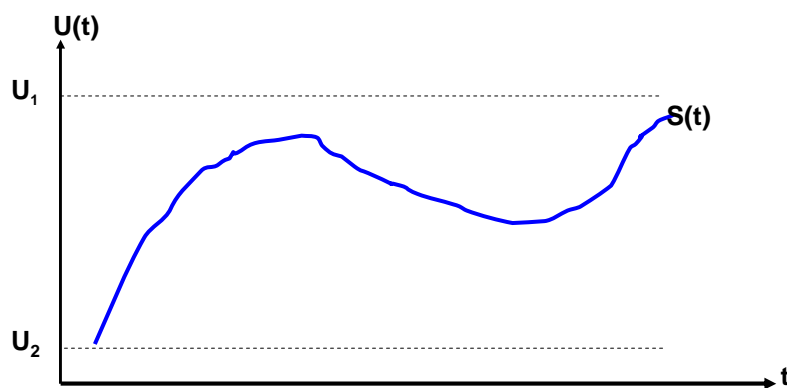


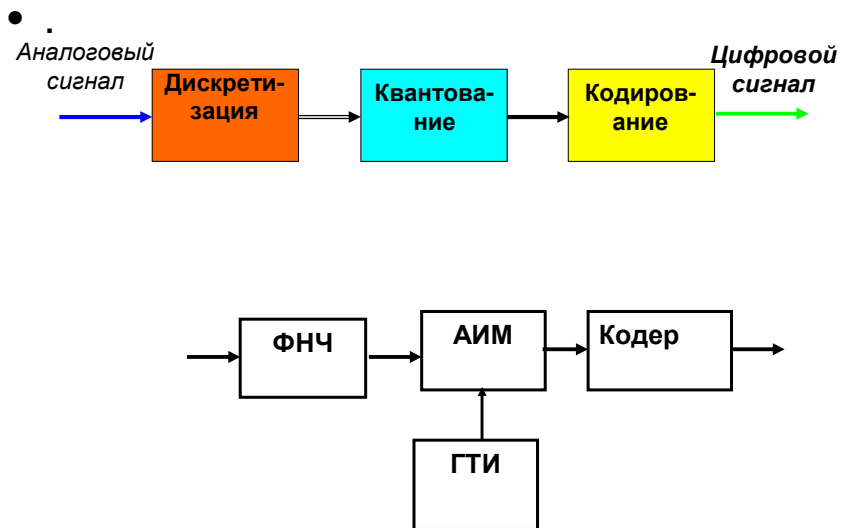
Методы цифровой обработки и сжатия сигналов изображения и звука. Стандарты MPEG

Копылов Анатолий Михайлович
Доцент кафедр «Инфокоммуникаций»
и «Радиовещания и акустики» МТУСИ
E-mail: amkop@yandex.ru; mppk@mtuci2.ru
Т. +7-499-1916673
28-30 октября 2008 г.

