



- Наиболее ярко особенности кодирования сигналов динамических изображений проявляются в системах вещательного телевидения. Многочисленные исследования в области кодирования сигналов динамических изображений привели к разработке новых принципов построения устройств и систем, стандартов кодирования и передачи видеосигналов
- 1988 г. сформирована международная группа экспертов по Движущимся изображениям - *Motion Pictures Experts Group (MPEG)* при Объединенном техническом комитете по информационным технологиям
- За последующие годы с целью хранения, обработки и передачи динамических изображений в мультимедийных приложениях этой группой были разработаны стандарты MPEG-1 MPEG-2, MPEG-4 и MPEG-7.

# Стандарт MPEG-1

- Сигналы от источников сообщений (видео, звука и данных), подвергнутые сжатию, в телевидении мультиплексируются в общий цифровой поток (называемый системным в MPEG-1, программным или транспортным в MPEG-2). Синтаксис этих потоков установлен соответствующим стандартом MPEG.
- Цифровой поток подвергается помехоустойчивому кодированию, вид которого зависит от используемого канала связи.
- Стандарт *MPEG-1* принят в 1993 г. и получил в международной классификации индекс *ISO/IEC11172*.
- *MPEG-1* не определяет схему и конструкцию кодера и декодера, он лишь описывает средства, используемые для обработки сигнала, определяет синтаксис (правила построения последовательности символов) цифрового потока.
- В качестве кодера может быть аппаратное или программное устройство любой сложности, дающее на выходе синтаксически правильный цифровой поток.

# ИЗБЫТОЧНОСТЬ

- **Избыточность является следствием определенных корреляционных связей.**
- **Наличие корреляции означает, что некоторый элемент изображения находится в определенной зависимости от соседних элементов в пространстве и во времени. Избыточность по восприятию (психовизуальная) связана с особенностями зрения человека.**
- **Например, цветовое разрешение человеческого зрения ниже яркостного, и эта особенность учтена во всех стандартных аналоговых системах цветового кодирования. В NTSC, PAL, SECAM цветовое разрешение существенно понижено по отношению к яркостному.**

# Алгоритм обработки видеоданных

- При разработке стандарта были приняты следующие ограничения, определившие его область применения:
- размер изображения по горизонтали  $< 768$  пикс;
- размер изображения по вертикали  $< 576$  строк;
- число макроблоков  $< 396$ ;
- частота кадров  $< 30$  Гц;
- развертка прогрессивная;
- скорость цифрового потока  $< 1,856$  Мбит/с.
- Ограничивающим параметром стандарта *MPEG-1* для кодирования сигналов вещательного телевидения является число макроблоков в видеокадре. Для обработки сигнала изображения Телевидения стандартной четкости (720 x 576 пикс.) надо иметь  $(720:16) \times (576:16) = 1620$  макроблоков/кадр.
- Второе ограничение - отсутствие чересстрочной развертки, принятой сегодня во всех телевизионных системах стандартного качества (уже и для аналогового ТВ).

- Разрядность квантования входного сигнала должна быть 8 бит. При поступлении на вход кодера *MPEG-1* 10-битового потока он игнорирует два *LSB* {*Least Significant Bit*} - младших бита, огрубляя отсчеты.
- Алгоритм работает следующим образом.
- Видеокадр разбивается на макроблоки размером 16 x 16 пикселей, каждый из которых содержит по 4 блока отсчетов яркости размером 8x8 пикселей и по одному блоку отсчетов сигналов цветности  $C_B$  и  $C_R$  (с учетом понижающей дискретизации).
- При кодировании макроблока в составе *I-видеокадра* вычисляются коэффициенты *ДКП* (дискретное косинусное преобразование), затем они квантуются с использованием таблицы квантования, имеющейся в памяти кодера и декодера и применяемой «по умолчанию». Кодер может изменить элементы таблицы, тогда он сообщает об этом декодеру.

# Структурная схема типичного алгоритма сжатия

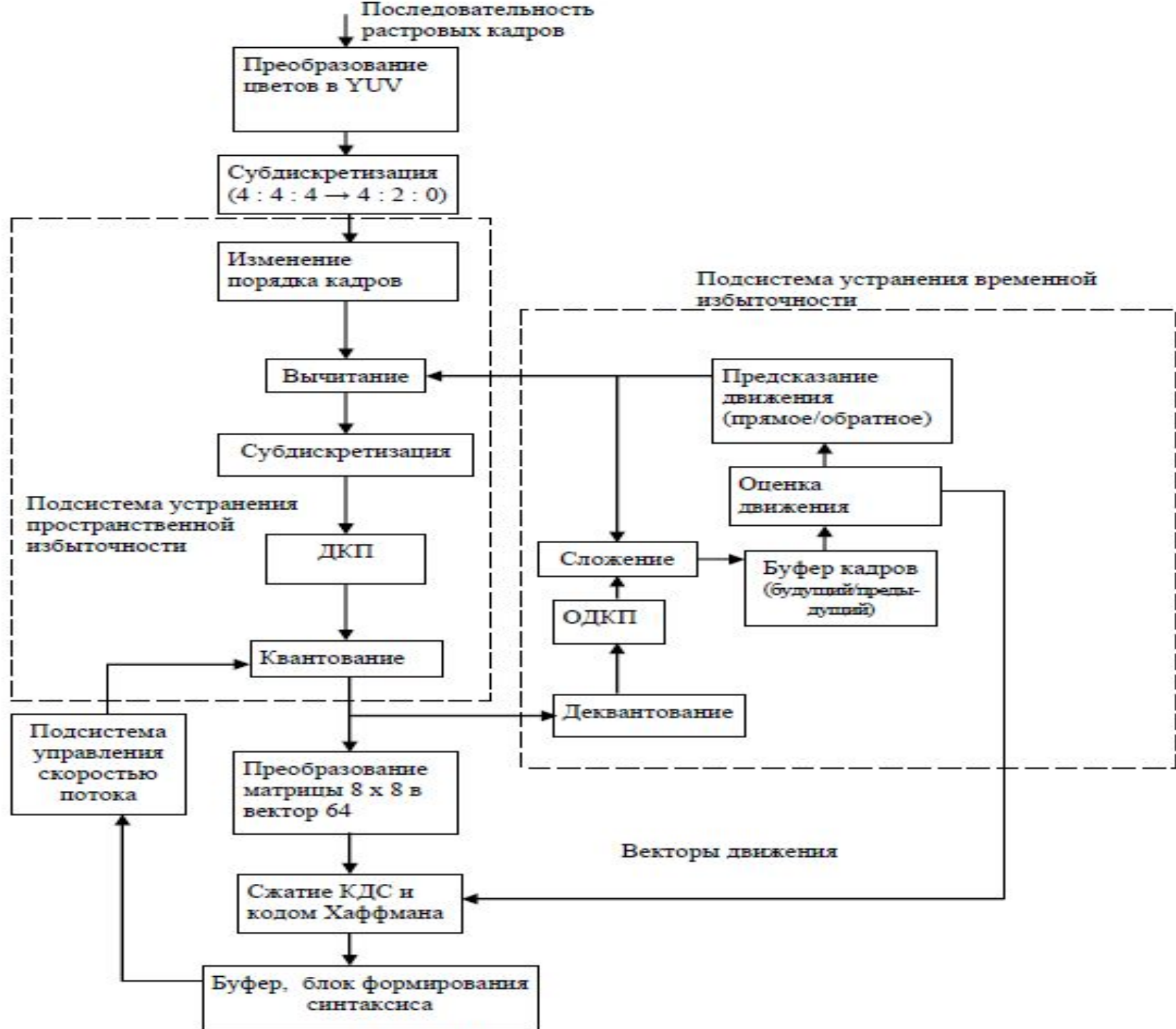
## • пример MPEG

Можно выделить три основные подсистемы кодера

- - подсистема устранения пространственной избыточности
- - подсистема устранения временной избыточности
- - подсистема управления скоростью потока.

