



Министерство образования и науки, молодежи и спорта Украины
Севастопольский национальный технический
университет

КОНСПЕКТ ЛЕКЦИЙ

по дисциплине

«Системы и сети телевидения»

(Цифровое телевидение, часть 1)

для студентов специальности 7.050901 — «Радиотехника»

дневной и заочной форм обучения

Севастополь

2012

УДК 621.397

Конспект лекций по дисциплине «Системы и сети телевидения. Цифровое телевидение, ч. 1» для студентов специальности 7.050901 «Радиотехника» дневной и заочной форм обучения / Ю.П. Михайлюк. — Севастополь: Изд-во СевНТУ, 2012. — 128 с.

Целью данной методической разработки является оказание помощи студентам очной и заочной форм обучения в изучении дисциплины «Системы и сети телевидения».

Рассмотрено и утверждено на заседании научно-методического семинара кафедры радиотехники и телекоммуникаций (протокол № 12 от 2 июля 2011 г.)

Допущено учебно-методическим центром СевНТУ в качестве методических указаний.

Рецензент:

доцент кафедры РТ, канд. техн. наук, доцент Слёзкин В.Г.

Ответственный за выпуск:

заведующий кафедрой РТ, д-р техн. наук, проф. Гимпилевич Ю.Б.

СОДЕРЖАНИЕ

Введение	5
1. Преобразование аналогового сигнала в цифровой	6
1.1. Основные понятия	6
1.2. Импульсно-кодовая модуляция	12
1.3. Обобщенная структурная схема системы цифрового телевидения	17
1.4. Дискретизация телевизионного сигнала	18
1.5. Квантование телевизионного сигнала	23
1.6. Формирователи цифровых телевизионных сигналов	30
1.7. Цифровые интерфейсы в телевидении	32
1.8. Кодирование телевизионного сигнала	41
1.9. Дискретно-косинусное преобразование	44
1.10. Квантование коэффициентов ДКП	48
1.11. Кодирование коэффициентов ДКП	50
1.12. Кодирование с предсказанием	51
1.13. Предфильтрация и шумоподавление	59
2. Стандарт сжатия движущихся изображений и звукового сопровождения	
<i>MPEG-2</i>	61
2.1 Общие сведения	61
2.2. Кодлируемые кадры	69
2.3. Компенсация движения	72
2.4. Использование ДКП в стандарте кодирования <i>MPEG-2</i>	75
2.5. Профили и уровни стандарта <i>MPEG-2</i>	84
2.6. Сжатие звуковых данных	93
2.7. Алгоритмы сжатия звукоданных кодерами первого и второго уровней	99
2.8. Алгоритм сжатия звуковых данных кодерами третьего уровня	103

3. Формирование цифровых потоков видео и звука	108
3.1. Мультиплексирование, ремultipлексирование и статистическое мультиплексирование	115
3.2. Таблицы программно-зависимой информации	117
3.3. Транспортировка пакетов <i>MPEG-2</i> в составе ячеек <i>ATM</i>	122
3.4. Принципы обеспечения синхронизации цифровых потоков в стандарте <i>MPEG-2</i>	127
Библиографический список	129

ВВЕДЕНИЕ

Быстрый переход от аналогового к цифровому телевизионному вещанию, обладающему важными преимуществами, приводит к существенному изменению технической базы телевидения, внедрению принципиально новых, цифровых технологий и методов формирования и передачи ТВ сигналов, основанных на методах цифрового сжатия. Этот переход требует не только замены значительной части парка оборудования, но и подготовки большого числа специалистов, работающих в области ТВ вещания.

С момента возникновения цифрового телевизионного вещания как самостоятельной отрасли в телевизионной технологии и до настоящего времени опубликовано множество теоретических и прикладных работ, в том числе известные книги, как монография под ред. М.И. Кривошеева [1], книга М. Птачека [2], а также работы Г.В. Мамчева, в частности, [3].

К сожалению, эти книги издавались давно и стали незначительными тиражами и стали недоступными для большинства студентов, особенно обучающимися по системе заочного обучения. Конспект является попыткой частично восполнить недостаток учебной литературы в области цифрового телевидения. При изложении материала уделяется особое внимание физической сущности процессов, происходящих в трактах цифрового телевидения.

Материал учебного пособия сгруппирован в трех частях, в которых рассматриваются:

- преобразование аналогового сигнала в цифровой;
- стандарт сжатия движущихся изображений и звукового сопровождения *MPEG-2*;
- формирование цифровых потоков видео и звука.

1. ПРЕОБРАЗОВАНИЕ АНАЛОГОВОГО СИГНАЛА В ЦИФРОВОЙ

1.1. Основные понятия

Прежде чем рассматривать методы преобразования видеосигнала в цифровую форму, приведем вкратце основные сведения о телевизионном изображении и аналоговом видеосигнале.

Телевизионным изображением называют отображение пространственных и временных изменений яркости, цвета и других физических параметров исходного изображения на конечной площади, обычно прямоугольной. В общем случае изображение можно представить в трехмерной системе координат (рисунок 1.1), где функции пространственных координат x , y описывают изменения параметров в плоскости экрана в фиксированный момент времени, а временная ось t отображает процесс во времени.

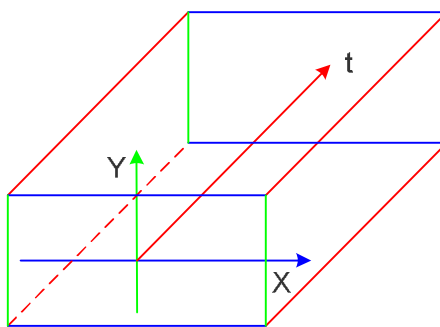


Рисунок 1.1 — Система координат ТВ изображения

Точка изображения, характеризующаяся определенным набором координат (x, y, t) , вместе с ее малой окрестностью называется элементом изображения, или отсчетом. Иногда по аналогии с компьютерной техникой ее еще называют *пикселем* (*pixel* — *picture element*). Понятие пикселя относится к дискретизированным областям пространства и здесь, строго говоря, неприменимо. Тем не менее, с определенными оговорками им пользуются и в телевизионной технике.

Для преобразования изображения в одномерную функцию времени применяют развертку — быстрое сканирование (обычно по горизонтали слева направо со смещением по вертикали сверху вниз) электронным лучом плоскости изображения за время одного кадра ($1/25$ — $1/30$ с), меньшее периода мерцания, различимого глазом. Развертка может быть прогрессивной, когда строки сканируются подряд, или чересстрочной, когда сканируются сначала нечетные, затем четные строки, образуя два полукадра, называемые полями. Развертка дискретизирует изображение только в вертикальном направлении, в горизонтальном сигнал остается аналоговым. Получившийся видеосигнал описывает изменение какого-либо параметра, например, яркости изображения в зависимости от времени. Для передачи по вещательным каналам к нему добавляют сигналы цветности, синхронизации, вводят звуковое сопровождение, телетекст и т.д. Чтобы телевизионный приемник смог правильно воспринять эти сигналы и преобразовать их в изображение и звук, все параметры сигналов должны быть унифицированы, стандартизованы. Говоря о ТВ стандартах, стоит сказать несколько слов об органах международной стандартизации в области телевидения.

Стандартизацией занимается целый ряд международных организаций, области деятельности которых частично перекрываются. Наиболее общие стандарты, охватывающие широкие области применения, разрабатывает Международная организация стандартизации: МОС (ISO), зачастую совместно со своими органами — Международной электротехнической комиссией — МЭК (IEC), Международной светотехнической комиссией — МСК (CIE) и другими. В частности, МОС разработала и приняла стандарты цифрового сжатия телевизионных сигналов семейства MPEG.

Регламентацию международной деятельности в области электросвязи осуществляет Международный союз электросвязи — МСЭ (ITU) через свои подразделения: сектор радиосвязи МСЭ-Р, сектор стандартизации связи МСЭ-Т и др. МСЭ принимает *Рекомендации*, являющиеся обязательными документами для его членов.

Большая работа по стандартизации ведется на региональном уровне. В Европе этим занимается Европейский союз вещания (ЕСВ), который выпускает Технические Рекомендации, обязательные для его членов. В 1993 году ЕСВ слился с Международной организацией по телевидению и радиовещанию восточноевропейских стран ОИРТ, и теперь Россия также является членом союза. ЕСВ тесно сотрудничает с Европейским институтом стандартов электросвязи (ETSI), который принял всеохватывающий набор стандартов по цифровому вещанию, разработанных организацией. Хотя Россия и не входит в Европейский Союз, эти стандарты де-факто признаются в нашей стране.

Одной из старейших нормотворческих организаций в области телевидения является Американское общество Кино и ТВ инженеров (*SMPTE — Society of Motion Picture and Television Engineers*), которое давно уже переросло свои национальные рамки, включает многие организации из разных стран, так что его стандарты по своей авторитетности приравниваются к международным. Формальную легализацию стандартов *SMPTE* для США осуществляет Американский национальный институт стандартов.

Стандартом ТВ сигнала называют совокупность определяющих его основных характеристик, таких как способ разложения изображения, число строк и кадров, формат кадра, длительность и форма синхронизирующих и гасящих импульсов, полярность сигнала, разнос между несущими частотами изображения и звукового сопровождения, метод представления и кодирования цветовой информации (компонентный или композитный) и др. Для черно-белого телевидения существует 10 стандартов, которые принято изображать латинскими буквами *B, D, G, H, I, K, K1, L, M, N*.

По способу передачи сигналов цветности различают три системы цветного телевидения: SECAM, PAL, NTSC. Каждая из трех систем может применяться с любым из 10 стандартов черно-белого ТВ вещания, давая 30 возможных комбинаций. На практике применяются девять разновидностей

PAL, шесть — *SECAM* и один стандарт из группы *NTSC*. По стандарту разложения для ТВ сигнала стандартной четкости (ТСЧ) наиболее распространены сочетания 525/59,94/2:1 (*NTSC*, *PAL-M*), и 625/50/2:1 (все остальные разновидности *PAL* и *SECAM*). В аналоговом телевидении высокой четкости использовались форматы 1125/60 (Япония) и 1250/50 (Европа), но они не нашли широкого распространения.

При международном обмене ТВ программами часто приходится преобразовывать изображение из одного стандарта в другой. Преобразование с различным числом строк, но с той же частотой кадров и коэффициентом чересстрочности называется *повышающим* или *понижающим преобразованием стандарта*. Если изменяется частота кадров, это перекрестное преобразование. Частный случай преобразования системы цветности без изменения стандарта ($PAL \leftrightarrow SECAM$ или $NTSC \leftrightarrow PAL-M$) принято называть *транскодированием*.

Информация о яркости и цвете объекта в телевидении передается сочетанием трех основных цветов — красного (*R*), зеленого (*G*), синего (*B*). Первоначально цветовые видеосигналы формируются в виде компонентных *RGB*-сигналов, но при дальнейшей обработке для обеспечения совместимости с черно-белым телевидением переходят к другому набору компонентных сигналов — яркостному и двум цветоразностным, получаемым из сигналов *RGB* путем матрицирования. С учетом пониженной разрешающей способности цветового восприятия человеческого глаза полосы частот цветоразностных сигналов выбирают вчетверо меньше полосы сигнала яркости. В зависимости от размаха цветоразностных сигналов различают стыки Y, P_B, P_R ; $Y, R-Y, B-Y$; YUV . Для целей наземного вещания используются композитные сигналы *NTSC*, *PAL*, *SECAM*, в которых сигнал цветности объединяется тем или иным способом с сигналом яркости, совместимым с монохромным телевидением, для передачи в одном частотном канале. В системах *NTSC* и *PAL* цветоразностные сигналы передаются методом квадратурной модуляции соответствующей поднесущей

частоты; в системе *SECAM* используется метод частотной модуляции двух поднесущих 4,206 и 4,43 МГц.

Для передачи в эфир к собственно видеосигналу добавляются сигналы синхронизации, управляющие перемещением электронного луча в приемном кинескопе. *Синхросмесь* содержит кадровые и строчные синхронизирующие и гасящие импульсы. На периоде кадрового гасящего импульса, в течение которого луч перемещается из нижней в верхнюю часть экрана, умещаются 25 периодов строки, в течение которых видеосигнал не передается, и эти строки обычно используют для уплотнения ТВ сигнала. В строках 17, 18, 330, 331 передаются сигналы испытательных строк, вертикальный временной код занимает строки с номерами 19, 24, 332 и 334, в строки 20, 21, 333, 334 вводится телетекст. Использование строк различается в национальных стандартах.

Аналоговый телевизионный сигнал в соответствии с его природой повторяет распределение яркости и цветности на пути, по которому производится развертка изображения, то есть он действительно является электрическим аналогом изображения. Разработчикам телевизионных систем пришлось столкнуться с ограничениями аналоговых методов, серьезно сужающими возможности дальнейшего развития телевидения.

Одной из главных причин этих ограничений следует считать слабую помехозащищенность аналогового сигнала, который подвергается в каждом из устройств телевизионного тракта воздействию шумов и других помех. Современная же вещательная ТВ система представляет собой весьма длинную цепь устройств преобразования и передачи сигналов.

В любом звене этой сложной цепи возникает неизбежная потеря качества изображения. Связано это с тем, что в каждом устройстве, при любом из преобразований, которому подвергается сигнал, на него воздействуют помехи. При аналоговых методах усиления и обработки ТВ сигнала эти помехи накапливаются от звена к звену, и тем сильнее, чем больше в ТВ системе процессов обработки, переприема или перезаписей

сигнала. Пока этих преобразований немного, суммарные искажения еще могут быть незаметны. Но с развитием телевидения число преобразований очень быстро возрастает.

В таких системах проблема обеспечения необходимой помехоустойчивости становится главенствующей. Существенно уменьшить искажения от помех при формировании телевизионной программы, ее консервации или передаче позволяют цифровые методы. Поэтому в последние годы основное внимание уделяется развитию цифрового телевидения.

Цифровое телевидение — область телевизионной техники, в которой операции обработки, консервации и передачи телевизионного сигнала связаны с его преобразованием в цифровую форму.

Цифровые методы помимо обеспечения высокого качества изображения при воздействии значительных помех обладают и другими достоинствами. Так, при одинаковой пропускной способности канала они позволяют передавать большее число программ по сравнению с аналоговым телевидением. Лучшая помехозащищенность цифрового сигнала позволяет снизить требования к его мощности в процессе доставки к приемным устройствам. Повторные изображения, характерные для аналогового телевидения при многолучевом приеме, цифровыми методами могут быть практически полностью исключены.

Можно представить системы цифрового телевидения двух типов. В системе *первого типа*, полностью цифровой, преобразование передаваемого изображения в цифровой сигнал и обратное преобразование цифрового сигнала в изображение на приемном экране осуществляются непосредственно в преобразователях свет-сигнал и сигнал-свет. Во всех звеньях тракта передачи изображения информация передается в цифровой форме. В перспективе создание таких преобразователей вполне реально. Однако в настоящее время они еще не существуют, а поэтому целесообразно рассматривать цифровые ТВ системы *второго типа*, в которых с датчиков

получается аналоговый ТВ сигнал, затем он преобразуется в цифровую форму, подвергается всей необходимой обработке, передаче или консервации, а затем снова приобретает аналоговую форму. При этом используются существующие датчики аналоговых ТВ сигналов и преобразователи сигнал-свет в телевизионных приемниках. В этих системах на вход тракта цифрового телевидения поступает аналоговый ТВ сигнал, затем он кодируется, т.е. преобразуется в цифровую форму.

1.2. Импульсно-кодовая модуляция

Преобразование аналогового сигнала в цифровую форму представляет собой комплекс операций, наиболее существенными из которых являются дискретизация, квантование и непосредственно кодирование.

Дискретизация — замена непрерывного аналогового ТВ сигнала $u(t)$ последовательностью отдельных во времени отсчетов этого сигнала. Наиболее распространена равномерная дискретизация, имеющая постоянный период, основанная на теореме Котельникова-Найквиста.

Согласно этой теореме любой непрерывный сигнал, имеющий ограниченный спектр частот (рисунок 1.2,а), может быть представлен значениями этого сигнала в $u(t_n)$, взятыми в дискретные моменты времени (отсчеты) $t_n = nT$ (см.рисунок 1.2,б), где $n = 1, 2, 3, \dots$ — целые числа; T — период или интервал дискретизации, выбранный из условия теоремы Котельникова-Найквиста: $T \leq 0,5/f_{cp}$. Здесь f_{cp} - максимальная частота спектра исходного сигнала $u(t)$. Величина, обратная периоду дискретизации, называется частотой дискретизации.

Минимально допустимая частота дискретизации

$$f_0 = 2f_{cp}.$$

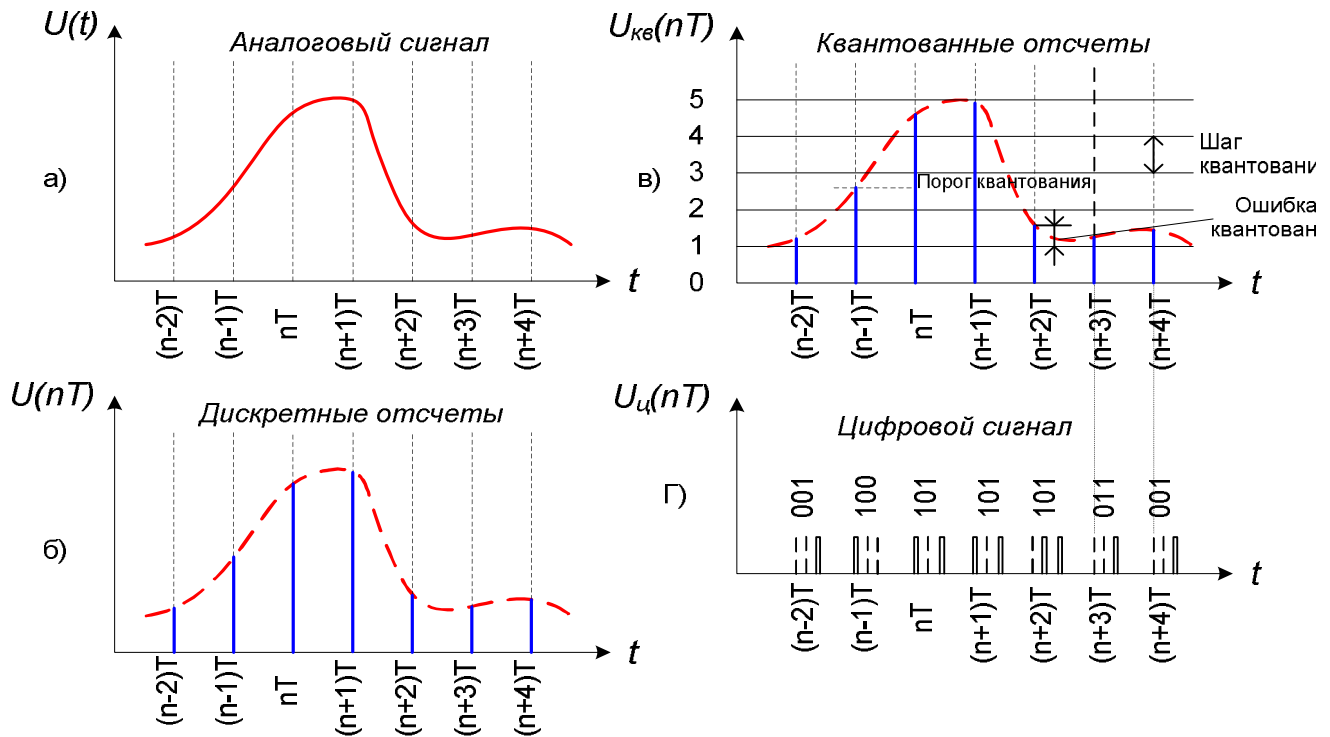


Рисунок 1.2 — Преобразование сигнала из аналоговой формы в цифровую

Аналитическое выражение теоремы Котельникова-Найквита имеет вид

$$u(t) = \sum_{n=-\infty}^{\infty} u(nT) \frac{\sin 2\pi f_{гр}(t - nT)}{2\pi f_{гр}(t - nT)} \quad (1.1)$$

Предполагается, что отсчеты $u(nT)$ являются δ -импульсами (бесконечно короткими). Для восстановления исходного аналогового сигнала $u(t)$ из последовательности отсчетов $u(nT)$ последние необходимо в соответствии с (1.1) пропустить через идеальный фильтр нижних частот (ФНЧ) со срезом на частоте $f_{гр}$.

Множитель $\frac{\sin 2\pi f_{гр}(t - nT)}{2\pi f_{гр}(t - nT)}$ представляет собой реакцию такого фильтра

на единичный импульс $u(nT)$. Из теоремы следует, что для точного восстановления исходного сигнала необходимо наличие бесконечно большого числа отсчетов. На практике же сигнал, всегда имеющий конечную длительность, описывается конечным числом отсчетов. Несмотря на несоответствие условиям теоремы, такой способ восстановления сигнала

широко используется в цифровом телевидении, и точность восстановления при соблюдении определенных требований оказывается достаточной.

За процессом дискретизации при преобразовании аналогового сигнала в цифровую форму следует процесс квантования. *Квантование* (этот термин заимствован из атомной физики) заключается в замене полученных после дискретизации мгновенных значений отсчетов ближайшими значениями из набора отдельных фиксированных уровней (рисунок 1.2,в). Квантование также представляет собой дискретизацию ТВ сигнала, но осуществляемую не во времени, а по уровню сигнала $u(t)$. Для устранения путаницы между этими понятиями и введена разная терминология.

Фиксированные уровни, к которым «привязываются» отсчеты, называют *уровнями квантования*. Разбивая динамический диапазон изменения сигнала $u(t)$ уровнями квантования на отдельные области значений, называемые *шагами квантования*, образуют *шкалу квантования*. Последняя может быть как линейной, так и нелинейной, в зависимости от условий преобразования. Округление отсчета до одного из двух ближайших уровней (верхнего или нижнего) определяется положением *порогов квантования* (рисунок 1.2,в).

Возможность восстановления в зрительном аппарате человека исходного изображения по его квантованному приближению (в теореме эта операция не предусматривается) вытекает из ограниченности контрастной (и цветовой) чувствительности зрительной системы.

Строго говоря, дискретизированный и квантованный сигнал в $u_{кв}(nT)$ уже является цифровым. Действительно, если амплитуда импульсов дискретизированного сигнала $u(nT)$ может принимать любые произвольные значения в пределах исходного динамического диапазона сигнала $u(t)$, то операция квантования привела к замене всех возможных значений амплитуды сигнала ограниченным числом значений, равным числу уровней квантования. Таким образом, квантованная выборка сигнала выражается некоторым числом в системе счисления с основанием m , где m - число

уровней квантования. Но цифровой сигнал в такой форме по помехозащищенности мало выигрывает по сравнению с аналоговым, особенно при большом m . Для увеличения помехозащищенности сигнала его лучше всего преобразовать в двоичную форму, т.е. каждое значение уровня сигнала записать в двоичной системе счисления. При этом номер (значение уровня) будет преобразован в кодовую комбинацию символов 0 или 1 (рисунок 1.2,г). В этом и состоит третья, заключительная операция по преобразованию аналогового сигнала $u(t)$ в цифровой, называемая *операцией кодирования*.

Кодирование представляет собой преобразование квантованного значения отсчета $u_{\text{кв}}(nT)$ в соответствующую ему кодовую комбинацию символов $u_{\text{ц}}(nT)$. Наиболее распространенный способ кодирования ТВ сигнала — представление его дискретных и проквантованных отсчетов в натуральном двоичном коде. Этот способ получил название *импульсно-кодовой модуляции* (ИКМ). На рисунке 1.2,г показан результат преобразования фрагмента исходного сигнала $u(t)$ в последовательность комбинаций двоичного трехразрядного кода.

Часто всю совокупность перечисленных операций — дискретизации, квантования и кодирования для краткости называют *кодированием* телевизионного сигнала. Это имеет определенные технические основания, поскольку все эти три операции выполняются одним техническим устройством — *аналого-цифровым преобразователем* (АЦП). Обратное преобразование цифрового сигнала в аналоговый производится в устройстве, называемом *цифро-аналоговым преобразователем* (ЦАП). Аналого-цифровые и цифро-аналоговые преобразователи — неперенные блоки любых цифровых систем передачи, хранения и обработки изображений.

Оценим скорость передачи цифрового потока ИКМ сигнала по каналу связи.

При непосредственном кодировании телевизионного сигнала методом ИКМ кодовые комбинации создаются с частотой, равной частоте отсчетов, то есть частоте дискретизации.

Каждая кодовая комбинация соответствует определенному отсчету и содержит некоторое число k двоичных символов (битов).

Скоростью цифрового потока C называется число передаваемых двоичных знаков в единицу времени. За единицу скорости принимается 1 бит в секунду. Таким образом, скорость передачи ТВ сигнала в цифровой форме равна произведению частоты дискретизации f_d и числа двоичных символов k в одном дискретном отсчете

$$C = f_d k. \quad (1.2)$$

Для количественной оценки скорости передачи ИКМ сигнала необходимо обосновать выбор k и f_d . Число двоичных символов в кодовой комбинации одного отсчета связано с числом уровней квантования m исходного сигнала соотношением

$$k = \log_2 m \approx 3,3 \lg m \quad (1.3)$$

Выбор числа уровней квантования определяется требованием к минимизации ошибок (ошибок квантования), возникающих из-за замены истинных значений отсчетов сигнала их квантованными (приближенными) значениями. Ошибки квантования носят случайный характер. Поэтому искажения, вызываемые этими ошибками, часто называют шумами квантования. На изображении они могут проявляться по-разному, в зависимости от свойства кодируемого сигнала.

1.3. Обобщенная структурная схема системы цифрового телевидения

Подлежащий преобразованию аналоговый сигнал поступает на вход цифровой ТВ системы (см. рисунок 1.3).

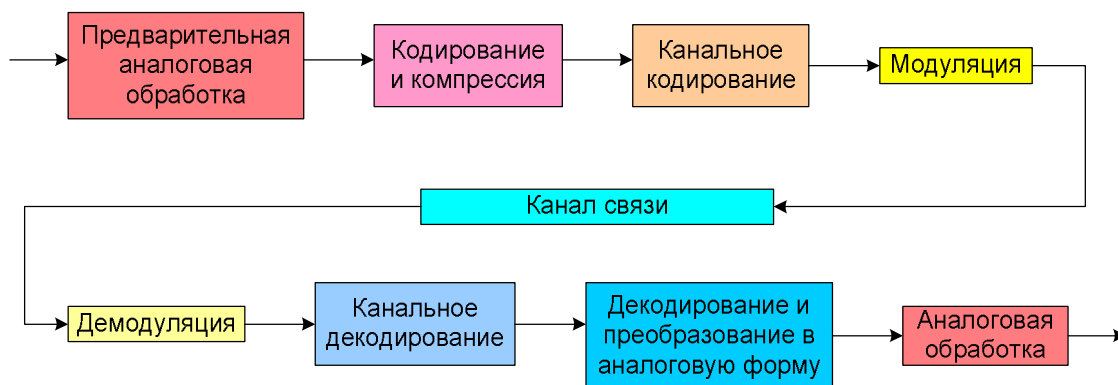


Рисунок 1.3 — Обобщенная структурная схема ЦТС

Этот сигнал подвергается предварительной обработке для упрощения последующих цифровых преобразующих устройств. Например, полный цветовой сигнал разделяется в устройстве предварительной обработки на сигнал яркости и цветоразностные сигналы с тем, чтобы цифровые преобразования производились с каждым из трех сигналов отдельно. Можно ввести в аналоговый сигнал определенные предсказания для улучшения субъективного качества выходного изображения и т.п. Несмотря на то, что многие из этих предварительных операций по обработке могут быть сделаны и в цифровой форме, на определенном этапе развития технически проще их выполнять в аналоговой форме. Далее, подготовленный для преобразования аналоговый сигнал поступает на кодирующее устройство, в котором он дискретизируется, квантуется и предварительно кодируется. Как указывалось, в полученном таким образом сигнале содержится значительная избыточность, которая может быть в определенной степени сокращена дополнительным, более эффективным кодированием перечисленными выше

методами компрессии. Затем цифровой сигнал подвергается так называемой прямой коррекции ошибок, выполняемой в устройстве канального кодирования, и, наконец, поступает на выходной преобразователь (например, на модулятор передающего устройства). В приемном устройстве осуществляются обратные операции.