Аналоговый сигнал

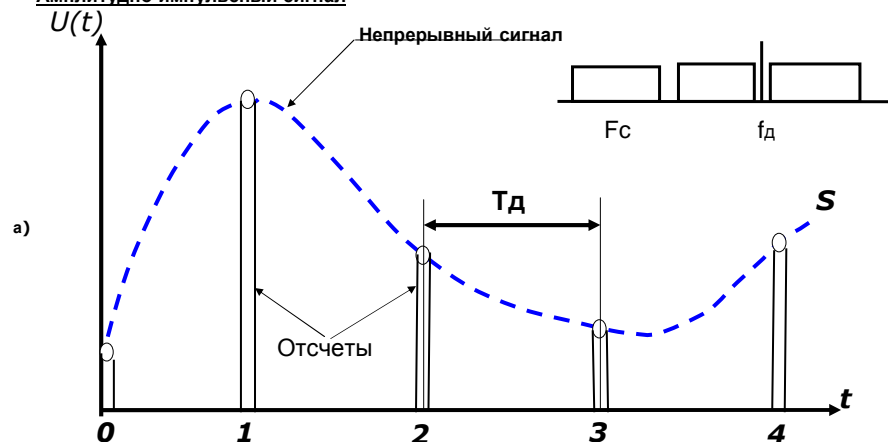


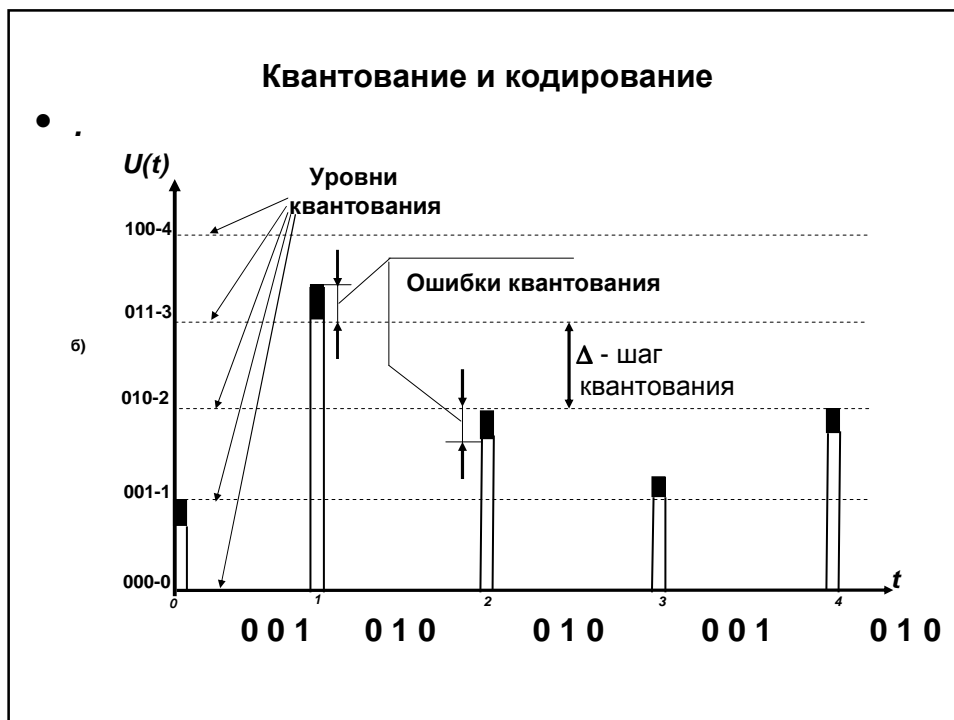
Последовательность операций преобразования аналогового сигнала в цифровой.



Представление непрерывного сигнала в цифровой форме: дискретизация во времени

- Амплитудно-импульсный сигнал





Частота дискретизации $F_d = 1/T_d > 2F_d$ и системные требования.

Звуковое вещание:

И трактах формирования программ $F = 48$ кГц;

В лазерных проигрывателях и видеоманитофонах - 44,1 кГц;

В трактах распределения программ - 32 кГц.

(кратна частоте дискретизации телефонного канала 8 кГц)

Верхняя частота ТВ изображения - 6,375 МГц (ГОСТ 7845-92)

Для сигнала яркости выбрана частота - 13,5 МГц

Она кратна строчным частотам двух стандартов разложения

625/50 и 525/59,94. Рекомендация 601МСЭ-Р для

производства ТВ программ стандартной чёткости.

Для цветоразностных сигналов $F = 6,75$ МГц,

Для формата изображения 16Х9 частота дискр. - 18 МГц

Для ТВ пониженной чёткости и мультимедиа применяют

формат по яркости 352Х288, по цветности - 176Х144 для

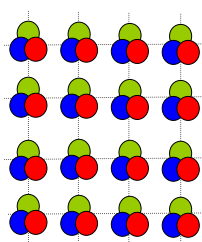
625 строчного разложения.

Рекомендации ITU-R BT 601

Дискретизация

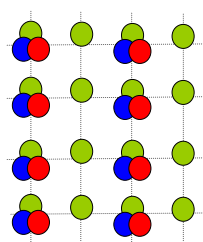
- Формат дискретизации обозначает:
 - 4:4:4 – оба цветоразностных сигнала передаются в каждой строке и дискретизируются с той же частотой, что и яркостной сигнал
 - 4:2:2 – Частота дискретизации цветоразностных сигналов в два раза меньше, чем частота яркостного сигнала.
 - 4:2:0 – каждый цветоразностный сигнал имеет частоту дискретизации в 2 раза ниже частоты дискретизации яркостного сигнала и передается через строку.
 - 4:1:1 – оба цветоразностных сигнала передаются в каждой строке, с частотой дискретизации в 4 раза меньше частоты дискретизации сигнала яркости.

Цифровые видео форматы



4 : 4 : 4

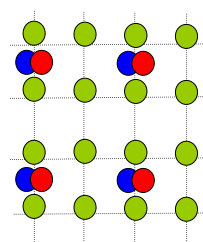
Каждый отсчет (пиксель) описывается – Y, Cr и Cb
Применяется в студийной работе



4 : 2 : 2

Каждый отсчет описывается Y. И каждый второй отсчет в строке описывается Cr и Cb
Применяется в студийной работе

