Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»
Кафедра сетей и устройств телекоммуникаций

СЖАТИЕ И ШИФРОВАНИЕ ВИДЕОДАННЫХ *В ФОРМАТЕ MPEG*

Методические указания
к лабораторной работе по курсу
«Цифровая обработка речи и изображений»
для студентов специальности «Сети телекоммуникаций»
дневной формы обучения

Минск 2004

УДК 621.391.25 (075.8)
ББК 32.811.4 я 73
С33

Составители:
А.А. Борискевич, А.Л. Гурский, Ю.Г. Кочубеев

Сжатие и шифрование видеоданных в формате MPEG: Метод. С33 указ. к лаб. работе по курсу «Цифровая обработка речи и изображений» для студ. спец. «Сети телекоммуникаций» дневной формы обуч. / Сост. А.А. Борискевич, А.Л. Гурский, Ю.Г. Кочубеев. – Мн.: БГУИР, 2004. – 24 с.: ил.
ISBN 985-444-667-0

В данных методических указаниях рассмотрены основные особенности сжатия и шифрования видеоданных в формате MPEG и критерии оценки качества восстановленного изображения для управления соотношением качество/сжатие. Представленные сведения могут быть использованы для решения задач обработки и защиты подвижных изображений.

УДК 621.391.23 (075.8)
ББК 32.811.4 я 73

ISBN 985-444-667-0

ЦЕЛЬ РАБОТЫ

Изучение особенностей сжатия видеоданных в формате MPEG1 и их шифрования для решения задач обработки и комплексной защиты видеоинформации.

1. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

1.1. Введение

Цифровое видео уже давно применяется в различных областях человеческой деятельности. Основные его достоинства – удобство хранения, воспроизведения и редактирования видеоинформации. Кроме того, цифровое видео можно передавать в виде битового потока по цифровым каналам связи, что позволяет избежать ухудшения качества при передаче.

Исходный видеопоток на практике использовать нецелесообразно ввиду большой избыточности. Например, студийный сигнал, известный как цифровое видео "D-1" или "CCIR 601" (858 x 525 точек/кадр x 30 кадров/с), оцифровывается со скоростью 270 Мбит/с. Час видео без звука в таком формате будет занимать более 120 Гбайт. Решением этой проблемы является сжатие видеоданных.

Наиболее распространенным из-за своей универсальности и открытости является семейство алгоритмов сжатия MPEG.

1.2. Особенности сжатия видеоданных в формате MPEG

Рассмотрим структурную схему алгоритма сжатия стандарта MPEG 1 (рис. 1). Можно выделить три основные подсистемы кодера: подсистема устранения пространственной избыточности, подсистема устранения временной избыточности и подсистема управления скоростью потока.

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.

В первой подсистеме блок преобразования цветового пространства и субдискретизации (уменьшения разрешения компонентов цветности) предназначен для первичного устранения психовизуальной избыточности кадра. Блок смены порядка обработки кадров необходим для работы системы компенсации движения, блок дискретного косинусного преобразования – для

получения спектральных коэффициентов и устранения пространственной корреляции пикселей, блок квантования – для устранения вторичной психовизуальной избыточности, блок преобразования матрицы коэффициентов ДКП (дискретного косинусного преобразования) в вектор зигзагообразным сканированием – для увеличения эффективности кода Хаффмана. Данная подсистема позволяет получить коэффициент сжатия до 10 : 1.

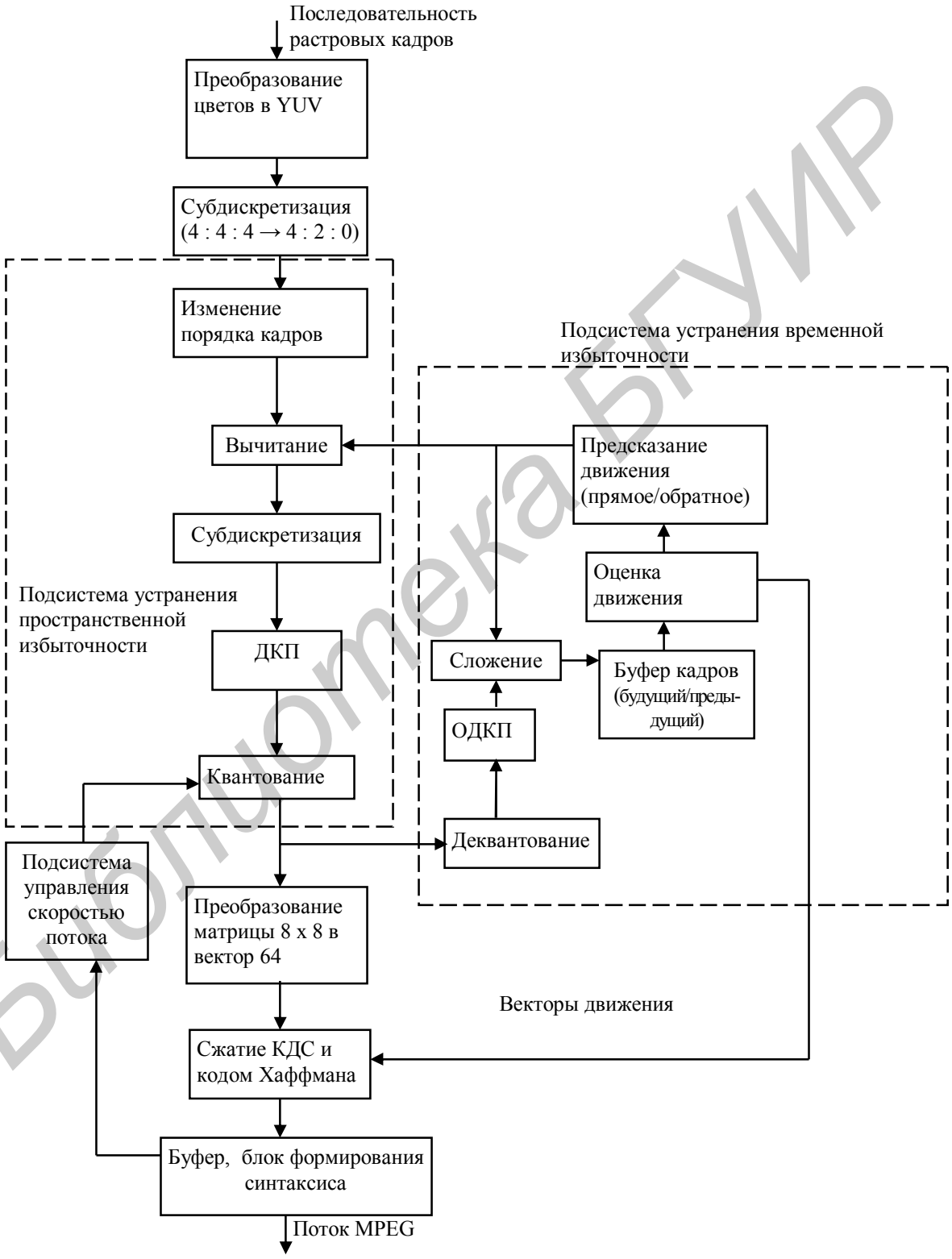


Рис.1. Структурная схема сжатия в стандарте MPEG 1

Во второй подсистеме устранение временной избыточности производится на основе межкадровой обработки потока, объединения кадров в группы, предсказания межкадровых сдвигов и компенсации ошибок предсказания. Блоки деквантования и обратного ДКП (ОДКП) необходимы потому, что в алгоритме компенсации движения совпадающие области ищутся не в оригинальном видеопотоке, а в кадрах, уже подвергшихся сжатию и распаковке, так как у распаковывающего модуля нет доступа к исходным видеокадрам. Данная подсистема позволяет получить коэффициент сжатия до 30 : 1.

