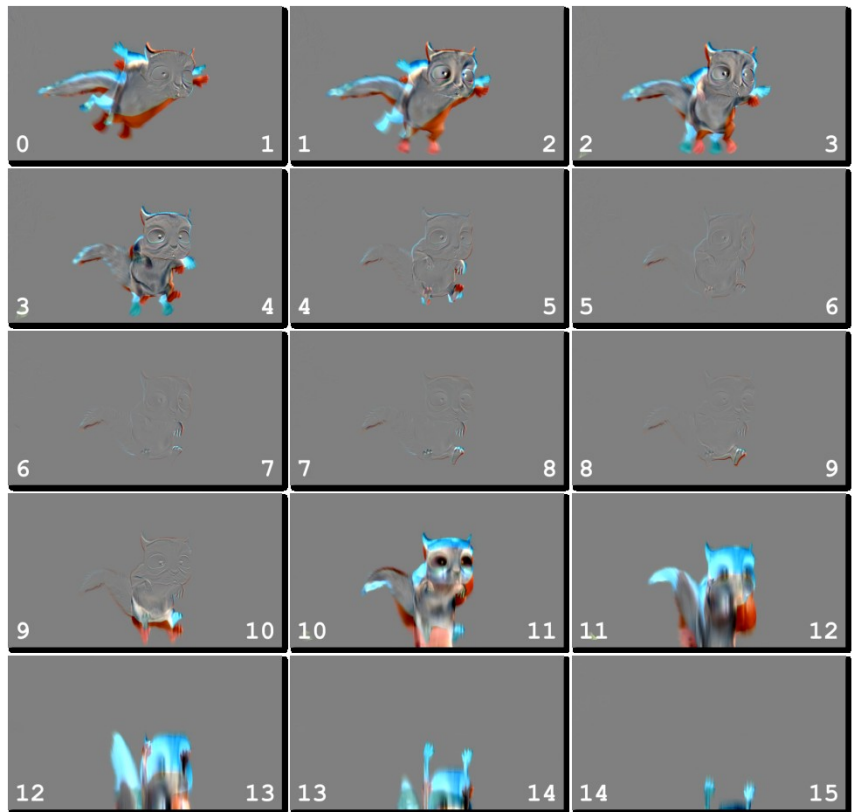
Если отличия равны нулю, то весь макроблок сжимается до одного слова.

Компенсация фона

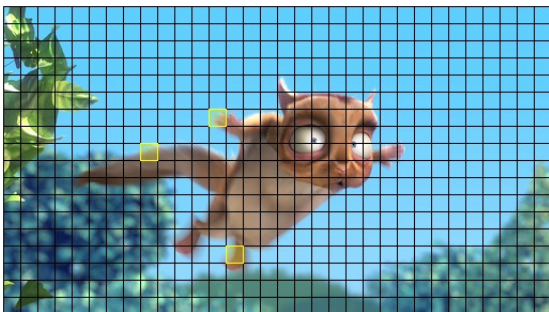
Тестовая последовательность (Big Buck Bunny)



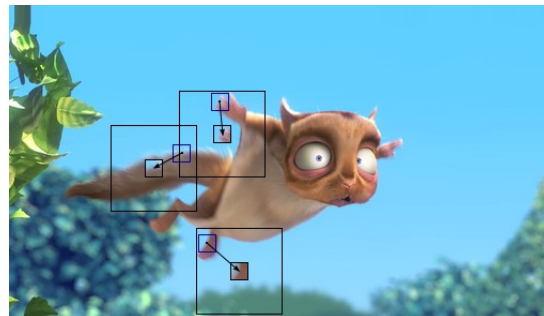
Межкадровая разница без применения алгоритмов компенсации движения