- В первой подсистеме (подсистема устранения пространственной избыточности) блок преобразования цветового пространства и субдискретизации (уменьшения разрешения компонентов цветности) предназначен для первичного устранения психовизуальной избыточности кадра.

– Данная подсистема позволяет получить коэффициент сжатия до 10 : 1.





**Во второй подсистеме (подсистема устранения временной избыточности)**

- **устранение временной избыточности производится на основе межкадровой обработки потока,**
- **объединения кадров в группы,**
- **предсказания межкадровых сдвигов**
- **компенсации ошибок предсказания.**

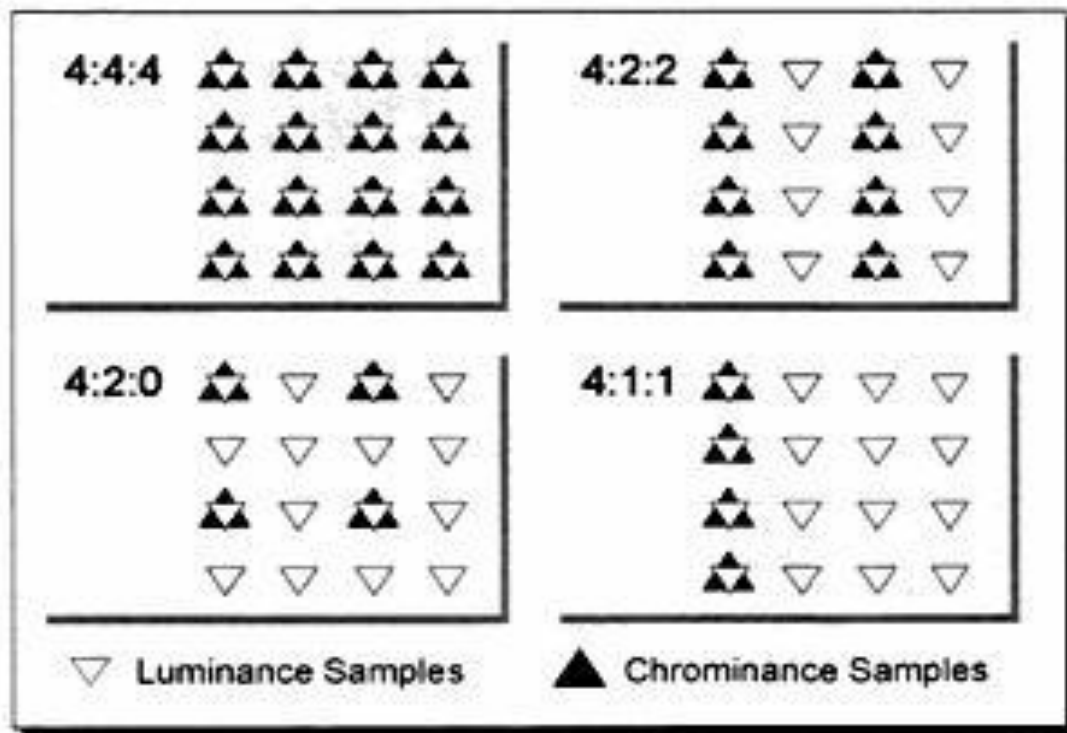
**Данная подсистема позволяет получить коэффициент сжатия до 30 : 1].**

# Преобразование цветового представления растрового изображения

- Цветное изображение содержит как яркостную, так и цветовую информацию. Яркостная информация существенно важнее для качества изображения, чем цветовая, поскольку зрение человека значительно сильнее реагирует на небольшие изменения яркости, чем на небольшие изменения цветового тона

- **RGB не используется для кодирования в телевидении**
- **Одна из этих причин – необходимость совместимости с черно-белыми телевизорами. Сам телевизионный сигнал является составным (композитным) – то есть представляет собой смесь нескольких сигналов. В него входит 3 составляющих:**
  - **Y (luma) сигнал яркости**
  - **U – 1 цветоразностный сигнал**
  - **V – 2 цветоразностный сигнал**
  - **Y – определяет яркость точки, которая может изменяться в диапазоне от чёрной до белой. Таким образом, сигнал яркости полностью формирует чёрно-белое изображение). Сигнал яркости Y в сочетании с цветоразностными сигналами U и V формирует цветное изображение.**
  - **Сигнал яркости Y формируется из RGB сигнала по следующей формуле:**
  - **$Y = 0.299R + 0.587G + 0.114B$**

- U сигнал формируются по формуле:
- $U = R - Y$
- V сигнал формируются по формуле:
- $V = B - Y$
- Для получения RGB представления изображения из YUV :
- $R = Y + U$
- $B = Y + V$
- $G = Y - 0.509U - 0.194V$



Преобразование цветового пространства RGB в цветовое пространство YUV представляется с помощью матрицы перехода в виде:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.229 & 0.587 & 0.144 \\ 0.5 & -0.4187 & -0.0813 \\ 0.1687 & -0.3313 & 0.5 \end{pmatrix} * \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix}$$

Обратное преобразование осуществляется с помощью обратной матрицы перехода в виде:

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1.402 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.772 & 0 \end{pmatrix} * \left( \begin{pmatrix} Y \\ U \\ V \end{pmatrix} - \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix} \right)$$

- Теоретически каждый элемент изображения требует 3 байта. Такое представление, когда и яркость, и компоненты цветности имеют равное число независимых значений, обычно обозначают как схема 4 : 4 : 4.
- Способ сжатия цветовой информации называется субдискретизацией и заключается в объединении цветовой информации для соседних элементов изображений. При использовании данного способа значения яркости  $Y$  запоминаются для каждого элемента изображения.
- Для значений  $U$  и  $V$  вычисляется и запоминается только среднее значение для четырех (схема субдискретизации 4 : 2 : 2), восьми (схема субдискретизации 4 : 1 : 1) и 16 элементов (схема субдискретизации 16:1:1).
- Требуемый объем памяти на каждый пиксель при таких способах составляет соответственно 12, 10 и 9 битов. Таким образом, количество данных на элемент изображения можно сократить с 24 до 12, 10 или 9 битов соответственно.