Подсистема управления обеспечивает постоянство скорости видеопотока, управляя числом частотных коэффициентов и длиной группы кадров.

Исходные видеоданные представляются в виде потока растровых изображений. Растр – форма представления изображения в виде элементов (пикселей), упорядоченных в строки и столбцы. Название “пиксель” образовано как сокращение от английского pixel (picture cell – ячейка изображения) – это наименьший элемент, из множества которых создается растровое изображение. Несмотря на четкую структуру пиксельной сетки, тоновые переходы в растровых картинках не выглядят дискретными. Это обусловлено, во-первых, тем, что пиксели обычно очень малы и при отображении без увеличения практически неразличимы. Во-вторых, плавные изменения цвета или света передаются за счет постепенного изменения значений соседних пикселей от одного к другому.

При записи изображения обычно используется по 8 бит (1 байт) для представления 256 уровней яркости красного, зеленого и синего цветов (RGB). Таким образом, для хранения одного элемента изображения (пикселя) требуется 3 байта памяти. Например, стандартный видеокадр формата 352 x 288 пикселей требует 304 128 байтов.

При записи изображений традиционно используется RGB-представление, когда на каждую цветовую составляющую приходится по одному байту. Также существуют и другие способы записи изображений.

1.2.1. 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. Это обусловлено тем, что в сетчатке глаза плотность палочек, пригодных для восприятия только яркостной информации, гораздо выше, чем колбочек.

В связи с этим в схеме кодирования MPEG (см.рис.1, блок преобразования в YUV) видеоданные переводятся в YUV-представление, также называемое YCrCb (Y – компонент яркости, U (или Cr) и V (или Cb) – два компонента цветности, где Cr – Chromatic red (хроматический красный), Cb –

Chromatic blue (хроматический синий)). Это необходимо потому, что при представлении изображений в RGB–виде нельзя различить яркостную и цветовую информацию.

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

$$\begin{vmatrix} Y \\ U \\ V \end{vmatrix} = \begin{vmatrix} 0{,}229 & 0{,}587 & 0{,}144 \\ 0{,}5 & -0{,}4187 & -0{,}0813 \\ 0{,}1687 & -0{,}3313 & 0{,}5 \end{vmatrix} \cdot \begin{vmatrix} R \\ G \\ B \end{vmatrix} + \begin{vmatrix} 0 \\ 128 \\ 128 \end{vmatrix}. \quad (1)$$

Обратное преобразование осуществляется с помощью обратной матрицы перехода в виде

$$\begin{vmatrix} R \\ G \\ B \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1{,}402 \\ 1 & -0{,}34414 & -0{,}71414 \\ 1 & 1{,}772 & 0 \end{vmatrix} \cdot \left(\begin{vmatrix} Y \\ U \\ V \end{vmatrix} - \begin{vmatrix} 0 \\ 128 \\ 128 \end{vmatrix} \right). \quad (2)$$

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 пиксел 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.

1.2.2. Дискретное косинусное преобразование

Для уменьшения корреляции соседних пикселей применяются различные обратимые преобразования, представляющие исходные данные в виде некоррелированных 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 В точках нечетного периодического расширения исходного ряда вперед и назад (на стыках

сегментов) имеет место разрыв амплитуд (рис.2, б). В этом случае из-за скачков амплитуды Фурье-спектр *B'(p)* убывает пропорционально *1/p*, где *p* – индекс спектрального коэффициента. Четное расширение приводит к разрыву не амплитуд, а первой производной (рис.2, в), вследствие чего частотный спектр ДКП *B''(p)* убывает пропорционально $1/p^2$. Таким образом, из-за сужения спектра для восстановления сигнала с заданной точностью требуется меньшее число коэффициентов. ДКП представляет изображение в виде набора спектральных компонентов, что позволяет вести дальнейшую обработку изображения, в частности уменьшение точности представления коэффициентов с учетом особенностей визуальной системы человека. Данное преобразование в стандарте MPEG выполняется поблочно.

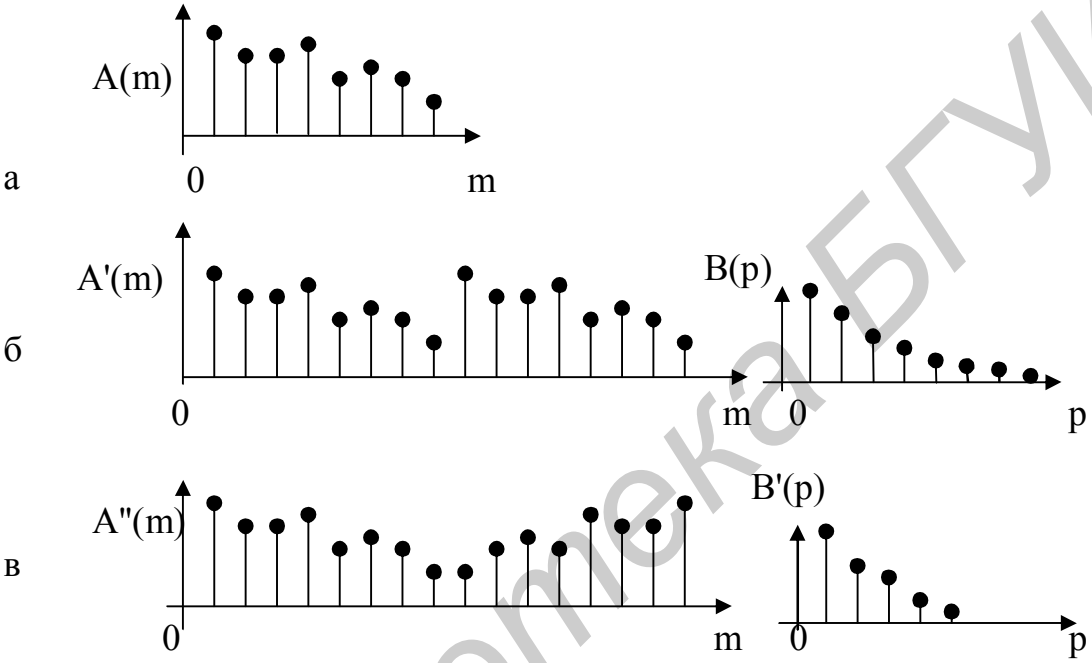


Рис.2. Сравнение спектральных особенностей ДКП и ДПФ:
а – последовательность исходных отсчетов *A(m)*; б – расширение последовательности для преобразования Фурье *A'(m)* и ее частотный спектр *B(p)~1/p*; в – расширение последовательности для косинусного преобразования *A''(m)* и ее частотный спектр $B'(p) \sim 1/p^2$

Прямое двухмерное ДКП блока $\|A\|$ размером *M* x *N* исходного изображения определяется следующим образом:

$$B[p,q] = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A[m,n] \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}, \qquad (3)$$

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

$\alpha_p = \begin{cases} 1/\sqrt{\mathrm{M}}, & \mathrm{p}=0, \\ \sqrt{2/\mathrm{M}}, & 1 \le \mathrm{p} \le \mathrm{M}-1; \end{cases}$ $\alpha_\mathrm{q} = \begin{cases} 1/\sqrt{\mathrm{N}}, & \mathrm{q}=0 \\ \sqrt{2/\mathrm{N}}, & 1 \le \mathrm{q} \le \mathrm{N}-1 \ ; \end{cases}$ *m,n* – координаты

пикселей в исходном блоке изображения; p,q – индексы коэффициентов в преобразованном блоке; $A[m,n]$ – значения пикселей в исходном блоке $\|A\|$.