Приведенная на рисунке 1.3 схема является обобщенной. В зависимости от задач, стоящих перед цифровой системой, она может видоизменяться. Например, система вообще не будет содержать аналоговых звеньев, если использовать преобразователи свет-сигнал и сигнал-свет, генерирующие и преобразующие сигнал в цифровом виде. В другом случае могут отсутствовать устройства, повышающие помехоустойчивость сигнала в каналах связи. Это допустимо при отсутствии протяженных линий связи и, в частности, при цифровой обработке сигнала внутри одного телецентра. В том же случае не обязательны в устройства, устраняющие в ТВ сигнале избыточность и сокращающие цифровой поток.

1.4. Дискретизация телевизионного сигнала

Дискретизация — первая операция из всего комплекса преобразований аналогового сигнала в цифровой. Рассмотрим факторы, определяющие выбор частоты дискретизации.

Согласно теореме Котельникова: «частота дискретизации f_d должна быть в два раза больше граничной (максимальной) $f_{гр}$ ». Второе важное соображение, которое необходимо учитывать — частота дискретизации должна быть целым кратным от частоты строк, чтобы на периоде строки помещалось целое число отсчетов сигнала.

С учетом этих соображений в Рекомендации ITU-R BT.601, описывающей студийный стандарт компонентного цифрового видеосигнала, выбраны полоса частот сигнала яркости 5,75 МГц и частота дискретизации для этого сигнала 13,5 МГц. Эта частота равна 858-й гармонике строчной частоты 15625 Гц стандарта 625/50/2:1 и 864-й гармонике частоты 15734 Гц стандарта 525/59,49/2:1. Для цветоразностных сигналов выбрана частота

дискретизации 6,75 МГц. При этом достигается фиксированная относительно телевизионная раstra структура отсчетов. Они располагаются в узлах прямоугольной решетки образуя так называемую ортогональную структуру дискретизации (см. рисунок 1.4). Полное число отсчетов сигнала яркости в строке 864, а число отсчетов каждого цветоразностного сигнала — 432, а в активной части строки соответственно 720 отсчетов сигнала яркости и по 360 отсчетов сигналов цветности.

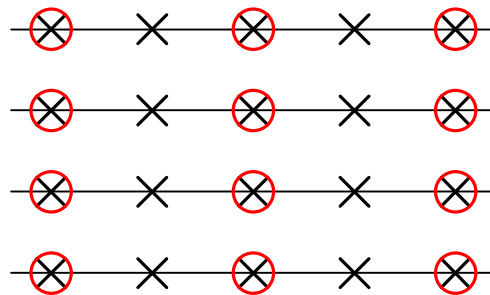


Рисунок 1.4 — Ортогональная структура дискретизации
(X — отсчеты сигнала яркости; O — отсчеты сигнала цветности)

Исходный сигнал $u(t)$ после дискретизации можно представить в виде суммы

$$u(nT) = \sum_{n=-\infty}^{\infty} u(t) \delta(t - nT) \quad (1.4)$$

где δ — дельта-функция;

T — период дискретизации.

Если (1.4) подвергнуть преобразованию Фурье, то

$$S_o(f) = \sum_{n=-\infty}^{\infty} S(f - nf_o). \quad (1.5)$$

$S(f)$ и $S(f_o)$ — спектры исходной и дискретной функций.

Из (1.5) следует, что спектр дискретизированного сигнала представляет собой сумму исходного спектра ($n = 0$) и «побочных» или дополнительных спектров того же вида, но сдвинутых один относительно другого на f_o , $2f_o$, ...и т.д. (см. рисунок 1.5).

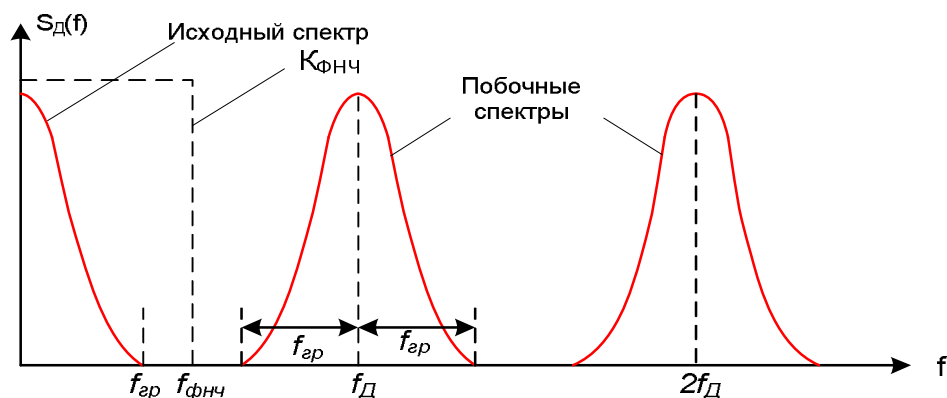


Рисунок 1.5 — Спектр сигнала после дискретизации

Из рисунка видно, что с помощью идеального фильтра нижних частот (ФНЧ) с частотой среза $f_{\text{ФНЧ}}$ можно выделить спектр исходного сигнала, если выполняются условия

- 1) $f_{\text{Д}} \geq 2f_{\text{гр}};$
- 2) $f_{\text{гр}} \leq f_{\text{ФНЧ}} \leq f_{\text{Д}} - f_{\text{гр}}.$

Таким образом, при полосе частот яркостного сигнала, составляющей 5,75 МГц, частота дискретизации должна быть выбрана не менее 12,5 МГц, а для цветоразностных сигналов — 6,75 МГц.

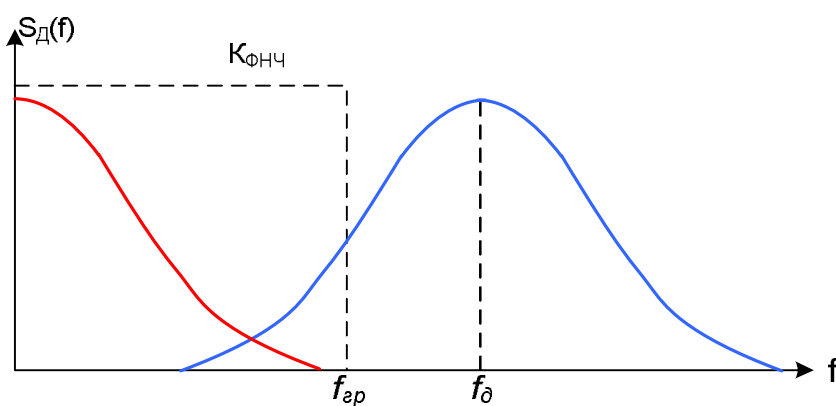


Рисунок 1.6 — Перекрывание спектров при $f_{\text{Д}} < 2f_{\text{гр}}$

По аналогии с аналого-цифровым преобразованием композитных сигналов *PAL* и *NTSC*, где частота дискретизации выбиралась равной учетверенной частоте цифровой поднесущей, частоту 13,5 МГц также называется «четверкой», а частоту 6,75 МГц — «двойкой». Таким образом, полный цифровой компонентный видеосигнал из Рекомендации ВТ.601 описывается формулой «4:2:2».

Компоненты его называются Y , C_B , C_R и связаны с исходными Y , $R-Y$, $B-Y$ следующими соотношениями:

$$Y = 0,299 R + 0,587 G + 0,114 B;$$

$$C_R = 0,713 (R - Y);$$

$$C_B = 0,546 (B - Y)$$

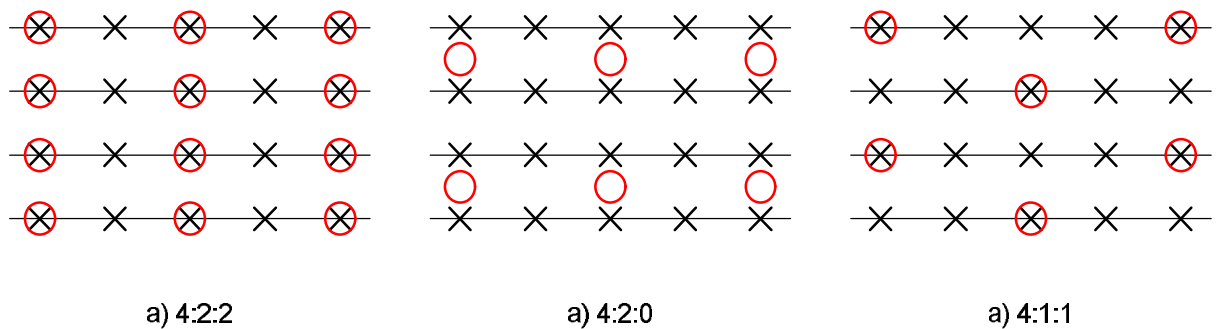


Рисунок 1.7 — Структура отсчетов компонентного сигнала Y , C_B , C_R (X — отсчеты сигнала яркости; O — отсчеты сигнала цветности)

На рисунке 1.7 показаны примеры часто используемых форматов дискретизации. Цифры в названиях форматов определяют соотношение яркостной и цветовой четкости по вертикали и горизонтали.

4:2:2 — на четыре отсчета сигнала яркости приходится по два отсчета сигнала цветности в четных и нечетных строках.

4:2:0 — на четыре отсчета сигнала яркости приходится по два отсчета сигнала цветности в четных строках и ни одного в нечетных.

Как видно из рисунков в форматах 4:2:2 и 4:2:0 цветовая четкость по горизонтали в два раза меньше, чем яркостная. В форматах 4:2:2 и 4:1:1 цветовая четкость по вертикали такая же, как и яркостная. В формате 4:2:0

цветовая четкость по вертикали в два раза меньше, чем яркостная. Ясно, что в формате 4:4:4 яркостная и цветовая четкость равны в обоих направлениях, а формат 4:0:0 — черно-белое телевидение. На рисунке 1.7 отсчеты CR и CB совпадают по горизонтали с нечетными отсчетами сигнала яркости. Общее число отсчетов в цифровой строке составляет 1440, в том числе 720 отсчетов яркости и по 360 — каждого из цветоразностных сигналов. Когда к Рекомендации BT.601 добавили часть «В», описывающую формат изображения 16:9, частоту дискретизации для этого сигнала пришлось увеличить в $4/3$ раз, до 18 МГц.

Благодаря усилиям 11-й Исследовательской Комиссии *ITU* удалось стандартизовать цифровой сигнал не только стандартного, но и высокого качества. Основой части II принятой «Рекомендации BT.709-3 является «Единый формат изображения» (*Common Image Format — CIF*), устанавливающий некоторые общие значения параметров ТВЧ сигнала для американского и европейского стандартов с частотой полей соответственно 50 и 60 Гц.

Несомненным достижением является тот факт, что в стандарте зафиксированы основные параметры разложения (1920×1080 пикс., формат 16:9), частоты дискретизации сигналов яркости (74,25 МГц) и цветности (37,125 МГц), скорость цифрового потока 1,485 Гбит/с.

Выбор числа уровней квантования видеосигнала в значительной степени определяется характеристиками зрительной системы человека. Максимальное воспринимаемое отношение яркостей 100:1 соответствует примерно 230 градациям яркости. Ближайшая степень двойки $256 = 2^8$, и в первой редакции Рекомендации BT.601 было принято 8-битовое квантование. Однако получающееся при этом предельное отношение сигнал-шум 56,8 дБ не удовлетворяет многим высококачественным применениям, и в дополнение к 8-битовому квантованию было введено 10-битовое. Результирующая скорость цифрового потока сигнала «4:2:2» с 10-битовым кодированием и

форматом кадра 4:3 составляет $13,5 \cdot 10^6 \cdot 10 + 2 \cdot 6,75 \cdot 10^6 \cdot 10 = 270$ Мбит/с (для формата 16:9 соответственно 360 Мбит/с).

1.5. Квантование телевизионного сигнала

Квантование сигнала по уровню заключается в нахождении для каждого отсчета сигнала ближайшего к нему уровня квантования из используемого набора уровней квантования или, говоря проще, округлении значений отсчетов сигнала до ближайших уровней квантования. Количество уровней квантования N связано с числом двоичных разрядов АЦП n соотношением $N_{\text{КВ}} = 2^n$.

На рисунке 1.8,а показан дискретный сигнал, содержащий 256 отсчетов (отдельные отсчеты в этом масштабе не различимы), значения уровня сигнала измеряются в вольтах и изменяются непрерывно, т.е. квантования нет. На рисунке 1.8,б показан тот же сигнал после квантования с числом двоичных разрядов $n = 4$, т.е. имеющий 16 уровней квантования. На рисунке 1.8,в показаны значения величины ошибки Δu , вносимой в сигнал операцией квантования. При этом предполагается, что квантование осуществляется путем усечения значения сигнала до ближайшего снизу уровня квантования. Величина ошибки квантования Δu изменяется в пределах от 0 до h .

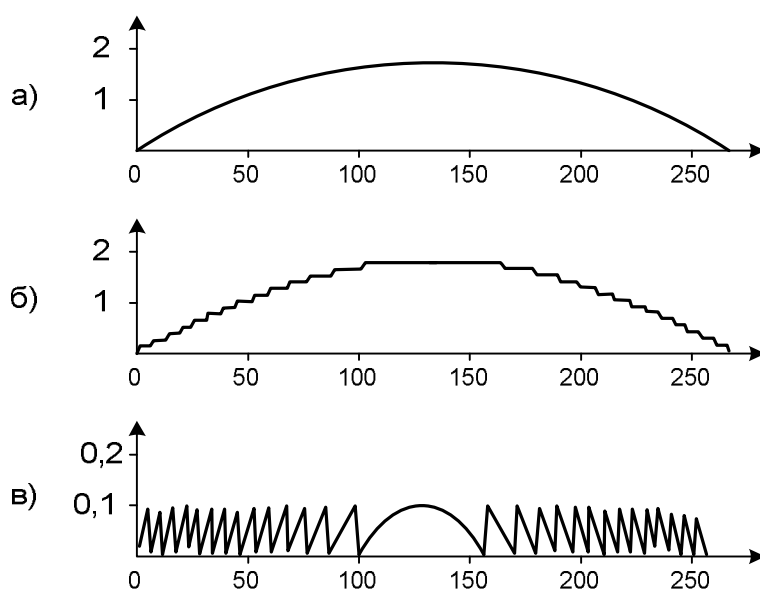


Рисунок 1.8 — Сигнал до квантования (а), сигнал после квантования (б), ошибка квантования (в)

Ошибка квантования является случайной величиной, поэтому ее часто называют шумом квантования. На изображении шум квантования может проявляться различным образом в зависимости от характера изменений яркости или цвета на данном участке изображения.

На участках, состоящих из мелких деталей, квантование приведет к случайным изменениям их яркости или цвета. На участках изображения с плавным изменением уровня видеосигнала квантование может привести к возникновению ложных контуров по тем линиям, на которых уровень видеосигнала переходит с одного уровня квантования на другой. Равномерное квантование телевизионного сигнала, вообще говоря, не является наилучшим. Это обусловлено в первую очередь свойствами зрительного аппарата человека. В соответствии с экспериментальным законом Вебера-Фехнера пороговое превышение яркости объекта над фоном, при котором объект различается наблюдателем, равно $1,5...2,0\%$ от величины яркости фона B_0 . Таким образом, с ростом яркости фона растет и порог $\Delta B_{\text{пор}}$. Отсюда следует, что в области значений телевизионного сигнала, близких к уровню черного, шаг квантования должен быть меньше, чем в области, близкой к уровню белого.

Однако техническая реализация неравномерного квантования существенно сложнее, чем равномерного. Вместо использования переменного шага квантования обычно выполняют предварительное нелинейное преобразование видеосигнала — гамма-коррекцию. При этом решается одновременно две задачи. Во-первых, корректируется нелинейность передаточной характеристики кинескопа и обеспечивается оптимальная форма передаточной характеристики всего тракта телевизионной системы «от света до света». Во-вторых, уменьшается влияние ошибок квантования при малых уровнях яркости изображения.

Передаточная характеристика гамма-корректора описывается соотношением

$$(U_{\text{ВЫХ}}/U_{\text{ВЫХ М}}) = (U_{\text{ВХ}}/U_{\text{ВХ М}})^{\gamma},$$

где U_{BX} и $U_{BЫX}$ — напряжения сигналов на входе и на выходе гамма-корректора; $U_{BX\ M}$ и $U_{BЫX\ M}$ — максимальные значения диапазонов напряжений сигналов на входе и выходе гамма-корректора; $\gamma = 0,42 \dots 0,48$ — показатель гамма-коррекции.

График передаточной характеристики гамма-корректора для γ , равного 0,45, показан на рисунке 1.9.

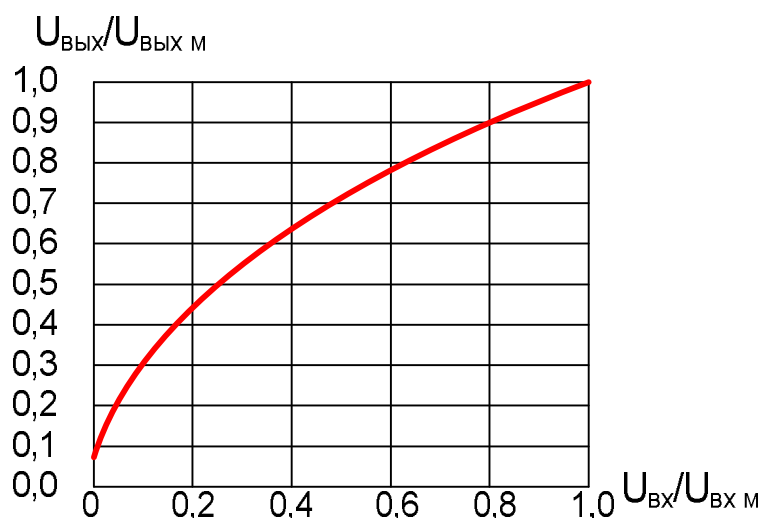


Рисунок 1.9 — Передаточная функция гамма-корректора

В системах цифрового телевидения, как правило, применяется равномерное квантование прошедших гамма-коррекцию сигналов с числом двоичных разрядов АЦП $n = 8$, что дает количество уровней квантования $N_{KB} = 256$. При этих условиях шум квантования на изображении практически незаметен. В последние годы все чаще применяется квантование с числом двоичных разрядов $n = 10$, что позволяет еще улучшить качество изображения.

Произведя «нумерацию» уровней квантования, можно передавать не сами уровни, а их значения по шкале уровней в двоичном коде. Полученная в результате этого преобразования импульсная последовательность является групповым ИКМ сигналом.

На выходе АЦП полученный номер уровня квантования представляется в виде двоичного числа, т.е. выполняется кодирование (оцифровка). При числе разрядов АЦП, равном 8, значения цифрового сигнала представляются числа ми от 00000000 до 11111111 в порядке нарастания их величины.

Число уровней квантования в цифровом сигнале может изменяться. Пусть, например, цифровой сигнал имел число двоичных разрядов $n = 8$, а в результате повторного квантования этого сигнала число разрядов уменьшается до $n = 4$. При выполнении этой операции сначала производится округление каждого 8-разрядного числа до ближайшего 8-разрядного числа, у которого 4 младших разряда равны 0. Так, число 10101010 округляется до 10110000, а число 11010011 — до числа 11010000. Затем 4 младших разряда отбрасываются, и остаются числа 1011 и 1101, соответственно.

Обратную операцию, при которой число двоичных разрядов в цифровом сигнале увеличивается, будем называть деквантованием. Самый простой вариант выполнения деквантования — добавить справа к каждому числу, входящему в цифровой сигнал, нужное количество нулей. Тогда числа из приведенного выше примера превратятся снова в 10110000 и 11010000, но точные значения их младших разрядов, потерянные при повторном квантовании, восстановить нельзя. Поэтому следует запомнить, что квантование вносит необратимые искажения в сигнал. Отметим также, что в последующем изложении повторное квантование, как правило, будет называться просто квантованием.

В Рекомендации 601 для всех сигналов предусмотрено число разрядов квантования $n = 8$, что дает число уровней квантования $N_{\text{кв}} = 256$. При этом уровню черного сигнала яркости соответствует 16-й уровень квантования, а номинальному уровню белого — 235-й уровень квантования. 16 уровней квантования снизу и 20 уровней квантования сверху образуют резервные зоны на случай выхода значений аналогового сигнала яркости за пределы номинального диапазона. Особые назначения имеют 0-й и 255-й уровни квантования. С помощью соответствующих им кодов передаются сигналы

синхронизации. Аналого-цифровое преобразование сигнала яркости описывается соотношением

$$Y = \text{Round} (219 E'_Y) + 16,$$

где E'_Y — аналоговый сигнал яркости, изменяющийся в диапазоне от 0 до 1 В (штрих, как это принято в технической литературе по телевидению, означает, что сигнал прошел гамма-коррекцию); Y — цифровой сигнал яркости, изменяющийся в диапазоне от 16 до 235; $\text{Round}(X)$ — операция округления числа X до целого.

При квантовании цветоразностных сигналов также предусматриваются резервные зоны — по 16 уровней квантования сверху и снизу. На АЦП поступают не сами цветоразностные сигналы E'_{R-Y} , E'_{B-Y} , а компрессированные цветоразностные сигналы, формируемые в соответствии с соотношениями

$$E_{CR} = 0,713E'_{R-Y}; \quad E_{CB} = 0,564E'_{B-Y}$$

Уровни сигналов E_{CR} и E_{CB} изменяются в диапазоне от $-0,5$ В до $0,5$ В.

Аналого-цифровые преобразования цветоразностных сигналов, в результате которых получаются цифровые цветоразностные сигналы C_R и C_B , выполняются в соответствии со следующими соотношениями

$$C_R = \text{Round} (224 E'_{CR}) + 128,$$

$$C_B = \text{Round} (224 E'_{CB}) + 128.$$

Так как цветоразностные сигналы являются двуполярными, и максимальные отклонения от нулевого значения в положительную и отрицательную стороны примерно одинаковы, 128-й уровень квантования должен соответствовать нулевому значению этих сигналов. Старший бит в цифровом цветоразностном сигнале показывает полярность аналогового цветоразностного сигнала.

На рисунке 1.10 показано соответствие между уровнями аналоговых телевизионных сигналов и уровнями квантования для обычного тестового изображения в виде восьми цветных полос.

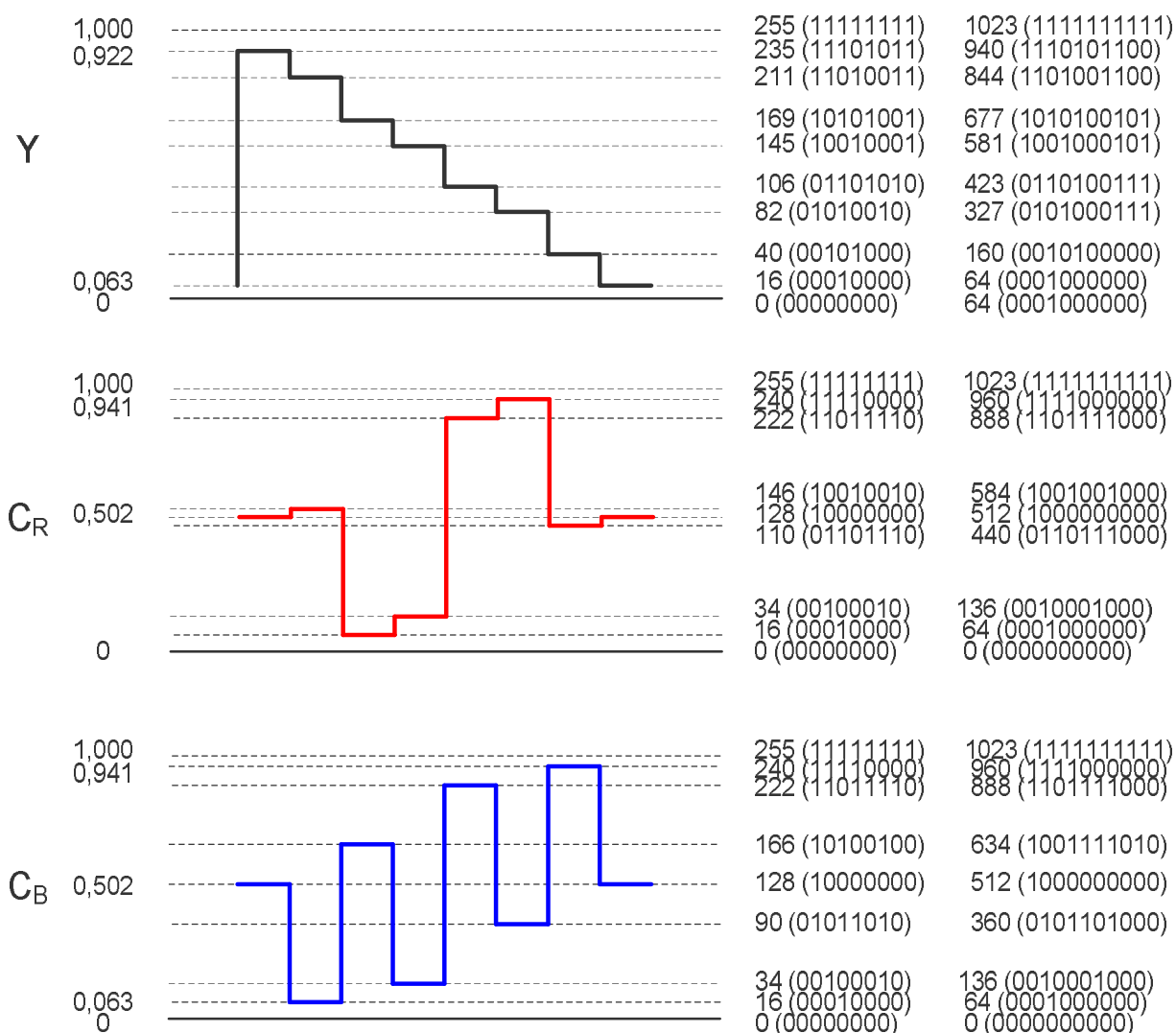


Рисунок 1.10 — Соответствие между уровнями аналоговых телевизионных сигналов и уровнями квантования по Рекомендации ITU-R BT 601

В состав цифрового телевизионного сигнала согласно Рекомендации ITU-R BT 601 входят синхросигналы. Перед началом активного участка каждой строки в конце строчного гасящего импульса передается синхросигнал «Начало активной строки — НАС» (*SAV — Start Active Video*), а после окончания активного участка каждой строки в начале строчного гасящего импульса передается синхросигнал «Конец активной строки — КАС» (*EAV — End Active Video*) (рисунок 1.10). Каждый из сигналов содержит 4 символа, в том числе один символ единиц, два символа нулей (уровни квантования соответственно 1024 и 0, не используемые для видеоданных) и информационный символ, указывающий на тип сигнала и перемежение. Распознавание синхросигналов производится по 8 старшим

битам. При выбранной частоте дискретизации 13,5 МГц и стандарте разложения 625 строк на периоде строки укладываются 864 периода, из которых для передачи полезных данных используются только 720, называемые *активной частью строки*. Начало цифровой строки совпадает с серединой переднего фронта строчного синхроимпульса. Чтобы передать 720 отсчетов сигнала яркости и по 360 отсчетов сигналов C_R и C_B , тактовая частота выбирается равной 27 МГц (для формата сигнала 16:9 частота дискретизации 18 МГц и тактовая частота 36 МГц). В стандарте 525 строк число циклов несколько меньше, но активная часть строки содержит те же 720 периодов, и формат передачи данных одинаков.