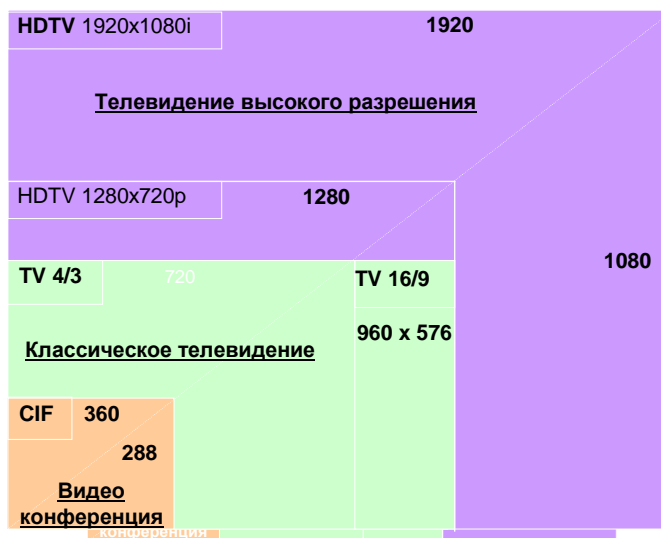
Основной формат



4 : 2 : 0

Каждый отсчет описывается Y. И только каждый второй отсчет и на 2 строки описывается Cr и Cb.
Применяется в широковещательной трансляция

Форматы телевизионных экранов по рекомендации (ITU-R 601 Rec)



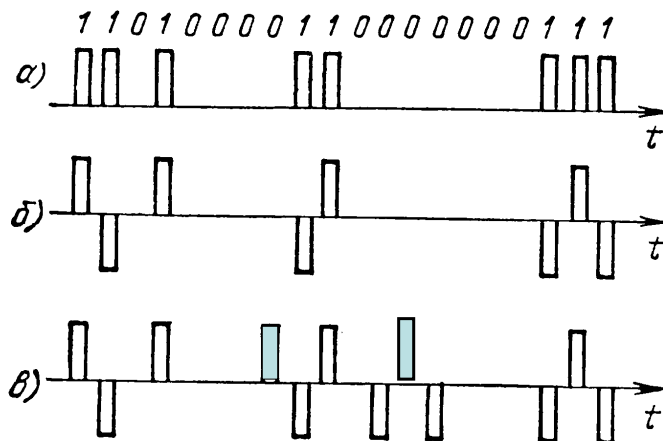
- $n_{\max} = 2^m$; $m = \log_2 n$;
- $m = 8$; $n = 256$; $m = 16$, $n = 65536$
- Симметричные коды + \rightarrow 1; - \rightarrow 0.
- $P_{\text{шк}} = \Delta^2 / 12$; $P_{\text{шп}} = \Delta^2 / 4$;
- Сигнал звукового вещания
 $V = m \cdot f_{\text{д}} \cdot n = (16 \cdot 48) = 768 \cdot 2 = 1536$ кбит/с;
- Телефонный сигнал: $V = 8 \cdot 8 = 64$ кбит/с.
- Сигнал ТВ изображения
 $V = 13,5 \cdot 10^6 \cdot 10 + 2 \cdot 6,75 \cdot 10^6 \cdot 10 = 270$ Мбит/с
 Для формата 16x9 $V = 360$ Мбит/с
 Для цифрового сигнала высокого качества (ТВЧ). Частота дискретизации сигналов цветности
 – 74,25 МГц, скорость цифрового потока -1,485 Гбит/с.

Ширина полосы частот сигналов

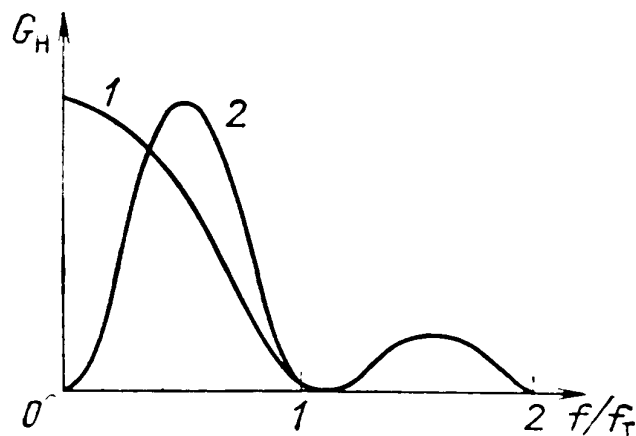
- $B_n = 2F_v$ – при двухполосной передаче;
- $B_n = F_v - F_n$ – при однополосной передаче;
- При частотной модуляции несущей $B_n = 2(\Delta f_d + F_v)$,
где Δf_d – девиация частоты.
- Для систем с цифровой манипуляцией несущей МСЭ–Р рекомендует определять значение НШПЧ исходя из условия попадания в эту полосу 99% средней мощности излучения.
- ШПЧ можно вычислить по формуле: $B_n = RK$,
- где R - скорость передачи, K – эмпирический весовой коэффициент, выбираемый в соответствии с рекомендацией МСЭ–Р. Значения весовых коэффициентов K приняты равными: 0,91 (система GSM), 1,03 (система DECT), 1 (система CDMA 450).
- Если 1 бит/Гц необходимо сжатие цифрового потока 270 Мбит/с примерно в 35 –40 раз, полоса 8МГц. $K = 1,2 \dots 1,3$