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).

Обратное ДКП определяется следующим выражением:

$$A[m,n]=\sum_{p=0}^{M-1}\sum_{q=0}^{N-1}\alpha_p\,\alpha_q B\,[p,q]\cos\frac{\pi(2m+1)p}{2M}\cos\frac{\pi(2n+1)q}{2N}. \qquad (4)$$

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

$$B[p,q]=\alpha_p\,\alpha_q\sum_{m=0}^{7}\sum_{n=0}^{7}A[m,n]\cos\frac{\pi(2m+1)p}{16}\cos\frac{\pi(2n+1)q}{16}, \qquad (5)$$

где $\alpha_p=\begin{cases}1/\sqrt{8}, & p=0,\\ 1/2, & 1\le p\le 7;\end{cases}$ $\alpha_q=\begin{cases}1/\sqrt{8}, & q=0,\\ 1/2, & 1\le q\le 7.\end{cases}$

Матричная форма выражения (4) имеет следующий вид:

$$B = TMT', \qquad (6)$$

где $T[i,j]=\begin{cases}1/\sqrt{N} & при\ \ i=0,\\ \sqrt{2/N}\cos\left[\dfrac{(2j+1)i\pi}{2N}\right] & при\ \ i>0\end{cases}$ – значения элементов матрицы T; M – сдвинутый блок исходного изображения; T' – транспонированная матрица T.

В выражении (6) матрица M получена из исходной матрицы путем вычитания 128 из каждого элемента пикселя, так как пиксельные значения черно-белого изображения изменяются от 0 до 255 (чисто черный цвет представляется 0, чисто белый цвет – 255), а ДКП работает со значениями пикселей от –128 до +127.

Восстановленный блок вычисляется по формуле $A' = round\,(T'BT) + 128$, где $round(x)$ – функция округления до целого.

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 двухмерным спектром изображения по горизонтальному и вертикальному направлениях кадра. Изображение базисных функций преобразования ДКП (3) представлено на рис. 3. 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). Если яркость изображения меняется по косинусоиде вдоль горизонтальной оси, то его блок ДКП будет содержать два коэффициента (0,0) и (2,0). Численные значения коэффициентов соответствуют яркостям соответствующих составляющих. Более сложные изображения представляются большим количеством коэффициентов ДКП.

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.

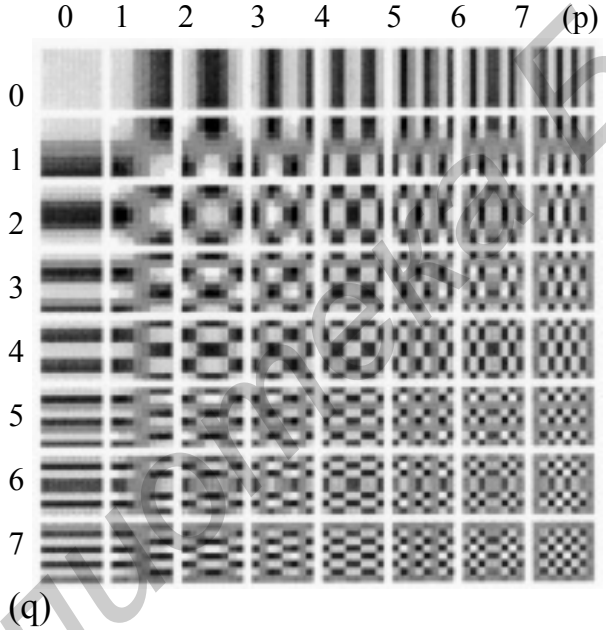
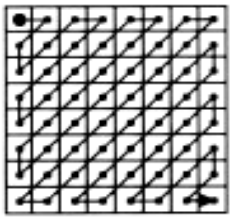


Рис. 3. Набор базисных функций ДКП

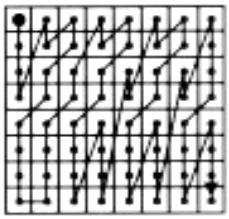
Важной операцией является перевод матричной формы представления коэффициентов в векторную. Матрица коэффициентов просматривается в определенном порядке с целью получения длинных последовательностей нулей и формирования вектора таким образом, чтобы с увеличением порядкового номера коэффициента возрастала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

Наиболее распространенный порядок просмотра матрицы – сканирование зигзагом. В стандарте MPEG предусмотрен еще один вид сканирования – альтернативное сканирование (рис.4). Оно применяется в случаях, когда видеосигнал подается с чересстрочной разверткой – метод развертки, при котором в первом полукадре (поле) воспроизводятся четные строки

видеоизображения, а во втором – нечетные, образуя в результате полное изображение.



Сканирование зигзагом



Альтернативное сканирование

Рис. 4. 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 х 8 пикселей) и использовании ДКП внутри каждого блока. Подобные алгоритмы имеют следующие недостатки:

- эффективно используется корреляция между пикселями внутри одного блока, а корреляция между пикселями соседних блоков практически не учитывается;
- при больших степенях сжатия на изображении становится видна блочная структура, этот эффект называется "блокинг-эффектом";
- энтропийное кодирование спектральных коэффициентов из различных блоков не зависит от информации в соседних блоках.

Вейвлет-сжатие не использует разбиения изображения каждого кадра на блоки и поэтому не подвержено "блокинг-эффекту" даже при очень больших степенях сжатия. Вейвлет-преобразование является последовательной операцией, в которой высокочастотная информация удаляется из изображения шаг за шагом. При больших степенях сжатия искажения изображения локализуются вблизи границ и не распространяются по всему блоку, как это имеет место в случае кодирования с использованием ДКП. Однако отсутствие операции разбиения на блоки не позволяет использовать в вейвлет-алгоритмах применяемый в MPEG метод компенсации движения.

1.2.3. Квантование как средство управления соотношением качество-сжатие

Динамический диапазон коэффициентов ДКП превышает в 8 раз динамический диапазон значений пикселей исходного изображения. Например, при 8-битовом представлении отсчетов изображения его динамический диапазон составляет 256 дискретных уровней (значения от 0 до 255), а динамический диапазон коэффициентов спектра ДКП – 2 048 значений уровней

(от 0 до 2 048 для коэффициентов постоянной составляющей DC и от – 1 023 до + 1 024 для переменных составляющих AC).

Кодирование коэффициентов ДКП в таком широком динамическом интервале потребует в последующих блоках MPEG кодера перехода от 8-битового к 11- битовому коду. Для предотвращения усложнения кодера после ДКП производится сжатие динамического диапазона сигналов коэффициентов ДКП за счет увеличения шага квантования в 8 раз. Эта операция сводится к делению полученных в матрице значений коэффициентов ДКП на 8. Результат деления затем округляется до ближайших целых значений уровней новой шкалы квантования. Например, если исходное значение коэффициента ДКП было 22, то после деления на 8 (22/8 = 2,75) и округления до ближайшего целого значения новое значение будет 3. При этом новый динамический интервал составит 256 дискретных уровней от – 128 до +127.