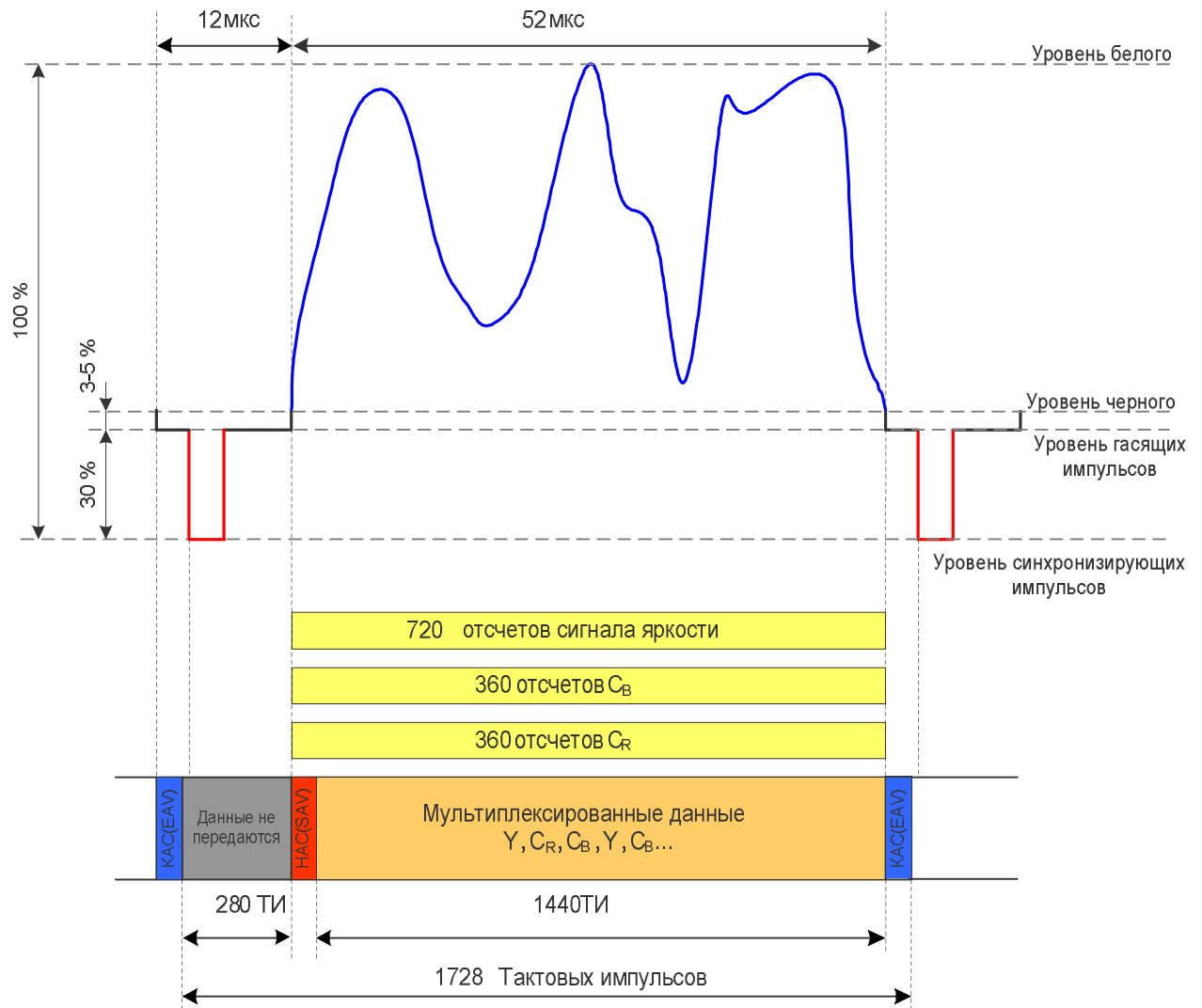


Рисунок 1.11 — Структура цифровой строки стандарта 625/50/2:1

1.6. Формирователи цифровых телевизионных сигналов

Рассмотрим два варианта структурной схемы формирователя цифрового телевизионного сигнала в соответствии с Рекомендацией *ITU-R BT 601*. В устройстве (см. структурную схему на рисунке 1.12,а) сигналы основных цветов E_R , E_G , E_B с источника телевизионных сигналов (телекамеры) вначале поступают на гамма-корректоры (ГК). Сформированные сигналы E'_R , E'_G , E'_B в кодирующей матрице (КМ) преобразуются в сигнал яркости E'_Y и цветоразностные сигналы E'_{R-Y} , E'_{B-Y} . далее эти сигналы преобразуются АЦП в цифровые сигналы Y , C_R , C_B , соответственно. На входах АЦП имеются дополнительные аналоговые узлы, выполняющие масштабирование и сдвиг сигналов. Число разрядов каждого АЦП, как правило, равно восьми. Синхроимпульсы (СИ) развертки источника ТВ сигналов поступают на формирователь цифровых СИ (ФЦСИ), вырабатывающий синхросигналы НАС и КАС. Кроме того, СИ используются для синхронизации генератора тактовых импульсов (ГТИ), который вырабатывает импульсы с частотами 27, 13,5 и 6,75 МГц, поступающие на другие узлы устройства. ГТИ содержит схему фазовой автоподстройки частоты (ФАПЧ), с помощью которой обеспечивается требуемое количество периодов тактовых импульсов за период строчной развертки источника ТВ сигналов.

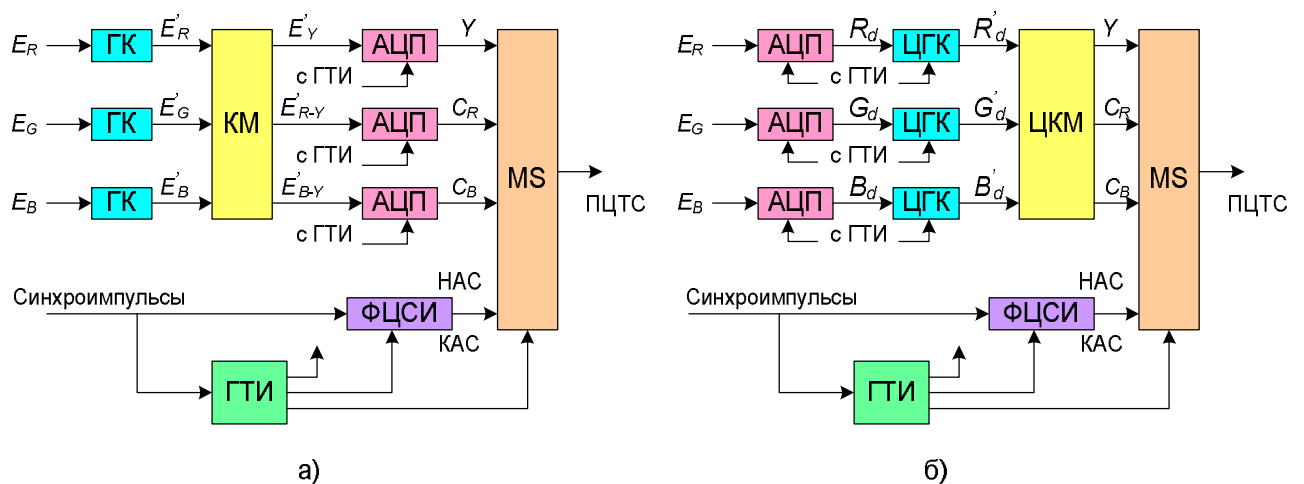


Рисунок 1.12 – Варианты структурной схемы формирователя цифрового телевизионного сигнала

Мультиплексор (MS) в заданной последовательности передает на выход цифровые сигналы Y , C_R , C_B и цифровые синхросигналы. В результате на выходе устройства оказывается сформированным цифровой телевизионный сигнал (ЦТС).

В другом варианте устройства (рисунок 1.12,6) сигналы основных цветов E_R , E_G , E_B сразу преобразуются в цифровые сигналы R_d , G_d , B_d , соответственно. При этом каждый АЦП должен иметь, по меньшей мере, 10, а лучше 12 двоичных разрядов.

Далее цифровые сигналы R_d , G_d , B_d поступают на цифровые гамма-корректоры (ЦГК), в которых выполняются нелинейные преобразования. Число двоичных разрядов прошедших гамма-коррекцию цифровых сигналов R'_d , G'_d , B'_d равно 8. Затем сигналы R'_d , G'_d , B'_d в цифровой кодирующей матрице (ЦКМ) преобразуются в цифровой сигнал яркости Y и цифровые цветоразностные сигналы C_R и C_B .

Формирование синхросигналов и тактовых импульсов и работа мультиплексора осуществляются аналогично первому варианту устройства. Выполнение гамма-коррекции цифровыми средствами обеспечивает более точное задание требуемой функции преобразования, но при этом необходимы имеющие больше двоичных разрядов и, следовательно, более дорогие АЦП.

Формирователи цифровых телевизионных сигналов в виде специализированных БИС, которые можно использовать, например, в устройствах оцифровки и ввода телевизионных сигналов в персональные компьютеры. Так, фирма *PHILIPS* производит БИС SAA7111A, в которой выполняются преобразование в цифровую форму ЦТС или отдельных сигналов яркости и цветности, цифровое декодирование сигнала цветности всех распространенных стандартов цветного телевидения и формирование выходных цифровых сигналов в соответствии с Рекомендацией 601 с форматами 4:2:2 и 4:1:1. Тактовая частота 27 МГц вырабатывается встроенным генератором с ФАПЧ и привязывается к частоте строчных

синхроимпульсов входного телевизионного сигнала. Гамма-коррекция не требуется, так как в цифровую форму преобразуется ПЦТС, пришедший с телевизионной студии, где он прошел гамма-корректор.

Перейдем к передаче цифрового телевизионного сигнала. Рекомендацией ITU-R BT 656 предусмотрены два варианта интерфейса для этой цели: параллельный видеостык и последовательный видеостык.

1.7. Цифровые интерфейсы в телевидении

Цифровым *интерфейсом*, или стыком, называется точка соединения двух приборов, в которой данные передаются от одного прибора к другому. Чтобы обеспечить быстрое и надежное соединение без подстройки параметров, оба прибора должны соответствовать единым требованиям к точке стыка. В интерфейсах, как ни в какой другой области цифровой техники важна стандартизация.

Соединение может быть однонаправленным, или *симплексным*, когда данные передаются только в одном направлении, и двунаправленным, или *дуплексным*, когда передача данных происходит в обоих направлениях. Возможна еще полудуплексная работа, когда данные передаются в обоих направлениях, но не одновременно, а с разделением по времени. Понятно, что сигналы реального времени (телевидение, звуковое вещание) могут передаваться только в симплексном режиме.

Для описания любого цифрового интерфейса необходимо указать:

- а) протокол обмена данными;
- б) электрические сигналы — напряжения, токи, входные и выходные сопротивления;
- в) физическое соединение — вид соединителя (штыревой или гнездный), крепление.

Поток цифровых данных, сформированный канальным передатчиком, не содержит каких-либо сведений об адресате, содержании и характеристиках кодирования сообщения. Принять его можно, только зная заранее указанные характеристики и настроив на них приемник. Если часть

передачи по каким-либо причинам не принята, нужно предусмотреть возможность сообщить об этом отправителю и запросить повторение поврежденного сообщения. Эти проблемы разрешает протокол — стандартизованная процедура установления, поддержания и прекращения цифрового соединения, определяющая процедуру передачи управляющей информации и данных, механизм выбора указанной процедуры из списка возможных, структуру и способ кодирования блоков данных.

Рассмотрим некоторые цифровые интерфейсы, наиболее часто используемые в цифровом вещании.

Параллельный интерфейс применяется в студийном оборудовании для коротких соединений, поэтому стоит остановиться несколько подробнее на основных свойствах параллельного компонентного стыка, стандартизованного в Рекомендации ВТ.656 и в аналогичном ей стандарте *SMPTE-125M*.

При выбранной частоте дискретизации 13,5 МГц и стандарте разложения 625 строк на периоде строки укладываются 864 периода, из которых для передачи полезных данных используются только 720, называемые *активной частью строки*. Начало цифровой строки совпадает с серединой переднего фронта строчного синхроимпульса. Чтобы передать 720 отсчетов сигнала яркости и по 360 отсчетов сигналов C_B и C_R , тактовая частота выбирается равной 27 МГц (для формата сигнала 16:9 частота дискретизации 18 МГц и тактовая частота 36 МГц). Отсчеты передаются в коде БВН в следующем порядке: C_B , Y, C_R , $Y...$ В старой версии стандарта разрядность квантования составляла 8 бит, и была предусмотрена возможность добавления двух младших разрядов после запятой. В ныне действующей 10-битовой версии по каналу всегда передаются 10 битов, и в случае 8-битового сигнала последние два бита обнуляются. Из общего числа 1024 уровня для передачи непосредственно отсчетов видеосигнала яркости используются 877 (с 64-го по 940-й), запас сохраняется как допуск на отклонения размаха исходного сигнала и выбросы фильтров. Сигналы

цветности передаются в пределах 897 уровней, центрированных относительно среднего уровня 512.

Компонентный цифровой сигнал «4:2:2» содержит 10 отдельных битовых потоков, которые передаются по отдельным парам многожильного кабеля специальной конструкции. По отдельной паре передаются сигналы тактовой синхронизации, еще несколько проводов используются для экранирования и заземления. Максимальная длина кабеля без коррекции амплитудно-частотной характеристики (АЧХ) 50 м, с коррекцией — не более 300 м, номинальный импеданс нагрузки составляет 110 Ом. Синхронизация осуществляется нарастающим фронтом тактового импульса, который должен соответствовать середине бита данных.

Параметры цифрового интерфейса *DVB-SPI (Synchronous Parallel Interface* — синхронный параллельный интерфейс), разработанного на основе компонентного параллельного стыка специально для передачи транспортного потока *MPEG-2*.

Из популярных параллельных интерфейсов следует упомянуть еще *RS422* и *M2P*. *RS422* имеет статус «рекомендованного стандарта» (*RS* — *Recommended Standard*), т.е. он не принят ни одним уполномоченным органом, поэтому разные производители оборудования по-разному трактуют некоторые его параметры. Обычно в стандартах группы *RS* нормированы электрические и электромеханические параметры, но не регламентированы жестко форматы передачи данных и протоколы стыка. В *RS422* передача осуществляется по балансным линиям, дифференциальное напряжение составляет от 2 до 7 В.

Область его использования заметно сокращается в пользу *SPI*. Схожие с *RS422* характеристики имеет нестандартизованный параллельный интерфейс компании *NDS* под названием «*NDS422* параллельный», у него также оригинальное распределение контактов и отсутствует сигнал подтверждения действительности данных.

Интерфейс *M2P* разработан и применяется компанией *Harmonic* (бывш. *DiviCom*), он использует тактовые, информационные и вспомогательные сигналы, аналогичные *SPI*, но уровни сигналов соответствуют *RS422*, а считывание данных производится по заднему фронту тактовых импульсов.

Последовательный интерфейс. При использовании *последовательного интерфейса* тактовая частота отдельно не передается, а выделяется на приеме из полезных данных.

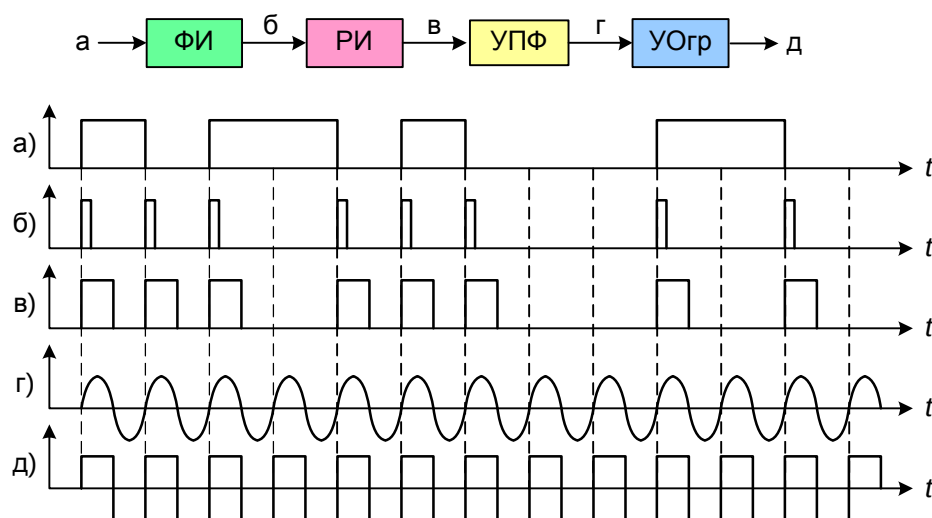


Рисунок 1.13 — Выделение тактовых импульсов из передаваемого последовательно цифрового сигнала

Один из простейших способов восстановления тактовой частоты при приеме цифрового сигнала, передаваемого по последовательному каналу, иллюстрируется структурной схемой, показанной на рисунок 1.13.

Принимаемый цифровой сигнал (рисунок 1.13,а), состоящий из последовательности «единичных» и «нулевых» уровней, поступает на формирователь импульсов (ФИ), вырабатывающий короткие импульсы на каждый положительный и отрицательный перепад напряжения в сигнале (рисунок 1.13,б). Расширитель импульсов (РИ) преобразует короткие импульсы в импульсы, длительность которых равна половине периода тактовой частоты (рисунок 1.13,в). Эти импульсы поступают на

узкополосный фильтр (УПФ), настроенный на тактовую частоту. На выходе фильтра выделяется синусоидальный сигнал тактовой частоты (рисунок 1.13,г), который затем преобразуется усилителем-ограничителем (УОгр) в прямоугольные импульсы (рисунок 1.13,д), используемые для тактирования принимаемого сигнала.

Как видно из временных диаграмм, в случае, если в принимаемом цифровом сигнале подряд идут несколько битов с одинаковыми значениями, то на выходе формирователя импульсов в течение этого отрезка времени нет сигналов, и выходные тактовые импульсы продолжают формироваться за счет наличия затухающего гармонического колебания на выходе узкополосного фильтра. Это накладывает ограничения на передаваемый сигнал, так как передача достаточно длинных последовательностей нулей или единиц приведет к прекращению формирования тактовых импульсов. Кроме того, в начале передачи цифрового сигнала амплитуда колебаний на выходе узкополосного фильтра нарастает постепенно, поэтому имеет место некоторая задержка до появления тактовых импульсов на выходе устройства синхронизации.

Для преодоления указанных трудностей в системах передачи цифровых сигналов по последовательным каналам связи выполняется дополнительное преобразование передаваемых данных (канальное кодирование), в результате которого количество передаваемых подряд нулей или единиц ограничивается.

В наиболее распространенном последовательном интерфейсе *SDI* (*Serial Digital Interface* — последовательный цифровой интерфейс), описанном в Рекомендации *BT.656* и стандарте *SMPTE-259M*, для повышения надежности выделения тактовой частоты используется инвертированный код БВН и сверточное скремблирование. В качестве исходного используется цифровой компонентный сигнал «4:2:2» в параллельном формате.

Входной 10-битовый сигнал записывается в параллельной форме в регистр сдвига с тактовой частотой 27 МГц, считывается последовательно с

тактовой частотой 270 МГц и поступает на скремблер. Генерация синхронной частоты 270 МГц осуществляется путем подстройки в схеме ФАПЧ частоты управляемого генератора по входной тактовой частоте 27 МГц. Специальный детектор синхрослов обнаруживает комбинации *SAV* и *EAV* и следит, чтобы они кодировались правильно независимо от разрядности передаваемого сигнала. Скремблирование осуществляется путем подачи на ЛЭ «исключающее ИЛИ» текущего бита и двух предшествующих, отстоящих от него на 5 и 9 тактовых интервалов. Соответствующий порождающий многочлен имеет вид: $g(x) = G_1 \cdot G_2$, где $G_1 = x^9 + x^4 + 1$, $G_2 = x + 1$. Второй сомножитель $G_2(x)$ описывает логическое инвертирование полученного сигнала, которое изменяет состояние сигнала при каждом появлении логической единицы (рисунок 1.14).

Информация в таком сигнале содержится в переходах, а не в самих импульсах, и потому не зависит от полярности сигнала. Это важное свойство *SDI*-сигнала, позволяющее использовать в канале как неинвертирующие, так и инвертирующие устройства.

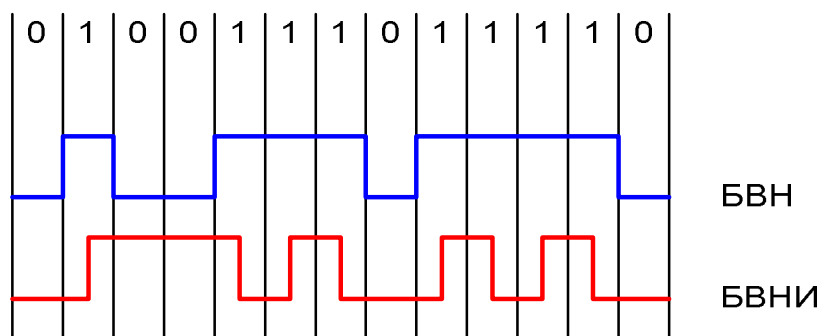


Рисунок 1.14 — Инвертирование кода БВН

На приемной стороне все операции производятся в обратном порядке. В приемном регистре сдвига обнаруживаются сигналы *SAV* и *EAV*, которые управляют процессом синхронизации и перезапускают последовательно-параллельный преобразователь.

Выделенный сигнал тактовой частоты делится на 10 и получившийся сигнал 27 МГц используется как сигнал цикловой синхронизации для считывания данных на выходе приемника. Как и в параллельном интерфейсе, интервал между *SAV* и *EAV* не используется для передачи видеоданных и может быть загружен другими видами информации, например, звукоданными.

Предшественником *SDI* был последовательный компонентный интерфейс с длиной кодового слова 8 бит, описанный в ранней версии Рекомендации BT.656. В этом интерфейсе 8-битовые слова кодировались в 9-битовые комбинации, которые затем передавались с тактовой частотой 243 МГц. Однако с постепенным переходом профессиональной цифровой видеотехники на 10-битовое кодирование он был вытеснен более удобным 10-битовым интерфейсом.

Интерфейс *SDI* широко используется для соединения различных устройств в пределах здания и между близко расположенными зданиями. Такой сигнал легко коммутировать, корректировать, контролировать. Соединение осуществляется по 75-омному коаксиальному кабелю с помощью простых и очень надежных соединителей *BNC*, обеспечивающих хорошее согласование в полосе частот до 600 МГц. С 75-омным кабелем можно использовать как 75-омные с тонким центральным проводником, так и 50-омные соединители *BNC*, обладающие большей механической прочностью, необходимо только следить, чтобы затухание несогласованности не падало ниже 15 дБ.

К недостаткам *SDI* можно отнести чувствительность к фазовому дрожанию. Оно измеряется в долях тактового интервала (ТИ), который для компонентного *SDI* сигнала составляет $1 / 270 \text{ МГц} = 3,7 \text{ нс}$. Низкочастотные (ниже 10 Гц) отклонения фронтов импульсов от номинальных временных позиций иногда называют *качанием* (*wander*) или *дрейфом* (*drift*), они в меньшей степени влияют на процесс синхронизации, так как стробирующий (тактовый) сигнал претерпевает те же смещения, что и стробируемая

сигнальная последовательность. В области частот выше 10 Гц различают фазовое дрожание двух типов — *общее дрожание (timing jitter)* как абсолютное смещение фронтов относительно номинала, и *высокочастотное дрожание, или дрожание синхронизации (alignment jitter)* — смещение фронтов импульсов относительно выделенного из сигнальной последовательности сигнала тактовой частоты. Нижняя частотная граница высокочастотного дрожания определяется полосой пропускания фильтра в цепи АПЧ. Рекомендация *ITU-P BT.1363* устанавливает для общего дрожания допустимое значение 0,2 ТИ (рассматривается возможность увеличения нормы до 1 ТИ) и для высокочастотного дрожания — не более 0,2 ТИ (740 пс). Увеличение высокочастотного дрожания свыше допустимого приводит к нарушению синхронизации и при приближении к 1500 пс к ее полному срыву. На рисунке 1.15 показаны пределы допустимого фазового дрожания на выходе *SDI* приемника.

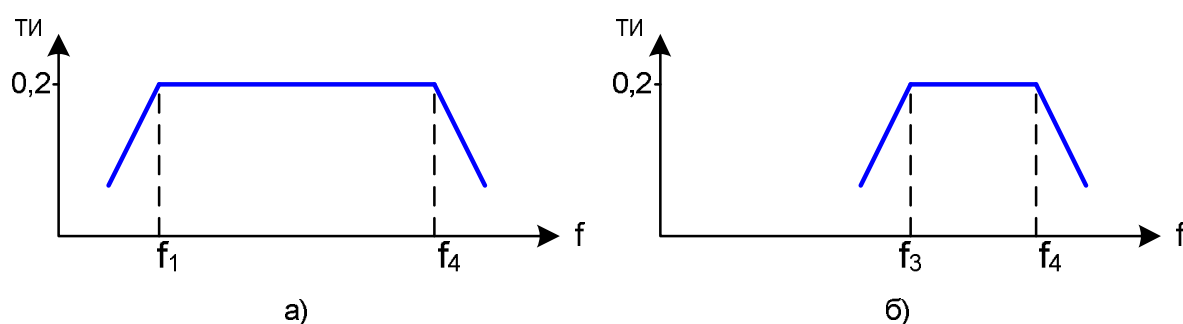


Рисунок 1.15 — Пределы допустимого фазового дрожания в сигнале *SDI*:
 а — общее дрожание; б — высокочастотное дрожание

Заслуживают упоминания еще несколько последовательных цифровых интерфейсов, используемых для передачи дополнительных данных, вводимых в цифровой телевизионный сигнал.

Асинхронный последовательный интерфейс *RS232*, как следует из его названия, имеет статус «рекомендованного стандарта» и может претерпевать небольшие отклонения в разных реализациях. Общепринятыми являются небалансное соединение, асинхронная дуплексная передача данных между

одним передатчиком и одним приемником, предельная скорость передачи данных 115,2 кбит/с. Практически используются скорости 19,2 и 38,4 кбит/с, максимальное расстояние 15 м. Логическая «1» передается напряжением отрицательной полярности амплитудой от 5 до 15 В, логический «0» таким же напряжением положительной полярности, соединитель 9-контактный типа *DB-9*, широко используемый в персональных компьютерах.

Синхронный последовательный интерфейс *RS422* также является «рекомендованным стандартом», но, в отличие от *RS232*, опирается на стандартизованный МСЭ электрический стык *V.11*. Передача осуществляется по балансной линии в дуплексном или полудуплексном режиме, разностное напряжение амплитудой от 2 до 6 В одной полярности соответствует логической «1», противоположной полярности — логическому «0». Максимальная скорость передачи данных 10 Мбит/с может поддерживаться на расстоянии до 15 м, со скоростью 100 кбит/с можно работать при длине кабеля до 1200 м. Предельная чувствительность приемника 200 мВ. В качестве соединителя может использоваться 9-контактный или 25-контактный разъем *D*-типа.