Необходимо сжатие цифровых сигналов

Линейные коды: а) однополярные импульсы, б) квазитроичный код, в) код с высокой плотностью единиц



Спектральная плотность мощности цифровых сигналов: 1) однополярных импульсов, 2) квазитроичного кода



Скремблирование

(scramble – перемешивать)

- При появлении длинных серий нулей может произойти сбой в синхронизации.
- Чтобы устранить это явление к двоичным символам цифрового потока добавляют другую последовательность:

(сумматоры по модулю 2)

Цифровой поток	0111000000000000000011
Скремблер	101010101010101010101
Выходной поток	110110101010101010110

- Длинные последовательности нулей исчезли.
- Чтобы вернуться к исходной последовательности, нужно сложить сигнал с такой же последовательностью

Достоверность цифровой передачи

- При использовании симметричного кода ошибка приводит к изменению полярности. Её заметность зависит от абсолютного значения сигнала.
- При линейной шкале квантования ошибка в старшем разряде приводит к погрешности, равной половине шкалы квантования, в младшем разряде погрешность равна шагу квантования.
- Если ошибка происходит в трёхразрядной комбинации, определяющей шкалу квантования при почти мгновенном компандировании, то с погрешностью будут восстановлены все 32 отсчёта 9(на 6 дБ, если ошибка в младшем разряде и связана с переходом к соседней шкале квантования; на 12 дБ, если ошибка во втором разряде, на 24дБ, если ошибка в старшем разряде.
- *При одиночных ошибках в младших разрядах их влияние может быть практически не заметно на слух.*
- *Ошибки в пяти - шести старших разрядах проявляются в виде щелчков, сильно ухудшающих качество звучания.*

- Считается, что допустимым является появление не более одного щелчка в час на канал, и что к щелчкам приводит половина всех ошибок, тогда

$$P_{\text{ош}} = 1/(2 \cdot 0,5 \cdot F_{\text{д}} \cdot 3600 \cdot t).$$

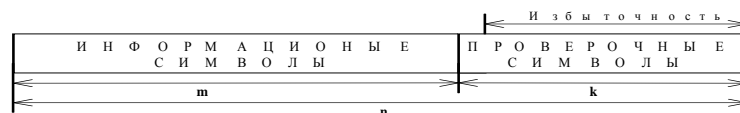
Для сигнала с гауссовском распределением требуемое отношение С/Ш на входе декодера должно быть не менее 24 ...26 дБ.

в реальных цифровых каналах связи с учётом их протяжённости и действия различных помех это трудно осуществить, средняя вероятность ошибки составляет 10-5 ...10-6.

Поэтому нужно принимать специальные меры по обнаружению и дальнейшему маскированию или исправлению ошибок, если $P_{\text{ош}} > 10^{-5}$.

Помехоустойчивое кодирование

Б Л О К О В Ы Й К О Д



- Введение избыточности в блочный код. К m информационным символам добавляются k проверочных символов.
- Относительная скорость кода:
$$R = m/n = (n - k)/n = 1 - k/n,$$
- где n – общее число символов.
- С целью упрощения техники кодирования и декодирования в цифровом телевидении применяются блочные циклические коды, в частности коды Хэмминга, Боуза, Чоудхури, Хэквиема (БЧХ) и их разновидности коды Рида – Соломона. $R = 1/2, 2/3, 3/4, 5/6, 7/8$.

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

- Сокращение статистической избыточности не более 3:1
Коды с переменной длиной кодового слова (VLC - Variable Length Coding), самый известный - код Хаффмана.

По принципу азбуки Морзе.

Алфавит «а», «в», «с», «d» с вероятностями 0,5;0,25;0,125;

Присвоим значения 00; 01; 10; 11; длина - 2 бита/символ.

Присвоим символу а значение 0, в - 10; с и d - значения 110 и 111.

В среднем для передачи одного символа потребуется:

$1 \times 0,5 + 2 \times 0,25 + 2 \times 3 \times 0,125 = 1,75$ бит. Вместо 2 бит.

Например, встречаемость букв на 1000 символов:

«о»-90; «а» и «и»- по 62; «щ» и «Э» - по 3; «ф» - 2.

Соответственно часто встречаемым присваиваются короткие последовательности (• -)

Редко встречаемым символам длинные (- - •••)

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

- **Квантование** - составная часть многих алгоритмов сжатия с потерей информации, обеспечивающих сжатие от 3:1 до 100:1. Один из таких алгоритмов - полосное кодирование. Если применять квантование с адаптивным распределением битов, учитывающим долю энергии сигнала в данной полосе, можно добиться существенного сжатия цифрового сигнала.