После выравнивания динамического диапазона над коэффициентами ДКП осуществляется взвешенное квантование для сокращения избыточности в высокочастотной области. Математически данное квантование представляет собой деление рабочей матрицы $\|B\|$ на матрицу квантования $\|Q\|$ поэлементно (6). Для каждого компонента (*Y, U и V*) в общем случае задается своя матрица квантования ($Q_Y[p,q]$, $Q_U[p,q]$, $Q_V[p,q]$), которая позволяет выбирать уровни сжатия и качества изображения:

$$B_Q[p,q] = Round\left(\frac{B[p,q]}{Q_Y[p,q]}\right). \qquad (7)$$

Матрицы квантования в формуле (7) задаются таким образом, чтобы обеспечить более точную передачу низкочастотной информации, так как чувствительность глаз в данной области наибольшая, и обнулить большинство высокочастотных коэффициентов. Это обеспечивает значительное сжатие потока данных, но приводит к снижению разрешения изображения и возможному появлению ложных деталей (в частности, на границах блоков). Чем грубее используемое квантование, тем больше степень сжатия, но ниже качество результирующего сигнала Точность кодирования коэффициентов зависит от шага квантования, который выбирается разным для различных коэффициентов матрицы ДКП.

Коэффициент, соответствующий постоянной составляющей ТВ сигнала, кодируется с использованием 10 бит, потому что при более грубом квантовании соседние блоки начинают отличаться по яркости. На экране это проявляется в виде шахматной структуры.

1.2.4. Структура потока данных MPEG

Поток MPEG разделен на несколько иерархических уровней (рис.5) для улучшения обработки ошибок и упрощения произвольного доступа и редактирования. Каждый слой имеет свой уникальный заголовок длиной в 32 бита.

Первый уровень – это собственно поток видео (Video sequence layer), второй уровень – групповой кадр (ГК, Group of Pictures – GOP), состоящий из нескольких кадров разного типа: *I* кадры (Intracoded – внутрикадровое кодирование) кодируются (сжимаются) без учета соседних кадров; предсказываемые *P* кадры (Predicted – кодирование с предсказанием) кодируются с учетом предыдущего *I* или *P* кадра; *B* кадры двунаправленного предсказания (Bidirectional – двунаправленное предсказание) кодируются с учетом предыдущего и последующего *I* или *P* кадров. Каждая GOP обязательно начинается с I и с определенной периодичностью содержит P-кадры. Ее структуру описывают как M/N, где M - общее число кадров в группе, а N - интервал между P-кадрами. Обычно группа кадров состоит из 12 кадров разного типа.

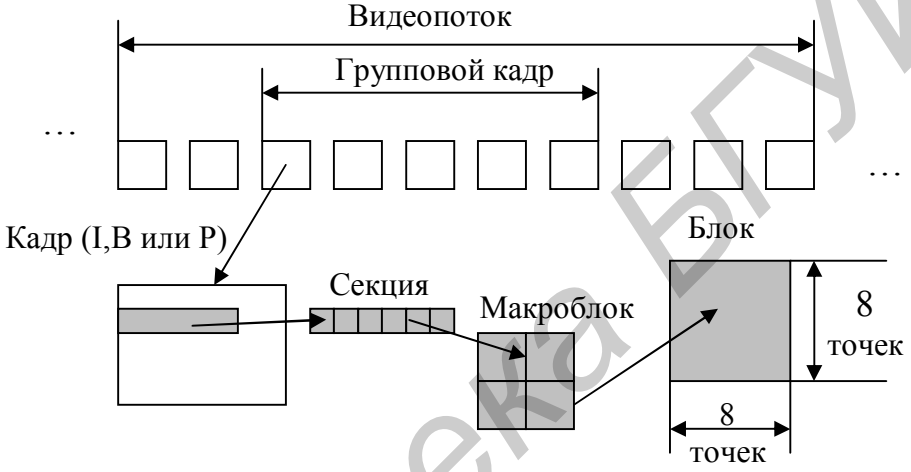


Рис. 5. Структура потока MPEG

Третий уровень потока данных – это слой отдельных кадров (Picture layer) того или иного типа, а четвертый уровень – секционный (Slice layer). Секция (обычно ее ширина равна ширине кадра) состоит из определенного количества макроблоков размером 16 x 16 пикселей. Пятый уровень потока данных – уровень макроблоков. В *I* кадре макроблоки должны быть закодированы как внутренние, т.е. без ссылок на предыдущие или последующие. Макроблоки в *P* кадре могут быть как внутренними *I* блоками, так и использовать данные предыдущих кадров. Алгоритмы кодирования *B* кадров зависят от динамики изображения.

В MPEG-2 предусмотрено пять способов кодирования. Первый – это компенсация движения и предсказание вперед по ближайшим предшествующим *I* или *P* кадрам. При появлении в кодируемом *B* кадре новых объектов применяется предсказание назад по ближайшим последующим *I* или *P* кадрам вместе с компенсацией движения. Третий алгоритм включает в себя компенсацию движения и двунаправленное предсказание по предшествующим и последующим *I* или *P* кадрам. Четвертый основан на внутрикадровом предсказании без компенсации движения (он чаще всего используется при резкой смене плана или высоких скоростях движения отдельных фрагментов картинки). Пятый способ

представления необходим, если в нескольких кадрах изменений нет или они незначительны. В этом случае блок пропускается, а декодер (декомпрессор) при восстановлении использует самый ранний вариант блока. В заголовке каждого макроблока есть элемент, определяющий его тип.

Рассмотрим формат битового потока (рис. 6). MPEG-файл состоит из одной или нескольких видеопоследовательностей. В заголовке видеопоследовательности основными параметрами являются следующие: старт-код, обозначающий начало видеопоследовательности; видеопараметры (ширина, высота, скорость кадров); параметры потока (битовая скорость, размер буфера, флаг стандартизированных параметров, наличие которого означает, что при кодировании применялись стандартные параметры и видео можно декодировать на большинстве декодеров); два типа таблиц квантования для кадров с внутрикадровым кодированием (*I* кадры) и для кадров с межкадровым кодированием (*P* и *B* кадры).

Заголовок группового кадра содержит следующие поля. Временной код – это битовое поле с временным кодом (часы, минуты, секунды, номер кадра). Параметры ГК – это поле, содержащее биты, описывающие структуру ГК.

Заголовок кадра включает следующие поля: тип кадра (*I*, *P* или *B*) и параметры буфера, показывающие, насколько полным должен быть буфер в момент начала декодирования.

Заголовок секции кадра содержит поля, включающие старт-код, отмечающий начало секции, код вертикальной позиции, показывающий, с какой линии начинается эта секция, и масштаб квантования, задающий матрицу квантования в данной секции.