Последовательный интерфейс *M2S*, предложенный в свое время компанией *DiviCom* (ныне *Harmonic*), применялся только в аппаратуре этой компании и сейчас практически вышел из употребления. Используя коаксиальную линию и тактовую частоту 270 Мбит/с, он допускает скорость передачи полезных данных до 108 Мбит/с.

Весьма перспективным специалисты считают предложенный компанией Sony и стандартизованный в *SMPTE-305M* интерфейс *SDTI* (*Serial Data Transport Interface* — последовательный интерфейс передачи данных). Он использует структуру сигнала *SDI* (рисунок 1.16), однако укладывает в область полезной нагрузки не только 10-битные слова видеоданных, но и любые блоки данных, состоящие из 8-битных или 9-битных слов, в том числе пакетированные блоки компрессированных данных фиксированной или переменной длины.

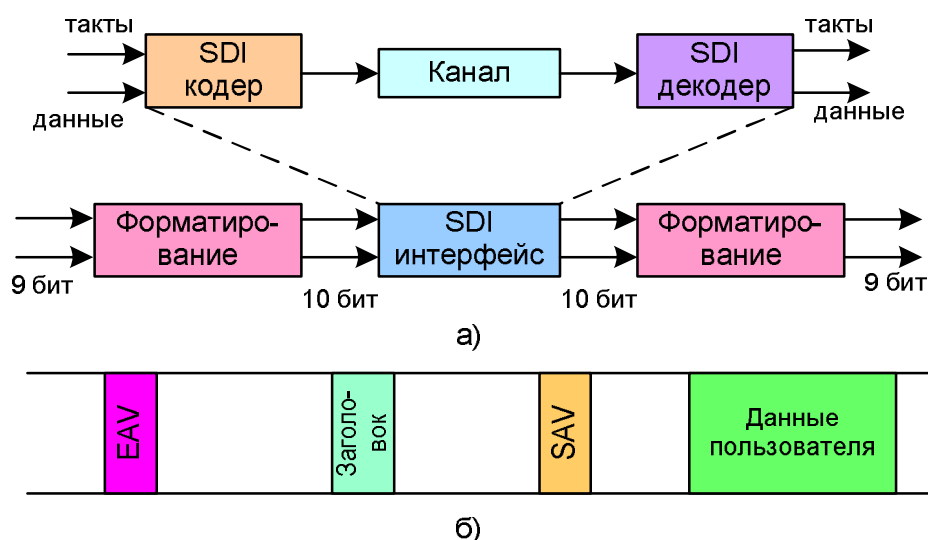


Рисунок 1.16 — Последовательный интерфейс *SDTI*:
а — структурная схема; б — формат цифровой строки

Каждый блок имеет заголовок, в котором указаны размер, тип блока, кроме того, в области дополнительных данных между *EAV* и *SAV* размещается заголовок строки, полностью описывающий ее структуру и состав размещаемых данных. Скорость передачи полезных данных достигает 200 Мбит/с в потоке 270 Мбит/с и 270 Мбит/с в потоке 360 Мбит/с. В 2000 году разработан новый стандарт *SMPTE-305M.2*, уточняющий и дополняющий базовый документ.

1.8. Кодирование телевизионного сигнала

Импульсно-кодовой модуляции в телевидении присуща значительная избыточность в передаваемой информации. Информационная избыточность заключена в специфике ТВ сигнала. Эту избыточность подразделяют на статистическую, физиологическую и структурную.

Статистическая избыточность определяется свойствами изображения, которое не является в общем случае хаотическим распределением яркости, а описывается законами, устанавливающими определенные связи (корреляцию) как между яркостями отдельных элементов, так и между их цветностями.

Визуальная или *физиологическая избыточность* ТВ сигнала обуславливается ограниченностью возможностей зрительного аппарата. Использовать физиологическую избыточность — значит, не передавать в сигнале ту информацию, которая не будет воспринята нашим зрением. Как следствие, устранение физиологической избыточности связано с частичной потерей информации, но той информации, которая не воспринимается зрением, и ее потеря не повлияет на визуальное качество изображения.

Структурная избыточность в соответствии с названием определяется структурой видеосигнала: наличием в аналоговом телевизионном сигнале гасящих импульсов. В цифровом потоке соответствующие им временные интервалы можно исключить или использовать для передачи другой информации, например звукового сопровождения.

Уменьшение цифрового потока ТВ сигнала за счет сокращения статистической и физиологической избыточности в изображении осуществляется в телевидении применением более эффективных методов кодирования по сравнению с ИКМ. При большом их многообразии наиболее широко распространены следующие виды эффективного кодирования: кодирование с предсказанием, кодирование с линейным ортогональным преобразованием, взвешенное квантование, энтропийное кодирование или кодирование с переменной длиной. Перечисленные виды кодирования могут использоваться как самостоятельно, так и в сочетании друг с другом, давая в последнем случае значительно больший эффект при компрессии цифровых сигналов.

В настоящее время существуют стандарты компрессии, которые определяют основные правила эффективного кодирования и декодирования цифровых потоков, как изображений, так и связанного с ними звукового сопровождения. Эти стандарты получили общее наименование MPEG, представляющее собой аббревиатуру от названия, разрабатывавшего их международного комитета — *Moving Pictures Experts Group* (Группа экспертов по движущимся изображениям).

Избыточность информации в потоке цифровых данных объясняется, прежде всего, спецификой ТВ изображения. Известно, что, несмотря на равновероятность любых из возможных значений яркости (цветности) для одного элемента изображения, содержание соседствующих с ним элементов мало отличается или не отличается вовсе. При поэлементной передаче яркости и цветности методом ИКМ в канал посылается одна и та же или мало отличающаяся по содержанию информация. Подобную избыточность информации в сигнале называют *пространственной* или *внутрикадровой*.

Значительная избыточность содержится и в передаче смежных во времени кадров изображения, в которых, несмотря на движение отдельных фрагментов, существенную роль играют общие для них неподвижный фон или задний план. Такого рода избыточность называют *временной* или *межкадровой*. И пространственная, и временная избыточность обусловлены статистическими свойствами телевизионного изображения.

Устранение избыточности информации в цифровом сигнале, в конечном счете, должно выразиться в уменьшении числа отсчетов сигнала и (или) уменьшении разрядности их двоичных кодовых символов. Реализация таких процедур напрямую с ИКМ сигналом недопустима, так как изъятие из сигнала отдельных отсчетов или замена их значений на более грубо проквантованные соответствует изъятию в воспроизводимом изображении соответствующих элементов и их искаженному представлению. Попытка изъятия отдельных элементов из группы равноправных слагаемых изображения, как с информационной, так и с энергетической точки зрения приводит к необратимой потере качества изображения.

Решение задачи по устранению избыточности информации в цифровом сигнале возможно только путем его предварительной обработки, которая должна перераспределить вклад отдельных отсчетов преобразованного сигнала в общее содержание изображения. Перераспределив, таким образом, функциональную значимость между отдельными отсчетами, можно будет выделить «главные» отсчеты, несущие основной объем информации,

обеспечив им наилучшие условия передачи, а на остальных отсчетах «сэкономить», не передавая их или передавая с минимальным числом градаций.

К настоящему времени предложено сравнительно много методов обработки сигнала с таким подходом.

1.9. Дискретно-косинусное преобразование

Методы преобразования изображения в общем случае основаны на том, что его цифровой эквивалент (сигнал ИКМ) приводится к виду, удобному для сокращения избыточной информации. В этом отношении наиболее эффективным является преобразование видеoinформации из временной области в спектральную. Это преобразование, как правило, предваряется разбивкой изображения на частичные подобласти, фрагменты (в терминологии *MPEG* — *блоки*), которые затем по отдельности подвергаются необходимой обработке. Результат преобразования представляет собой совокупность спектральных коэффициентов, которые характеризуют амплитуды пространственных частот изображения.

В основу преобразования изображений могут быть положены различные приемы. Наиболее часто используются методы линейных ортогональных преобразований. Линейность преобразований означает, что операции сложения, вычитания и умножения на скаляр действительны в после преобразований, а ортогональность — что преобразуемый фрагмент представляется ограниченным набором ортогональных функций. Линейные преобразования можно осуществлять как с непрерывным, так и с дискретным сигналом. В первом случае процессу преобразования соответствует интегральная форма записи, во втором — матричная.

Из различных ортогональных преобразований, позволяющих эффективно выявлять избыточную информацию, стандартом *MPEG* рекомендовано использовать *дискретно-косинусное преобразование* (ДКП), являющееся частным случаем двумерного преобразования Фурье. Как известно, преобразование Фурье — это метод обработки, который,

анализируя изменения сигнала во времени, выражает их в виде частотного спектра. Любой сигнал можно разложить на частотные гармонические составляющие, и затем по известным значениям амплитуды и фазы этих составляющих их линейным суммированием восстановить исходный сигнал. Последнюю операцию называют *обратным преобразованием Фурье*. В цифровых системах сигнал выражается последовательностью дискретных отсчетов. При использовании преобразования Фурье для фрагмента цифрового сигнала из некоторого ограниченного числа отсчетов последний можно разложить на такое же число дискретных частот. Это преобразование называют дискретным преобразованием Фурье.

Поскольку любое изображение или его фрагмент можно рассматривать как функцию изменения яркости (цветности) как по оси X , так и по оси Y , то дискретное ортогональное преобразование Фурье будет представлять собой замену массива отсчетов изображения соответствующего фрагмента на массив коэффициентов, соответствующих амплитудам частотных составляющих Фурье.

Объем расчетов для нахождения этих коэффициентов весьма значителен. Поэтому преобразования осуществляются над небольшими по размеру фрагментами, обычно 8×8 элементов. Дискретно-косинусное преобразование Фурье в определенной степени минимизирует объем этих вычислений использованием в качестве набора преобразующих (базисных) функций только косинусных составляющих. В результате массиву исходных значений сигнала соответствует массив из такого же числа коэффициентов, представляющих собой амплитуды этих косинусных составляющих.

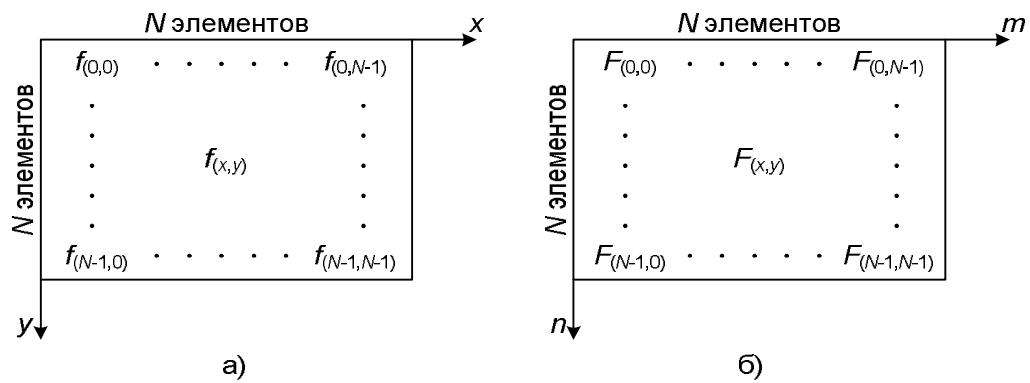


Рисунок 1.17 — Преобразование блока изображения $f_{(x,y)}$ в блок ДКП коэффициентов $F_{(m,n)}$:

а — блок изображения; б — блок коэффициентов ДКП

Аналитически двумерное дискретно-косинусное преобразование описывается следующим образом (рисунок 1.17):

$$F_{(m,n)} = \frac{2}{n} C_{(m)} C_{(n)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f_{(x,y)} \cos \frac{pm(2x+1)}{2N-1} \cos \frac{pn(2y+1)}{2N-1},$$

где $C_{(m)} = 1$ при $m \neq 1$; $C_{(m)} = 1/\sqrt{2}$ при $m = 0$; $C_{(n)} = 1$ при $n \neq 1$; $C_{(n)} = 1/\sqrt{2}$ при $n = 0$; $f_{(x,y)}$ — отсчеты изображения с пространственными координатами x, y (от 0 до $N-1$); N — размер блока изображения ($N \times N$ элементов); $F_{(m,n)}$ — коэффициенты, характеризующие изображение в спектральной плоскости m, N (от 0 до $N-1$).

ДКП является обратимым: по распределению $F_{(m,n)}$ обратным преобразованием однозначно восстанавливается $f_{(x,y)}$.

Очевидно, что поскольку число отсчетов преобразуемого сигнала равно числу отсчетов преобразованного сигнала, устранения избыточности информации в результате такого преобразования не происходит. Однако следует обратить внимание на значительное изменение содержания блока коэффициентов ДКП по отношению к блоку преобразуемого изображения (рисунок 1.18).

59	59	59	60	60	65	64	64
63	62	62	62	62	61	61	62
137	123	111	101	96	89	88	86
237	236	235	233	231	216	213	208
225	229	232	232	231	237	238	239
193	195	197	198	199	204	204	205
182	182	181	181	181	180	180	180
183	182	181	180	179	178	178	177

а)

1249	19	3	1	1	1	0	1
-381	14	3	2	2	0	0	1
-318	-14	3	1	-1	0	1	-2
31	-45	-4	-3	-5	0	2	-4
154	-7	-8	-2	-2	0	-1	0
38	20	-3	2	2	0	-2	2
-39	11	8	3	0	1	1	0
-42	3	10	1	-1	1	1	-1

б)

Рисунок 1.18 — Пример дискретно-косинусного преобразования для некоторого произвольного сюжета:
а — блок изображения; б — блок коэффициентов ДКП

По физическому смыслу блок коэффициентов ДКП представляет собой совокупность значений амплитуд пространственных косинусоидальных гармоник с частотами m и n . При этом значение $F_{(0,0)}$ пропорционально среднему уровню (постоянной составляющей) в блоке и может достигать при 256 уровнях квантования значения 2040. (Чтобы ошибки от округления коэффициентов ДКП не сказывались существенным образом на точности преобразования, их значения на этапе преобразования увеличены в восемь раз по сравнению с их действительными значениями).

Компоненты $F_{(0,1)}$ и $F_{(1,0)}$ характеризуют плавное изменение яркости в блоке вдоль строки и поперек строк соответственно. Разночастотные изменения яркостей пикселей с диагональными структурами характеризуются диагональными спектральными компонентами $F_{(1,1)}$, $F_{(1,2)}$, $F_{(1,3)} \dots$

Обычно для большинства блоков изображения лишь малая часть коэффициентов имеет значительную величину. Это объясняется небольшими размерами блока, внутри которого яркость меняется мало, и поэтому относительно большие величины имеют только постоянная составляющая и несколько низкочастотных компонентов, расположенных в левом верхнем углу матрицы коэффициентов ДКП.

Мелким деталям изображения, как известно, соответствуют высокие пространственные частоты, и коэффициенты ДКП характеризующие их

амплитуды, располагаются правее и ниже. Поскольку, как правило, мелкие детали изображения выражены энергетически слабо относительно среднего уровня, то и соответствующие им коэффициенты ДКП по сравнению с $F_{(0,0)}$ малы или вообще равны нулю.

Таким образом, если передавать вместо значений яркости изображения коэффициенты ДКП, то сокращение скорости передачи данных может быть достигнуто уже хотя бы за счет исключения нулевых коэффициентов. Однако эта задача решается уже вне процедуры ДКП.

1.10. Квантование коэффициентов ДКП

Выгодное для компрессии различие в амплитудах между отдельными коэффициентами ДКП может быть еще более усилено за счет устранения психофизической избыточности в изображении. Как следствие, увеличится число нулевых коэффициентов и коэффициентов с малыми значениями. Эта задача решается в процессе квантования коэффициентов, полученных после ДКП.

Установлено, что глаз более чувствителен к ошибкам передачи яркости и цветности на больших площадях, в то время как при передаче контуров и мелких деталей остаются незамеченными более серьезные ошибки. Отсюда вытекает возможность определенного огрубления значений коэффициентов ДКП, отвечающих за передачу мелких деталей и контуров, без возникновения заметных для глаза искажений в изображении.

С этой целью производят процедуру квантования коэффициентов ДКП блока на разное число уровней: коэффициенты, расположенные в левом верхнем углу блока, квантуются на максимально большое число уровней (особенно это касается коэффициента $F_{(0,0)}$, отвечающего за среднюю яркость блока); остальные коэффициенты передаются с меньшей точностью, а значит, квантуются на меньшее число уровней. Для тех же из них, что располагаются в правом нижнем углу, шкала квантования может содержать всего несколько уровней. Практическая реализация процесса квантования

достигается поэлементным делением матрицы коэффициентов ДКП на матрицу квантования.

В приемном устройстве, прежде чем осуществить обратное дискретно-косинусное преобразование для восстановления исходного изображения, матрица ДКП умножается на матрицу квантования. Эта операция называется *деквантованием*. Очевидно, что после деквантования возвратиться к исходному, неквантованному блоку ДКП уже нельзя. Ошибки, возникающие от округления квантуемых величин, в связанные с ними искажения в изображении необратимы. Отсюда вытекает необходимость отыскания таких матриц квантования, которые не приводили бы к визуально заметным искажениям.

На рисунке 1.19 показана матрица квантования, используемая стандартом *MPEG*. Учитывая, что значения большинства коэффициентов ДКП в блоке весьма малы, деление их на числа, характеризуемые почти двумя порядками, приводит или к обнулению многих коэффициентов, или к сильному их уменьшению (рисунок 1.20). Это в свою очередь позволит при передаче проквантованных значений коэффициентов ДКП по каналу связи значительно уменьшить скорость цифрового потока.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Рисунок 1.19 – Матрица взвешенного квантования коэффициентов ДКП

156	1	0	0	0	0	0	0
-24	1	0	0	0	0	0	0
-17	-1	0	0	0	0	0	0
1	-2	0	0	0	0	0	0
7	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0
-2	0	0	0	0	0	0	0
-2	0	0	0	0	0	0	0

Рисунок 1.20 – Значение коэффициентов ДКП, полученные делением матрицы рисунка 1.18,б на матрицу квантования рисунка 1.19

1.11. Кодирование коэффициентов ДКП

Следующим шагом после квантования коэффициентов ДКП является преобразование матрицы этих коэффициентов в одномерную последовательность. Именно здесь окончательно реализуется процесс устранения избыточности, подготовка к которому проводилась на рассмотренных выше этапах ДКП и взвешенного квантования. Данное преобразование предусматривает объединение коэффициентов матрицы в определенные группы и применение затем так называемого энтропийного кодирования.

Алгоритм группирования (упорядочивания) коэффициентов ДКП существенно влияет на эффективность компрессии: в процессе сканирования преобразуемой во временной последовательность чисел матрицы нулевые коэффициенты объединяются в максимально длинные серии. Тогда их описание может сводиться к лаконичной записи длины серии и ее местоположения в матрице. Одним из вариантов такого алгоритма группирования является *зигзагообразное сканирование*, при котором преобразование начинается с левого верхнего угла матрицы и заканчивается в ее правом нижнем углу (рисунок 1.21).

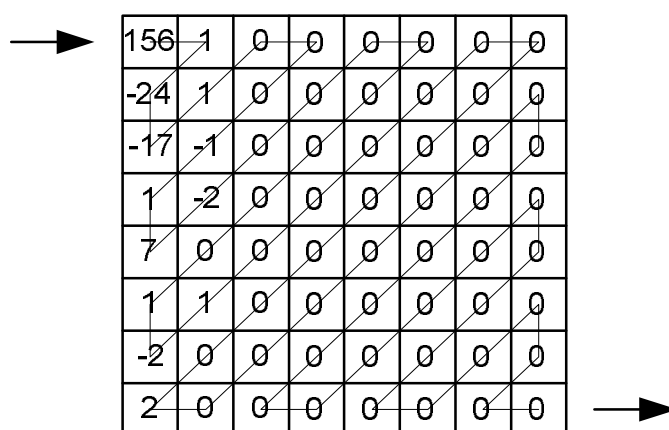


Рисунок 1.21 — Зигзаг-сканирование коэффициентов ДКП

Поскольку именно в правом нижнем углу сосредоточено большинство нулевых коэффициентов, такой порядок сканирования обеспечивает формирование наиболее длинных серий нулей, а, следовательно, и самую компактную форму их передачи.

Полученная в результате сканирования последовательность чисел подвергается упомянутому выше энтропийному кодированию или кодированию с переменной длиной слова. Наиболее употребимым из энтропийных кодов является код Хаффмана. Он основывается на том, что коды символов, обладающих большей вероятностью, описываются меньшим числом бит, чем коды символов с меньшей вероятностью. Как было показано, после взвешенного квантования матрицы ДКП в последней преобладают числа с малыми амплитудами, и их целесообразно кодировать короткими словами. Большие амплитуды, характерные для левого верхнего угла матрицы, по сравнению с другими значениями коэффициентов встречаются реже, и им можно приписать символы с большим числом разрядов.

Эффективность энтропийного кода Хаффмана повышается также за счет того, что не требуется разделителей между символами. И хотя последние имеют различную битовую длину, они декодируются единственным образом.

1.12. Кодирование с предсказанием

Принцип действия систем кодирования с предсказанием заключается в передаче вместо истинного значения сигнала закодированной разности истинного и предсказанного значений. В соответствии с этим принципом подобные системы кодирования получили еще одно название — системы с *дифференциальной импульсно-кодовой модуляцией* ДИКМ.

Дифференциальная импульсно-кодовая модуляция. ДИКМ представляет собой еще один из методов эффективного кодирования цифрового телевизионного сигнала и применяется при сокращении его временной

избыточности. Как известно, в большинстве случаев содержание двух соседних во времени телевизионных кадров мало отличается друг от друга. Отличие заключается только в положении отдельных подвижных фрагментов изображения. В основном же в этих кадрах содержится очень много одинаковых областей, информацию о которых многократно передавать нецелесообразно. Можно ограничиться передачей значений только одного кадра, а содержание последующих, используя статистические законы, предсказать, т.е. вычислить на приемном конце системы.

Однако как бы ни были близки по содержанию соседние кадры, как бы ни был совершенен их статистический анализ, предсказание может нести в себе ошибку. Эта ошибка для каждого элемента изображения обязательно учитывается и корректируется. Только при этом условии на приемном конце системы воссоздается изображение, соответствующее оригиналу.

Отсюда вытекает следующий принцип кодирования сигнала с предсказанием: передача в каждом отсчете кодированной разности между текущим истинным значением сигнала и предсказанным. Эту разность называют ошибкой предсказания. Чем точнее сделано предсказание, тем меньший объем данных будет содержаться в подлежащем передаче разностном сигнале. Логично ожидать, что этот объем будет меньше, чем в полном отсчете. Для формирования сигнала ошибки предсказания, на передающем конце системы устанавливается устройство предсказания и вычитающий каскад (рисунок 1.22).

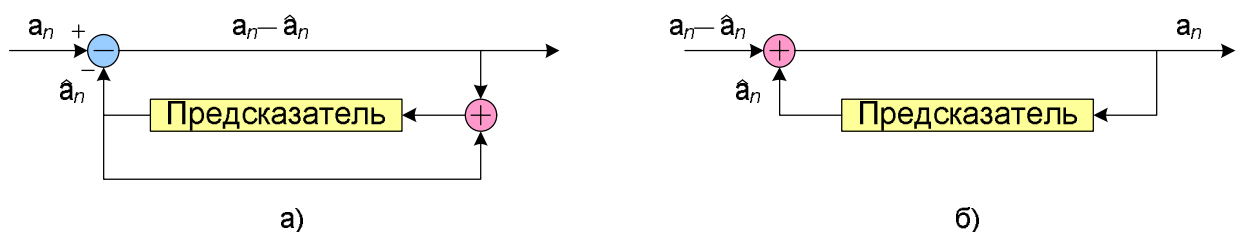


Рисунок 1.22 — Структура системы ДИКМ: а — кодер; б — декодер;

a_n — текущее значение сигнала;

$\hat{a}_n$ — предсказанное значение сигнала; $(a_n - \hat{a}_n)$ — сигнал ошибки

На приемном конце системы исходный сигнал восстанавливается сложением предсказанного и принятого сигнала ошибки предсказания. Устройства предсказания на обоих концах системы для правильного восстановления в приемнике исходных значений сигнала должны быть одинаковыми.

Предсказание, применяемое в системе, может быть организовано различными способами и с разной степенью сложности. Однако вполне хорошие результаты по глубине достигаемой компрессии реализуются весьма простым способом, когда в качестве сигнала предсказания используются значения предыдущего кадра.

Статистические исследования показали, что свойства телевизионного изображения, обусловленные межкадровыми связями, в целом аналогичны пространственным свойствам в неподвижном изображении. А коэффициенты корреляции в соседних кадрах получаются зачастую даже большими, чем для соседних пикселей в одном кадре. Отсюда следует вывод о целесообразности обработки разностного сигнала ДИКМ рассмотренными ранее способами для внутрикадровой обработки, к каковым относятся ДКП, взвешенное квантование и энтропийное кодирование. В этом случае структура системы ДИКМ приобретает вид, изображенный на рисунке 1.23 и работает следующим образом.

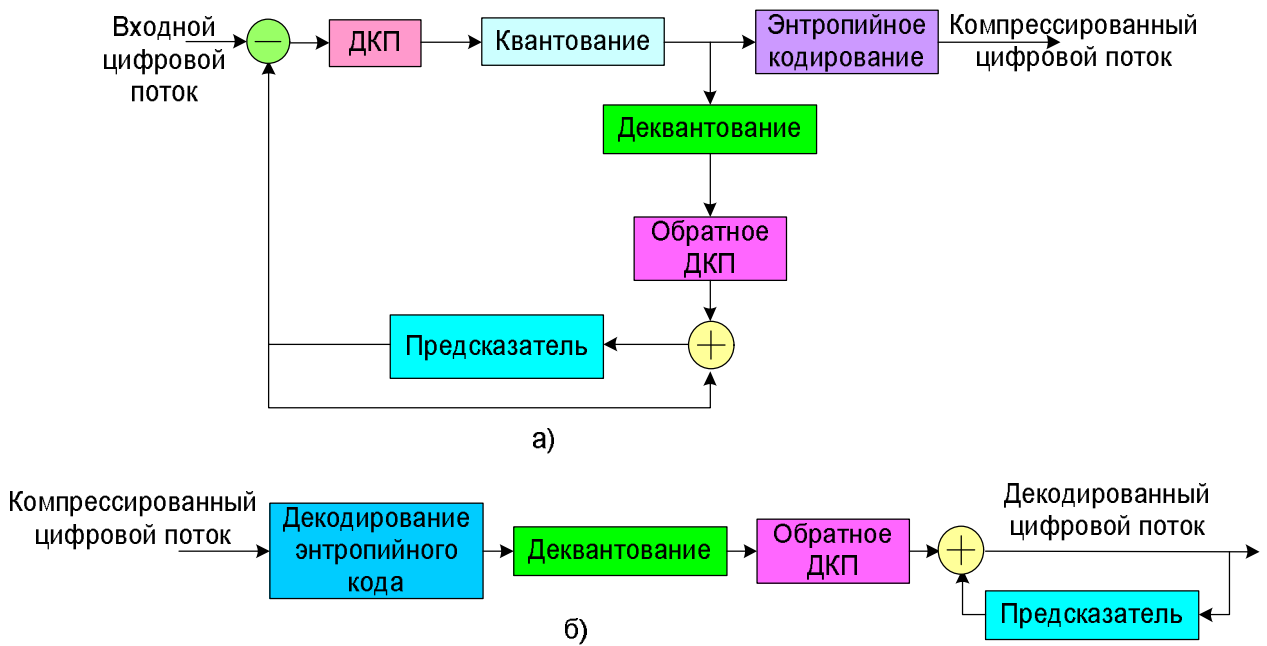


Рисунок 1.23 — Структура ДИКМ, совмещенная с процедурами ДКП и квантования: а — кодер; б — декодер

Для первого кадра из последовательности кадров, подлежащих преобразованию, предсказание не может быть выполнено из-за отсутствия каких-либо априорных сведений (отсутствия предшествующих ему кадров). Поэтому предсказанные значения отсчетов для первого кадра можно приравнять пулю, и тогда сигнал ошибки предсказания будет представлять собой непосредственно отсчеты этого первого кадра. Следующее за вычитающим каскадом устройство осуществляет ДКП отсчетов первого кадра, т.е. выявляет в нем пространственную избыточность. Дальнейшие процедуры над коэффициентами ДКП описаны в предыдущих разделах (взвешенное квантование и энтропийное кодирование). Компрессированный цифровой поток через соответствующие устройства передачи направляется в приемное устройство (декодер), где осуществляются обратные операции.