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

- **ДИКМ** - передаётся разность между соседними элементами в строке, в соседних строках и смежных кадрах. Разновидность ДИКМ - кодирование с межкадровым предсказанием, при котором передаётся разница между текущим кадром и кадром, предсказанным в кодере по одному или нескольким предшествующим кадрам. При хорошем предсказании эта разность значительно меньше, чем различие между двумя последовательными кадрами.

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

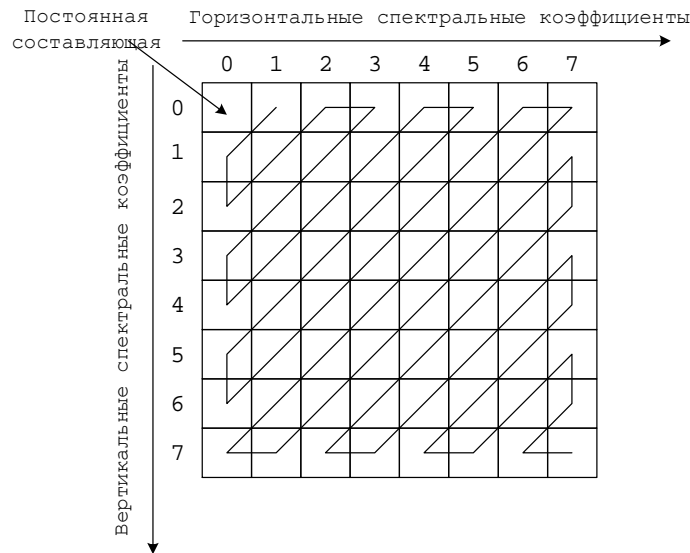
- **ДКП - дискретно-косинусное преобразование.** Избыточность сокращается *кодированием с преобразованием*, при котором набор статистически зависимых отсчётов изображения во временной области преобразуется в наборов независимых коэффициентов в спектральной области. Энергия в спектральной области сосредотачивается вблизи НЧ компонент спектра, а коэффициенты в ВЧ области оказываются малыми или вообще нулевыми.

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

- Пространственная избыточность обусловлена наличием в видеокадре значительных по размеру однотонных участков, отсчеты сигнала в соседних точках практически одинаковы или слабо изменяются, то есть присутствуют только НЧ составляющие двумерного пространственного спектра. Изображение разбивается на блоки 8 x 8 пикселей. С помощью ДКП 64 отсчётам цифрового сигнала ставится в соответствие 64 коэффициента
- Зачастую большую часть коэффициентов после квантования можно отбросить.

Сканирование значений матрицы спектральных коэффициентов

в порядке возрастания пространственных частот.



В результате зигзагообразного сканирования образуется последовательность из 63 спектральных коэффициентов, в которой некоторые коэффициенты имеют нулевые значения, причем нулевые коэффициенты могут образовывать последовательности. Количество нулей и код ненулевых коэффициентов с помощью специальных кодовых таблиц преобразуются в код, который передается по каналу связи.

- При кодировании по Хаффмену используется распределение вероятности появления значений коэффициентов, то есть производится энтропийное кодирование.

Восстановленные значения сигнальных матриц коэффициентов преобразуются в последовательность цифровых сигналов, которые после цифро-аналогового преобразования вновь представлены в аналоговой форме

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

- Применяя эффективные методы кодирования отдельно для контуров (резкие перепады яркости) и текстуры (всё, что не контуры) можно получить очень высокую степень сжатия, стандарт MPEG-4.

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

- Временная избыточность изображения проявляется в относительно малом отличии двух последовательных кадров ТВ изображения. Можно сократить объём передаваемой информации, если содержание опорного кадра передавать целиком, а вместо следующего кадра передавать только информацию о переместившихся объектах. Предсказание кадра на основе одного или нескольких предыдущих и последующих кадров и передача разницы между текущим и предсказанными кадрами.

МЕТОДЫ ЦИФРОВОГО СЖАТИЯ

- Для улучшения результатов предсказания применяют процесс, называемый **компенсацией движения**. Изображение в текущем кадре разбивают на блоки и для каждого такого блока ищут его наиболее вероятное положение. В реальных кодерах отсчёты сигнала, соответствующие одному видеокадру, сначала группируются в блоки 16 x 16 пикселей, производится компенсация движения, далее макроблоки разбиваются на блоки отсчётов яркости и цветности размером 8 x 8, для них вычисляются межкадровые разности, которые подвергаются ДКП для устранения пространственной избыточности. Это позволяет получить коэффициент сжатия 100:1 и более.