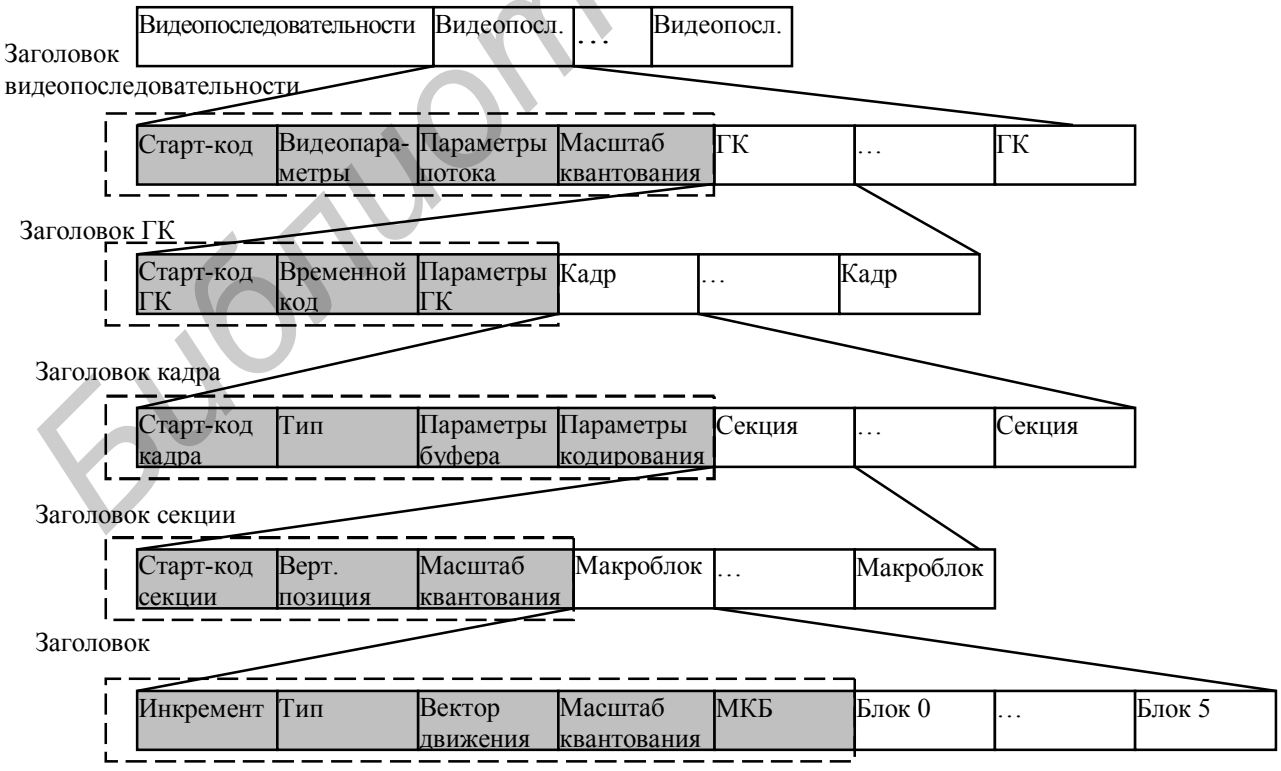


Рис. 6. Структура битового потока данных в формате MPEG

Заголовок макроблока включает поля, содержащие инкремент (количество пропускаемых макроблоков), тип, показывающий, используются ли векторы движения и какого типа, масштаб квантования, задающий матрицу квантования в данном макроблоке, и массив кодированных блоков (МКБ), являющийся битовой картой, показывающей, какой блок закодирован, а какой пропущен.

1.2.5. Межкадровое предсказание движения

Компенсация движения является одной из важнейших составных частей стандартов MPEG 1 и MPEG 2 (рис.7). Метод компенсационного предсказания движения позволяет значительно уменьшить временную избыточность видеопотока. Если следующий кадр содержит сдвинутые части предыдущего кадра, то в этом случае выгодно передавать не весь кадр, а только информацию о движении и изменении сдвинутого пикселя. Ввиду высокой пространственной корреляции достаточно передавать один общий вектор движения для макроблока размером 16 x 16 пикселей.

Для каждого блока первого кадра производится поиск наиболее похожего блока во втором и вычисляется вектор движения, указывающий направление движения блока от первого кадра ко второму. Поправка компенсации движения *E[m,n]* вычисляется вычитанием значений пикселей найденного сдвинутого блока из значений пикселей исходного блока первого кадра:

$$E[m,n] = A_i[m,n] - A_j([m,n] + M_{ij}),$$

где *E[m,n]* – значения ошибок предсказания для каждого пикселя; $A_i[m,n]$ – значения *(m,n)*-го пикселя в блоке *i*-го кадра; $A_j[m,n]$ – значения *(m,n)*-го пикселя в блоке *j*-го кадра; $M_{i,j}$ – вектор движения для макроблока в *i*-м кадре относительно макроблока в *j* - м кадре, состоящий из двух значений, задающих смещение макроблока по вертикали и горизонтали.

Зона поиска должна быть достаточно большой, чтобы быстро движущийся макроблок изображения первого кадра не вышел из зоны поиска второго кадра. Размеры зоны поиска ограничиваются объемом вычислений, которые необходимо выполнить в реальном масштабе времени.

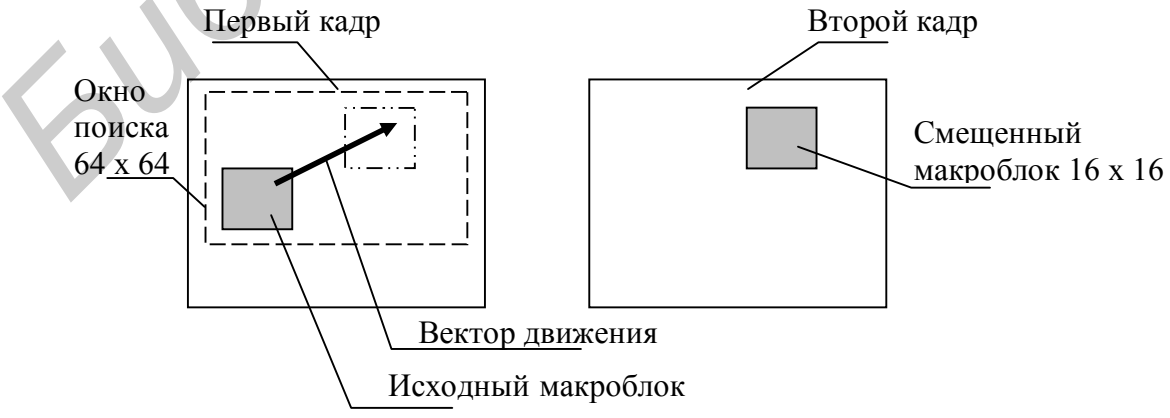


Рис.7. Иллюстрация принципа компенсации движения

Чаще всего размеры зоны поиска задаются в 4 раза больше размеров отдельного макроблока (64 х 64). Например, во время формирования *P* кадра надо определить координаты движения при предсказании вперед (см. рис.7). Для этого берется макроблок текущего *P* кадра и ищется его новое положение в зоне поиска предыдущего *I* или *P* кадра, затем вычисляются межкадровые разности точек. Положение макроблока, при котором суммарное значение модулей межкадровых разностей макроблока получается наименьшим, принимается за его реальное перемещение, после чего координаты вектора движения рассчитываются как смещение макроблока по вертикали и горизонтали относительно его начального положения.

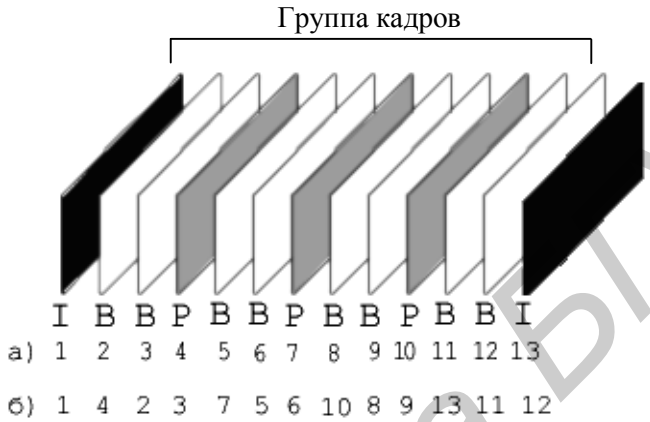


Рис.8. Порядок обработки кадров в группе кадров:
а – порядок передачи и демонстрации кадров, б – порядок декодирования кадров

Для корректного восстановления видеопотока последовательность декодирования кадров (рис.8, б) должна отличаться от последовательности их передачи и показа зрителю (рис.8, а). Так как внутри группы кадров, состоящей обычно из 12 кадров, каждый *B* кадр восстанавливается по окружающим его *P* кадрам (в начале и конце группы – по *I* и *P*), а в свою очередь каждый *P* кадр - по предыдущему *P* (или *I*) кадру, *I* кадры могут быть восстановлены независимо от других, они являются опорными для всех *P* и *B* кадров группы. Соответственно у *I* наименьшая степень компрессии, а у *B* - наибольшая. По размеру типичный *P* кадр составляет 1/3 от *I*, а *B* – 1/8. Два последовательных *B* кадра, формирующиеся по одному алгоритму и использующие одни и те же опорные кадры, не являются одинаковыми, так как представляют разные моменты времени видеопоказа.