Второй и последующие кадры передаются по алгоритму ДИКМ. Для этого в кодере формируется сигнал предсказания. В качестве него можно использовать отсчеты предыдущего кадра. Однако следует отметить, что истинные значения предыдущего кадра, полученные в кодере, например,

задержкой на время кадра, не пригодны в качестве предсказания. Дело в том, что предыдущий кадр, используемый в качестве предсказания в декодере, отличается от своего истинного значения, так как в процессе перечисленных выше преобразований он реконструируется в декодере с определенной погрешностью. А поскольку предсказываемые значения на обоих концах системы должны быть одинаковы, то формирование предсказания в кодере проводится по тому же алгоритму, что и в декодере. С этой целью в петлю обратной связи кодера, где формируется предсказание, включается деквантователь и процессор обратного преобразования Фурье. Энтропийное кодирование в общем случае не вносит ошибок, а значит, не влияет на формирование сигнала предсказания. Предсказатель, таким образом, представляет собой задержку на время кадра отсчетов, Воспроизводимых на приемном конце системы.

Первый кадр рассмотренной преобразуемой последовательности кадров называется опорным. Из него устранена только пространственная избыточность. По сравнению с ним последующие кадры компрессированы значительно сильнее. Как следствие, они менее помехоустойчивы. Для поддержания необходимой помехоустойчивости системы и сведения к минимуму возможных сбоев в ее работе в последовательность разностных сигналов ДИКМ периодически врезаются опорные кадры. Их периодичность в стандарте MPEG составляет около 0,5 с.

Разностный сигнал и сигнал опорного кадра после ДКП квантуется по разным алгоритмам. Если для опорного кадра используется матрица квантования для внутрикадрового кодирования, то для разностного сигнала применяется другая, более однородная матрица. Это объясняется тем, что разностный сигнал в основном несет информацию о небольших изменениях в отдельных фрагментах изображения (в процессе их перемещения). И высокие частоты (коэффициенты ДКП правого нижнего угла), ответственные за детализацию изображения, столь же важны, как и низкие. Конкретный вид матрицы в зависимости от сюжета может меняться, но по умолчанию

используется однородная матрица квантования с постоянным значением 16 для всех частот, включая нулевую.

Компенсация движения в динамических изображениях. Из принципа работы ДИКМ следует, что чем ближе по содержанию соседние кадры, тем меньше в разностном сигнале информации, а значит, тем эффективнее компрессия. Всякое изменение положения отдельных фрагментов изображения в результате их собственного движения или перемещения передающей ТВ камеры снижает эффективность кодирования. Сохранить эффективность работы ДИКМ можно соответствующим анализом характера движения. Целью анализа является возможность установления соответствия между предыдущим по времени блоком изображения в блоком в текущем кадре, смещенным в пространстве на координаты Δx , Δy . Под соответствием понимается достижение максимально точного равенства яркостей пикселей некоторого блока в текущем кадре яркостям пикселей блока, найденного в предшествующем кадре:

$$b_{(x+\Delta x, y+\Delta y, t+\Delta t)} = b_{(x, y, t)}, \quad (1.6)$$

где b — яркость пикселей;

Δt — время между сравниваемыми кадрами.

Разницу в координатах ($\Delta x, \Delta y$) этих двух блоков называют *вектором движения*. Нахождение вектора движения, другими словами, есть определение нового положения подвижного фрагмента изображения.

Под *компенсацией движения* понимается использование в качестве предсказания блока из предшествующего (или последующего) кадра, но при этом смещенного в плоскости изображения на величину вектора движения. Такое предсказание позволяет значительно уменьшить ошибку предсказания по сравнению с обычной ДИКМ. Это иллюстрируется рисунком 1.24, где показана последовательность из трех кадров изображения с перемещающимся серым квадратом на фоне гладкого белого фона (рисунок 1.24,а).

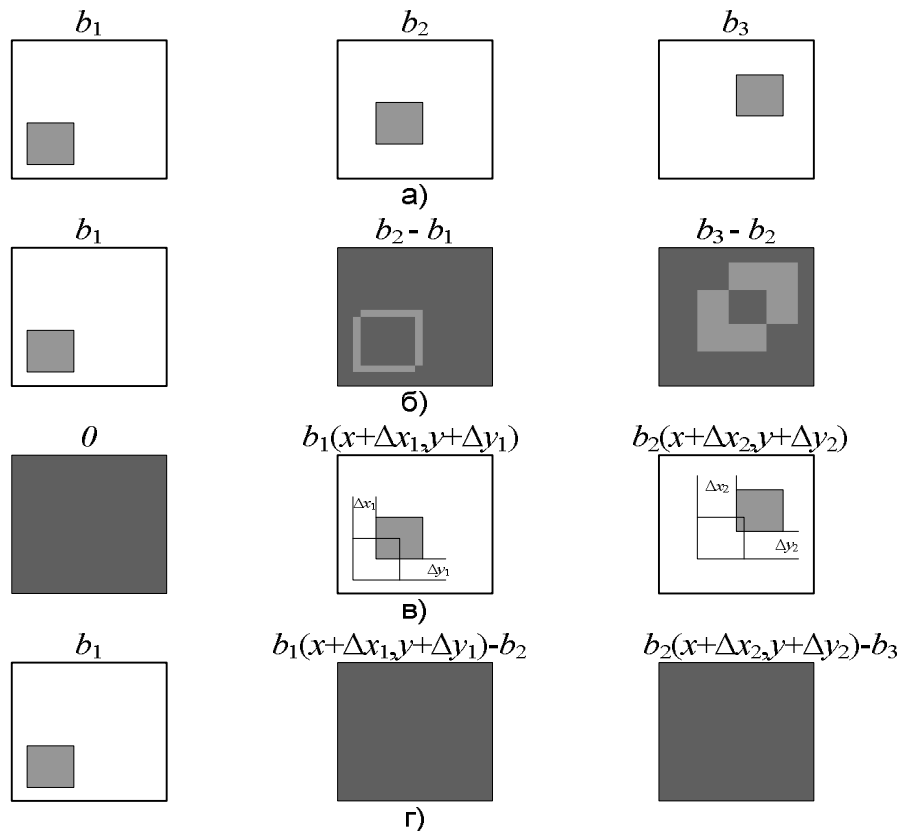


Рисунок 1.24 — Сравнение обычной ДИКМ и ДИКМ с компенсацией движения:

а — исходная последовательность изображений; б — ошибка предсказания при обычной ДИКМ; в — предсказание с учетом вектора движения; г — ошибка предсказания с компенсацией движения

Первый кадр b_1 в этой последовательности является опорным. Как отмечалось, предсказание для него не делается (сигнал предсказания равен нулю для всего изображения), и он подвергается только внутрикадровой компрессии. Предсказанием для второго кадра b_2 является реконструируемый по алгоритму ДКП и взвешенного квантования первый кадр b_1 . Ошибка предсказания для второго кадра в обычной ДИКМ равна $b_2 - b_1$, для третьего кадра соответственно $b_3 - b_2$ и т.д. Для простоты следует пренебречь небольшой разницей между истинными значениями отсчетов $b_1, b_2, b_3, \dots$ и их реконструируемыми значениями $b_1, b_2, b_3, \dots$. Тогда ошибки предсказания для обычной ДИКМ будут иметь вид рисунок 1.24,б, где черный фон соответствует нулевым значениям, а серый остальному множеству отсчетов. Если бы в предложенном примере движение отсутствовало, то ошибка

предсказания для второго и последующих кадров равнялась бы нулю для всего изображения. Именно такого результата для данного примера с движущимся фрагментом изображения, а не статическим удастся добиться в ДИКМ с компенсацией движения.

В этом случае предсказание для первого (опорного) кадра, как и при обычной ДИКМ, равно нулю, а для второго и последующих кадров соответственно — $b_1(x + \Delta x_1, y + \Delta y_1)$, $b_2(x + \Delta x_2, y + \Delta y_2)$. Здесь Δx_1 , Δy_1 , Δx_2 , Δy_2 и т.д. — изменение координат движущегося фрагмента (векторы движения) в последовательности кадров (рисунок 1.24,в). В результате ошибка предсказания для приведенного примера с подвижным фрагментом изображения в ДИКМ с компенсацией движения обращается в ноль для всего множества отсчетов изображения (рисунок 1.24,г), что иллюстрирует более эффективную компрессию.

При реализации компрессии с компенсацией движения сигнал ошибки предсказания должен дополняться передачей данных о векторах движения. Последние, так же как и коэффициенты ДКП, компрессируются использованием кодов переменной длины.

Абсолютное равенство в (1.6) может быть достигнуто только при выполнении следующих условий: движение объекта в изображении исчерпывается только плоскопараллельным перемещением, его яркость постоянна, градиент яркости фона равен нулю, и все объекты в блоке движутся с постоянной скоростью. Именно этим условиям отвечает пример изображения, приведенный на рисунке 1.24, и соответственно предельно достижимый результат компрессии (сигнал ошибки в системе с компенсацией движения равен нулю). Нарушение перечисленных выше условий уменьшает эффективность компрессии. К возможным факторам, влияющим на эффективность кодирования, следует отнести и осложнения глобального порядка, когда превосходит резкое изменение сюжета (например, при коммутации одновременно работающих камер). При этом изменения в изображении охватывают все пространство в кадре, и блоков,

соответствующих друг другу с удовлетворительной точностью, в смежных кадрах не существует. Эффективность компрессии в таких экстремальных случаях может снижаться настолько, что целесообразно отказываться от процедуры ДИКМ, и для таких изображений данные передавать только на основе внутрикадровой компрессии. Решение о замене способа кодирования должно приниматься в кодере автоматически, путем оперативного анализа возможных результатов компрессии.

Самым трудоемким при реализации алгоритма компенсации движения является *согласование блоков*, то есть нахождение блоков в разных кадрах, наилучшим образом соответствующих условию (1.6). Для этого текущий блок сравнивается со всеми блоками предыдущего кадра в пределах некоторой области поиска: ± 128 шагов (пикселей) при точности определения вектора движения в один пиксель или ± 64 шага при точности $1/2$ пикселя.

Стандарт не ограничивает методы согласования, но наиболее точным принято считать так называемый полный поиск (full search), который сводится к расчету ошибки предсказания для всех возможных векторов и определению положения (т.е. нахождения x и y) блока с минимальной ошибкой (например, по среднеквадратическому отклонению для всех отсчетов блока). Существуют и другие методы, отличающиеся большей скоростью определения векторов движения, но несколько меньшей точностью. К ним, в частности, относятся логарифмический и телескопический методы.

1.13. Предфильтрация и шумоподавление

Для эффективного сжатия необходимо устранить из исходного сигнала все факторы, препятствующие сокращению избыточности телевизионного изображения. Сюда относятся в первую очередь шумы видеотракта и шумы киноплёнки, проявляющиеся на всем изображении. Источником шумов в видеотракте могут быть тепловой «белый» шум электронных устройств, пороговые шумы спутниковых ЧМ приемников, помехи декодирования аналоговых композитных сигналов, шумы видеозаписи (шумы и выпадения

ленты). Один из крайне неприятных для цифровой компрессии дефектов исходного изображения — неподавленный остаток несущей цветности после аналого-цифрового преобразования, он проявляется как раз в высокочастотной области, которая обычно слабо заполнена в компонентном сигнале. Отрицательно сказываются дефекты киноплёнки — царапины, загрязнения, шумы зернистости, проявляющиеся при увеличении кинокадров. Необходимо избегать нежелательных движений, добавляющих в сигнал помехи, например дрожания плохо закрепленной видеокамеры или неравномерного движения киноплёнки в телекинопроекторе из-за поврежденной перфорации. Влияние этих составляющих дает себя знать в появлении после ДКП большого числа ненулевых коэффициентов в области высоких пространственных частот, которые кодер не может отличить от элементов изображения. В этих условиях необходимо подвергать входные сигналы кодеров сжатия специальной предобработке. Она называется *предфильтрацией* и включает шумоподавление и различные виды пространственной фильтрации. В некоторых случаях при высокой степени сжатия оказывается выгоднее осуществить двумерную низкочастотную фильтрацию (прореживание отсчетов), несколько потерять в четкости изображения, но существенно снизить количество артефактов (дефектов изображения) из-за влияния нежелательных высокочастотных компонентов на входе. В общем случае хорошие результаты дает сочетание рекурсивной адаптивной фильтрации и трехмерной медианной фильтрации. Первая эффективно подавляет «белый» шум путем усреднения близких значений видеосигнала в последовательных кадрах, вторая усредняет отсчеты, расположенные выше, ниже, слева, справа, до и после текущего отсчета видеосигнала, и таким образом хорошо справляется с импульсными и другими помехами.

2. СТАНДАРТ СЖАТИЯ ДВИЖУЩИХСЯ ИЗОБРАЖЕНИЙ И ЗВУКОВОГО СОПРОВОЖДЕНИЯ *MPEG-2*

2.1. Общие сведения

Стандарты сжатия движущихся изображений *MPEG* (*Moving Picture Experts Group*) вырабатываются и принимаются имеющей такое же название группой экспертов при Международной организации стандартизации *ISO*. Стандарт *MPEG-1*, используемый в основном при записи видеопрограмм на компакт-диски, был окончательно утвержден в 1993 г., а стандарт *MPEG-2*, предназначенный в первую очередь для телевизионного вещания, был принят в ноябре 1994 г.

Стандарты *MPEG-1* и *MPEG-2* имеют много общего, но между ними есть и различия. В данном разделе в основном излагается содержание стандарта *MPEG-2*, и указываются его отличия от *MPEG-1*.

Метод кодирования движущихся изображений, используемый в стандартах *MPEG-1* и *MPEG-2*, называется *гибридным*, так как в нем сочетаются *внутрикадровое* (*intraframe*) кодирование, направленное на уменьшение в основном психофизиологической избыточности в отдельных кадрах, и *межкадровое* (*interframe*) кодирование, с помощью которого уменьшается избыточность, обусловленная межкадровой корреляцией.

Все семейства стандарта *MPEG* не определяют схему и конструкцию кодера и декодера, а лишь описывают средства, используемые для обработки сигнала, определяют синтаксис (правило построения последовательности символов) совместимого цифрового потока и дают примеры реализации декодера. Состав и построение кодера оставлено на усмотрение разработчика. Это может быть аппаратное или программное устройство любой сложности, дающее на выходе синтаксически правильный цифровой поток.

На рисунке 2.1 показан один из вариантов устройств кодирования по стандарту *MPEG-2*.

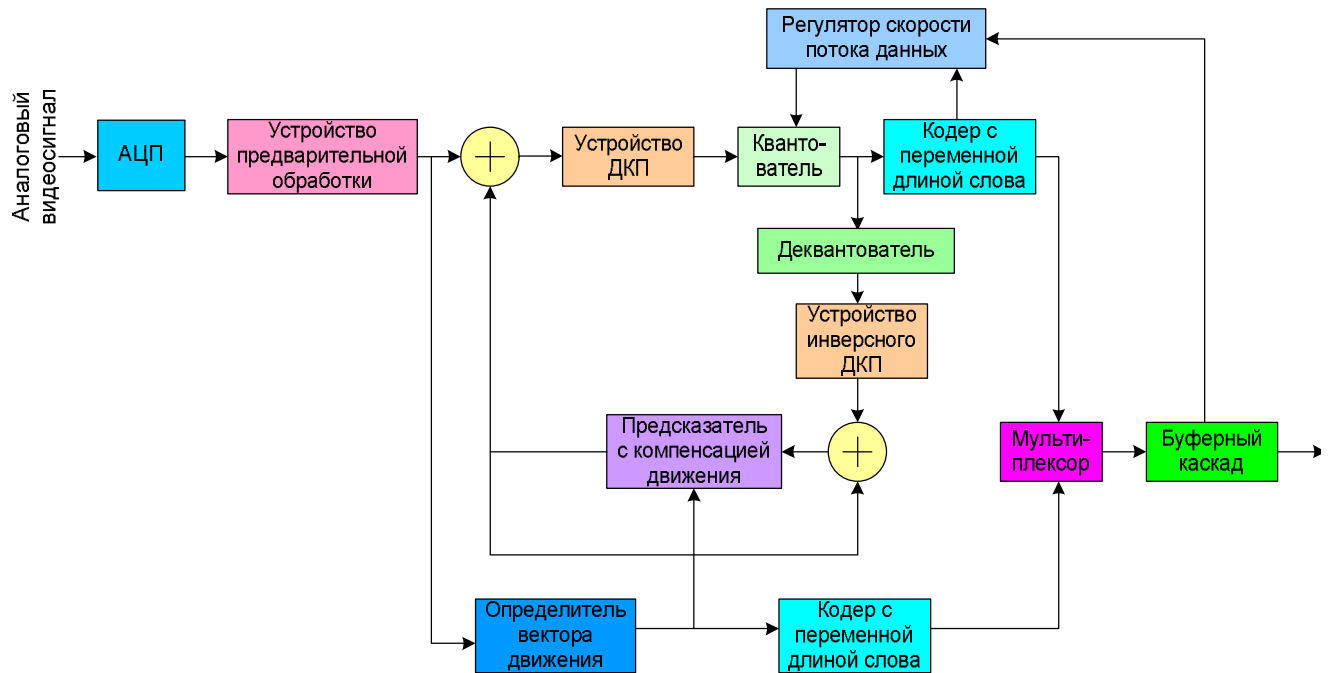


Рисунок 2.1 — Функциональная схема устройства кодирования с информационным сжатием по стандарту *MPEG-2*

Итак, в *MPEG-2* применены известные, давно апробированные методы: сокращения избыточности. Вместе с ними использованы и новые подходы. В особенности это относится к совокупности согласованных алгоритмов сокращения статистической избыточности. Здесь особо эффективными оказались два метода: кодирование ТВ отсчетов с предсказанием и ДКП.

Кодирование с предсказанием реализуется с помощью дифференциальной импульсно-кодовой модуляции (ДИКМ). При кодировании с предсказанием вычисляется разность между истинным и предсказанным значением отсчета. Затем разность квантуется по уровню. От точности предсказания зависит среднее число бит, необходимых для передачи разностной информации. Предсказание может быть экстраполяционным. В этом случае (его часто называют предсказанием вперед) по предшествующим значениям отсчетов ТВ сигнала оцениваются

последующие отсчеты. Интерполяционное (двунаправленное) предсказание означает, что оценка среднего по положению отсчета ТВ сигнала выполняется по известным значениям предшествующих и последующих отсчетов. Такое предсказание наиболее точно оценивает текущие отсчеты. Однако за точность приходится расплачиваться возросшим объемом вычислений и соответственно памяти, необходимой при реализации. При этом эффект не окупает затраты.

Предсказание выполняется по соседним с предсказываемым отсчетам, причем под соседними надо понимать отсчет, расположенные «до и за» рассматриваемым. К ним надо добавить соседей в предшествующей и последующей строках, полях и кадрах. Таким образом, возможно построчное, межстрочное, внутрислоевое, внутрикадровое, межслоевое, межкадровое предсказание. Это полный набор возможных направлений корреляций. Но уже подчеркивалось, что предсказание вдоль отдельной строки по предшествующему и последующему элементу не эффективно. По этому же критерию можно отсеять и некоторые другие возможные направления.

При простейшем внутрислоевом предсказании вперед предшествующий отсчет ТВ строки принимается как ожидаемый уровень последующего отсчета. Фактически это означает вычеркивание постоянной составляющей или, что тоже самое, выделение разностной информации. Такой метод предсказания особенно эффективен, когда передаются крупные, не содержащие мелких деталей, фрагменты изображения, где яркость постоянна или изменяется медленно.

Другой способ — межкадровое предсказание вперед. В этом случае текущий отсчет оценивается по отсчету с теми же координатами, но предыдущего кадра. Это очень эффективный метод предсказания для неподвижных изображений. Ситуация усложняется, когда изображение содержит движущиеся объекты или изменяющиеся в целом. В этом случае отсчеты, принадлежащие однозначным элементам изображения, от кадра к

кадру будут смещаться. Возникает разностная информация, даже если в остальном никаких изменений не происходит. Это можно ослабить, если ввести компенсацию движения. Для этого необходимо определить векторы перемещения движущихся частей изображения при последовательном переходе от кадра к кадру. Векторы движения позволяют определить положение кодируемого отсчета в новом кадре (скомпенсировать его перемещение) и, таким образом, сохранить высокую точность предсказания.

Таким образом, стандарт *MPEG-2* фактически не регламентирует методы сжатия видеосигнала, а только определяет структуру битового потока кодируемого видеосигнала. Поэтому конкретно используемые алгоритмы зависят от собственных разработок фирм-производителей оборудования. При этом ярко выражены общие принципы построения стандарта *MPEG-2*, в соответствии с которыми процесс сжатия цифрового видеосигнала может быть разбит на ряд последовательных операций: преобразование аналогового сигнала в цифровую форму, предварительная обработка, ДКП, квантование, кодирование (рисунок 2.1).

После АЦП производится предварительная обработка сигнала, которая включает в себя следующие преобразования.

1. Удаление избыточной информации. Например, если фон изображений состоит из идентичных символов (пикселей), то совершенно не обязательно их все передавать. Достаточно описать один пиксель и послать его с сообщением о том, как часто и где он повторяется в изображении.
2. Если исходное изображение передается в виде чересстрочных полей, то они преобразуются в кадры с прогрессивной разверткой.
3. Цветоделенные сигналы E_R , E_G , E_B преобразуются в цветоразностные сигналы U и V и сигнал яркости Y .
4. Изображение достраивается до кратного 16 количества пикселей по строкам и столбцам, чтобы обеспечить разбиение изображения на целое число макроблоков.

5. Производится преобразование стандарта цифрового представления ТВ сигнала 4:4:4 в стандарт 4:2:2 (горизонтальная передискретизация цветоразностных компонентов) или 4:2:0 (горизонтальная и вертикальная передискретизация цветоразностных компонентов).

6. Изображение разбивается на последовательность макроблоков, количество которых 1620. Каждый из них, в свою очередь, состоит из нескольких блоков размером 8×8 пикселей. (Исследования проводились по разбиению на блоки размером 4×4 , 8×8 , 16×16 пикселей, которые показали, что разбиение 8×8 является наилучшим компромиссом между точностью преобразования, то есть минимальной среднеквадратичной ошибкой, и необходимым объемом вычислений.) В этом случае каждый блок представляет собой квадратную матрицу, содержащую 64 отсчета ТВ сигнала и называемую сигнальной. Следовательно, макроблок несет информацию о компонентах яркости определенного участка изображения и пространственно соответствующих им компонентах цветности. Поэтому любой макроблок состоит из трех прямоугольных матриц, содержащих восьмибитовые отсчеты, а именно: матрицы яркости Y_D и двух матриц цветности C_r и C_b . Цифровые отсчеты матриц Y_D , C_r и C_b непосредственно связаны с первичными значениями красного, зеленого и синего (R, G, B) соответствующих точек изображения. Значен. первичных цветов гамма — предкорректированы. Причем величина гамма-коэффициента стандартом не определяется, но обычно находится в диапазоне 2,2...2,8.

Стандартами цифрового кодирования предусматриваются три формата цветности, каждому из которых соответствует свой порядок следования блоков в макроблоке:

4:2:0 — макроблок состоит из шести блоков. Четыре блока, несущие информацию о яркости, образуют матрицу Y_D размером 16×16 пикселей. Два блока цветности, определяющие цветоразностные компоненты U и V , соответствующие матрицам C_r и C_b с размерами 8×8 пикселей (см. рисунок 2.2).

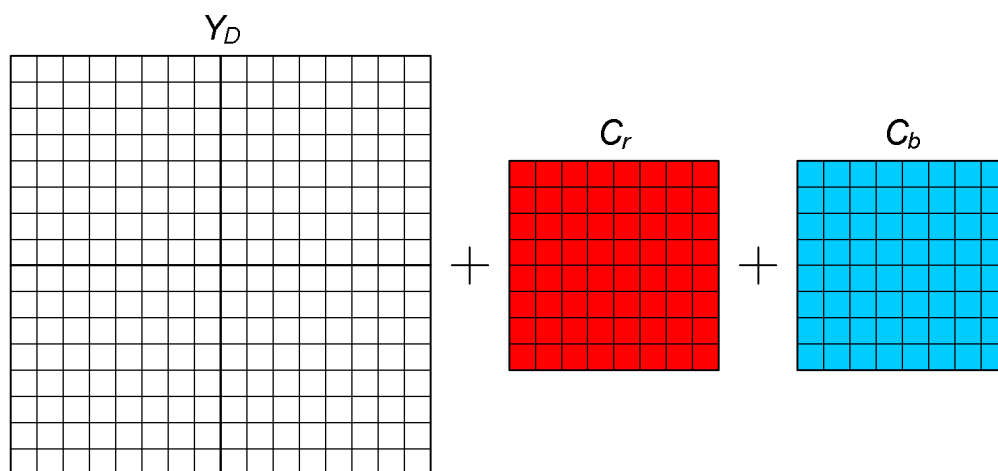


Рисунок 2.2 — Схема разбивки изображения на блоки при реализации формата цветности 4:2:0

4:2:2 — макроблок состоит из восьми блоков. Четыре блока яркости образуют матрицу Y_D . Четыре блока цветности соответствуют матрицам C_r и C_b , каждая из которых содержит по два блока.

4:4:4 — макроблок состоит из двенадцати блоков. Он содержит четыре блока яркости и восемь блоков цветности.

Внутренняя организация макроблоков различна при кодировании полей и кадров. Пары полей могут кодироваться как отдельно, так и вместе как единый кадр. При кодировании полей блоки яркости группируются по полям: верхние — из первого полукадра, нижние — из второго. Блоки цветности располагаются в порядке следования кадров для обоих типов кодирования. Два полукадра, составляющие один кадр, всегда следуют один за другим в битовом потоке. При кодировании кадрами чересстрочных изображений каждый кадр набирается из двух полукадров (через строку) и кодируется как единое изображение.

Следующие друг за другом макроблоки объединяют в независимые, друг от друга серии (*Slice* — *слайс*). Серия является основным элементом синхронизации для восстановления данных, составляющих изображение, и обычно состоит из всех блоков в горизонтальном направлении изображения с

интервалом 16 строк, то есть имеет толщину в один макроблок. Порядок макроблоков в серии тот же, что и в обычном сканировании раstra в телевидении: слева направо и сверху вниз. Представление информации сериям удобно для коррекции ошибок. Когда появляется ошибка в потоке данных, декодер может обратиться к началу следующей серии. Соответственно число таких серий влияет на эффективность передачи. Каждая серия (слайс) должна содержать, по крайней мере, один макроблок. *Первый и последний макроблоки в серии не должны быть пропущенными*. Серии не должны перекрываться и не должно быть интервалов между сериями. Положение серий (слайсов) может меняться от изображения к изображению. Первая серия начинается с первого макроблока, а последняя серия заканчивается последним макроблоком изображения. Кадр делится на несколько смежных серий.