СТАНДАРТ ЦИФРОВОГО СЖАТИЯ MPEG-1

- **MPEG** означает Motion Pictures Experts Group
- (Группа экспертов по движущимся изображениям)
- Полное название группы - ISO/IEC JTC1/SC29/WG11
- Группа создана в 1988 г. для выработки международного стандарта движущихся изображений и звуковых сигналов до скорости порядка 1,5 Мбит/с с целью записи на CD-ROM с качеством бытового стандарта VHS. Исходные скорости - 270 (360) Мбит/с.

СТАНДАРТ ЦИФРОВОГО СЖАТИЯ MPEG-1

- Стандарт принят в 1993 г. и получил индекс ISO/IEC 11172.
- Стандарты MPEG не определяют схему и конструкцию кодера и декодера, они лишь описывают средства, используемые для обработки сигнала, определяют правила построения последовательности символов совместимого цифрового потока и дают примеры реализации декодера.
- Состав и построение кодера оставлены на усмотрение разработчика. Декодер должен декодировать любой цифровой поток, совместимый со стандартом MPEG-1.

MPEG-1 - методы компрессии:

- предсказание, когда передаются не сами элементы изображения, а разность между элементами в строке, в соседних строках и смежных кадрах;
- внутрикадровое и межкадровое кодирование,
- дискретно-косинусное преобр. Компенсация движения,
- Определены три типа видеокадров:
- *I* - *видеокадры* (от *intra*-внутри)-внутрикадровое кодир.
- *P* - *видеокадры* (*predicted*) первый *P* кадр предсказывается по *I* - кадру, второй и последующие по предыдущему *P* кадру;
- *B* - *видеокадры* кодер рассчитывает как прямое, так и обратное предсказание и посылает декодеру данные, имеющие наименьший объём (отсюда - *bidirectional*, то есть двунаправленный). Польза *B* - кадров проясняется при рассмотрении задней границы движущегося объекта, для передачи открывающегося фонового изображения выгоднее воспользоваться данными более позднего кадра.

СТАНДАРТ ЦИФРОВОГО СЖАТИЯ MPEG-1

- Ограничения, определяющие область применения:
- размер изображения по горизонтали < 768 пикс.;
- размер изображения по вертикали < 576 строк;
- частота кадров < 30 Гц;
- развёртка прогрессивная (в ТВ системах стандартного качества - черезстрочная);
- число макроблоков < 396 (для ТСЧ надо иметь (720: 16) x (576: 16) = 1620 макроблоков /кадр);
- скорость цифрового потока < 1,856 Мбит/с.
396 макроблоков соответствует формату разложения CIF (352x480), не для вещательного телевидения.

СТАНДАРТ ЦИФРОВОГО СЖАТИЯ MPEG - 2

- Стандарт MPEG-2 принят в 1996 г., дополнен в 1997 г. и получил индекс ISO/IEC 13818.
- Стандарт MPEG-2 значительно улучшен по сравнению с MPEG-1, имеется возможность обработки черезстрочных изображений, набор уровней и профилей, масштабируемый синтаксис, системный уровень с программным и транспортным потоками, новые средства кодирования звука и др.
- Стандарт MPEG-2 значительно сложнее MPEG-1 (кодер примерно на 50%), охватывает более широкий круг применений, включая вещательное телевидение. MPEG-2 может использоваться и для ТВЧ, поэтому были прекращены работы над стандартом MPEG-3, для систем ТВЧ.
- Стандарт MPEG-2 называется «Информационные технологии – Обобщённое кодирование движущихся изображений и сопровождающей звуковой информ.»

СТАНДАРТ ЦИФРОВОГО СЖАТИЯ MPEG-2

- В стандарте приняты 5 основных и один дополнительный профессиональный Профиль "4:2:2", введённый позднее.
- Внутри каждого Профиля выделено 4 уровня, определяющие допустимые пределы изменения параметров цифрового потока.