1.2.6. Сжатие информации после компенсации движения

Вектор коэффициентов ДКП сжимается с помощью алгоритма кодирования длин серий КДС (RLE – Run Length Encoding). При этом получаются пары типа “пропустить”, “число”, где “пропустить” является счетчиком пропускаемых нулей, а “число” – значение, которое необходимо поставить в следующую ячейку. Так, вектор (42, 3, 0, 0, 0, –2, 0, 0, 0, 0, 1, ...)

будет свернут в пары (0, 42) (0, 3) (3, –2) (4, 1) ... . Для обозначения конца вектора применяется маркер EOB (End Of Block).

Последний из алгоритмов сокращения избыточности связан с кодами переменной длительности. При этом те коэффициенты ДКП, которые повторяются наиболее часто, кодируются короткими кодовыми комбинациями, а редкие значения коэффициентов – более длинными. В MPEG применяется код Хаффмана с фиксированными таблицами, определенными в стандарте. Например, для последовательности (0, 8) (0, 4) (0, 4) (0, 2) (0, 2) (0, 2) (0, 1) (0, 1) (0, 1) (0, 1) (12, 1) (42 нуля) коды Хаффмана будут иметь значения, показанные в таблице.

Пример значений кодов Хаффмана для коэффициентов ДКП

Количество нулей	Коэффициент	Код
-	8 (DC)	110
0	4(AC)	0000
0	4(AC)	0000
0	2(AC)	0100 0
0	2(AC)	0100 0
0	2(AC)	0100 0
0	1(AC)	110
0	1(AC)	110
0	1(AC)	110
0	1(AC)	110
12	1	0010
EOB	EOB	10

Следует отметить, что 12 нулей представлены всего лишь 9 битами, а последние 43 нуля вообще отброшены и заменены двумя битами флага конца блока (EOB – End Of Block). Учитывая, что блок коэффициентов 8 x 8 занимает 512 бит, а после кодирования всего 61, можно говорить о сжатии на этом этапе с коэффициентом примерно 8 : 1.

1.2.7. Управление скоростью и качеством потока

Управление степенью сжатия осуществляется с помощью изменения степени квантования. Для реализации эффективного управления применяется двухпроходное кодирование, при котором каждый кадр просматривается дважды. Во время первого прохода анализируется сложность изображения, чтобы обнаружить момент смены плана, выбрать тип кадра и задать матрицу квантования для наилучшего кодирования данного изображения. При втором проходе создается бинарный поток и выполняется оптимизация распределения бит с учетом заданной скорости потока на выходе компрессора.

Задавая матрицу квантования с большими коэффициентами, можно получить больше нулей и, следовательно, большую степень сжатия. Это

необходимо в тех случаях, когда при кодировании сложной сцены со слабокоррелированными кадрами резко возрастает исходящий поток и переполняется буфер на выходе компрессора. Тогда компрессор увеличивает коэффициент квантования, тем самым уменьшая количество обрабатываемых частотных коэффициентов изображения и соответственно увеличивая степень сжатия. Это делается для того, чтобы у зрителя не возникал неприятный с точки зрения визуального восприятия эффект «зависания» кадров. В стандарте MPEG задается минимальное значение объема памяти для буфера декодера, которое является максимальным для буфера кодера. Это сделано для уменьшения ошибок восстановления потока, вызванных разностями скоростей кодера (компрессора) и декодера (декомпрессора).

1.3. Оценка качества восстановленного изображения

К наиболее употребляемым объективным оценкам качества изображения относятся оценки, основанные на вычислении разности между исходным некомпрессированным изображением и его восстановленной копией. Основными оценками качества восстановления изображения являются следующие:

среднеквадратичная погрешность (Mean Square Error):

$$MSE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]-A'[m,n]\right]^2, \tag{8}$$

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

максимальная среднеквадратичная погрешность (Peak Mean Square Error):

$$PMSE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]-A'[m,n]\right]^2 / \left[\max\left\{A(m,n)\right\}\right]^2; \tag{9}$$

нормированная абсолютная погрешность (Normalized Absolute Error):

$$NAE = \sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]-A'[m,n]\right] / \sum_{m=1}^{M}\sum_{n=1}^{N}\left|A[m,n]\right|; \tag{10}$$

нормированная среднеквадратичная погрешность:

$$NMSE = \sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]-A'[m,n]\right]^2 / \sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]\right]^2 \quad ; \tag{11}$$

отношение сигнал/шум:

$$SNR = 10\log_{10}\left(\sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]\right]^2 / \sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]-A'[m,n]\right]^2\right). \tag{12}$$

Частным случаем критерия качества (12) для полутонового изображения с 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):

$$PSNR = 10\log_{10}\frac{255^2 MN}{\sum_{m=1}^{M}\sum_{n=1}^{N}\left[A[m,n]-A'[m,n]\right]^2}. \quad (13)$$

Один из объективных способов измерения визуальных искажений состоит в оценке MSE или RMSE между исходным и восстановленным изображениями. Параметры MSE и RMSE достаточно объективны и повторяемы. Однако объективных оценок недостаточно, значения MSE и RMSE не показывают, какие ошибки зритель заметит в первую очередь. Мера, которую сейчас чаще используют на практике, - это PSNR. Данная мера аналогична среднеквадратичному отклонению и ей присущи те же недостатки, что и среднеквадратичному отклонению, однако пользоваться ею удобнее из-за логарифмического масштаба шкалы.

1.4. Методы шифрования видеоданных в формате MPEG

В стандарте MPEG не предусмотрено решение проблем защищенности передаваемого видео. В настоящее время все большее применение находят технологии потокового цифрового видеовещания. С развитием Интернета широкое распространение получают такие услуги, как видеоконференции, видео-по-запросу, цифровое вещание по подписке. Все эти услуги требуют защищенности как с точки зрения пользователя, так и поставщика. Например, для услуги видео-по-запросу (VOD – Video On Demand) необходимо обеспечить уверенность в том, что просмотреть требуемый контент (содержание) сможет только тот, кто оплатил данную услугу. Для этого применяются различные методы шифрования видео. В то же время, возможно, что поставщик видеоуслуг захочет, чтобы неподписавшиеся пользователи тоже смогли смотреть его передачи, только в ухудшенном качестве. В этом случае шифровать достаточно не весь видеопоток, а только некоторую часть. При передаче же, например, медицинских или других персональных видеоданных требуется обеспечить более высокий уровень их защищенности.

Наиболее надежное шифрование всего потока можно осуществить стандартным блочным алгоритмом шифрования, например DES (Data Encryption Standard) или AES (Advanced Encryption Standard), или потоковым алгоритмом (например RC4) (рис. 9).