7. Производится разбиение потока кадров изображения по типам, для них находятся векторы движения, которые необходимы для повышения сказуемости величин элементов изображения. Векторы движения обеспечивают компенсацию перемещений в прошедших и последующих кадрах.

8. Для блоков с использованием компенсации движения находятся разностные ошибки предсказания движения.

В соответствии со стандартом *MPEG-2* декодере выполняются декодирование кодов переменной длины деквантования обратное ДКП, компенсация движения и восстанавливается исходная последовательно кадров (рисунок 2.3).

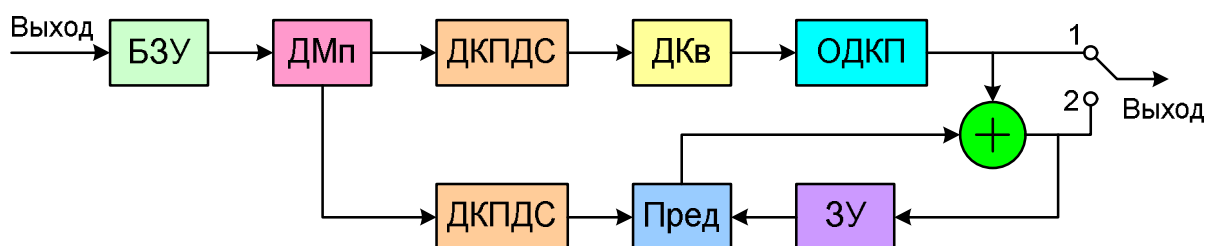


Рисунок 2.3 — Структурная схема видеodeкодера *MPEG-2*

Декодер содержит буферное запоминающее устройство (БЗУ) демультиплексор (ДМп), декодеры кодов с переменной длиной кодовых слов (ДКПДС), а также деквантователь (ДКв), блок обратного дискретно-косинусного преобразования (ОДКП), Предсказатель (Пред) и ЗУ, аналогичные соответствующим блокам кодера. Тактовая частота 27 МГц восстанавливается с использованием данных из декодируемого потока.

БЗУ на входе декодера выполняет функцию согласования постоянной скорости передачи двоичных символов во входном потоке данных с процессами в декодере, при которых данные из БЗУ считываются неравномерно во времени. С выходов демультиплексора кодированные данные изображения и значения параметра квантования поступают на ДКПДС и далее на деквантователь, а векторы движения поступают на ДКПДС и далее на предсказатель.

Так же как и в кодере, в декодере имеются два режима работы. При приеме *I*-кадров и передаваемых с внутрикадровым кодированием макроблоков *P*-кадров и *B*-кадров на выходе блока обратного ДКП формируются блоки изображения. Переключатель на структурной схеме при этом находится в положении 1, и сигнал с блока обратного ДКП направляется на выход. При приеме макроблоков *P*-кадров и *B*-кадров, кодируемых в межкадровом режиме, переключатель находится в положении 2. В этом случае формирование выходного сигнала происходит путем поэлементного сложения поступающих с блока обратного ДКП значений разностей с предсказанным макроблоком, формируемым из элементов ранее декодированных изображений с использованием декодированных векторов движения.

Реализация декодера аппаратными программными или аппаратно-программными средствами существенно проще, чем реализация кодера, так как в декодере не надо выполнять поиск соответствующих областей в опорных изображениях, а именно этот поиск требует наибольшего количества вычислений.

2.2. Кодлируемые кадры

Базовым объектом кодирования в стандарте *MPEG-2* является кадр ТВ изображения. При этом очевидно, что для ТВ сигналов, в которых смешаны различные сюжеты с разными типами движений «от ничего до много» простое предсказание, в принципе не обеспечит высокую эффективность. По той причине в стандарте *MPEG-2* используются три вида предсказаний: внутрикадровое, межкадровое предсказание вперед с компенсацией движения, межкадровое двунаправленное предсказание также с компенсацией движения.

Формат видеоинформации в стандарте *MPEG-2* содержит три типа кадров (*I*, *P*, *B*). Основные, так называемые кадры *I*-кадры (Intraframes) обрабатываются только с применением внутрикадрового предсказания. Они кодируются независимо от других кадров, так как обрабатываются с использованием собственной информации, т.е. по принципу случайного доступа к сжатым видеоданным. Они применяют кодовое преобразование блоков элементов изображения и обеспечивают умеренное сжатие. Это первый этап сжатия видеоданных, Зато при восстановлении ТВ изображения по *I*-кадрам оно менее всего деградирует и зависит от ошибок кодирования и передачи видеоданных по каналу связи. *I*-кадры служат опорными при межкадровом предсказании *P* и *B* кадров.

P-кадры (*Predicted Frames*), то есть кадры с предсказанием, с компенсацией движения. Кодирование осуществляется с учетом ближайших предшествующих *I* или *P*-кадров. Этот способ называется с предсказанием вперед, так как используется «разностная» схема сжатия, при которой сохраняются только отличия от предшествующего кадра. В *P*-кадрах, если сравнивать их с *I*-кадрами, в три раза выше достижимая степень сжатия видеоданных.

Обработка видеоданных в *P*-кадре выполняется по макроблокам. Каждый макроблок обрабатывается с использованием алгоритмов компенсации движения и предсказания вперед, пока в блоке не появится

новый объект. С этого момента процесс кодирования переключается на алгоритмы, используемые в *I*-кадрах, т.е. на внутрикадровое предсказание. *P*-кадры являются опорными для последующих *P* или *B*-кадров. Необходима высокая точность восстановления исходного изображения при декодировании опорных *P*-кадров. Дело в том, что ошибки опорного кадра распределяются по всем кадрам, связанным с опорным. При компенсации движения, применяемой к макроблокам *P*-кадров, вырабатывается два вида информации: векторы движения (разница между базовыми и кодированными макроблоками) и значения ошибок (разница между предсказанными величинами и действительными результатами). Если макроблок в *P*-кадре не может быть описан с использованием компенсации движения, что случается при появлении некоторого неизвестного объекта, то он кодируется тем же способом, что и макроблок в *I*-кадре.

B-кадры (*Bi-Directional Frames*), то есть кадры с двунаправленным предсказанием, с компенсацией движения. Для формирования *B*-кадров также используется «разностная» схема сжатия аналогично *P*-кадрам, однако, в качестве «базовых» кадров используются оба соседних кадра: предыдущий и последующий. Этот способ называется двунаправленным предсказанием.

Алгоритмы кодирования *B*-кадров зависят от характера ТВ изображения. Предусмотрено четыре способа кодирования. В одном применяет компенсация движения и предсказание вперед по ближайшим предшествующим опорным *I* или *P*-кадрам, в другом — компенсация движения и обратное предсказание по ближайшим последующим *I* или *P*-кадрам. Обратное предсказание используется в тех случаях, когда в кодируемом *B*-кадре появляются новые объекты изображения. Третий алгоритм — компенсация движения и двунаправленное предсказание, при котором опорными являются предшествующий или последующий *I* или *P*-кадры. И, наконец, внутрикадровое предсказание без компенсации движения. Такое кодирование нужно при резкой смене передаваемых сюжетов, а также при больших скоростях перемещения объектов ТВ изображения. С *B*-

кадрами связано наиболее глубокое сжатие видеоданных. Поскольку высокая степень сжатия снижает точность восстановления исходного ТВ изображения, *B*-кадры не используются в качестве опорных. Ошибки при их декодировании не распределяются по другим кадрам.

Очевидно, что точность кодирования должна быть максимальной для *I*-кадров, ниже для *P*-кадров и минимальной для *B*-кадров. В стандарте *MPEG-2* порядок записи информации о ТВ кадрах не совпадает с порядком кодирования и декодирования. Так, для декодирования текущего *B*-кадра требуется обработать следующий кадр, поэтому необходимо иметь дополнительный буферный блок памяти в декодирующем устройстве для хранения информации об очередном кадре. Рассмотренное усложнение декодирующего устройства компенсируется улучшением субъективного качества воспроизводимого изображения за счет *B*-кадров на 20%.

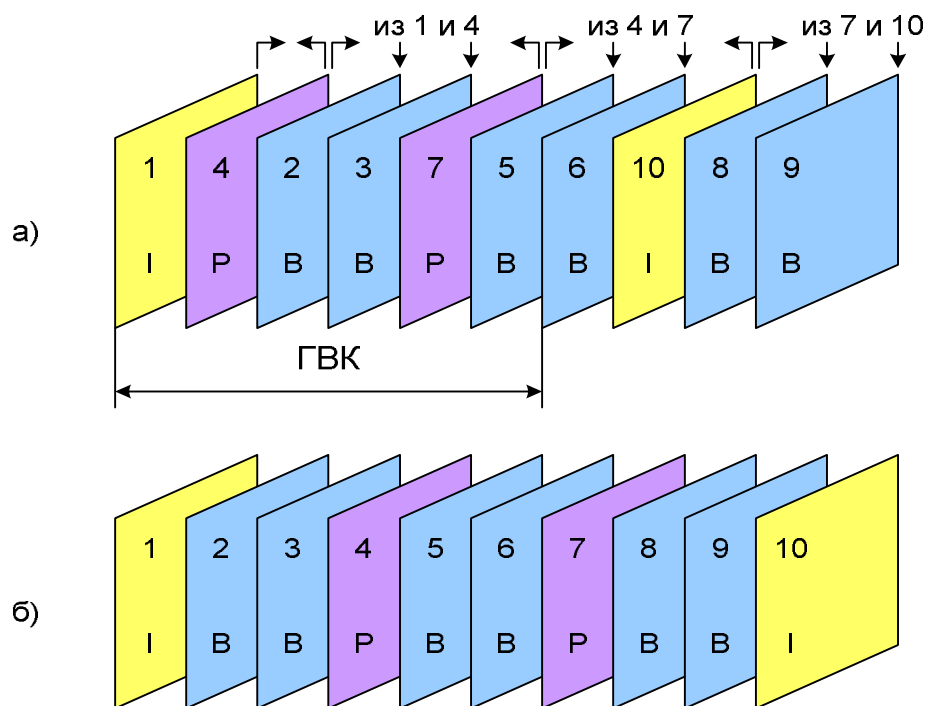


Рисунок 2.4 — Цикл кодирования и декодирования по стандарту *MPEG-2*:

а — порядок кодирования и декодирования изображений

б — порядок воспроизведения изображений

Изображения различных типов объединяются в повторяющиеся серии, называемые *группами видеокадров (ГВК)*. Порядок кодирования

декодирования и воспроизведения видеокадров указанных трех типов может быть Различным для примера на рисунке 2.4 изображен один из вариантов сочетания видеокадров, обеспечивающих минимизацию энтропии, то есть количества битов на элемент изображения. Группа начинается с *I*-кадра, образующего опорный сигнал для предсказания при кодировании изображений других типов. Группа изображений должна быть достаточно большой, если необходимо добиться высокой степени компрессии ГВК заканчивается непосредственно перед появлением следующего *I*-кадра. ГВК определяет границы межкадрового кодирования, Многочисленные исследования эффективности различных сочетаний *P* и *B*-кадров в ГВК показали, что последовательности длинных ГВК целесообразно использовать только для высококачественных незашумленных изображений. Однако большое число *B*-кадров означает значительную задержку, так как *I* или *P*-кадры, окружающие кадры *B*-типа, уже должны быть в декодере к моменту декодирования *B*-кадра.

В начале каждого сюжета должен стоять *I*-кадр, в конце — *P*-кадр. Увеличивать долю *B*-кадров можно только в рамках одного сюжета, иначе возникнут большие ошибки предсказания и компенсации движения. Поскольку типичная длительность группы кадров (во временном представлении примерно 0,5 с) значительно меньше характерного расстояния между границами сюжетов, то в большинстве случаев жесткое задание структуры не приводит к существенным визуальным ошибкам из-за того, что смена сюжета попадает внутрь группы кадров. При передаче по каналу связи порядок следования *I*, *P* и *B*-кадров меняется.

2.3. Компенсация движения

Точность предсказания при передаче изображений движущихся объемов можно увеличить за счет оценки вектора движения и компенсации этого движения, которая уменьшает ошибку предсказания.

В стандарте *MPEG-2* используется метод компенсации движения, основанный на макроблоках. Два смежных кадра содержащих только

активны строки сигнала яркости (576 активных строк), разбиваются на макроблоки более крупные зоны поиска. Размеры макроблока должны быть согласны со структурой дискретизации кадра ТВ изображения. В стандарте MPEG-2 ТВ кадр разбивается на целое число зон. По вертикали (576 активных строк / 16) — это 36 зон, по горизонтали (704 активных отсчета / 16) — 44 зоны.

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. Эти размеры также должны быть согласованы с принятой структурой дискретизации ТВ кадра. Обычно, они в 4 раза больше размеров отдельного макроблока. Иными словами, размеры зоны поиска — это 64×64 . Таким образом, в ТВ кадре создается $576 / 64 = 9$ зон поиска по вертикали и $704 / 64 = 11$ зон по горизонтали.

К примеру, надо определить координаты движения при предсказании вперед. Для этого берется макроблок отсчетов первого кадра и ищется его новое положение в зоне поиска второго кадра, вычисляются межкадровые разности отсчетов. 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.

Компенсация движения, применяемая при компрессии *P* и *B*-кадр, улучшает фактор сжатия в 3 раза по сравнению с внутрикадровым кодированием, где сохраняется временная избыточность.

Компенсация движения, примененная к макроблокам *B*-кадра, осуществляется с использованием как предшествующего, так и последующего опорного кадра.

Кодирование по стандарту *MPEG-2* требует запоминания одного или двух кадров опорного макроблока, обеспечивающего данные для предсказания с компенсацией движения. Большой объем компьютерных вычислений занимает поиск и оценка движения для подбора макроблоков в двух кадрах, чтобы найти направление и расстояние, определяющие движение макроблока между кадрами, то есть вектор движения.

Определяемый вектор движения используется для предсказания с компенсацией движения. В кодере вычисляется ошибка предсказания, т.е. разность между фактическим и предсказанным с использованием вектора движения блоками изображения, что сокращает временную избыточность. Ошибка предсказания подвергается ДКП, квантуется и кодируется в кодере с переменной длиной слова. Такой процесс уменьшает и временную, и пространственную, и психофизическую избыточность. Вектор движения также кодируется словами переменной длины. Кодированная ошибка предсказания объединяется с кодами вектора движения, после чего формируется передаваемый цифровой поток.

Чем быстрее движутся снимаемые объекты, тем больше отличаются текущий макроблок и опорный макроблок из предыдущего кадра и тем больше данных об ошибке предсказания должно быть передано, а это снижает эффективность компрессии. Компенсация движения при формировании предсказания уменьшает ошибку предсказания, которая на практике не может быть сведена к нулю. Этому мешают и изменения размеров движущегося объекта, его поворота, неточности в определении вектора движения и т.п. Однако предсказание с компенсацией движения позволяет значительно повысить эффективность ДИКМ при видеокомпрессии.

В декодере выполняется инверсное квантование, инверсное ДКП, в результате чего формируется ошибка предсказания. Ошибка предсказания складывается с декодированным изображением предыдущего кадра, образуя декодированное изображение текущего кадра.

Предсказатели с компенсацией движения в современных системах видеокомпрессии стандарта *MPEG-2* могут использовать целый ряд методов. Например, макроблок может предсказываться на основе предыдущего изображения, на основе последующего изображения, а также на основе и предыдущего, и последующего. В чересстрочных системах поля кадра могут предсказываться отдельно с использованием разных векторов движения или вместе с использованием общего вектора. Существует также возможность нулевого предсказания (если не найден подходящий опорный макроблок). При этом сам текущий макроблок будет кодироваться вместо ошибки предсказания, что означает отказ от межкадрового кодирования и переход к внутрикадровому. Для каждого текущего макроблока кодер выбирает метод предсказания, обеспечивающий наивысшее качество декодированного изображения с учетом ограничений на скорость передачи данных. Сведения о методе предсказания включаются общий поток и передаются декодеру для верного восстановления изображения.

Оценка вектора движения и определение наилучшей стратегии предсказания требуют применения сложных вычислительных процедур, которые должны осуществляться в реальном времени. Поэтому кодер гораздо сложнее декодера, т.е. в стандарте *MPEG-2* система компрессии на базе ДИКМ с компенсацией движения является асимметричной.

2.4. Использование ДКП в стандарте кодирования *MPEG-2*

Цель внутрикадрового кодирования сокращение пространственной избыточности в пределах кадра (или поля) ТВ изображения. Эта избыточность вызвана сильными корреляционными связями между элементами изображения. Если найти соответствующее ортогональное преобразование, то можно преобразовать массив отсчетов изображения в матрицу коэффициентов, которые уже не будут коррелированы друг с другом. К этим некоррелированным коэффициентам можно применить энтропийное кодирование и добиться сокращения цифрового потока. Наиболее часто используются методы линейных ортогональных

преобразований. Линейность ортогонального преобразования означает, что операции сложения, вычитания и умножения на скаляр действительны и после преобразования, а ортогональность — что преобразуемый фрагмент представляется ограниченным набором ортогональных функций.

Линейные ортогональные преобразования характеризуются тем, что между элементами изображения устраняются статистические зависимости, и распределение энергии в преобразованном спектральном фрагменте является неравномерным. Эти особенности используются непосредственно в процессах кодирования. На практике можно подобрать такое ортогональное преобразование, что для типичных изображений большая часть коэффициентов матрицы будет иметь практически нулевые значения. Исключая эти нулевые коэффициенты, можно также сокращать цифровой поток. Среди возможных ортогональных преобразований наиболее широко используется ДКП, основанное на применении ортогональной системы дискретных косинусных функций возрастающей частоты. Преобразования данного типа хорошо согласуются с параметрами ТВ сигнала, что является необходимым условием эффективного кодирования видеoinформации.

Стандарт *MPEG-2* подразумевает применение ДКП составляющих видеосигнала. Например, ДКП действительной последовательности отсчетов $x(n)$ при $n = 0, 1, \dots, N-1$ определяется соотношением:

$$X_c(k) = (1/\sqrt{N}) \sum_{n=0}^{N-1} c(k)x(n) \cos[(2n+1)kp / 2N],$$

где коэффициенты $c(k) = \begin{cases} 1 & \text{при } k = 0 \\ 2 & \text{при } k = 1, 2, \dots, N-1 \end{cases}$.

Обратное ДКП, восстанавливающее исходную последовательность отсчетов по последовательности коэффициентов преобразования $X_c(k)$, задается выражением

$$x(n) = (\sqrt{N})^{-1} \sum_{k=0}^{N-1} c(k)X_c(k) \cos[(2n+1)kp / 2N],$$

где $k = 0, 1 \dots N-1$.

ДКП сокращает избыточность и высокочастотную информацию в пределах кадра. Это позволяет получить высокое качество кодированных изображений со сжатием.

ДКП выполняется поблочно, для чего ТВ изображение разбивается на блоки. При этом в ТВ кадре создается: $576 / 8 = 72$ зоны по вертикали $704 / 8 = 88$ зон по горизонтали, что в общей сложности дает $72 \times 88 = 6336$ блоков, подлежащих ДКП в реальном масштабе времени. В результате ДКП исходная сигнальная матрица $8 \times 8 = 64$ ТВ отсчетов преобразуется в матрицу частотных коэффициентов ДКП такого же размера $8 \times 8 = 64$ (рисунок 2.5).

Каждый коэффициент характеризует амплитуду определенной частотной составляющей кадра, причем коэффициенты в матрице располагаются по возрастанию частот в вертикальном и горизонтальном направлениях.

Поскольку положение ТВ отсчетов сигнальной матрицы определяется двумя координатами, то частотные коэффициенты матрицы ДКП являются функциями этих двух переменных.

а)

U_{11}	U_{12}	U_{13}	U_{14}	U_{15}	U_{16}	U_{17}	U_{18}
U_{21}	U_{22}	U_{23}	U_{24}	U_{25}	U_{26}	U_{27}	U_{28}
U_{31}	U_{32}	U_{33}	U_{34}	U_{35}	U_{36}	U_{37}	U_{38}
U_{41}	U_{42}	U_{43}	U_{44}	U_{45}	U_{46}	U_{47}	U_{48}
U_{51}	U_{52}	U_{53}	U_{54}	U_{55}	U_{56}	U_{57}	U_{58}
U_{61}	U_{62}	U_{63}	U_{64}	U_{65}	U_{66}	U_{67}	U_{68}
U_{71}	U_{72}	U_{73}	U_{74}	U_{75}	U_{76}	U_{77}	U_{78}
U_{81}	U_{82}	U_{83}	U_{84}	U_{85}	U_{86}	U_{87}	U_{88}

б)

Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}	Y_{16}	Y_{17}	Y_{18}
Y_{21}	Y_{22}	Y_{23}	Y_{24}	Y_{25}	Y_{26}	Y_{27}	Y_{28}
Y_{31}	Y_{32}	Y_{33}	Y_{34}	Y_{35}	Y_{36}	Y_{37}	Y_{38}
Y_{41}	Y_{42}	Y_{43}	Y_{44}	Y_{45}	Y_{46}	Y_{47}	Y_{48}
Y_{51}	Y_{52}	Y_{53}	Y_{54}	Y_{55}	Y_{56}	Y_{57}	Y_{58}
Y_{61}	Y_{62}	Y_{63}	Y_{64}	Y_{65}	Y_{66}	Y_{67}	Y_{68}
Y_{71}	Y_{72}	Y_{73}	Y_{74}	Y_{75}	Y_{76}	Y_{77}	Y_{78}
Y_{81}	Y_{82}	Y_{83}	Y_{84}	Y_{85}	Y_{86}	Y_{87}	Y_{88}

$U_{11} \div U_{88}$ — совокупность отсчетов ТВ сигнала

$Y_{11} \div Y_{88}$ — совокупность отсчетов ТВ сигнала

Рисунок 2.5 — ДКП отсчетов ТВ изображения:
а — сигнальная матрица отсчетов ТВ изображения;
б — матрица частотных коэффициентов ДКП.

Следует отметить, что матрица частотных коэффициентов ДКП уже не имеет прямой геометрической связи с положением отсчетов ТВ сигнала на

ТВ растре, а представляет собой только удобную форму математической записи, при которой частотные коэффициенты, ДКП можно трактовать как двумерный спектр ТВ изображения в горизонтальном и вертикальном направлениях ТВ кадра. Каждый частотный коэффициент ДКП содержит информацию не об одном отсчете ТВ изображения и сигнальной матрицы, а обо всех 64 элементах.

При этом справедливо и обратное положение — знание матрицы коэффициентов, т.е. значений амплитуд базисных косинусоидальных функций разных частот позволяет сформировать блок элементов изображения (сигнальную матрицу). Таким образом, ДКП является обратимым.

Спектр ДКП имеет очень важную, если ее оценивать с позиций компрессии видеоданных особенность: для большинства блоков типичных изображений основная энергия частотных составляющих этого спектра концентрируется в небольшой области около нулевых частот. Например, постоянная составляющая и несколько низкочастотных базисных функций, имеющих значимые величины, отображаются совокупностью коэффициентов ДКП в левом верхнем углу матрицы. Амплитуда высокочастотных составляющих или мала, или просто равна нулю, поэтому их потеря почти не сказывается на качестве изображения. Передаче подлежат только те частотные коэффициенты матрицы ДКП, величины которых превышают принятые пороговые значения, коэффициенты ниже порогового значения считаются нулевыми, что и приводит к желаемой компрессии.

Введение пороговой (нелинейной) обработки, строго говоря, ведет к потерям информации и, соответственно, к снижению качества восстановленного в декодере ТВ изображения. Однако при разумном выборе величины порога это ухудшение окажется практически незаметным или же допустимым.

Следует отметить, что при внутрикадровом кодировании динамический интервал коэффициентов ДКП возрастает в 8 раз. Так, при уровне

кодировании видеосигнала восемью битами его динамический интервал равен 0 —250 дискретных уровней. При этом динамический интервал коэффициентов спектра ДКП составит от 0 до 2040 и от -1020 до +1020 дискретных уровней для коэффициентов постоянной и переменных составляющих ДКП, соответственно.

Вычисление коэффициентов ДКП с большей точностью, чем имеют значения отсчетов яркости изображения, объясняется необходимостью пренебречь ошибками их определения для исключения возможных неточностей при ДКП.

Кодирование коэффициентов ДКП в таком широком динамическом интервале потребует в последующих узлах кодера перехода от 8 битового к 11-битовому коду. Чтобы избежать этого, после ДКП производится масштабирование (сжатие) динамического интервала 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. Так, например, если исходное значение коэффициента ДКП было 22, то после деления на 8 и округления до ближайшего целого значения ($22 / 8 = 2,75$) новое значение будет 3. При этом новый динамический интервал составит от -255 до +255 дискретных уровней.

В обычной ИКМ уменьшение разрядности приводит к возрастанию шумов квантования на всех частотах. Коэффициенты ДКП определяют энергию сигнала на различных частотах, поэтому появляется возможность изменять параметры квантования дифференцировано на разных частотах, учитывая различную чувствительность зрения к разным пространственным частотам.

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 они проявляются в виде шахматной структуры.

На предельно большое число уровней квантуются и амплитуды нескольких самых низкочастотных косинусоидальных волн, располагающихся в левом верхнем углу блока ДКП. По мере увеличения частоты косинусоидальных компонент (при удалении от левого верхнего угла блока ДКП) количество уровней квантования уменьшается, достигая для самых больших частот значений в нескольких единиц. Например, самая высокочастотная компонента, располагающаяся в правом нижнем углу, может квантоваться всего на два уровня. Это означает, что для ее передачи можно использовать одноразрядные двоичные числа.

Практически квантование выполняется путем поэлементного деления матрицы коэффициентов ДКП на матрицу квантования (взвешивания), значения элементов которой возрастают по мере удаления от левого верхнего угла и приближения к правому нижнему углу. При этом важно отметить, что для квантования сигнала яркости и цветоразностных сигналов используются разные матрицы. Построчные матрицы квантования могут задаваться кодером, но по умолчанию стандарт *MPEG-2* предполагает использование следующих эффективных матриц (таблицы 2.1, 2.2).

При декодировании на приемной стороне коэффициенты матрицы ДКП, значения которых были переданы по каналам связи, умножаются на элементы матрицы квантования, что восстанавливает верные значения коэффициентов, но с ошибкой округления, значение которой мало для низкочастотных компонентов блока отсчетов изображения, но велико для высокочастотных.

Массив коэффициентов, извлекаемых из матрицы ДКП, содержит некоторое количество нулевых значений.