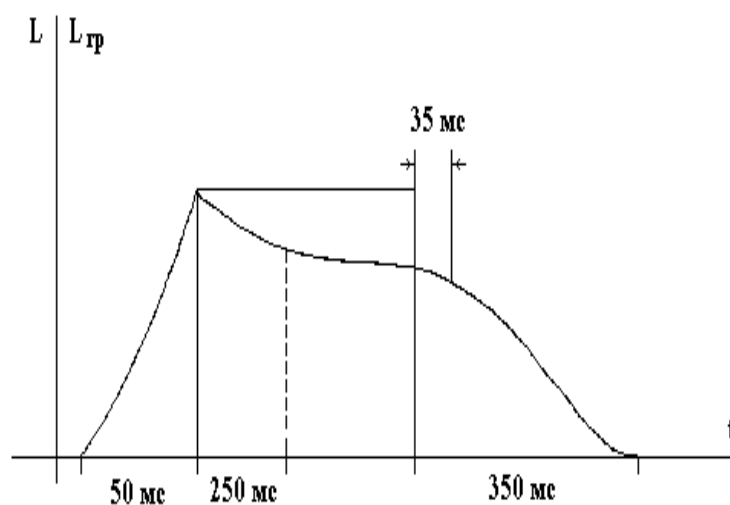
Уровни основного профиля MPEG-2

- Высокий **MP@HL** Число отсчётов на строку 1920 Число строк на кадр 1152 Частота кадров, Гц 60 Частота отсчётов сигнала яркости, отсч/с 62668800 Скорость потока Мбит/с 80 Размер буфера, бит 9781248
- Высокий **MP@H** Число отсчётов на строку 1440 Число строк на кадр 1152 Частота кадров, Гц 60 Частота отсчётов сигнала яркости, отсч/с 47001600 Скорость потока Мбит/с 60 Размер буфера, бит 7340032
- Основной **MP@ML** Число отсчётов на строку 720 Число строк на кадр 576 Частота кадров, Гц 30 Частота отсчётов сигнала яркости, отсч/с 10368000 Скорость потока Мбит/с 15 Размер буфера, бит 1835008
- Низкий **MP@LL** Число отсчётов на строку 352 Число строк на кадр 288 Частота кадров, Гц 30 Частота отсчётов сигнала яркости, отсч/с 3041280 Скорость потока Мбит/с 4 Размер буфера, бит 475136

Редукция (сокращение) звуковых данных, обусловленное психоакустическими особенностями

- В присутствии громкого сигнала кривая слышимости слуха поднимается выше.
- Ухо не в состоянии различать звуки с близкими частотами (частотное маскирование).
- Насчитывается 25 полос частот, называемые критическими, в пределах которых действует маскирование.
- В области НЧ ширина критической полосы слуха менее 100 Гц, в районе 2кГц – 300 Гц, в области ВЧ – 4 кГц.
- Громкий звук маскирует слабый раньше громкого сигнала на 5 – 50 мс (явление предмаскировки), во время громкого сигнала и после его прекращения до 50-200 мс.
- Цифровые данные можно сжимать более, чем в 10 раз.

Характер восприятия по громкости



Сжатие звуковых сигналов (стандарты MPEG)

- **Уровень 1** – $F_v = 15$ кГц, $n = 4$, $t_z = 20$ мс, $v = 192$ кбит/с.

Рассчитан на упрощенный кодер. Используется полосное кодирование (32 полосы), слуховая модель создаётся по уровням в самих частотных полосах.

- **Уровень 2** – $F_v = 15$ кГц, $n = 6$, $t_z = 40-50$ мс, $v = 128$ кбит/с.

Тот же набор фильтров и отдельный частотный анализатор, который создаёт более точную слуховую модель. Масштабирующие коэффициенты создаются более сложным образом, учитывается сходство этих коэффициентов в соседних кадрах звукозаписи.

Уровень 3 – $F_v = 15$ кГц, $n = 12$, $t_z =$ более 50 мс, $v = 64$ кбит/с.

Намного сложнее. Попытка создать точную модель слухового Восприятия. Моделируется ширина критических полосок слуха. Применяется код Хаффмена.

Кодирование звука в MPEG-2

- В MPEG-2, совместимом с MPEG-1 используются трёхуровневые системы кодирования звука. Различие между стандартами начинается при переходе от двухканального звука к многоканальному звуку (5+1).
- Одной из разновидностей многоканального звука является многоязычное звуковое сопровождение. Оно может осуществляться :
 - либо передачей отдельного цифрового потока для каждого языка,
 - либо добавлением нескольких (до 7) языковых каналов 64 кбит/с к многоканальному потоку 384 кбит/с.
- В дополнение к основному режиму с частотами дискретизации 32, 44,1 и 48 кГц в MPEG-2 введён низкочастотный режим (Low Sampling RATE) с пониженными вдвое частотами дискретизации: 16, 22,05 и 24 кГц. Это позволяет субъективно повысить качество звучания речевого сигнала. Отбрасывание верхних частот почти не влияет на качество речи, а высвобождающиеся ресурсы битов используются кодером для более точной передачи нижней части спектра.

MPEG4 Кодирование видеообъектов

В отличие от MPEG-1, MPEG-2 в которых применяется фиксированный алгоритм кодирования, в MPEG-4 используется целый набор методов кодирования, включающий как алгоритмы сходные с применяемыми в MPEG-1 и MPEG-2, так и принципиально новые методы кодирования основанные на понятии видеообъекта.

Выбор метода кодирования в конкретном случае определяется характером изображения и требуемым коэффициентом сжатия информации.