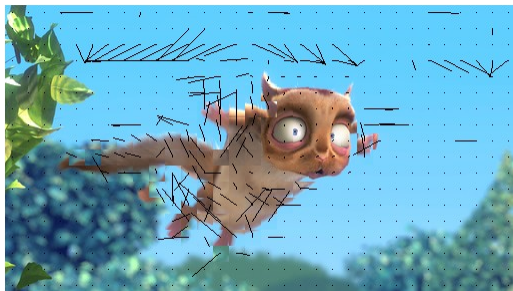
Компенсация фона



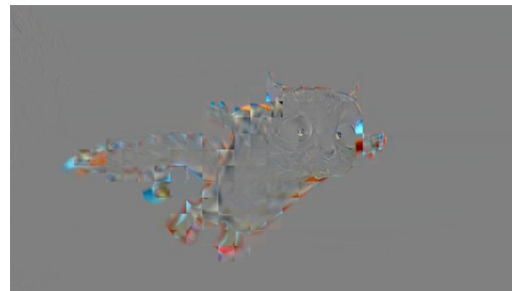
Текущий кадр (3 блока)



Вектора смещения 3 блоков



Все вектора смещения



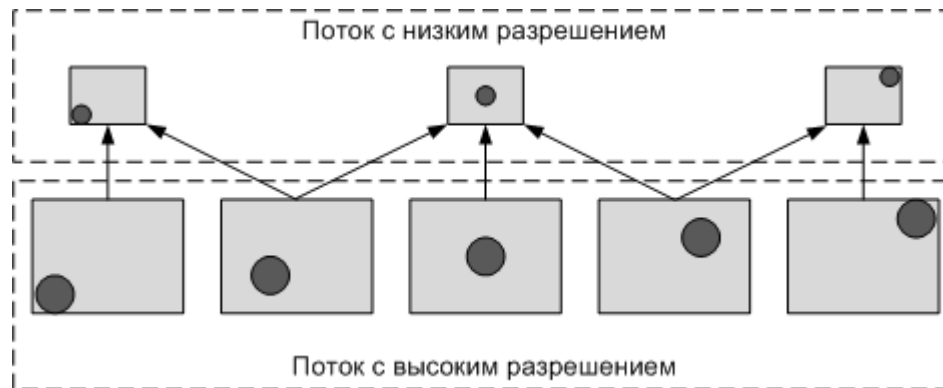
Кадр с компенсацией

Версии MPEG

- MPEG 1 (1993) – сжатие 360x240, 15 к/с на компакт диск (~1 Mb/s) со звуком.
- MPEG2 (1995) – поддержка цифрового ТВ (стандартное разрешение и ТВЧ).
MPTS как протокол передачи данных !!!
- MP3 = MPEG1 Layer 3 – сжатие звука
- MPEG4 (1998) – сжатие в узкие каналы (сеть, телефон)
AVC = MPEG4 Layer 10 = H.264
- Современный кодер HEVC = H.265

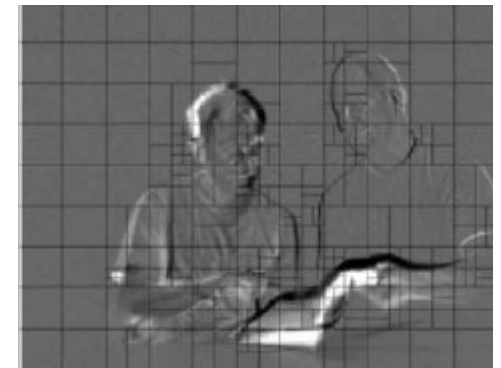
MPEG-2

- Разрешения:
 - SD 4:3 (720x576 25к/с, 720x480 30к/с)
 - SD 16:9 (1024x576 25к/с)
 - HD Ready (1280x720 50/60 к/с)
 - True HD (1920x1080 25/30 к/с)
- Несколько потоков с разным разрешением



Особенности AVC

- Внутрикадровое предсказание (опора)
- Много опорных кадров (32 ссылки)
- Вектор смещения до $\frac{1}{4}$ пикселя
- Арифметическое кодирование
- Макроблоки разных размеров 16/8/4x16/8/4
- Deblocking-фильтры