- На каждый пиксел отводится определенное фиксированное число бит. MPEG-1 использует цветовую схему YCbCr, где Y - это яркостная плоскость, Cb и Cr - плоскости цветовые. Эти плоскости кодируются с разным разрешением. Существуют несколько вариантов кодирования, которые можно представить с следующим виде:

Вариант кодирования	Отношение разрешений Cb/Y (Cr/Y) по горизонтали	Отношение разрешений Cb/Y (Cr/Y) по вертикали
4:4:4	1:1	1:1
4:2:2	1:2	1:1
4:2:0	1:2	1:2
4:1:1	1:4	1:1
4:1:0	1:4	1:4

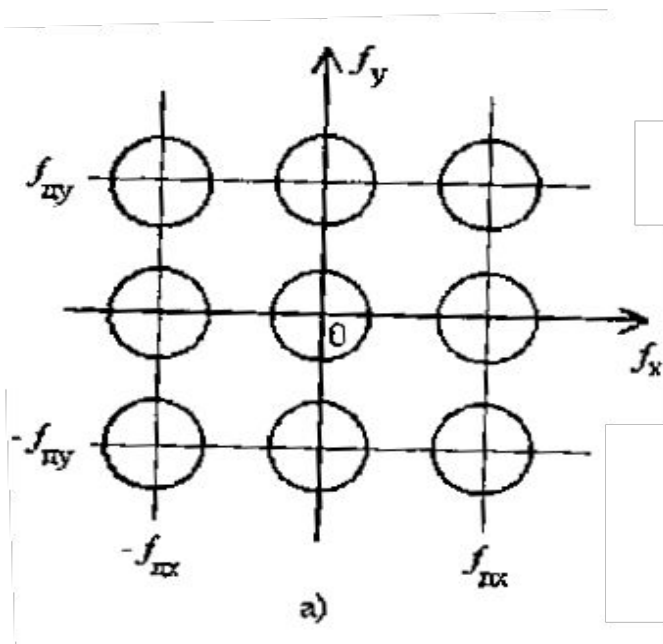
- **Перед началом кодирования происходит анализ видеоинформации, выбираются ключевые кадры, которые не будут изменяться при сжатии, а так же кадры, при кодировании которых часть информации будет удаляться. Всего выделяется три типа кадров:**
- **Кадры типа I - Intra frame. Ключевые кадры, которые сжимаются без изменений.**
- **Кадры типа P - Predirected frame. При кодировании этих кадров часть информации удаляется. При воспроизведении P кадра используется информация от предыдущих I или P кадров.**
- **Кадры типа B - Bidirectional frame. При кодировании этих кадров потери информации еще более значительны. При воспроизведении B кадра используется информация уже от двух предыдущих I или P кадров. Наличие B кадров в видеоролике - тот самый фактор, благодаря которому MPEG-1 имеет высокий коэффициент сжатия (но и не очень высокое качество).**



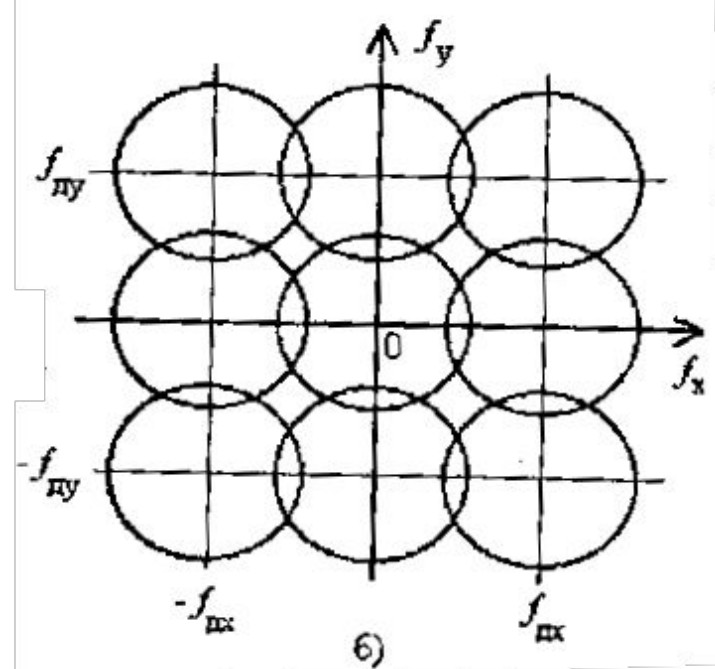
# Подготовка

- I кадры при подготовке к кодированию - кадр разбивается на блоки. В MPEG-1 блоки имеют размер 8x8 пикселей.
- При подготовке кадров типа P и B используется алгоритм предсказания движения.
- входная информация алгоритма предсказания движения получает блок 16x 16 пикселей текущего кадра и аналогичные блоки от предыдущих кадров (I или P типа).
- На выходе данного алгоритма имеем следующую информацию о вышеуказанном блоке:
  1. Вектор движения текущего блока относительно предыдущих
  2. Разницу между текущим и предыдущими блоками, которая собственно и будет подвергаться дальнейшему кодированию.
- Вся избыточная информация подлежит удалению, благодаря чему и достигается столь высокий коэффициент сжатия, невозможный при сжатии без потерь.

- В статических сценах Р и В кадры, по сути, представляют собой копии I кадров, потерь практически нет, но и сжатие информации незначительно. - Качество выше
- При корректном срабатывании алгоритма предсказания движения, объемы кадров разного типа в байтах соотносятся друг с другом примерно следующим образом - I:P:В как 15:5:2.
  
- Далее кодирование – 3 стадии
- Discrete Cosine Transformation - DTC, дискретно- косинусное преобразование, преобразование Фурье.
- Quantization - квантование. Перевод данных из непрерывной формы в прерывистую, дискретную.
- Преобразование полученных блоков данных в последовательность, то есть преобразование из матричной формы в линейную.



**.Пространственные спектры дискретизированного изображения в случае выполнения условий аналога теоремы Котельникова для двумерных сигнал**