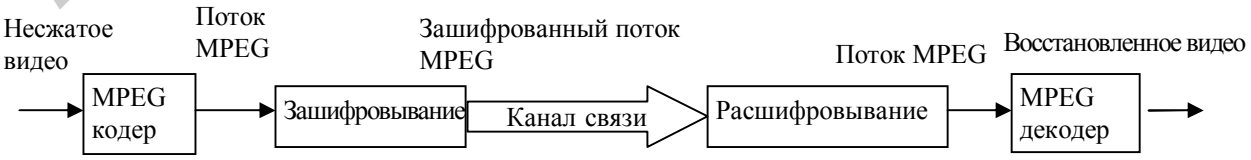


Рис.9. Схема прямого шифрования без учета особенностей MPEG

В этом случае поток MPEG рассматривается как обычный бинарный поток и его специфическая структура не учитывается. Такой метод является наиболее защищенным от криптографических атак, поскольку современные стандартные алгоритмы обеспечивают высокую безопасность. Недостатком данного метода являются высокие требования по производительности в связи с необходимостью шифровать большие объемы информации. Процесс полного шифрования требует высокой производительности сервера, а в сетевых приложениях, когда один сервер обслуживает тысячи клиентов, требования по производительности выполнить практически невозможно, включая аппаратные решения.

Современные пользовательские системы достигли скоростей, достаточных для одновременного полного расшифровывания и декомпрессии потока MPEG, однако производительность мобильных устройств – телефонов, карманных компьютеров (PDA – Personal Digital Assistant) – еще не достаточно велика, поэтому данный метод не всегда является оптимальным.

1.4.1. Метод случайной перестановки коэффициентов ДКП

Данный метод заключается в использовании преобразования блока коэффициентов ДКП 8 х 8 в вектор 1 х 64 в случайном порядке вместо преобразования в «зигзагообразном» порядке. Ключом алгоритма является матрица, представляющая собой набор номеров коэффициентов, задающий последовательность выбора коэффициентов из блока при формировании вектора. Данная матрица формируется с помощью случайных перестановок из исходной матрицы, задающей зигзагообразный порядок выбора коэффициентов. Преимуществом такого метода является высокая скорость шифрования. Схема шифрования коэффициентов ДКП на основе случайного выбора схемы алгоритма представлена на рис. 10.

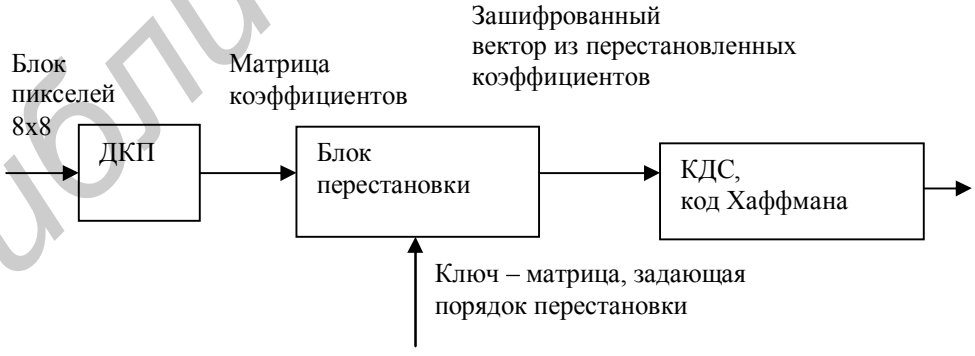


Рис. 10. Схема шифрования коэффициентов ДКП
на основе случайного выбора

Данный алгоритм неустойчив к криптоатакам с использованием как открытого текста, так и только шифротекста. Если криптоаналитик имеет открытый текст и соответствующий ему закрытый шифротекст, то порядок

перестановки можно найти, и криптоаналитик получает доступ к любому потоку, зашифрованному с использованием данной перестановки. При наличии только шифротекста вскрытие также возможно, поскольку коэффициенты, как правило, сосредоточены в верхнем левом углу матрицы, и, зная это, можно найти их нужное местоположение. Для увеличения криптостойкости алгоритма восемь младших бит коэффициента DC разделяют на два числа по четыре бита и второе число записывают в последний, наименее значимый для качества изображения коэффициент AC. Это позволяет скрыть коэффициент DC, иначе его легко обнаружить, поскольку его значение обычно намного больше, чем значения коэффициентов AC. Для дальнейшего повышения криптостойкости может применяться алгоритм, состоящий из группирования коэффициентов DC нескольких последовательных блоков, шифрования их традиционным алгоритмом, например AES, и возврата соответствующих зашифрованных бит обратно в поток. Данный метод не является достаточно криптостойким, поскольку не обладает свойством рассеивания. Кроме того, видеоинформацию можно распознать при задании одного DC для всех блоков потока и правильном восстановлении двух-трех первых коэффициентов AC для каждого блока.

1.4.2. Метод селективного шифрования

Существует несколько криптографических решений, основанных на многоуровневой структуре MPEG, которые выполняют селективное шифрование. Базовый метод селективного шифрования основан на наличии *I,B* и *P* типов кадров в стандарте MPEG (рис. 11). Он заключается в шифровании ключевых *I* кадров, поскольку, теоретически, *P* и *B* кадры бесполезны без соответствующих *I* кадров. При этом шифрованию подвергается около десяти процентов потока, а это снижает требования к вычислительным ресурсам.

Рис. 11. Схема селективного шифрования на уровне кадров MPEG

Данный метод имеет следующие недостатки. В *P* и *B* кадрах часто содержатся *I* макроблоки, что делает видимой довольно большую часть изображения. Кроме того, большая межкадровая корреляция также способствует проявлению части скрытой информации. Таким образом, шифрование только *I* кадров не является достаточным. При шифровании всех *I* макроблоков тоже возникают ряд проблем. Во-первых, идентификация *I* макроблока в потоке MPEG – задача ресурсоемкая, поскольку требуется анализировать поток побитово. Во-вторых, существуют потоки, либо

состоящие только из *I* кадров, либо содержащие количество *I* макроблоков того же порядка, что и количество *I* кадров. В этих случаях шифрование *I* кадров и *I* макроблоков в *P* и *B* кадрах по объему шифруемых данных (до 90 % всего потока) и соответственно по требуемой вычислительной мощности приближается к полному шифрованию.

Существует вариант реализации метода селективного шифрования SECMPEG, не совместимый со стандартным MPEG из-за дополнительной информации в заголовках и требующий поэтому специализированного декомпрессора. SECMPEG использует DES или RSA и позволяет выбрать один из четырех уровней защиты: первый уровень – шифруются все заголовки; второй уровень – шифруются заголовки, коэффициент DC и нижние коэффициенты AC в *I* кадрах; третий уровень – шифруются *I* кадры и *I* макроблоки в *P* и *B* кадрах; четвертый уровень – полное шифрование потока.

1.4.3. Метод шифрования на основе изменяемых кодовых таблиц

Методы селективного шифрования обладают недостатком: поскольку шифрование происходит до сжатия, то оно может увеличивать размер сжатого видео. Для устранения этого недостатка используется схема шифрования со сжатием, осуществляющая одновременно шифрование и сжатие (рис.12).