Исходный кадр



Восстановленный кадр
без фильтрации



Восстановленный кадр
с фильтрацией

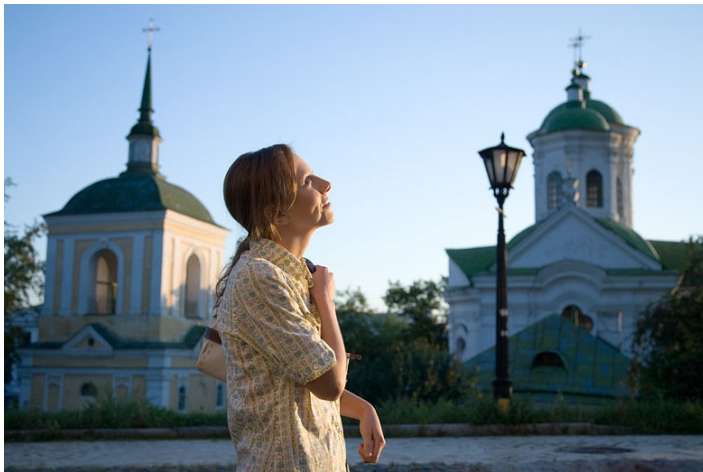
Особенности HEVC

- Для высоких разрешений (4K, HD)
- Образцовый декодер с точностью до бита
- Блоки до 64x64
- Вектор смещения до $\frac{1}{8}$ пикселя
- Адаптивный выбор матрицы квантования
- Кодирование параметров с предсказанием (например, величина вектора смещения)



Кодирование контента

- «Голливудский фильм»
- «Говорящая голова»
- «Футбол»



Кодирование контента

- «Голливудский фильм»
- «Говорящая голова»
- «Футбол»



Кодирование контента

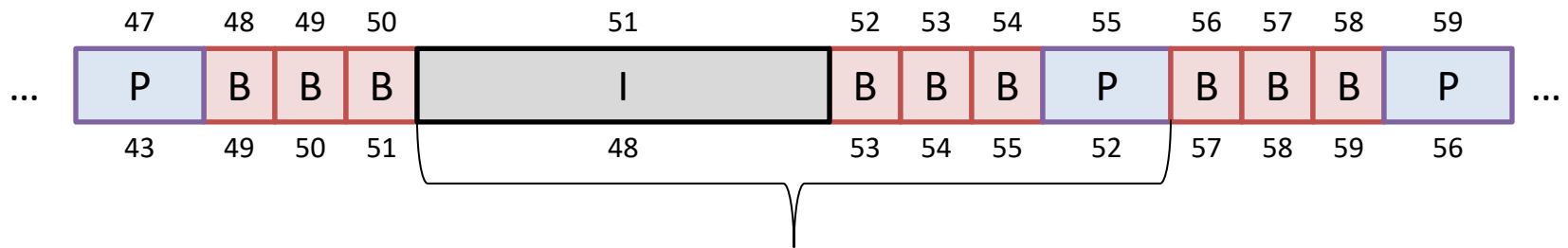
- «Голливудский фильм»
- «Говорящая голова»
- «Футбол»



Задержка в цифровом ТВ

Один кадр кодируется в среднем
за время одного кадра

Кодер должен сначала сжать Р-кадр,
и только затем сжимать В-кадры



Входной буфер (на 5 кадров)

Разница в размере сжатых данных

- Для GOP из 25 кадров:

In:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Out:	I	B	B	P	B	B	P	B	B	P	B	B	P	B	B	P	B	B	P	B	B	P	B	B	P