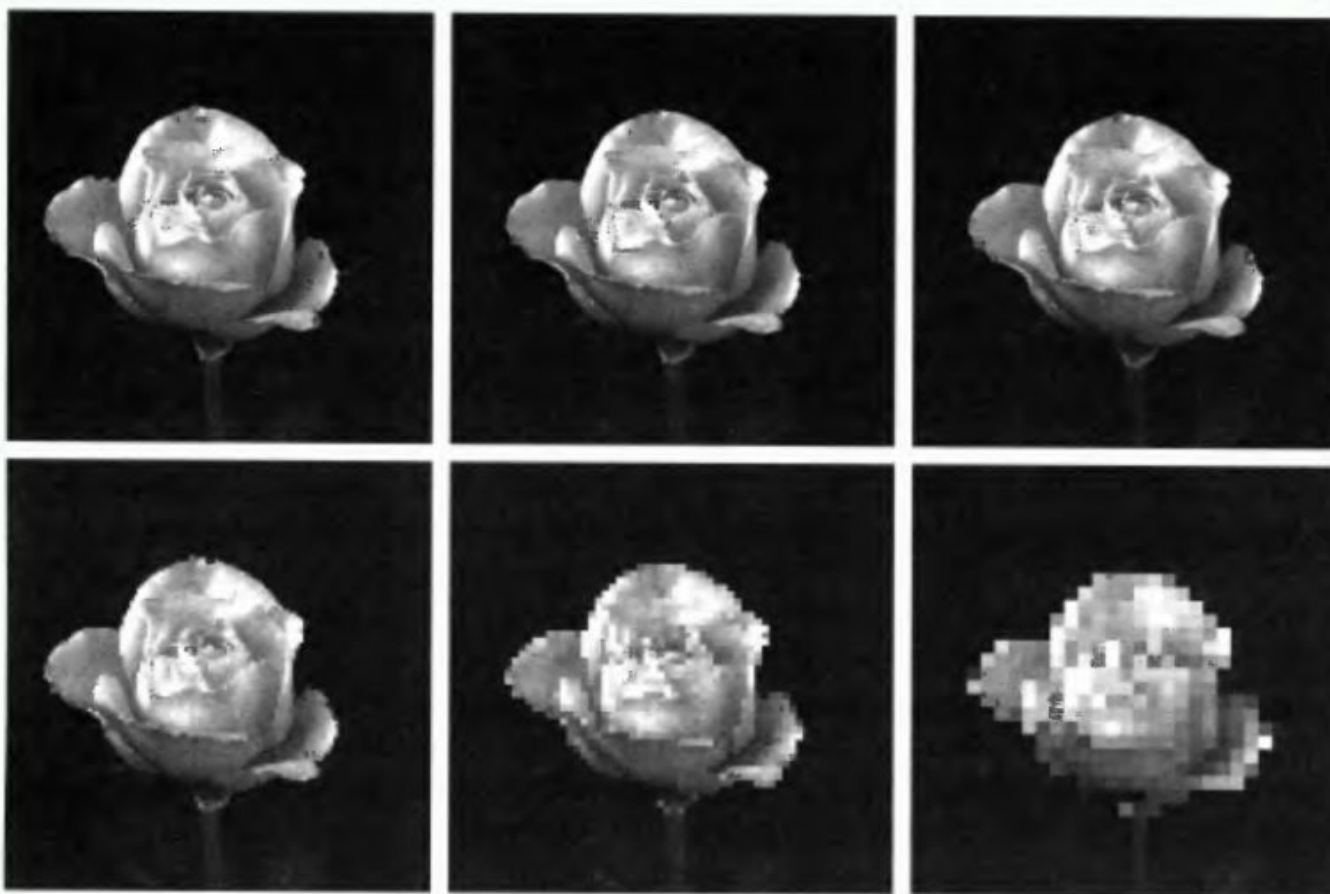


**Пространственные спектры дискретизированного изображения в случае невыполнения условий аналога теоремы Котельникова для двумерных сигналов**

**Когда побочные спектры перекрываются со спектром исходного изображения, что является следствием недостаточно больших частот дискретизации по пространственным координатам.**

**При этом восстановление исходного изображения по дискретизированному без искажений невозможно.**

- Если частота дискретизации изображения недостаточна и возникают перекрытия спектров, то в восстановленном изображении, даже при идеальных характеристиках восстанавливающих фильтров, синтезируются ложные пространственные гармоники, которые приводят к появлению ложных узоров (муар-эффект). Для уменьшения этих нежелательных эффектов используют низкочастотную фильтрацию исходного изображения.



а б в  
г д е

Восьмибитное изображение 1024x1024. (б) 512x512. (в) 256x256. (г) 128x128. (д) 64x64. (е) 32x32. Изображения (б)-(е) увеличены до размеров 1024x1024 дублированием строк и столбцов.

Сравнивая рис. (а) с изображением рис. (б) можно заметить, что их практически невозможно различить. Потеря степени детализации слишком мала, чтобы её можно было увидеть при масштабе рисунка. Следующее изображение на рис. (в) демонстрирует очень слабую ступенчатость на границах между лепестками цветка и черным фоном. Начинает также появляться слегка более выраженная зернистость по всему полю изображения. Эти эффекты становятся ещё более заметными в изображении на рис.10 (г) и совершенно отчетливы в изображениях 64x64 и 32x32.

- **Чтобы искажений, вызванных дискретизацией, не возникало, необходимо выбирать достаточно большие пространственные частоты для дискретизации по обеим координатам.**
- **Однако если рассматривать этот процесс в телевидении (при передаче изображения в цифровой телевизионной системе осуществляется двумерная дискретизация;**
- **при этом дискретизация по вертикальной координате выполняется уже в оптико-электронном преобразователе путём разложения передаваемого кадра на строки, т.е. эта операция имеется уже в аналоговом телевидении;**
- **дискретизация по горизонтальной координате выполняется путём дискретизации видеосигнала во времени), эти пространственные частоты фактически predetermined параметрами, задаваемыми в используемой стандартом разложения, т.е. количеством строк и количеством элементов в каждой строке**

- **Для согласования пространственного спектра изображения с указанными параметрами во многих случаях приходится ограничивать верхние граничные пространственные частоты изображения перед дискретизацией.**
- **Эта операция выполняется с помощью специальных оптических рассеивающих элементов, располагаемых перед ПЗС-матрицей (прибор с зарядовой связью), или просто путём небольшой расфокусировки объектива в телевизионной камере.**

# Дискретно косинусное преобразование.

- Для уменьшения корреляции соседних пикселей применяются различные обратимые преобразования, представляющие исходные данные в виде не коррелированных коэффициентов. ДКП является одним из эффективных преобразований для решения задачи декорреляции и концентрации энергии в спектральных составляющих
- Косинусное преобразование в отличие от преобразования Фурье применяется только для симметричных функций. При использовании частотно-временных преобразований используется понятие периодического расширения функции, заключающееся в следующем:
- Если преобразуется дискретный ряд отсчетов, то его спектр становится периодичным, а в случае преобразования частотного спектра периодически продолжается восстановленный дискретный ряд данных.



- Прямое, двухмерное ДКП блока  $\|A\|$  размером  $M \times N$  исходного изображения определяется следующим образом:

$$B[p, q] = a_p a_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A[m, n] \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}$$

где  $B[p, q]$  – значения спектральных коэффициентов в преобразованном блоке;

Обратное ДКП определяется следующим выражением:

$$A[m, n] = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} a_p a_q B[p, q] \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}$$

- Для стандартного блока 8 x 8, который используется в стандарте MPEG,  $M = N = 8$ , поэтому ДКП имеет вид:

