



Современные системы цифрового телевидения

Старт 2-клик
Стоп - 1 клик

Лекция 3

Основы сжатия видеоинформации

ФИО преподавателя: Смирнов

Александр Витальевич

e-mail: 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.



Задача видеокомпрессии

Фильмы и другие виды видеоинформации характеризуются объемами в десятки и сотни Гбайт. Их скорости потоков измеряются в Гбит/с.

Поэтому одной из важнейших задач в области цифрового телевидения является задача значительного сокращения скорости цифрового потока и, соответственно, требуемой полосы частот канала связи, а также объема памяти, необходимой для записи телевизионных программ или отдельных изображений. Данная задача актуальна и для звуковой информации.

Задача сжатия может быть решена путем уменьшения избыточности информации, передаваемой в телевизионном сигнале. Методы сжатия можно разделить на два класса: без потерь информации и с частичной потерей информации.



Основные термины видеокомпрессии

Межкадровое кодирование – предсказание изображения по изображениям ранее кодированных кадров и передача ошибок предсказания. Использует повторяемость изображений кадров.

Внутрикадровое кодирование – удаление информации, не воспринимаемой или плохо воспринимаемой зрением. К такой информации относятся мелкие детали и текстура, а также детали быстро движущихся объектов.

К этому же режиму относится **внутрикадровое предсказание**.

Изображение – кадр или поле

Макроблок – Единица межкадрового кодирования. В MPEG-2 и H.264 размер 16x16 пикселей. В H.265 - до 64x64 пикселя.

Блок – квадрат 8x8 или 4x4 пикселей. Единица внутрикадрового кодирования. Макроблок содержит блоки Y , C_R и C_B .

Слайс (срез) – группа макроблоков, сжимаемых в одном режиме или с одинаковыми параметрами.



Дискретное косинусное преобразование

1. Поблочное дискретное косинусное преобразование (ДКП) – Разложение блока на пространственно-частотные гармоники.

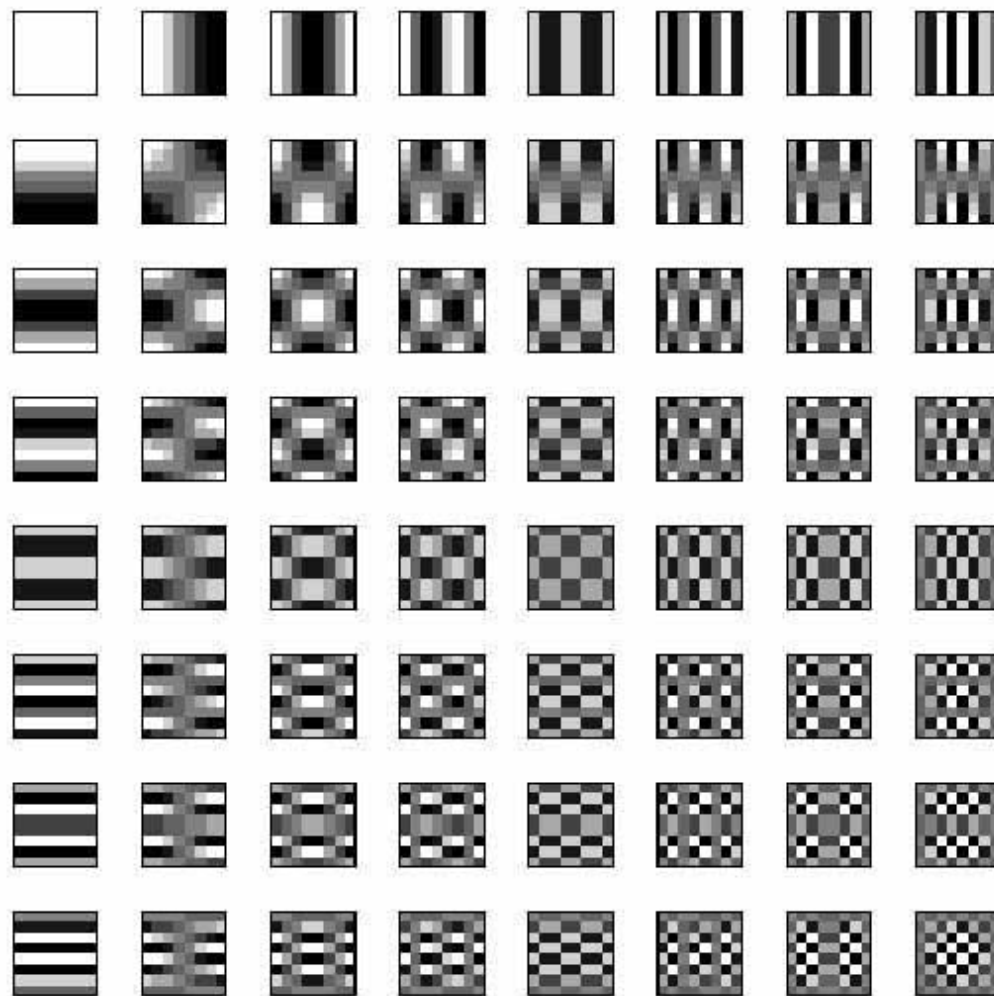
$$x(m, n) = \frac{1}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k, l) \cos\left(\frac{2m+1}{2M} \pi k\right) \cos\left(\frac{2n+1}{2N} \pi l\right).$$