- Объем входных данных $V_{inTotal} = 25 * V_{in}$ (1)
- Объем выходных данных
 $V_{outTotal} = 1 * V_{outI} + 8 * V_{outP} + 16 * V_{outB}$ (2)
- Сжатие видео в 100 раз $= V_{inTotal} / V_{outTotal}$ (3)
- Сжатие JPEG(I-frame) в 10 раз $= V_{in} / V_{outI}$ (4)
- Пусть $V_{outP} = K * V_{outB}$ (5)

Разница в размере сжатых данных

- $100 * V_{outTotal} = V_{inTotal} = 25 * V_{in}$ (из 1 и 3)
- $V_{outTotal} = 0.25 * V_{in} = 1 * V_{outI} + 8 * V_{outP} + 16 * V_{outB}$ (+2)
- $V_{outI} = V_{in} / 10 = 0.1 * V_{in}$ (из 4)
- $0.25 * V_{in} = 0.1 * V_{in} + 8 * V_{outP} + 16 * V_{outB}$
- $0.15 * V_{in} = 8 * V_{outP} + 16 * V_{outB}$
- $V_{outP} = K * V_{outB}$ (из 5)
- $0.15 * V_{in} = 8 * K * V_{outB} + 16 * V_{outB} = (8 * K + 16) * V_{outB}$
- Сжатие В-кадра = $V_{in} / V_{outB} = (8 * K + 16) / 0.15$
для $K=3$ сжатие = 267 раз

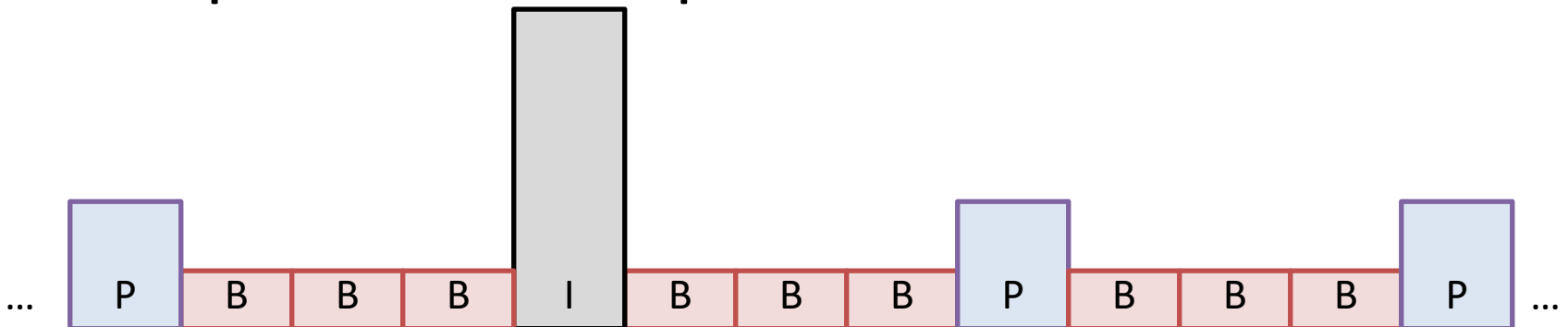


Задержка в цифровом ТВ

- CBR (Constant Bit Rate) – передача данных с фиксированной скоростью



- VBR (Variable Bit Rate) – передача данных с переменной скоростью



Задержка в цифровом ТВ

Для $GOP=25$ и $K=3$:

- Объем I-кадра = 40%
- Объем P-кадров = 36%
- Объем B-кадров = 24%

то есть 40% времени GOP-а передается первый кадр в режиме CBR (например, через спутник)

- Плюс задержки в декодере на буферизацию, сглаживание разного времени доставки пакетов и другие проблемы (синхронизация времени, проверка наличия следующих данных и т.д.)

Задержка в интернет ТВ

Форматы DASH или HLS работают с файлами:

- Плеер постоянно скачивает текущий плейлист
- Плейлист содержит несколько файлов (до 10)
- Длительность одного файла = несколько GOP (обычно 5-10 секунд)
- Для «живого» потока время начала определено с точностью до одного файла