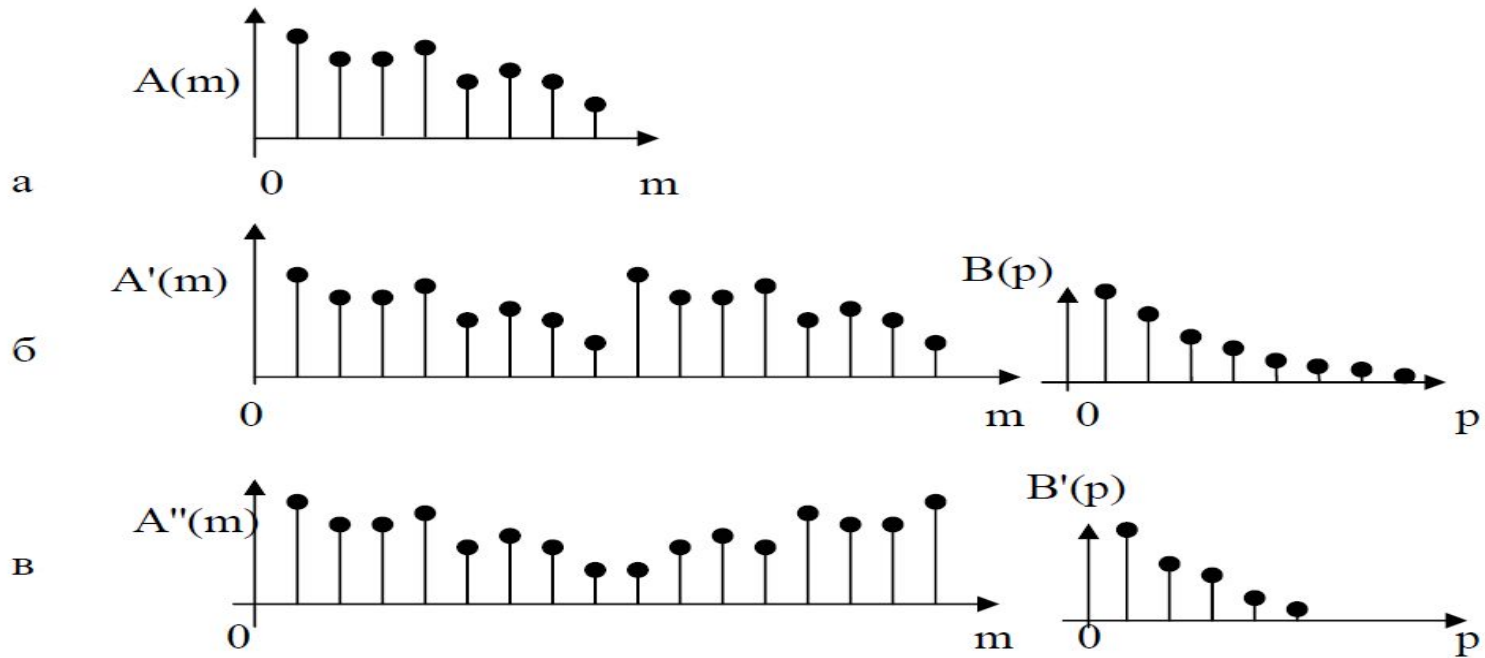
$$B[p, q] = a_p a_q \sum_{m=0}^7 \sum_{n=0}^7 A[m, n] \cos \frac{\pi(2m+1)p}{16} \cos \frac{\pi(2n+1)q}{16}$$

$$a_p = \begin{cases} 1/\sqrt{M}, & p = 0, \\ \sqrt{2/M}, & 1 \leq p \leq M-1; \end{cases} \quad a_q = \begin{cases} 1/\sqrt{N}, & q = 0, \\ \sqrt{2/N}, & 1 \leq q \leq M-1; \end{cases}$$

Матричная форма выражения для  $B[p, q]$   $B = TMT'$

$$T[i, j] = \begin{cases} 1/\sqrt{N}, & i = 0, \\ \sqrt{2/N} \cos\left[\frac{(2j+1)i\pi}{2N}\right], & i > 0; \end{cases}$$

$T[i, j]$  - значения элементов матрицы  $T$ ;  $M$  –сдвинутый блок исходного изображения;  $T'$  – транспонированная матрица  $T$ .



**Сравнение спектральных особенностей ДКП и ДПФ:**

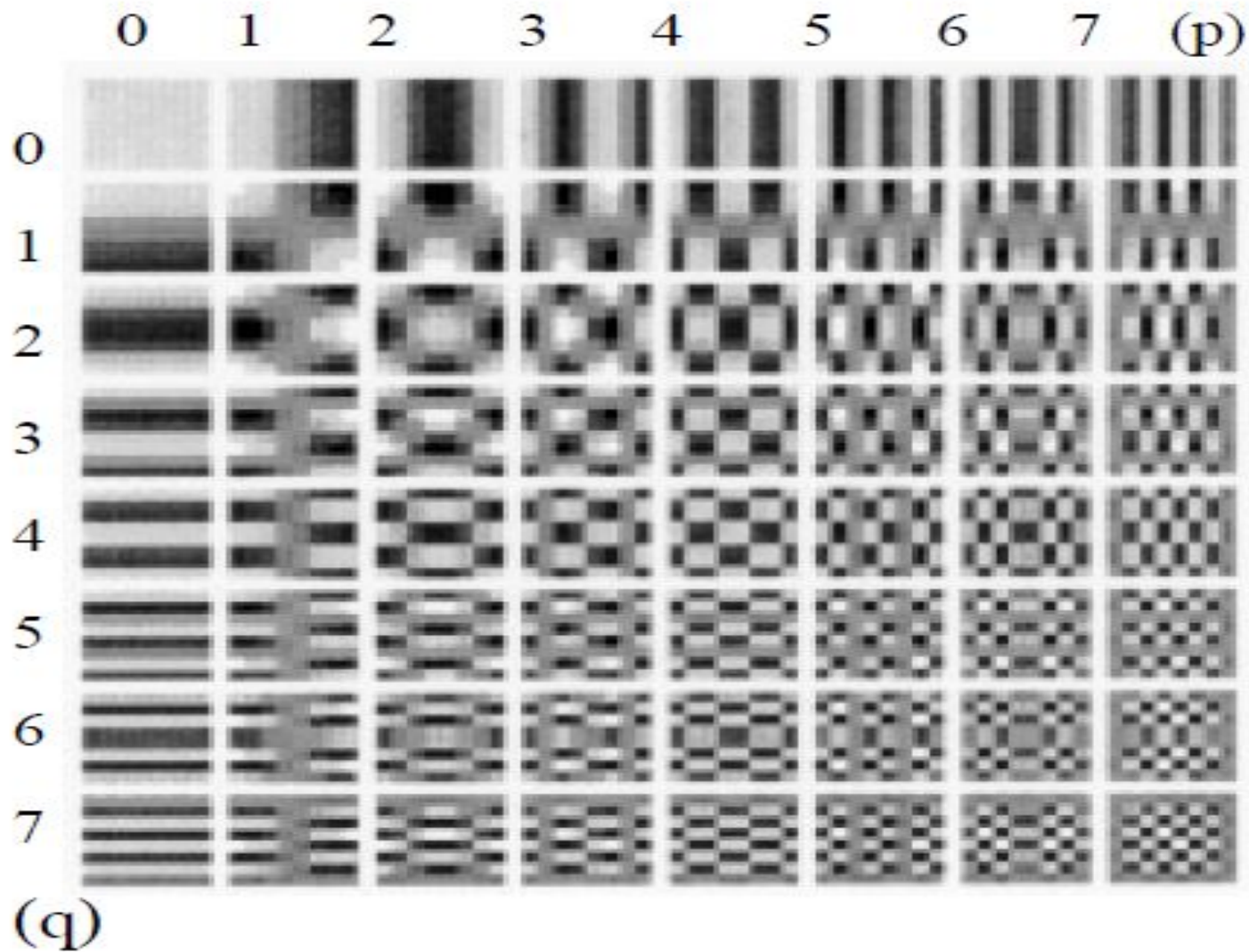
**а – последовательность исходных отсчетов  $A(m)$ ;**

**б – расширение последовательности для преобразования Фурье  $A'(m)$  и ее частотный спектр  $B(p) \sim 1/p$ ;**

**в – расширение последовательности для косинусного преобразования  $A''(m)$  и ее частотный спектр  $B'(p) \sim 1/p^2$ .**

- В результате исходный блок точек преобразуется в матрицу частотных коэффициентов ДКП такого же размера.
- Наиболее важным коэффициентом является коэффициент с координатами  $(0,0)$ , поскольку он представляет собой среднее значение всей матрицы и является постоянной составляющей сигнала DC (Direct Current).
- Все остальные коэффициенты являются переменными составляющими AC (Alternating Current).
- Матрица частотных коэффициентов ДКП не имеет прямой геометрической связи с положением пикселей видеосигнала на растре, а представляет собой форму математической записи, при которой частотные коэффициенты ДКП являются двумерным спектром изображения по горизонтальному и вертикальному направлениях кадра.