$F(k, l)$ – коэффициенты ДКП.

В MPEG-2 $M, N = 8$. В H.264 $M, N = 4$.



Двумерные гармоники 8x8 пикселей





Квантование коэффициентов ДКП

2. Квантование коэффициентов ДКП с переменным шагом квантования: НЧ – точно, ВЧ – грубо.

$$F_q(m, n) = \text{Round} \left(\frac{F(m, n) \cdot 16}{f \cdot Q(m, n)} \right),$$

где $m, n = 0..7$; $F(m, n)$ - коэффициенты ДКП;

$F_q(m, n)$ - их квантованные значения;

$Q(m, n)$ - элементы весовой матрицы квантования;

f - коэффициент для регулировки степени сжатия; $\text{Round}(x)$ - операция округления числа x до ближайшего целого значения.

3. Кодирование квантованных коэффициентов ДКП: Используется факт, что многие гармоники после квантования равны нулю.



Нормальное сжатие изображения



Несжатый



Сжатый в 6,5 раз



Сравнение увеличенных изображений



Несжатый



Сжатый в 6,5 раз



Сильное сжатие



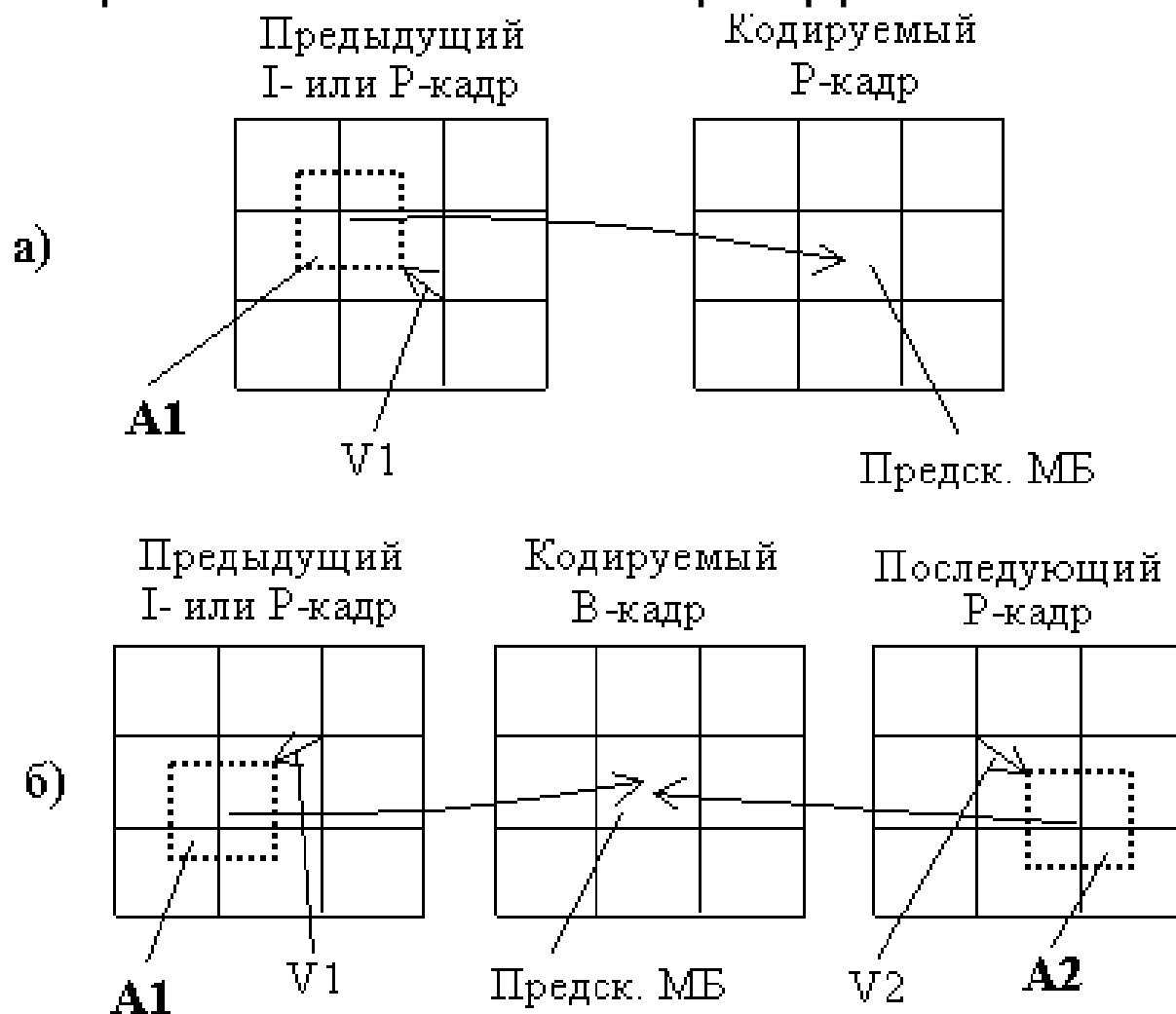


Межкадровое кодирование

1. Поиск для каждого макроблока наиболее похожего фрагмента 16x16 пикселей в одном (для Р изображения) или двух (для В изображения) опорных изображениях.
2. Формирование предсказанного изображения из найденных фрагментов путем их перемещения. Перемещение фрагмента характеризуется вектором движения.
3. Формирование ошибки предсказания как разности кодируемого и предсказанного макроблоков и ее кодирование внутрикадровым методом.
4. Кодирование векторов движения.