Таблица 2.1 — Матрица квантования яркостного сигнала для блоков изображений *I*-кадров

8	16	19	22	26	27	29	34
---	----	----	----	----	----	----	----

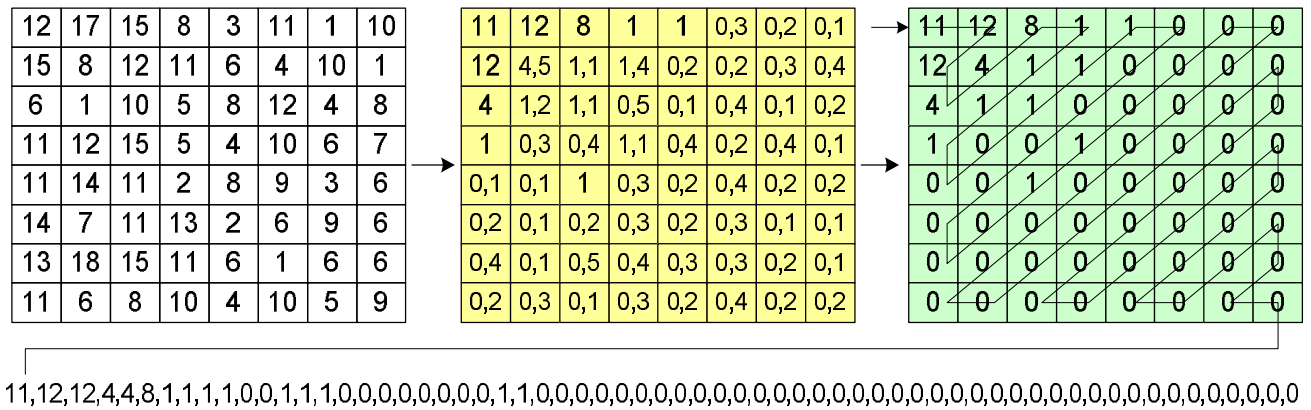
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Таблица 2.2 — Матрица квантования цветоразностного сигнала для блоков изображений *I*-кадров

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Для того чтобы способствовать объединению нулевых элементов в группы и преобразованию в одномерную последовательность, используется зигзагообразное сканирование матрицы, начиная с левого верхнего угла. В этом случае коэффициенты выстраиваются в порядке возрастания частот, причем если пространственные частоты одинаковы, то впереди следуют коэффициенты для меньших вертикальных частот. Зигзаг-сканирование (z-упорядочивание) позволяет сгруппировать самый большой массив нулевых коэффициентов, обычно располагающийся в правой нижней части блока ДКП.

Примером преобразования стандартной сигнальной матрицы 8×8 пикселей может служить процесс ДКП, квантования и зигзагообразного сканирования, графически представленный на рисунке 2.6.



Последний алгоритм сокращения избыточности связан с кодами, образующими комбинации переменной длительности. При этом те коэффициенты ДКП, которые повторяются наиболее часто, кодируются короткими кодовыми комбинациями, а редкие значения коэффициентов — более длинными.

Для экономного представления двоичных сообщений в стандарте *MPEG-2* используется код Хаффмана, позволяющий практически в 1,22 раза сократить поток данных по сравнению с равномерным кодированием. При кодировании символов сообщения комбинациями переменной длины обычно возникает проблема отделения одной комбинации от другой. Код Хаффмана обладает свойством префиксности, т.е. ни одна его кодовая комбинация не является началом другой комбинации, что позволяет обойтись в тексте кодированного сообщения без разделителей между комбинациями.

Энтропийный кодер в устройстве кодирования с информационным сжатием MPEG-2 должен иметь многокадровый буферный накопитель, в котором происходит накопление данных для их оптимального использования. Например, в случае, если содержание изображения изменяется с очень большой интенсивностью, возможно изменение скорости передачи данных для сохранения соответствующего качества изображения. Однако на практике при передаче и распределении ТВ программ колебания

скорости передачи данных могут быть недопустимы. В этом случае используется режим работы энтропийного кодера, при котором обеспечивается фиксированная скорость выходного потока данных. Для чего данные заносятся в буферный накопитель с переменной скоростью, а считываются с постоянной. Для предотвращения переполнения буферного накопителя или полного стирания информации в нем, что может привести к сбоям в работе системы видеокомпрессии, используется адаптивное квантование. Сведения о степени заполнения буферной памяти служат сигналом управления, регулирующим шкалу квантования. Если, например, кодируемое изображение характеризуется высокой детальностью, то число ненулевых элементов матриц коэффициентов ДКП увеличивается. Возрастает и объем передаваемых данных, поэтому буферный накопитель заполняется с повышенной скоростью. Благодаря обратной связи (через регулятор скорости потока данных, см. рисунок 2.1) квантование становится более грубым и скорость поступления данных в буферную память уменьшается, но за счет увеличения шумов квантования и ухудшения качества изображения.

Если кодируется простое по структуре изображение с малой детальностью, то число нулевых элементов сигнальных матриц коэффициентов ДКП увеличивается и скорость поступления данных в буферную память снижается по сравнению со средней величиной. Тогда квантование становится менее грубым (большое число коэффициентов ДКП квантуется на максимальное число уровней). Таким образом, скорость заполнения буферного накопителя в среднем поддерживается на постоянном уровне.

Упрощенная структура декодирования на приемном конце показана на рисунке 2.7.

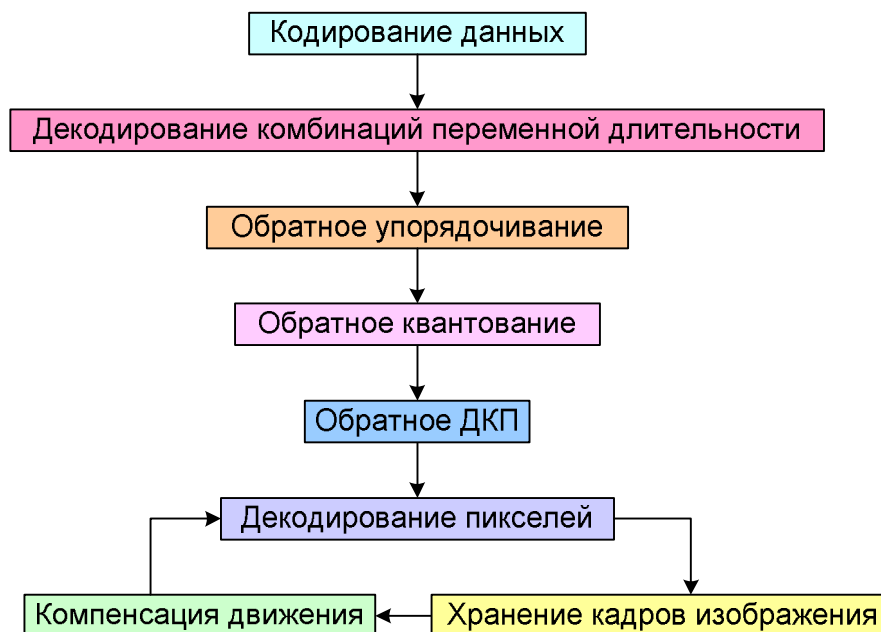


Рисунок 2.7 — Структура декодирования в декодерах *MPEG-2*

2.5. Профили и уровни стандарта *MPEG-2*

Даже в рамках одного стандарта, как показывает практика, передача сигналов телевидения — и цифрового здесь не исключение, ведется на разных уровнях качества. То же самое можно сказать и о ТВ приемниках, жесткие, а главное узкие допусковые интервалы не жизненны, поскольку лишают систему гибкости, приспособляемости к разным условиям функционирования с ориентацией на различные слои потребителей. При этом любая перспективная система должна иметь резервы для перехода на более высокие уровни качества. Эти и многие другие соображения и требования легли в основу очень важного документа: *ISO/IEC 13818-2*.

В этом документе определено, что стандарт *MPEG-2* — это целое семейство взаимосогласованных совместимых цифровых стандартов информационного сжатия ТВ сигналов с различной степенью сложности используемых алгоритмов. Поэтому в рамках стандарта *MPEG-2* была разработана система профилей и уровней. Профиль — это подмножество стандарта для специализированного применения, задающее алгоритмы и

средства компрессии. Уровни внутри каждого профиля связаны с параметрами компрессируемого изображения

Градации качества ТВ изображения для вещательных систем в стандарте *ISO/IEC 13818-2* устанавливаются введением четырех уровней для формата разложения строк ТВ изображения и пяти профилей для форматов кодирования сигналов яркости и цветности. Общая идеология построения стандарта *MPEG-2* поясняется таблицей 2.3.

Таблица 2.3 — Профили и уровни стандарта *MPEG-2*

Уровни	Скорость передачи битов, Мбит/с, для профильного вида				
	Простой профиль без <i>B</i> -кадров (<i>SP</i>) 4:2:0	Основной профиль без <i>B</i> -кадров (<i>MP</i>) 4:2:0	Профиль с масштабированием отношением С-III <i>B</i> -кадры (<i>SNR</i>) 4:2:0	Специальный масштабированный профиль <i>B</i> -кадры (<i>Spatial</i>) 4:2:0	Высший профиль <i>B</i> -кадры (<i>HP</i>) 4:2:0 или 4:2:2
Высокий Уровень (<i>HL</i>) 1920 × 1152 (активных)	×	80	×	×	100
Высокий уровень (<i>H-14</i>) 1440 × 1152 (активных)	×	60	60	60	80
Основной уровень (<i>ML</i>) 720 × 576 (активных)	15	15	15	15	20
Низкий Уровень (<i>LL</i>) 352 × 288 (активных)	×	4	4	4	×

Расположенный в нижней части таблицы уровень называется «низким уровнем» и ему соответствует новый класс качества ТВ изображения, которое вводится в стандарте MPEG-2 — телевидение ограниченной четкости. В этом случае в кадре ТВ изображения содержится 288 активных строк (в 2 раза меньше, чем в вещательном телевидении) и каждая строка дискретизируется на 352 отсчета.

Кодирование сигналов телевидения вещательного стандарта выполняется в соответствии с основным уровнем, т.е. с форматом разложения на 576 активных строк в кадре, которые кодируются с использованием 720 отсчетов на строку.

Высокий уровень — 1440 поддерживает ТВ изображения высокой четкости с разрешением 1440×1152 элементов.

Высокий уровень — 1920 поддерживает ТВ изображения высокой четкости широкого формата с разрешением 1920×1152 (HDTV-plus). В обоих «высоких» уровнях кадр ТВ изображения содержит 1152 активные строки (вдвое больше, чем в вещательном телевидении). Эти строки дискретизируются соответственно на 1440 или 1920 отсчетов.

В стандарте используются 5 профилей, которым соответствуют 5 наборов функциональных операций по обработке (компрессии) видеоданных.

Профиль, в котором используется наименьшее число функциональных операций по компрессии видеоданных, назван простым. В нем при компрессии видеоданных используется компенсация движения изображения и гибридное ДКП.

Следующий профиль назван основным. Он содержит все функциональные операции простого профиля и одну новую: предсказание по двум направлениям. Эта новая операция, естественно, повышает качество ТВ изображения.

Следующий за основным назван профилем с масштабируемым отношением сигнал-шум. Термин «масштабирование», в данном случае, означает возможность обмена основных показателей системы, способность

воспроизведения ТВ изображений из части полного потока видеоданных. Этот профиль к функциональным операциям основного профиля добавляет новую — масштабирование. Основная идея — повышение устойчивости цифрового телевидения и сохранение работоспособности при неблагоприятных условиях приема. Операция масштабирования позволит в рассматриваемом случае повысить устойчивость системы за счет некоторого снижения требований к допустимому уровню отношения сигнал-шум в воспроизводимом ТВ изображении.

При масштабировании потоки видеоданных разделяют на две части. Одна из них несет наиболее значимую часть информации — ее называют основным сигналом. Вторую часть, несущую менее значимую информацию, называют дополнительным сигналом. Декодирование только одного основного сигнала позволяет получить ТВ изображение с пониженным отношением сигнал-шум относительно исходного значения.

И все же, что можно извлечь из идеи деления потока данных на более и менее значимые части? А все дело в защите системы от ошибок. Помехоустойчивое кодирование требует введения дополнительных бит, что повышает общий поток информации. Задача упрощается, когда более мощная защита применяется только к части информации и тем самым соблюдается разумный баланс между уровнем потока видеоданных и степенью их защиты. При неблагоприятных условиях приема (например, при низкой напряженности радиополя, при приеме на комнатную антенну и т.п.) сохраняется возможность устойчивого декодирования более защищенного основного сигнала, а неустойчиво воспринимаемый дополнительный сигнал просто отключается. Это ведет к росту уровня шума, зато система остается работоспособной.

Бывают ситуации, когда сигналы приходится передавать по каналам с ограниченной пропускной способностью. Деление потока видеоданных на два, позволяет использовать и «плохие» каналы, ограничивая передачу основным сигналом.

Следующий, четвертый профиль назван специально масштабируемым профилем. Здесь, естественно, сохранены все операции предшествующего профиля и добавлена новая — разделение потока видеоданных по критерию четкости ТВ изображения. Этот профиль обеспечивает переходы между ныне действующими вещательными системами и ТВЧ. С этой целью видеоданные сигнала ТВЧ разделяются на три потока. Первый — это основной (значимый) поток видеоданных, например, по стандарту разложения на 625 строк. Второй поток несет дополнительную информацию об изображении с числом строк до 1250. Одновременное декодирование первого и второго потоков видеоданных позволяет получить ТВ изображение высокой четкости, но с пониженным отношением сигнал-шум. В третьем потоке сосредоточена менее значимая информация, его декодирование позволяет повысить отношение сигнал-шум в видеоканале до уровня, принятого в ТВЧ. Обычно первый поток видеоданных, представляющих сигнал 625-строчного ТВ, — это 6 Мбит/с, Дополняющий его до ТВЧ — 6 Мбит/с, а повышающий отношение сигнал-шум до уровня, когда шумы визуально незаметны 12 Мбит/с.

Стандартом *MPEG-2* потенциально предусмотрена масштабируемость по времени, позволяющая получать от одного источника видеoinформации ТВ изображения с двумя уровнями разрешающей способности по времени. Например, основной поток видеоданных обеспечивает воспроизведение ТВ изображения с частотой кадров 25 Гц и чересстрочной разверткой. Добавление дополнительного потока видеоданных к основному позволяет получить ТВ изображение с частотой кадров 50 Гц и прогрессивной разверткой.

Таким образом, стандарт *MPEG-2* предусматривает возможность организации потоков видеоданных как с масштабируемостью, так и без нее. Однако, масштабируемость, заложенная в стандарте *MPEG-2*, пока редко встречается в практических реализациях цифровых ТВ систем, но она является важной предпосылкой их дальнейшего развития.

В рассмотренных четырех профилях при кодировании сигналов яркости и цветности используется формат представления видеоданных 4:2:0, в котором число отсчетов сигналов цветности по сравнению с сигналом яркости уменьшается в два раза не только по горизонтальным, но и по вертикальным направлениям. Следующий, пятый профиль называется высшим и он включает в себя все функциональные операции специального профиля 4:2:2, при котором число Отсчетов сигналов цветности в вертикальных направлениях остается тем же, что и у сигнала яркости. В этом случае коэффициент компрессии минимален, а качество изображения наивысшее.

Приведенные в таблице 2.3 пять профилей и четыре уровня образуют 20 возможных комбинаций видеосигнала, из которых, вероятнее всего, только 11 будут необходимыми. Для этих комбинаций (согласованные точки) в таблице указаны максимальные значения скорости передачи видеоданных в Мбит/с.

Для всех стандартизованных точек указаны максимальные потоки видеоданных, которые позволяют получить ТВ изображение, свободное от каких-либо дефектов. В иных случаях они могут проявиться в процессе кодирования/декодирования видеосигнала. Используемые в конкретных кодерах потоки видеоданных могут быть меньше (в несколько раз) указанных значений. Выбор уровня компрессии и, в конечном итоге, уровня потока зависит от допустимой степени искажений ТВ изображения.

Таким образом, стандарт *MPEG-2* позволяет гибко менять скорость передачи видеоданных в очень широких пределах. Надо заметить, что системы кодирования стандарта *MPEG-2* могут работать как с чересстрочной, так и с прогрессивной развертками при частоте полей 50 или 60 Гц. Для каждой стандартизованной точки в таблице оговорено число отсчетов сигнала яркости на активной части строки. Рассмотренные комбинации параметров информационного кодирования Пригодны для работы с различными цифровыми трактами.

Стандарт *MPEG-2* принципиально нацелен в будущее. Большинство выпускаемых в настоящее время декодеров в интегральном исполнении относится к основному профилю и основному уровню (MP@ML) рассчитанных на ТВ изображение с чересстрочным разложением на 625 строк. Эта система принята для первого поколения цифровых телевизоров для НТВ со спутников, работающих в диапазоне 11/12 ГГц, и кабельной сети распределения.

Однако ряд особенностей основного профиля и основного уровня стандарта *MPEG-2*, например, низкое вертикальное разрешение в цветоразностных каналах, ограничивают его применение в условиях ТВ студий, в видеопроизводстве. Для достижения высоких качественных показателей в случае многократного кодирования-декодирования важно кодировать видеосигналы стандарта 4:2:2. Использование видеосигналов, кодированных в стандарте 4:2:0, совместно с основным профилем и основным уровнем MP@ML означает, что вертикальное разрешение в цветоразностных каналах уменьшается вдвое. Взаимное преобразование видеосигналов стандартов 4:2:2 и 4:2:0, необходимое для обеспечения совместимости в ТВ тракте в соответствии с Рекомендацией МСЭ-Р ВТ.601-5, требует в каждой точке преобразования вертикальные фильтры низких частот. Хорошо известно, что каскадное включение таких фильтров быстро приведет к «смягчению», т.е. к размытию цветовых границ. Для сохранения цветового вертикального разрешения лучше осуществлять компрессию видеоданных, кодированных по стандарту 4:2:2. Поэтому в рамках группы *MPEG-2* был разработан дополнительный стандарт 422 Profile @ Main Level (422 P@ML).

Стандарт 422 P@ML является подмножеством основного профиля и основного уровня MP@ML. в том смысле, что все значения параметров первого либо равны, либо превышают соответствующие значения второго. Принцип обратной совместимости, заложенный в *MPEG-2*, гарантирует, что декодеры 422 P@ML способны декодировать цифровые потоки MP@ML.

Основные возможности стандарта 422 P@ML, превосходящие соответствующие свойства основного профиля и основного уровня MP@ML, заключаются в следующем:

- допускается кодирование сигнала по стандарту 4:2:2, в то время как MP@ML ограничен кодированием сигналов способом 4:2:0;
- цифровой поток кодированных видеоданных может принимать любое значение до 50 Мбит/с, а в MP@ML — только 15 Мбит/с;
- вертикальное разрешение ограничено значением 512 ТВ линий в случае 525-строчных систем и 608 ТВ линий в случае 625-строчных систем, а в MP@ML — значениями 480 и 576 ТВ линий соответственно;
- в 625-строчных системах кроме 576 активных строк стандарт 422 P@ML обеспечивает возможность кодирования еще 32 строк в кадре как составной части видеосигнала. Это позволяет пропускать напрямую через систему цифрового сжатия важные строки полевого интервала гашения. Поэтому стандарт 422 P@ML гарантирует пропускание такой информации как полевой временной код и сигналы испытательных Строк, не требуя отдельной обработки этих строк.

После того как была сформулирована профессиональная версия *MPEG-2 422 P@ML*, используемая для студийного производства, он получил статус полноценного международного стандарта, который иногда называют *422 Studio Profile/ML*. В результате появилась реальная возможность применения стандарта сжатия *MPEG-2* на всех участках технологической цепочки создания ТВ программ: от съемки до телезрителя, включая доставку новостных сюжетов на телецентр, студийную компоновку программ, их распространение и передачу в эфир. Стандарт предоставляет возможности эффективной работы во всех этих звеньях. Профили *MPEG-2* определяют набор способов и технических приемов по сжатию видеоданных, а уровни — такие параметры, как размер изображения или скорость цифрового потока при выбранном способе кодирования. Профиль 422 характеризуется высокой скоростью цифрового потока и относительно короткими группами

изображений, что позволяет монтировать ТВ программу с достаточно высоким качеством

Стандарт *MPEG-2 4:2:2 P@HL* (профиль 4:2:2 на высоком уровне), ориентирован на использование в системах ТВЧ.

Для кодирования звуковых сигналов с целью их компрессии в *MPEG-2* используется стандарт информационного сжатия звуковых данных *MUSICAM (Masked Pattern for Adapted Universal Coding and Multiplexing)*, обозначаемый также как *MPEG Layer II*. Стандарт *MUSICAM* позволяет передавать моно, стерео, многоязыковый и surround (пространственный) звук.

MUSICAM позволяет снизить скорость потока данных, необходимую для воспроизведения звуковых сигналов, адекватного качеству, получаемому при воспроизведении компакт-дисков, до 128 кбит/с на каждый моноканал звукового сопровождения. Таким образом, для самого низкого уровня — двухканального стереофонического звукового сопровождения — потребуется скорость передачи цифровых данных, равная $128 \text{ кбит/с} \times 2 = 256 \text{ кбит/с}$. В основе стандарта лежат два психоакустических эффекта. Дело в том, что человеческое ухо не способно различать звуки с громкостью ниже определенного минимума, так называемого «порога тишины». Кроме того, более тихие звуки «маскируются» более громкими. Соответственно, алгоритм *MUSICAM* обеспечивает передачу только тех звуков, которые реально различаются человеком.

В случае воспроизведения пространственного (*Surround*) звука для передачи шести сигналов звукового сопровождения в отсутствие компрессии потребуется пропускная способность канала, достигаемая 5,18 Мбит/с ($48\,000 \times 18 \times 6 = 5,18 \text{ Мбит/с}$, где 18 кГц — частота дискретизации звуковых сигналов, 18 бит — разрядность квантования). По протоколу для передачи звукового сопровождения предусмотрен канал с пропускной способностью 384 кбит/с. В данном случае требуемый коэффициент сжатия, равный 13, сможет обеспечить цифровая система компрессирования *Dolby AC-3*,

основанная на спектральном анализе звуковых сигналов и удалении частотных компонент, не слышимых человеком.

2.6. Сжатие звуковых данных

2.6.1. Эффект маскирования и психоакустическая модель слуха

Из-за повышенной чувствительности уха к искажениям сжатие звуковых сигналов оказывается более сложной задачей, чем видеокомпрессия. В то же время некоторые психофизиологические особенности слухового восприятия позволяют даже при существенной компрессии (6:1 и более) обеспечить «прозрачный» канал, то есть канал, звучание выходного сигнала которого субъективно неотличимо от звучания сигнала источника. Для понимания этих особенностей познакомимся вкратце с механизмом слухового восприятия.

Частотный диапазон восприятия простирается от 20 Гц до 20 кГц; область наибольшей чувствительности находится в интервале между 1 кГц и 5 кГц. Как и во всякой системе с преобразованием из временной в частотную область, в слуховой системе имеется некоторая неопределенность в разделении двух событий: чем точнее известны характеристики сигнала во временной области, тем менее точно можно описать его спектральный состав, и наоборот. Несовершенное разделение по частоте приводит к тому, что ухо не в состоянии различить звуки с близкими частотами, этот эффект называется *частотным маскированием*. Граница восприятия вблизи маскирующего тона называется *порогом маскирования*.

В результате длительных исследований удалось измерить ширину и расположение частотных полос, в пределах которых действует маскирование. Эти полосы получили название *критических* и в полосе слухового восприятия их насчитывается 25. В области низких слышимых частот ширина критической полосы менее 100 Гц, в районе 2 кГц она равна 300 Гц и возрастает до 4 кГц в области высших воспринимаемых частот (рисунок 2.8).

Измеряется ширина критической полосы в Барках и рассчитывается по формуле

$$B = f/100, \quad \text{для } f \leq 500 \text{ Гц};$$

$$B = 9 + 4 \log_2(f/1000), \quad \text{для } f > 500 \text{ Гц}.$$

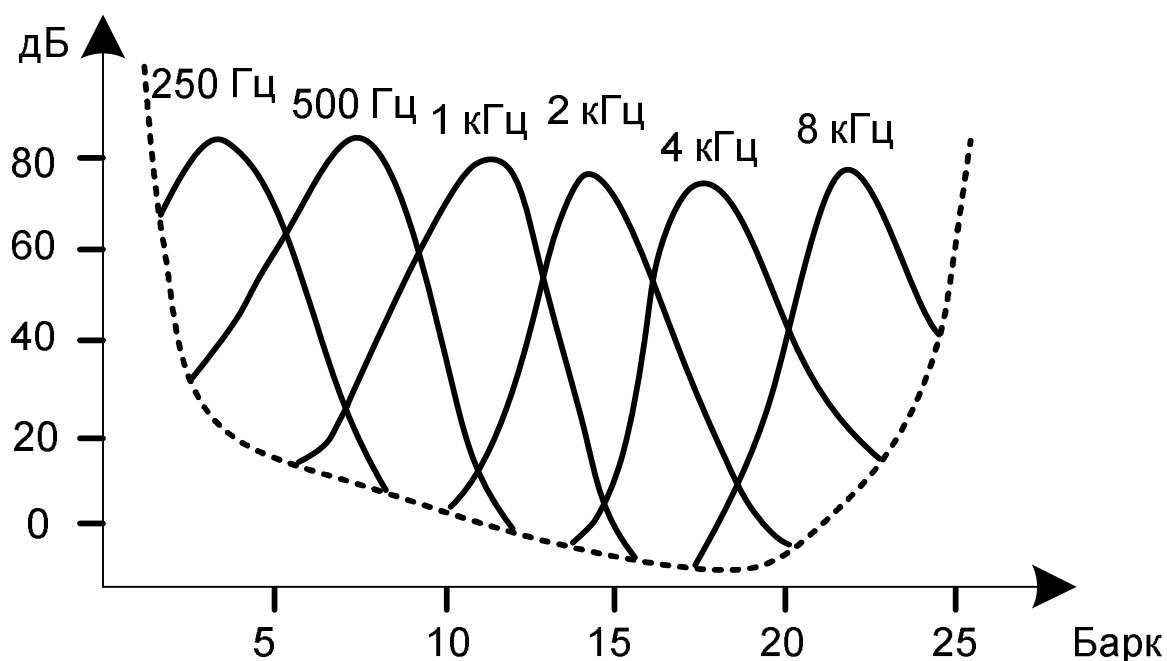


Рисунок 2.8 — Критические полосы человеческого слуха

Кроме частотного (статического), имеет место и временное (динамическое) маскирование. Дело в том, что слабый сигнал, возникающий сразу после окончания сильного сигнала, остается в течение некоторого времени незамеченным. С другой стороны, даже предшествующий слабый сигнал становится незаметным за несколько миллисекунд до появления сильного сигнала. Эти явления называются «маскированием вперед» и «маскированием назад» (рисунок 2.9). Несовершенство временного разделения связано с резонансным характером восприятия. Добротность резонансной системы уха такова, что нарастание колебаний происходит за время порядка (1...2) мс.

Использование эффекта маскирования позволяет существенно сократить объем звукоданных, сохраняя приемлемое качество звучания. Принцип здесь очень простой: «Если какая-то составляющая не слышна, то и передавать ее не надо». На практике это означает, что в области маскирования можно уменьшить число битов на отсчет до такой степени, чтобы шум квантования все еще оставался ниже порога маскирования. Таким образом, для работы звукового кодера необходимо знать пороги маскирования при различных комбинациях воздействующих сигналов. Вычислением этих порогов занимается важный узел в кодере — *психоакустическая модель слуха* (ПАМ).