- Изображение базисных функций преобразования ДКП представлено на рис



**Набор базисных функций ДКП.**

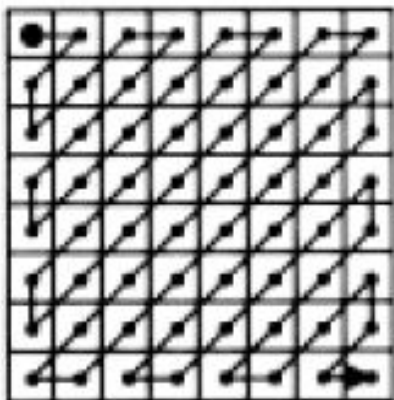
- **Графический смысл этого преобразования заключается в том, что блок изображения рассматривается как суперпозиция изображений синусоидальных колебаний разной частоты. Так, если изображение имеет постоянную яркость, то (для простоты рассматривается черно-белое изображение) его блок коэффициентов будет содержать только один коэффициент DC с координатами (0,0).**
- **Спектр ДКП имеет важную особенность для компрессии видеоданных: основная энергия частотных составляющих этого спектра концентрируется в небольшой области около нулевых частот. Амплитуда высокочастотных составляющих мала или равна нулю.**

- **Часть информации теряется за счет выравнивания сильно выделяющихся участков, которые не подчиняются корреляции.**
- **Алгоритм Quantization - квантование, который формирует Quantization matrix.**
- **Quantization matrix - это матрица квантования, элементами которой являются преобразованные из непрерывной в дискретную форму данные, то есть числа, которые представляют собой значения амплитуды частотных фурье-компонентов.**
- **После формирования quantization matrix происходит разбивка частотных коэффициентов на конкретное число значений. Точность частотных коэффициентов фиксирована и составляет 8 бит.**
- **После квантования многие коэффициенты в матрице обнуляются.**
- **И в качестве завершающей стадии происходит преобразование матрицы в линейную форму.**

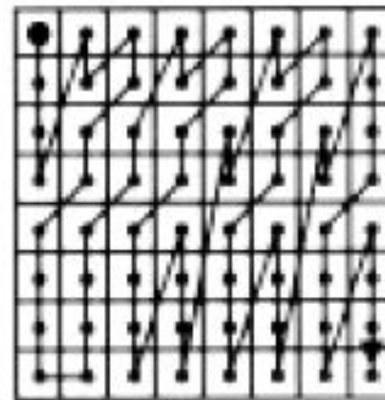
**Важной операцией является перевод матричной формы представления коэффициентов в векторную. Матрица коэффициентов просматривается в определенном порядке с целью получения длинных последовательностей нулей и формирования вектора таким образом, чтобы с увеличением порядкового номера коэффициента возрастала и соответствующая ему частота.**

**Это необходимо для правильного квантования.**





Сканирование  
зигзагом



Альтернативное сканирование

### **Способы формирования вектора коэффициентов.**

Большинство из существующих стандартных алгоритмов видеосжатия, например таких, как MPEG-1, MPEG-2, MPEG-4, H.261, H.263, divX, основаны на разбиении изображения на блоки (обычно размером 8 x 8 пикселей) и использовании ДКП внутри каждого блока.

- · Устраняется временная избыточность видео (учитывается только разностная информация).
- · Устраняется пространственная избыточность изображений путем подавления мелких деталей сцены.
- · Устраняется часть информации о цветности.
- · Повышается информационная плотность результирующего цифрового потока путем выбора оптимального математического кода для его описания.

В третьей подсистеме происходит управление скоростью потока.

- Важным звеном алгоритма кодирования остается обеспечение постоянства выходной скорости цифрового потока. Отсчеты сигнала от источника поступают с постоянной скоростью, однако на выходе квантователя скорость поступления битов может варьироваться в широких пределах. Она зависит от типа видеокadra (*I-кадр* требует значительно больше битов, чем *P- и B-кадры*), от его содержания («спокойный» кадр с однородными участками требует меньше битов, чем кадр с выраженной мелкозернистой структурой).

- Попытка кодировать все кадры одинаковым числом битов приведет к изменению качества изображения от кадра к кадру, а это крайне неприятно для телезрителя.
- Единственным средством выравнивания скорости потока является применение буферной памяти в кодере и декодере. Буфер может заполняться со стороны квантователя неравными порциями, а опустошаться в сторону канала равномерно, обеспечивая постоянную скорость цифрового потока.
- Однако быстрая смена сюжетов и связанное с этим увеличение доли *I*-кадров вскоре может привести к переполнению буфера, а длинный сюжет без движения - к его опустошению. Увеличение размера буфера вызывает возрастание задержки сигнала в кодере, что в некоторых случаях мешает телезрителям .
- В декодере восстанавливается опорный кадр путем выполнения операций, обратных кодированию - восстановления отсчетов, обратного ДКП и восстановления движения по векторам перемещения, т.е. тех же операций, которые осуществляются в кодере

- Для сохранения приемлемого качества изображения при постоянной скорости потока необходимо, во-первых, регулировать шкалу квантования и, во-вторых, адаптивно распределять имеющийся ресурс битов между разными типами кадров с учетом их сложности.
- Адаптивное управление процессом квантования осуществляется умножением всех элементов матрицы на масштабирующий множитель величиной от 8 до 1/4, общий для всех отсчетов макроблока, который также сообщается декодеру.
- Величина множителя больше 1 означает увеличение доли битов, выделяемой данному макроблоку или кадру, меньше 1 - соответствующее уменьшение. Перераспределение ресурса между *I*-, *P*- и *B*-кадрами производится кодером на основе оценки соотношения битов в предыдущих кадрах.
- Некоторые кодеры осуществляют кодирование «в два прохода», оценивая на первом проходе сложность видеокадра, выделяя ему на основе этой оценки определенный ресурс битов и уже на втором проходе кодируя отсчеты с учетом выделенного ресурса.

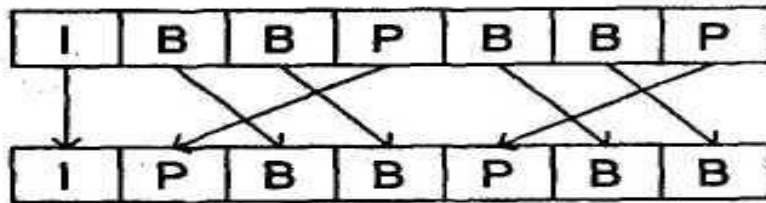
## Кодер MPEG-1 работает следующим образом

- В режиме *I*-кадров осуществляется только пространственное кодирование. Коэффициенты *ДКП* квантуются с переменной длиной шага, зависящей от состояния буфера, кодируются кодом Хаффмана и направляются в буфер, откуда считываются с постоянной скоростью.
- В режиме *с предсказанием* дискретно-косинусному преобразованию подвергается не весь текущий кадр, а только разность между ним и предыдущим, опорным кадром.
- В режиме *двунаправленного предсказания* формируются одновременно два опорных кадра - предыдущий и будущий, для чего используются две независимые схемы формирования.
- Кодер сравнивает предсказанные кадры с текущим кадром, кодирует разности и решает, какое из предсказаний требует меньше битов. Кадр с минимальным числом битов посылается декодеру.