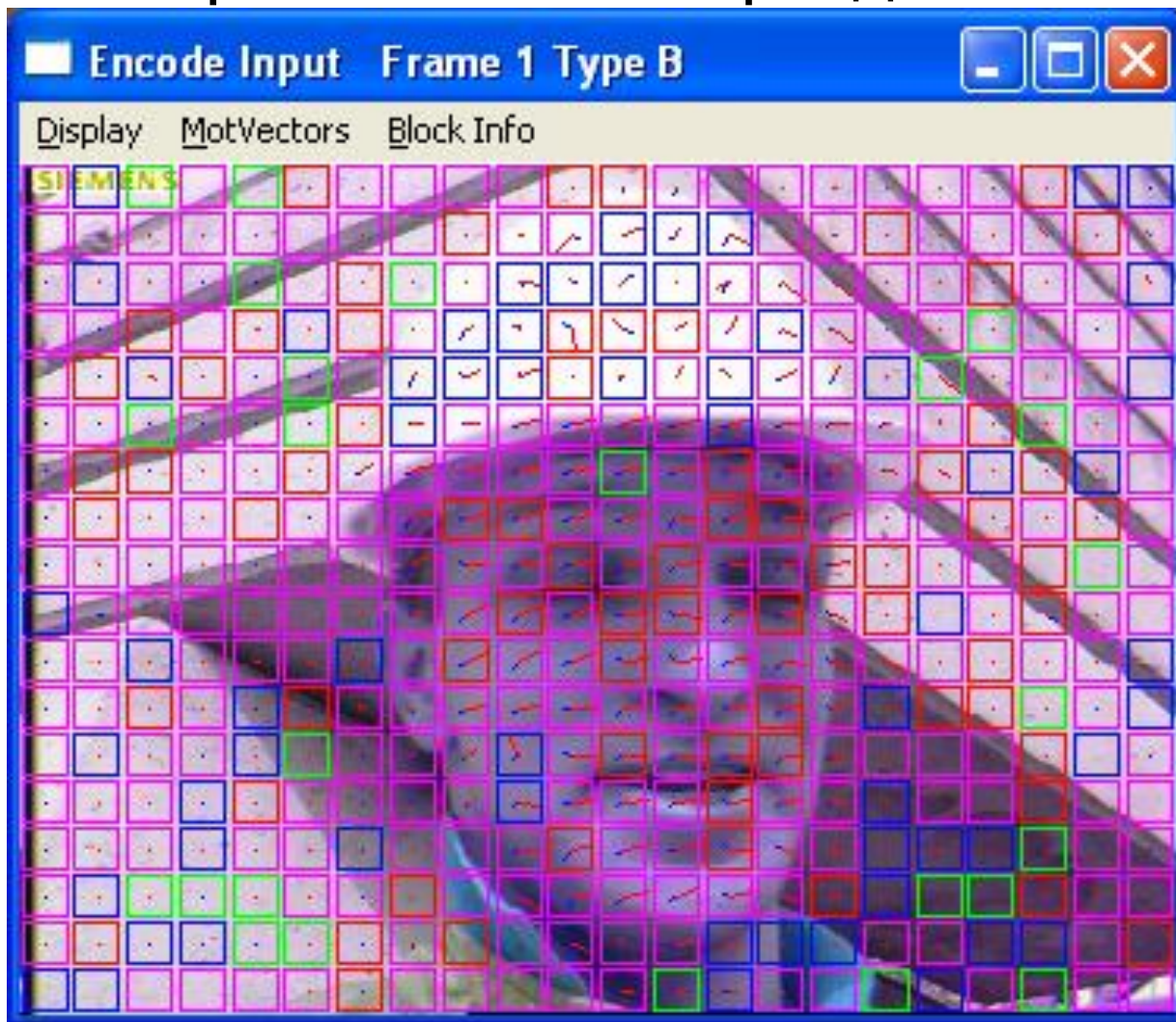


Оценка и компенсация движения





Макроблоки и векторы движения





Ошибка предсказания



Предсказанное изображение

Ошибка предсказания



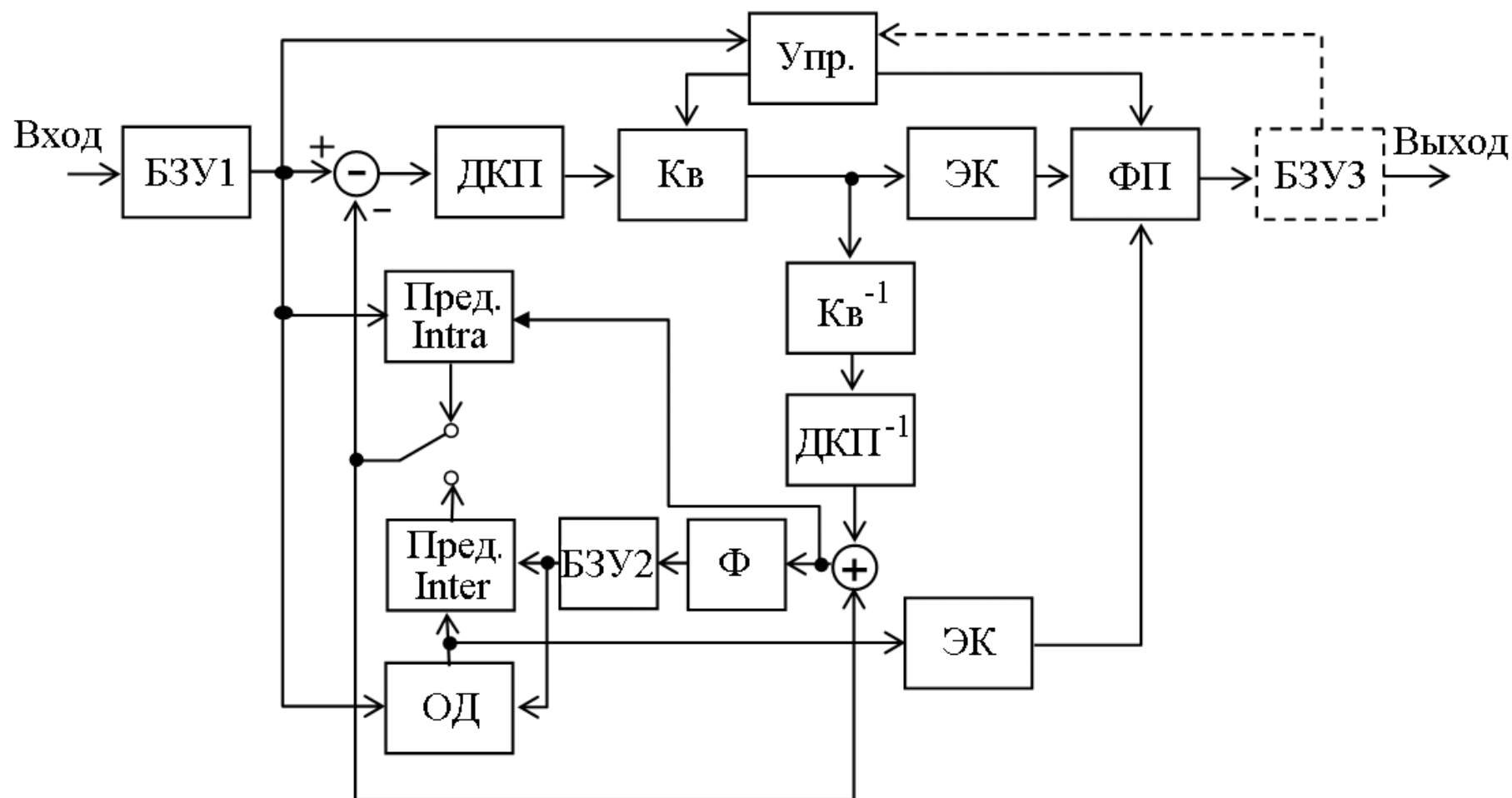
Внутрикадровое предсказание

В этом случае элементы кодируемого макроблока предсказываются по элементам ранее кодированных макроблоков того же изображения. Кодирование макроблоков производится в последовательности, аналогичной построчной развертке, то есть горизонтальными рядами, начиная сверху, и слева направо в каждом ряду. Поэтому для внутрикадрового предсказания можно использовать элементы изображения, расположенные левее и выше кодируемого макроблока. Для левого верхнего макроблока изображения внутрикадровое предсказание не применимо.

В стандартах AVC/H.264, HEVC/H.265 и VVC/H.266 внутрикадровое предсказание стало составной частью внутрикадрового кодирования.