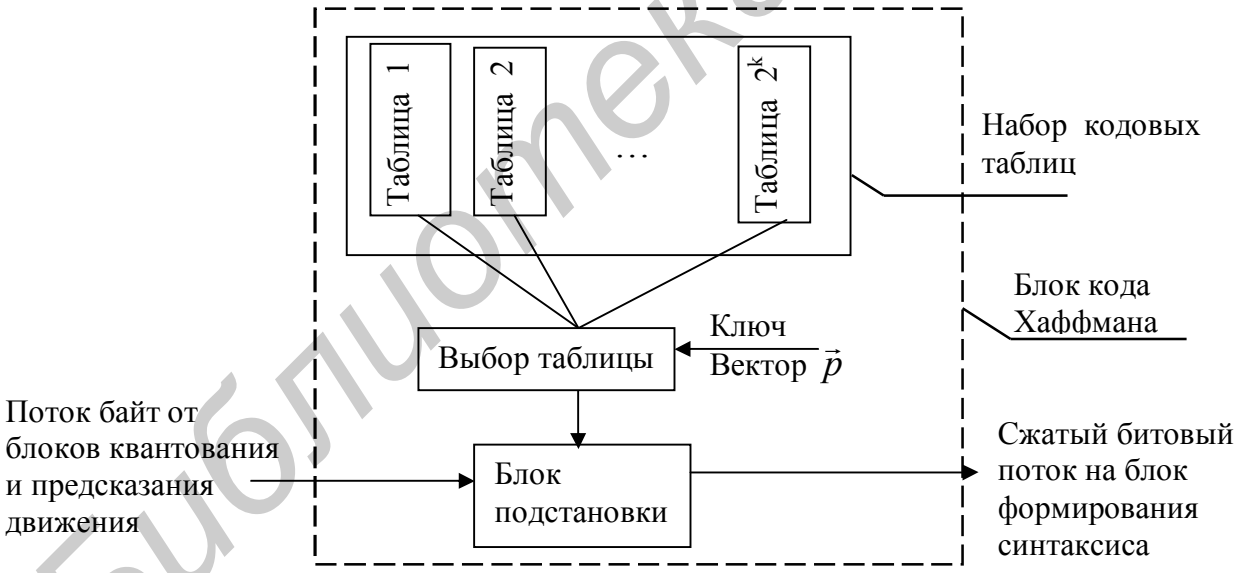


Рис. 12. Схема шифрования с изменяемыми кодовыми таблицами на этапе кодирования по Хаффману

Алгоритм заключается в создании 2^k таблиц Хаффмана и задании ключа в виде вектора $\vec{p} = (p_1..p_i..p_n)$, где p_i – число длиной k бит, являющееся номером одной таблицы из $n = 2^k$. Для каждого j-го входящего символа (байта) используется p_m–я таблица для кодирования и шифрования, где

$m = ((i-1) mod\ n)+1$, т.е. для каждого поступающего на вход кодера Хаффмана байта в соответствии с i-м элементом вектора $\vec{p}$ выбирается кодовая таблица, а затем из нее – кодовое слово, соответствующее данному байту.

В данной схеме ключом является набор таблиц и управляющий вектор $\vec{p}$. Быстродействие этого алгоритма очень высокое и практически не влияет на скорость работы MPEG кодека. Данный алгоритм небезопасный, так как кодирование Хаффмана – это простой подстановочный шифр с символами переменной длины, не обладающий свойством рассеивания. Поэтому данный шифр чувствителен к криптоатаке с выбранным открытым текстом, которая позволяет за некоторое конечное число попыток найти все кодовые таблицы. Вскрытие с известным открытым текстом затруднено из-за сложности для криптоаналитика синхронизировать открытый текст и шифротекст.

2. ЛАБОРАТОРНОЕ ЗАДАНИЕ

2.1. Изучите теоретическую часть.

2.2. Запустите Windows – приложение vc.exe.

2.3. Изучите режим сжатия MPEG.

2.3.1. Выберите тестовую видеопоследовательность test.avi.

2.3.2. Выберите параметры сжатия MPEG (разрешение, битовую скорость, компенсацию движения).

2.3.3. Запустите процесс компрессии и проанализируйте его результаты.

2.3.4. Оцените качество восстановленного изображения для каждого варианта преобразования матрицы коэффициентов ДКП (все AC=0; кроме первых трех AC; все AC=0; все ненулевые AC случайны).

2.4. Изучите режим шифрования MPEG.

2.4.1. Выберите тип шифрования (случайная перестановка коэффициентов ДКП, селективное шифрование).

2.4.2. Оцените уровень защищенности и проанализируйте оценки качества восстановленного изображения.

2.5. Оформите отчет и сделайте выводы.

3. СОДЕРЖАНИЕ ОТЧЕТА

3.1. Результаты выполнения работы.

3.2. Анализ результатов и выводы.

4. КОНТРОЛЬНЫЕ ВОПРОСЫ

4.1. Какими видами избыточностей обладает видеоинформация? Какие из них устраняются в стандарте MPEG?

4.2. Чем обусловлена возможность сжимать массивы для цветовых компонентов *Cr* и *Cb* с большим, чем для яркостного компонента, коэффициентом сжатия?

4.3. Чем объясняется использование ДКП в стандарте MPEG? Какой вид избыточности удаляется с помощью ДКП?

4.4. В чем состоит сущность принципа компенсации движения? Какой вид избыточности удаляется с помощью компенсации движения?

4.5. Каким образом происходит управление соотношением качество/сжатие в стандарте MPEG? Какие коэффициенты сжатия можно получить на разных стадиях сжатия в стандарте MPEG?

4.6. Перечислите требования, предъявляемые к выбору метода шифрования видео.

4.7. В чем сущность метода случайной перестановки коэффициентов ДКП?

4.8. В чем состоят достоинства и недостатки метода шифрования на основе случайной перестановки коэффициентов ДКП, метода селективного шифрования и метода шифрования с изменяемыми кодовыми таблицами?

4.9. Перечислите основные критерии качества изображения и их особенности. Для чего используют оценки качества восстановленного изображения?

ЛИТЕРАТУРА

1. Лаврус В.С. Практика измерений в телевизионной технике. – М.: Солон, 1996.

2. Красильников Н.Н. Теория передачи и воспроизведения изображений.– М.: Радио и связь, 1986.

3. Брайс Р. Справочник по цифровому телевидению. – М.: Изд-во «Эра», 1998.

4. Скляр Б. Цифровая связь. – М: Вильямс, 2003.

5. Lintian Qiao and Klara Nahrstedt. Comparison of MPEG Ecryption Algorithms // International Journal on Computers and Graphics, Special Issue: Data Security in Image Communication and Network, 1998, vol. 22.

6. Chung-Ping Wu, C. Jay Kuo. Efficient Multimedia Encryption via Entropy Codec Design // SPIE International Symposium on Electronic Imaging 2001 (San Jose, CA, USA), Proceedings of SPIE, Jan. 2001, vol. 4314.

Св. план 2004, поз. 106

Учебное издание

СЖАТИЕ И ШИФРОВАНИЕ ВИДЕОДАННЫХ В ФОРМАТЕ MPEG

Методические указания
к лабораторной работе по курсу
«Цифровая обработка речи и изображений»
для студентов специальности «Сети телекоммуникаций»
дневной формы обучения

Составители:
Борискевич Анатолий Антонович,
Гурский Александр Леонидович,
Кочубеев Юрий Георгиевич

Редактор Н.А. Бебель
Корректор Е.Н. Батурчик

Подписано в печать 23.06.2004.		Формат 60×84 1/16.
Бумага офсетная.	Печать ризографическая	Усл. печ.л. 1,63.
Уч.-изд.л. 1,3.	Тираж 70 экз.	Заказ 96.

Издатель и полиграфическое исполнение: Учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники»
Лицензия на осуществление издательской деятельности № 02330/0056964 от 01.04. 2004.
Лицензия на осуществление полиграфической деятельности № 02330/0133108 от 30.04. 2004.
220013, Минск, П. Бровки, 6.