•

# Видеопоследовательность в стандарте *MPEG-1*

- Видеопоследовательность называется любая последовательность видеок кадров произвольной длины. В
- заголовке видеопоследовательности передается информация о скорости цифрового потока, разрешающей способности изображения, формате цветности, формате кадра, частоте кадров и т.п.



Последовательность  
передачи видеок кадров

Видеопоследовательность подразделяется на *группы видеок кадров (ГВК)*, содержащие один *I-кадр* и, возможно, несколько *P-* и *B-кадров*. Расстояние между ближайшими *P-кадрами* обычно обозначается через  $M$ , общее число кадров в группе -  $N$ , группа записывается как  $(M, N)$ .

Типичная длина *ГВК* от 6 до 20 видеок кадров. Для вещания обычно используется  $M = 3$ ,  $N = 12$ . *ГВК* всегда начинается с *I-кадра*, и все предсказанные видеок кадры внутри группы опираются, как правило, на кадры внутри этой же группы.

- Привязка сборки ПЭП (*пакетированные элементарные потоки ПЭП*) к определенному моменту времени в этом стандарте осуществляется периодическим помещением в заголовок ссылок на системные часы (*SCR - System Clock Reference*) - 33-битового числа отсчитывающего периоды опорной частоты 90 кГц. Число периодов частоты 90 кГц подсчитывается многоразрядным счетчиком и вводится в заголовок потока.
- Основная задача, решаемая системными часами, обеспечение синхронизации тактовых частот кодера и декодера. Она решается с помощью схемы дискретной автоподстройки частоты в декодере .
- Привязка к временным меткам позволяет обеспечить одинаковую задержку кодера и декодера для всех составляющих программы при их независимой передаче с разбивкой на пакеты. Таким образом обеспечивается синхронизм элементарных потоков видео и аудио из которых формируется ТВ программа.

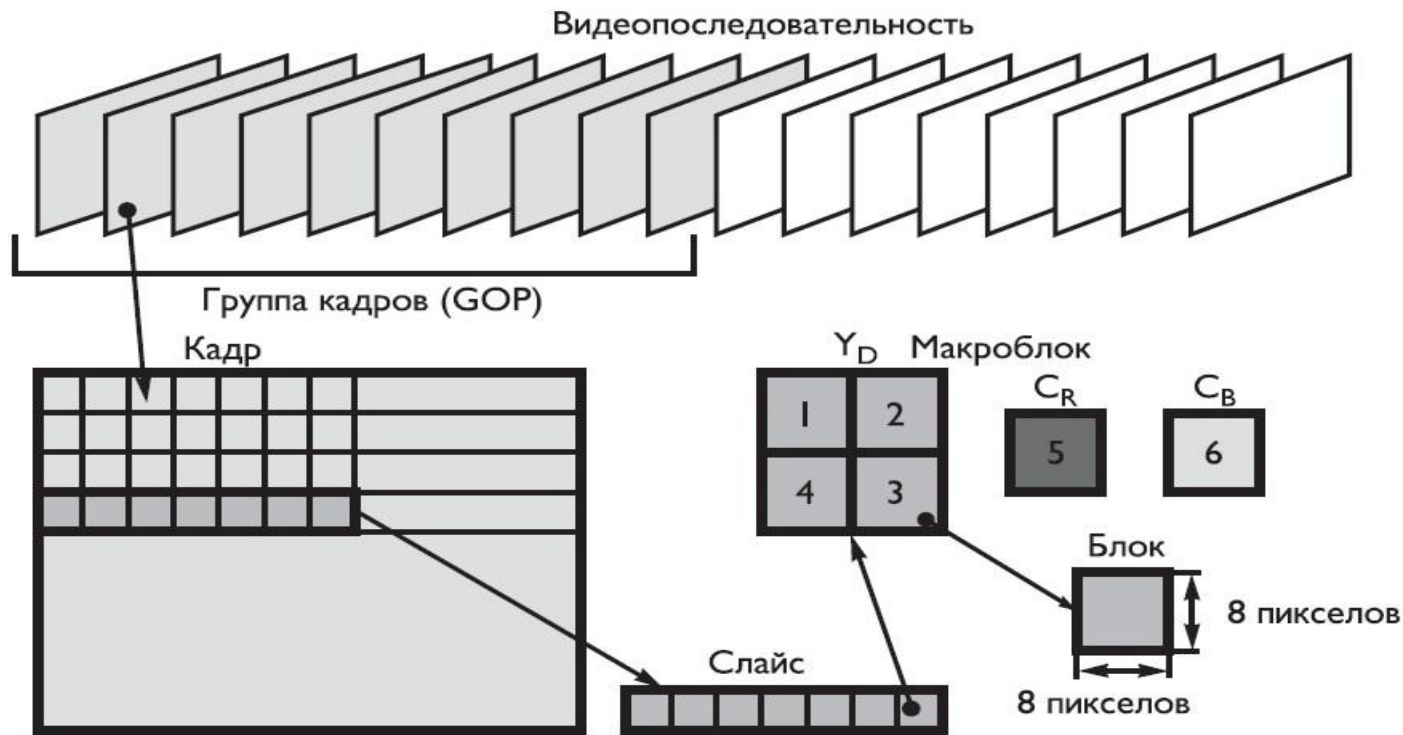
# Синтаксис кодирования

Синтаксис кодированного битового потока содержит шесть уровней:

- – уровень кадров видеопоследовательности,
- – уровень группы кадров,
- – уровень кадров,
- – уровень слайсов,
- - уровень макроблоков,
- – уровень блоков.

Видеопоследовательность разбивается на группы кадров, обработка которых осуществляется совместно. Перед началом кодирования производится анализ видеоинформации, в выбранной группе кадров определяются ключевые кадры - *I*, которые должны подвергаться только внутрикадровой обработке, а также кадры *P*; *B*, при кодировании которых часть информации будет удаляться за счет межкадровой обработки.





## Пример последовательности составляющих динамических изображений

Каждый кадр видеопоследовательности разбивается на макроблоки, в составе которых имеются 4 блока сигнала яркости и по одному блоку двух цветоразностных сигналов. Размер каждого квадратного блока равен  $8 \times 8$  пикселей, а размер макроблока сигнала яркости —  $16 \times 16$  пикселей.

Слайс — группа следующих друг за другом макроблоков. Число макроблоков в слайсе может быть произвольным, но слайсы в изображении не должны перекрываться. В заголовке слайса задается параметр квантования макроблоков, но этот параметр может быть изменен на уровне макроблока.