Обобщенная структура видеокодера





Пояснения к схеме видеокодера

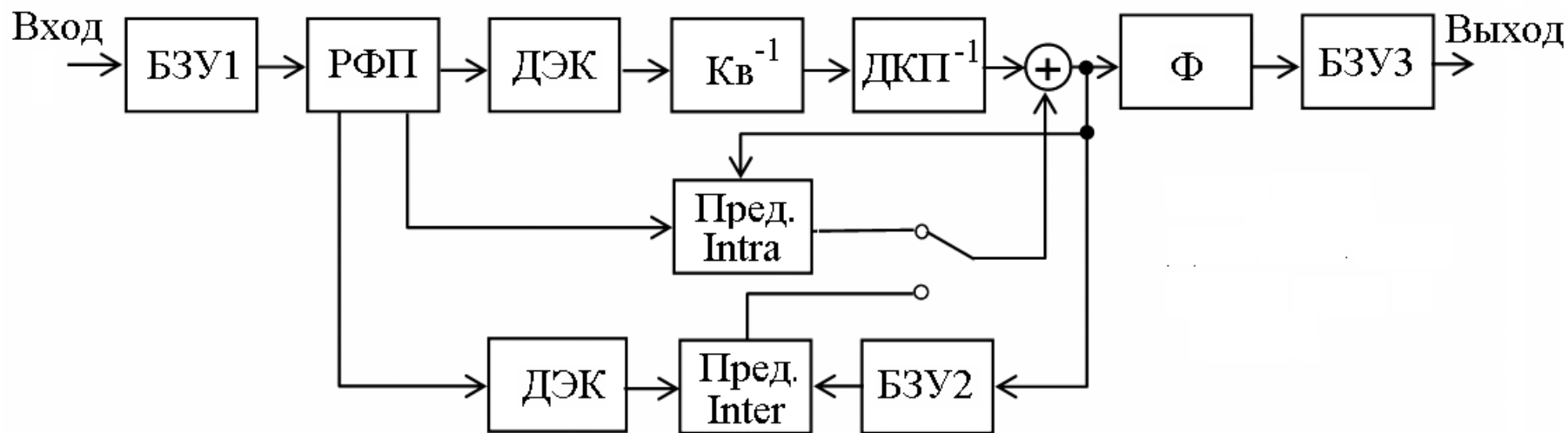
ДКП и ДКП⁻¹ – блоки прямого и обратного дискретного косинусного преобразования; Кв – квантователь; Кв⁻¹ – деквантователь; БЗУ – буферные запоминающие устройства; Пред.Intra и Пред.Inter – блоки, выполняющие внутрикадровое (Intraframe) и межкадровое (Interframe) предсказание; ОД – блок оценки движения и формирования векторов движения; ЭК – энтропийные кодеры; Ф – блок фильтрации, ФП – блок формирования потока; Упр. – блок управления. В кодере MPEG-2 блоки внутрикадрового предсказания и фильтрации отсутствуют.

Предсказание выполняется по видеоданным, прошедшим сжатие и содержащим ошибки, созданные квантованием. Эти ошибки попадают в следующие кадры со знаком "минус".

Управление сжатием осуществляется путем изменения коэффициента f в квантователе. Автоматическое управление сжатием поддерживает заданный битрейт на выходе.



Обобщенная структура видеодекодера



РФП - расформирование потока; ДЭК - энтропийные декодеры.



Масштабируемость (Scalability)

Масштабируемость определяется как возможность получения изображения из части полного потока видеоданных, который при этом состоит из двух или более слоев (layer). Базовый слой дает возможность получить изображение с некоторыми начальными параметрами качества. Дополнительные или улучшающие (enhancement) слои потока данных позволяют получить изображение улучшенного качества.

Может быть масштабируемость по пространственному разрешению, по времени (по частоте кадров), по отношению С/Ш, где под шумом понимают ошибки сжатия.