MPEG4

Видеообъекты, аудиообъекты и аудио-визуальный объектами

Видеообъектами (VO - visual object) могут быть изображения людей и предметов, перемещающиеся перед неподвижным фоном, и сам неподвижный фон.

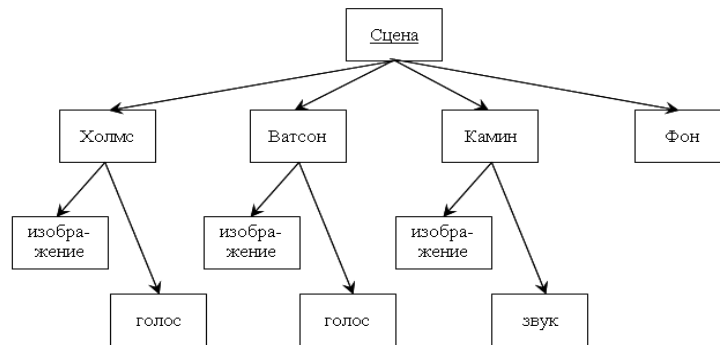
Аудиообъектами (AO - audio object) могут быть голоса людей, музыка, другие звуки.

Аудио-визуальный объектами (AVOs – audio-visual object) могут быть связанные видео и аудио объекты, например, изображение человека и его голос.

MPEG4

MPEG-4 содержит специальный язык описания сцен
– BIFS (Binary Format for Scenes)

Сцены MPEG-4

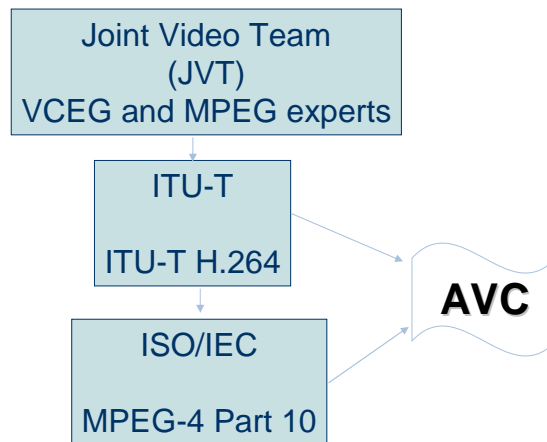


Пример структуры описания сцены

MPEG4 H.264

- Рекомендация является заметным шагом вперед по сравнению с версией стандарта MPEG4 (рассмотренная рекомендация должна стать расширением стандарта MPEG 4).
Формат кодирования позволяет добиваться эффективности представления информации, превосходящей эффективность, допустимую в рамках стандарта MPEG4, более чем на 30%.
Естественно, вычислительная сложность несколько выше вычислительной сложности стандартных решений, однако, это обстоятельство не должно стать препятствием для использования этой рекомендации.

5 названий на одно и то же:
H.264, MPEG-4 Part 10, AVC, H.26L, JVT



H.264 - профили и уровни

- Разные конфигурации - разные возможности
 - Профиль: набор алгоритмических особенностей
 - Level: уровень возможностей. 4 профиля H.264
- **Baseline** - наименее сложный (нет B-slices).
Области применения: видео телефония, видеоконференции, беспроводные системы
 - **Main** - наилучшее сочетание уровня компрессии и качества видео под выбранную скорость передачи.*Области применения:* широкоэ debateльное телевидение+DVD стандартного разрешения
 - **Extended** - потоковое видео с очень высокой скоростью передачи
Области применения: сервисы мобильного видео Mobile video services
 - **High** (High, High 10, High 4:2:2, High 4:4:4) - кодирование видео без потерь. *Области применения:* ориентирован на HDTV

Стандарты сжатия ТВ сигнала

Страна	Стандарт	Кол-во MUX	Кол-во каналов	Начало трансляции	Отключение аналогового вещания
Австрия	MPEG-2	1	4/--	2005	2010
Испания		5	20 / 6*	2005	2010
Великобритания		6	30 / 10*	2002	2012
Германия		4	12 / 14*	2003	2010
Дания		1	4 / --	2005	2009
Финляндия		4	12 / 10*	2002	2007
Италия		6	24 / 14*	2004	2008
США	MPEG-2/4			2002	2012
Франция		6	19 / 11*	2005	2010
Швеция		5	8 / 23*	1999	2008
Бельгия		2	5 / 6*	2003	2010
Литва	MPEG-4	4	-- / 27*	2006	2012
Эстония		2	10 / 12*	2007	2010
Россия	MPEG-2/4	?	?	2007	2015
Белоруссия	MPEG-4			2008	
Казахстан				2008	
Украина				2008	
Молдова				2008	
Армения				2009	
Грузия				2009	

* - платные каналы

Благодарю за внимание!