Восстановление тактовой частоты в приемнике **MPEG-1** по меткам **SCR**

## Выводы

- Для получения приемлемого качества изображения в стандарте MPEG-1 максимальное число элементов изображения было ограничено до 352 x 288 элементов, частота кадров до 30 Гц при прогрессивном разложении. Верхняя граница системного потока данных составляла 1,5 Мбит/с.
- Такие существенные ограничения стандарта не позволили использовать его для целей телевизионного вещания. Не предусматривалась также возможность работы с различными форматами субдискретизации сигналов цветности кроме 4:2:0 и с чересстрочными изображениями.
- MPEG-1 стандартизует запись динамических изображений и сопутствующего звука на компакт-диски (CD-ROM, VideoCD/DVD) и передачу видеоизображений пониженной четкости по сравнительно низкоскоростным каналам связи (скорость цифрового потока до 1,5–3 Мбита/с)
- Стандарт определяет принципы обработки информации и структуру цифрового потока

# Объективные оценки качества

\*Среднеквадратичная погрешность (Mean Square Error):

$$NMSE = \sum_{m=1}^M \sum_{n=1}^N [A[m, n] - A'[m, n]]^2 / \sum_{m=1}^M \sum_{n=1}^N [A[m, n]]^2$$

где  $A'[m, n]$  – значение  $(m, n)$ -го пикселя восстановленного кадра

\*Максимальная среднеквадратичная погрешность (Peak Mean Square Error):

$$PMSE = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N [A[m, n] - A'[m, n]]^2 / [\max \{A(m, n)\}]^2$$

\*Нормированная абсолютная погрешность (Normalized Absolute Error):

$$NAE = \sum_{m=1}^M \sum_{n=1}^N [A[m, n] - A'[m, n]] / \sum_{m=1}^M \sum_{n=1}^N |A[m, n]|$$

нормированная среднеквадратичная погрешность

$$NMSE = \sum_{m=1}^M \sum_{n=1}^N [A[m, n] - A'[m, n]]^2 / \sum_{m=1}^M \sum_{n=1}^N [A[m, n]]^2$$

- отношение сигнал/шум

$$SNR = 10 \log_{10} \left( \sum_{m=1}^M \sum_{n=1}^N (A[m, n])^2 / \sum_{m=1}^M \sum_{n=1}^N (A[m, n] - A'[m, n])^2 \right)$$

Частным случаем критерия качества для полутонного изображения диапазоном значений пикселей от 0 до 255 ( $\max\{A(m, n)\}=255$ ) является отношение пикового сигнала к шуму (Peak-to-Peak Signal-to-Noise Ratio – PSNR):

$$PSNR = 10 \log_{10} \frac{255^2 MN}{\sum_{m=1}^M \sum_{n=1}^N [A[m, n] - A'[m, n]]^2}$$

Данная мера аналогична среднеквадратичному отклонению и ей присущи те же недостатки, что и среднеквадратичному отклонению, однако пользоваться ею удобнее из-за логарифмического масштаба шкалы.

# Субъективные оценки качества

- **Международный союз электросвязи (ITU) . Метод измерения под названием Double Stimulus Impairment Scale ("Шкала деградации с двумя стимулами") заключается в следующем.**
- **Участники просматривают эталонную сцену (часто в форме видео стандарта 601), а затем - ту же сцену после ее обработки в тестируемой системе.**
- **Каждый участник оценивает деградацию изображения сцены после обработки одним числом по пятибалльной шкале:**
  - **1 - очень раздражает;**
  - **2 - раздражает;**
  - **3 - слегка раздражает;**
  - **4 - заметно, но не раздражает;**
  - **5 - незаметно.**
- **Баллы выставляются для некоторого количества различных сцен.**

# Звук в MPEG:

- **Форматы кодирования звука делятся на три части: Layer I, Layer II, Layer III (прообразом для Layer I и Layer II стал стандарт MUSICAM, .**
- **Принципы кодирования основаны на том факте, что человеческое ухо не совершенно и на самом деле в несжатом звуке (CD-audio) передается много избыточной информации.**
- **Принцип сжатия работает на эффектах маскировки некоторых звуков для человека (например, если идет сильный звук на частоте 1000 Гц, то более слабый звук на частоте 1100 Гц уже не будет слышен человеку, также будет ослаблена чувствительность человеческого уха на период в 100 мс после и 5 мс до возникновения сильного звука).**

- **Psycoacoustic (психоакустическая) модель используемая в MPEG разбивает весь частотный спектр на части, в которых уровень звука считается одинаковым, а затем удаляет звуки не воспринимаемые человеком**
- **В Layer III части разбитого спектра самые маленькие, что обеспечивает самое хорошее сжатие. MPEG Audio поддерживает совместимость Layer'ов снизу вверх, т.е. decoder (декодировщик) для Layer II будет также распознавать Layer I.**
- **Синхронизация и объединение звука и видео, осуществляется с помощью System Stream , который включает в себя:**
- **Системный слой, содержащий временную и другую информацию чтобы разделить и синхронизовать видео и аудио.**
- **Компрессионный слой, содержащий видео и аудио потоки.**



- **Видео поток содержит заголовок, затем несколько групп картинок (заголовок и несколько картинок необходимы для того, что бы обеспечить произвольный доступ к картинкам в группе в независимости от их порядка).**
- **Звуковой поток состоит из пакетов каждый из которых состоит из заголовка и нескольких звуковых кадров (audio-frame).**
- **Для синхронизации аудио и видео потоков в системный поток встраивается таймер, работающий с частотой 90 КГц (System Clock Reference -- SCR, метка по которой происходит увеличения временного счетчика в декодере) и Presentation Data Stamp (PDS, метка начала воспроизведения, вставляются в картинку или в звуковой кадр, чтобы объяснить декодеру, когда их воспроизводить.**
- **Размер PDS составляет 33 бита, что обеспечивает возможность представления любого временного цикла длиной до 24 часов).**

# Параметры MPEG-1 в аудио

- Параметры MPEG-1 (Утверждены в 1992)
- Параметры Аудио:
- 48, 44.1, 32 КГц,
- mono, dual (два моно канала),
- стерео,
- интенсивное стерео (объединяются сигналы с частотой выше 2000 Гц.),
- m/s stereo (один канал переносит сумму - другой разницу).
- сжатие
- скорость

Способ кодирования	Скорость передачи kbps (килобит в сек.)	Коэффициент сжатия
Layer I	192	1:4
Layer II	128..96	1:6..8
Layer III	64..56	1:10..12

- **Параметры Видео:** в принципе с помощью MPEG-1 можно передавать с разрешением вплоть до 4095x4095x60 fps (в этих границах кадр может быть произвольного размера).
- Так как существует **Constrained Parameters Bitstream (CPB, неизменяемые параметры потока данных;** другие стандарты для MPEG-1 поддерживаются далеко не всеми декодерами) которые ограничивают общее число макроблоков в картинке (396 для скорости  $\leq 25$  fps и 330 для скорости  $\leq 30$  fps) то MPEG-1 кодируется стандартом SIF /352\*240\*30 - (получено урезанием стандарта CCIR-601) или 352\*288\*25 - (урезанный PAL, SECAM) формат 4:2:0, 1.15 MBPS (мегабит в сек.), 8 bpp (бит на точку) - в каждой плоскости/.

- Существует более высокое разрешение для MPEG-1 - так называемый MPEG-1 Plus, разрешение как у MPEG-2 ML@MP (Main Level, Main Profile) - этот стандарт часто используется в Set-Top-Box для улучшения качества.
- Компрессия по стандарту MPEG-2 кардинально меняет положение вещей. Более 97% цифровых данных, представляющих видео сигнал дублируются, т.е. являются избыточными и могут быть сжаты без ущерба качеству изображения. Алгоритм MPEG-2 анализирует видеоизображение в поисках повторений, называемых избыточностью.
- В результате процесса удаления избыточности, обеспечивается превосходное видеоизображение в формате MPEG-2 при более низкой скорости передачи данных. По этой причине, современные средства поставки видеопрограмм, такие как цифровые спутниковые системы и DVD, используют именно стандарт MPEG-2.