Уровни и профили

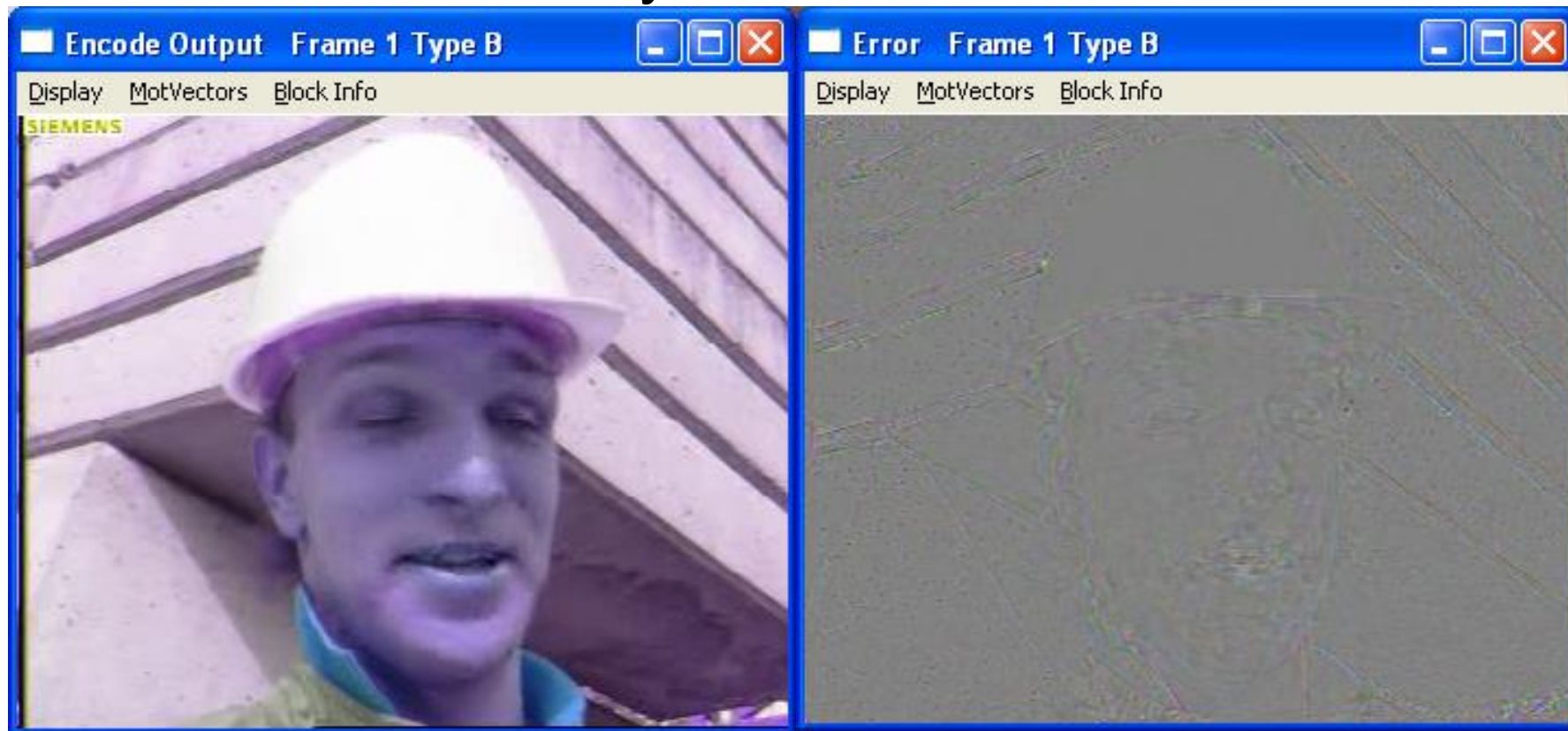
В стандартах видеокомпрессии используются понятия уровня (Level) и профиля (Profile).

Разные **уровни** соответствуют разным количествам строк в кадре и пикселей в строке.

Профиль показывает уровень сложности и функциональных возможностей применяемых средств кодирования. Могут быть более простые и более сложные профили, в частности, профили без масштабируемости и масштабируемые. Также есть специализированные профили для особых областей применения.



Результаты сжатия



Сжатый кадр

Ошибка сжатия

Выходной битрейт 1500 кбит/с



Визуальная оценка искажений

Балл	Оценка	Заметность искажений
5	отлично	незаметно
4	хорошо	заметно, но не мешает
3	удовлетвор.	заметно, немного мешает
2	плохо	мешает
1	очень плохо	сильно мешает

0	одинаково
-1	немного хуже
-2	хуже
-3	намного хуже



Количественная оценка искажений

PSNR (Pixel Signal-Noise Ratio)

$$PSNR = 10 \lg \left(\frac{D_{\max}^2 MN}{\sum_{i=1}^M \sum_{j=1}^N (x_{ij} - y_{ij})^2} \right) \text{ (дБ)}.$$

M и N - размеры изображения по вертикали и по горизонтали,

x_{ij} - значения отсчетов исходного изображения,

y_{ij} - значения отсчетов искаженного изображения,

D_{\max} - максимальное значение цифрового сигнала, например, 255 при 8 битах на отсчет для каждого цвета.



Спасибо за внимание!