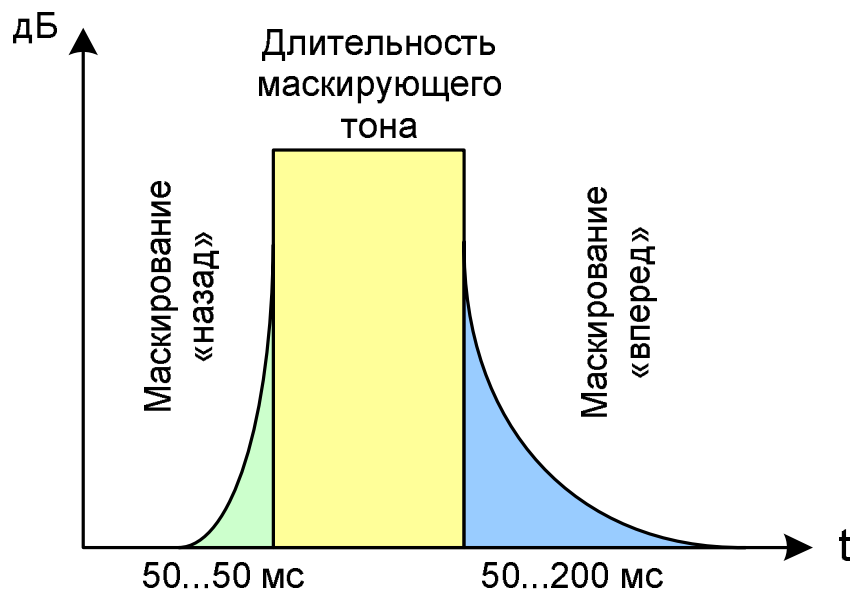


Рисунок 2.9 — Эффект временного маскирования

ПАМ анализирует входной сигнал в последовательные отрезки времени и определяет для каждого блока отсчетов спектральные компоненты и соответствующие им области маскирования. Входной сигнал анализируется в частотной области, для этого блок отсчетов, взятых во времени, с помощью дискретного преобразования Фурье преобразуется в набор коэффициентов при компонентах частотного спектра сигнала.

Разработчики кодеров компрессии имеют значительную свободу в построении модели, точность ее функционирования зависит от требуемой

степени сжатия. При невысоком сжатии можно обойтись и без ПАМ, однако на высоких уровнях компрессии она играет важную роль.

2.6.2. Полосное кодирование и блок фильтров

Наилучшим методом кодирования звука, учитывающим эффект маскирования, оказывается полосное кодирование. Сущность его заключается в следующем. Группа отсчетов входного звукового сигнала, называемая кадром, поступает на *блок фильтров* (БФ), который содержит, как правило, 32 полосовых фильтра. Учитывая сказанное ранее о критических полосах и маскировании, хорошо бы иметь в блоке фильтров полосы пропускания, по возможности, совпадающие с критическими. Однако практическая реализация цифрового блока фильтров с неравными полосами сложна и оправдана только в устройствах самого высокого класса. Обычно используется блок фильтров на основе *квадратурно-зеркальных фильтров* (КЗФ) с равными полосами пропускания, охватывающих с небольшим взаимным перекрытием всю полосу слышимых частот (рисунок 2.10).

При частоте дискретизации 48 кГц полоса пропускания секции фильтра составляет 750 Гц. Каждая секция фильтра позволяет проводить расщепление полосы пропускания на две равные части, не увеличивая при этом объем данных в каждой половине полосы, т.е. с одновременной понижающей дискретизацией. Расщепление на $N = 2^m$ полос производится каскадным включением m секций, содержащих соответственно 1, 2, 4... 2^{m-1} фильтров (рисунок 2.11).

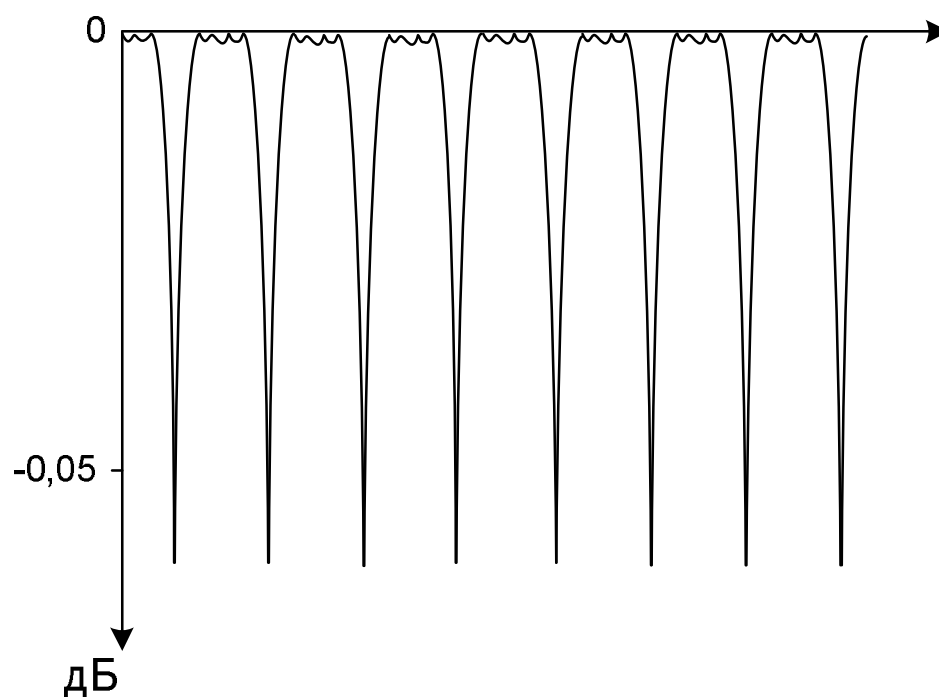


Рисунок 2.10 — Номинальная АЧХ блока фильтров

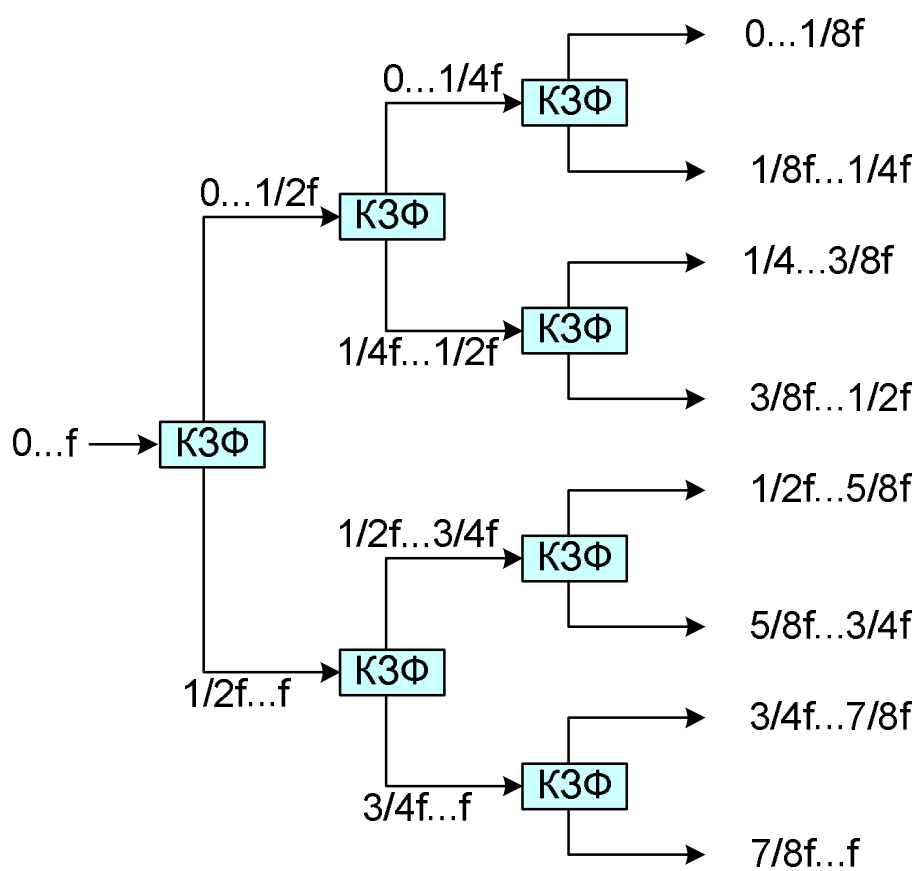


Рисунок 2.11 — Трехсекционный квадратурно-зеркальный фильтр

На выходе каждого фильтра оказывается та часть входного сигнала, которая попадает в полосу пропускания данного фильтра. Далее в каждой полосе с помощью ПАМ анализируется спектральный состав сигнала и оценивается, какую часть сигнала следует передавать без сокращений, а какая лежит ниже порога маскирования и может быть переквантована на меньшее число битов.

Поскольку в реальных звуковых сигналах максимальная энергия обычно сосредоточена в нескольких частотных полосах, может оказаться, что сигналы в других полосах не содержат различимых звуков и могут вообще не передаваться. Наличие, например, сильного сигнала в одной полосе означает, что несколько вышележащих полос будут маскироваться и могут кодироваться меньшим числом битов.

Для сохранения максимального динамического диапазона определяется максимальный отсчет в кадре и вычисляется масштабирующий множитель, который приводит этот отсчет к верхнему уровню квантования. На этот же множитель умножаются и все остальные отсчеты. Масштабирующий множитель передается к декодеру вместе с кодированными данными для коррекции коэффициента передачи последнего. После масштабирования производится оценка порога маскирования и осуществляется перераспределение общего числа битов между всеми полосами.

2.6.3. Квантование и распределение битов

Все описанные операции не сокращали заметно объем данных, они были как бы подготовительным этапом к собственно сжатию звукоданных. Как и при компрессии цифровых видеосигналов, основное сжатие происходит в квантователе. Исходя из принятых ПАМ решений о переквантовании отсчетов в отдельных частотных полосах, квантователь изменяет шаг квантования таким образом, чтобы приблизить шум квантования в данной полосе к вычисленному порогу маскирования. При этом на отсчет может понадобиться вместо 16...20 всего 4 или 5 битов! Возможны три стратегии распределения битов.

В системе с *прямой адаптацией* кодер производит все расчеты и посылает результаты декодеру. Преимущество данного способа в том, что алгоритм

распределения битов может обновляться и изменяться, не затрагивая работы декодера. Однако для пересылки дополнительных данных декодеру расходуется заметная часть общего запаса битов. Система с *обратной адаптацией* осуществляет одинаковые расчеты и в кодере, и в декодере, поэтому нет необходимости пересылать декодеру дополнительные данные. Однако сложность и стоимость декодера значительно выше, чем в предыдущем варианте, и любое изменение алгоритма требует обновления или переделки декодера. Компромиссная система с *прямой и обратной адаптацией* разделяет функции расчета распределения битов между кодером и декодером таким образом, что кодер производит наиболее сложные вычисления и посылает декодеру только ключевые параметры, затрачивая на это относительно немного битов, декодер проводит лишь несложные вычисления. В такой системе кодер не может быть существенно изменен, но настройка некоторых параметров допустима.

Обобщенная схема звукового кодера и декодера, выполняющих цифровое сжатие согласно описанному алгоритму с прямой адаптацией, приведена на рисунке 2.12,а.

Сигналы на выходе частотных полос объединяются в единый цифровой поток с помощью мультиплексора. В декодере процессы происходят в обратном порядке. Сигнал демультимплексируется, делением на масштабирующий множитель восстанавливаются исходные значения цифровых отсчетов в частотных полосах и поступают на объединяющий блок фильтров, который формирует на выходе поток звукоданных, адекватный входному, с точки зрения психофизиологического восприятия, звукового сигнала человеческим ухом.

Вариант схемы с обратной адаптацией показан на рисунке 2.12,б.

2.7. Алгоритмы сжатия звукоданных кодерами первого и второго уровней

Рассмотрим более подробно работу кодера первого уровня (рисунок 2.13). БФ обрабатывает одновременно 384 отсчета звукоданных и распределяет их с соответствующей субдискретизацией в 32 полосы, по 12 отсчетов в каждой полосе с частотой дискретизации $48/32 = 1,5$ кГц.

Длительность кадра при частоте дискретизации 48 кГц составляет 8 мс. Упрощенная ПАМ оценивает только частотное маскирование по наличию и «мгновенному» уровню компонентов сигнала в каждой полосе.

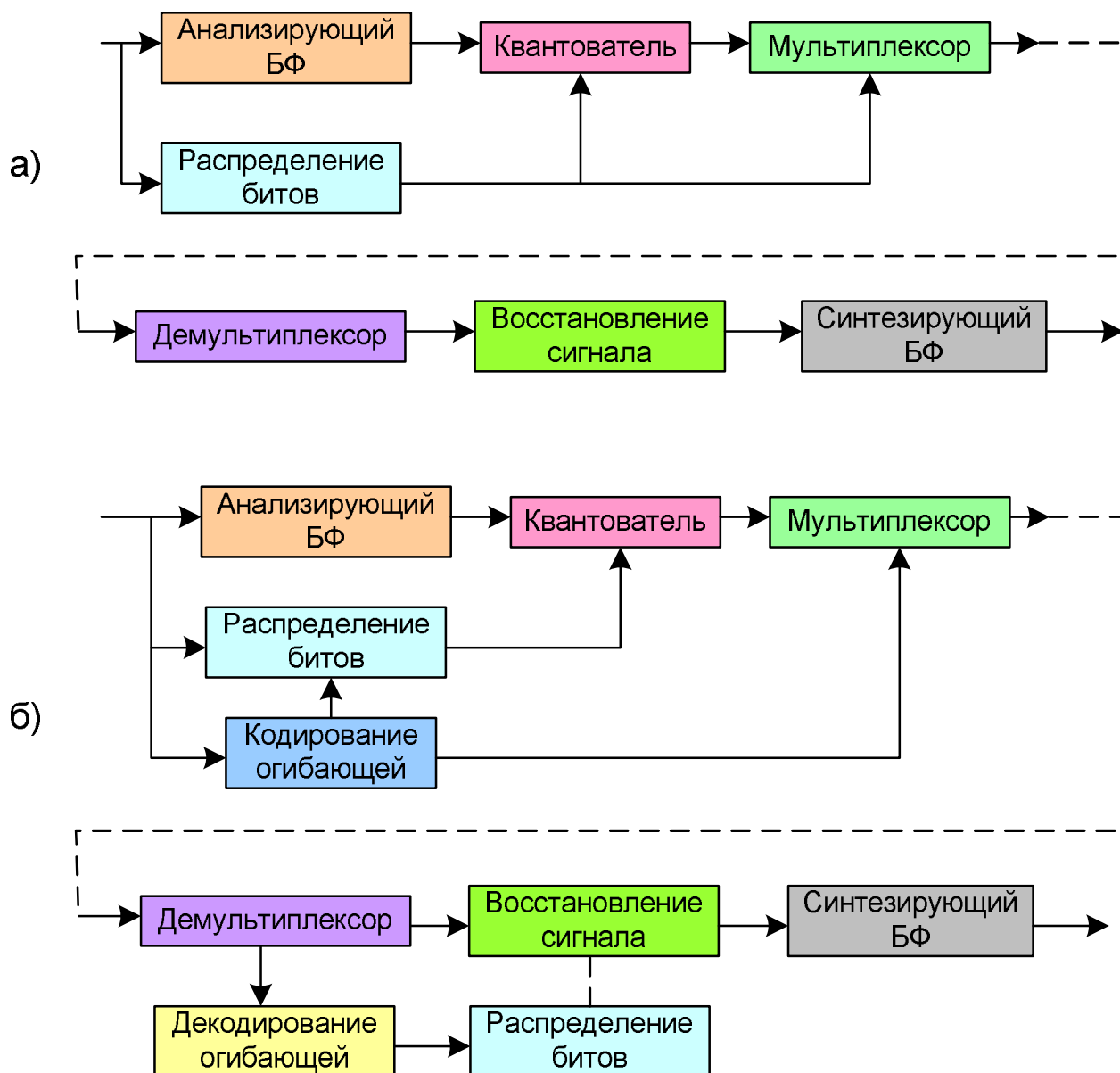


Рисунок 2.12 — Обобщенная структурная схема звукового кодера и декодера:
а — с прямой адаптацией; б — с обратной адаптацией

По результатам оценки для каждой полосы назначается как можно более грубое квантование, но так, чтобы шум квантования не превышал порога маскирования. Масштабирующие множители имеют разрядность 6

бит и перекрывают динамический диапазон 120 дБ с шагом 2 дБ ($2^6 = 64 \approx 120/2$). В цифровом потоке передаются также 32 кода распределения битов. Они имеют разрядность 4 бита и указывают на длину кодового слова отсчета в данной полосе после переквантования.

В декодере (см. рисунок 2.14) отсчеты каждой частотной полосы выделяются демультиплексором и поступают на перемножитель, который восстанавливает их первоначальный динамический диапазон. Перед этим восстанавливается исходная разрядность отсчетов — отброшенные в квантователе младшие разряды заменяются нулями. Коды распределения битов помогают демультиплексору разделить в последовательном потоке кодовые слова, принадлежащие разным отсчетам и передаваемые кодом с переменной длиной слова. Затем отсчеты всех 32 каналов подаются на синтезирующий БФ, который проводит повышающую дискретизацию и расставляет отсчеты должным образом во времени, восстанавливая исходную форму сигнала.

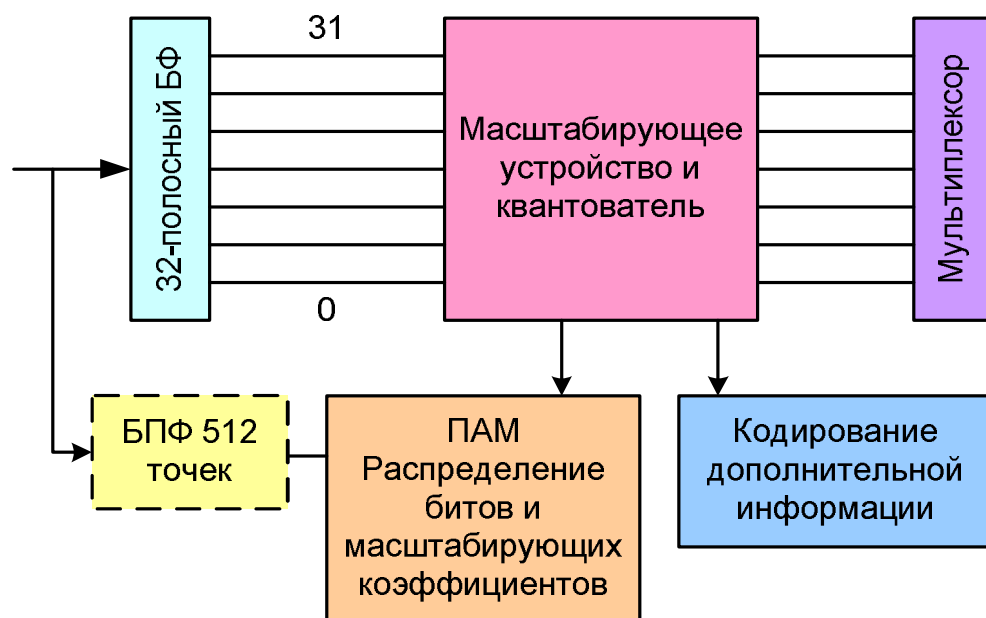


Рисунок 2.13 — Структурная схема звукового кодера MPEG-1 первого и второго уровней (пунктиром показан модуль быстрого преобразования Фурье (БПФ), добавляемый на втором уровне)

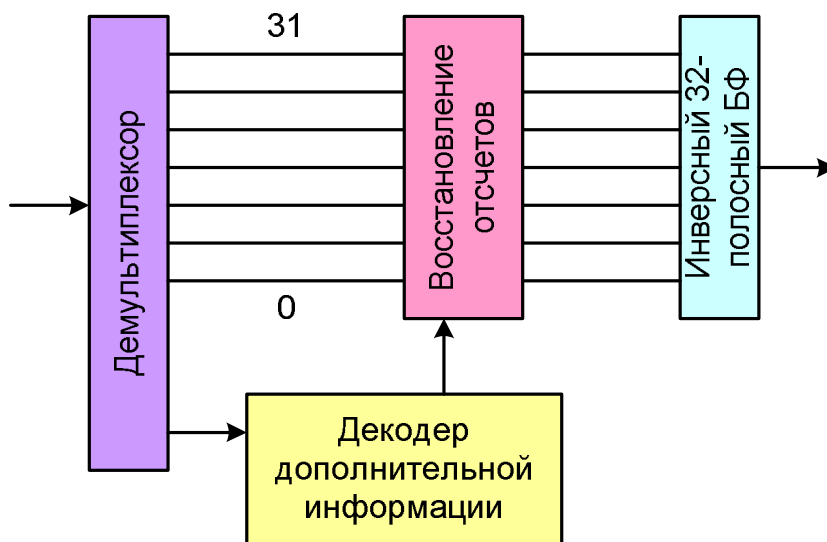


Рисунок 2.14 — Структурная схема звукового декодера MPEG-1 первого и второго уровней

В кодере второго уровня устранены основные недостатки базовой модели полосного кодирования, связанные с несоответствием критических полос слуха и реальных полос БФ, из-за чего в низкочастотных участках диапазона эффект маскирования практически не использовался. Величина кадра увеличена втрое, до 24 мс при дискретизации 48 кГц, одновременно обрабатываются уже 1152 отсчета (3 субкадра по 384 отсчета). В качестве входного сигнала для ПАМ используются не полосные сигналы с выхода БФ, а спектральные коэффициенты, полученные в результате 512-точечного преобразования Фурье входного сигнала кодера. Благодаря увеличению и временной длительности кадра и точности спектрального анализа эффективность работы ПАМ возрастает.

На втором уровне применен более сложный алгоритм распределения битов. Полосы с номерами от 0 до 10 обрабатываются с 4-разрядным кодом распределения (выбор любой из 15 шкал квантования), для полос с номерами от 11 до 22 выбор сокращается до 3 разрядов (выбор одной из 7 шкал), полосы с номерами от 23 до 26 предоставляют выбор одной из 3 шкал (2-битовый код), а полосы с номерами от 27 до 31 (выше 20 кГц) не передаются.

Если шкалы квантования, выбранные для всех блоков кадра, оказываются одинаковыми, то номер шкалы передается только один раз.

Еще одно существенное отличие алгоритма второго уровня в том, что не все масштабирующие множители передаются по каналу связи. Если различие множителей трех последовательных субкадров превышает 2 дБ не более чем в течение 10% времени, передается только один набор множителей и это дает экономию расходуемых битов. Если в данной полосе происходят быстрые изменения уровня звука, передаются два или все три набора масштабирующих множителей. Соответственно декодер должен запоминать номера выбранных шкал квантования и масштабирующие множители и применять их при необходимости к последующему субкадру.

2.8. Алгоритм сжатия звуковых данных кодерами третьего уровня

Кодер Уровня III использует усовершенствованный алгоритм кодирования с дополнительным ДКП. Структурная схема кодера показана на рисунке 2.15.

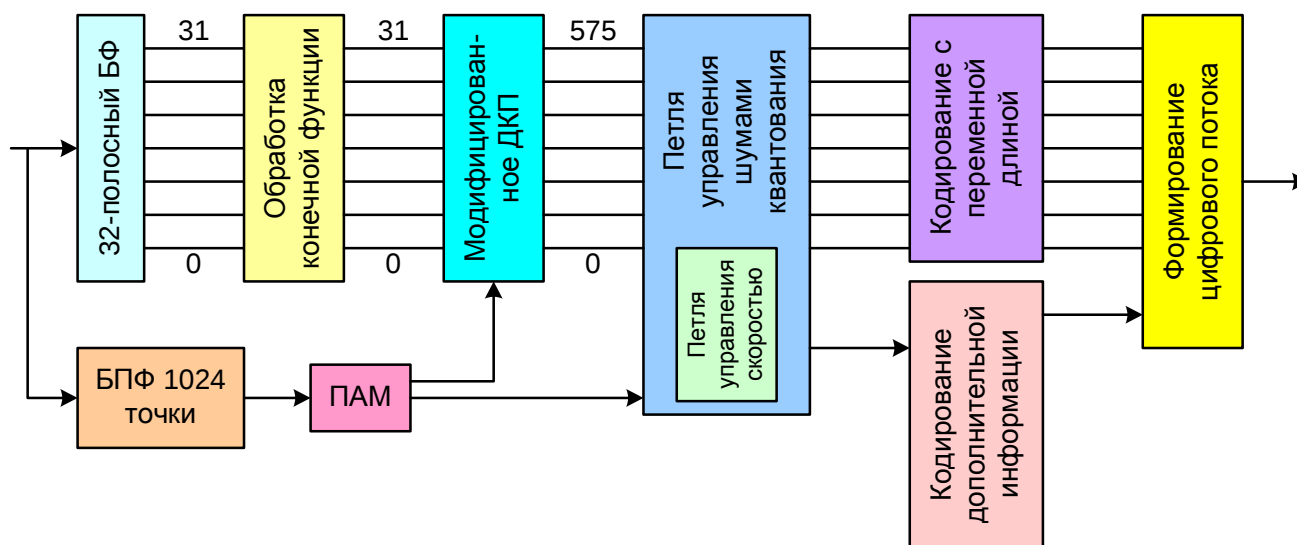


Рисунок 2.15 — Структурная схема звукового кодера третьего уровня

Основной недостаток кодеров второго уровня — неэффективная обработка быстро изменяющихся переходов и скачков уровня звука —

устраняется благодаря введению двух видов блоков ДКП — «длинного» с 18 отсчетами и «короткого» с 6 отсчетами.

Выбор режима осуществляется адаптивно путем переключения оконных функций в каждой из 32 частотных полос. Длинные блоки обеспечивают лучшее частотное разрешение сигнала со стандартными характеристиками, в то время как короткие блоки улучшают обработку быстрых переходов. В одном кадре могут быть как длинные, так и короткие блоки, однако общее число коэффициентов ДКП не изменяется, так как вместо одного длинного передаются три коротких блока.

Для улучшения кодирования применяются также следующие усовершенствования:

1. Неравномерное квантование (квантователь возводит отсчеты в степень $3/4$ перед квантованием для улучшения отношения сигнал-шум; соответственно, декодер возводит их в степень $4/3$ для обратной линеаризации).

2. В отличие от кодеров первого и второго уровней, на третьем уровне масштабирующие множители присваиваются не каждой из 32 частотных полос БФ, а полосам масштабирования — участкам спектра, не связанным с этими полосами и примерно соответствующим критическим полосам.

3. Энтропийное кодирование квантованных коэффициентов кодом Хаффмана.

4. Наличие «резервуара битов» — запаса, который кодер создает в периоды стационарного входного сигнала.

Улучшение частотного разрешения при введении дополнительного ДКП влечет за собой значительное ухудшение временного разрешения, проявляющееся в виде предэхо. В кодере имеются средства для борьбы с этим явлением. Во-первых, ПАМ модифицирована для обнаружения условий возникновения предэхо, во-вторых, имеется резервуар неиспользованных битов, из которого кодер может занять на короткое время необходимые ему

биты для снижения шумов квантования, в-третьих, кодер может переключаться в режим коротких блоков.

Существенное отличие кодера третьего уровня от кодеров нижних уровней сложный адаптивный алгоритм распределения битов. Он включает две вложенные одна в другую итерационные петли: внутреннюю — петлю скорости потока, и внешнюю петлю управления шумами квантования. Расчет параметров кодирования начинается со значения масштабирующего множителя, равного 1. Если в первый момент скорость потока на выходе кодера Хаффмана превышает заданную, увеличивается шаг квантования до тех пор, пока скорость не войдет в заданные пределы. Далее рассчитывается шум квантования в данной частотной полосе и сравнивается с порогом маскирования, сообщенным психоакустической моделью. Масштабирующий множитель изменяется таким образом, чтобы сблизить эти значения. Однако новое значение масштабирующего множителя означает изменение шага квантования и, следовательно, скорости потока, а потому внутренняя петля должна каждый раз отрабатывать и вычислять новое значение коэффициента усиления и шага квантования. Если итерационный процесс во внутренней петле всегда сходится, то во внешней петле он может расходиться. Итерационный процесс заканчивается в одном из трех случаев:

1. Шумы квантования во всех полосах масштабирования не превышают допустимых.
2. Следующая итерация приведет к возрастанию усиления в одной из полос выше допустимого.
3. Следующая итерация требует увеличения усиления во всех полосах масштабирования.

В реальных кодерах накладывается еще и временное ограничение.

Кодер третьего уровня более полно обрабатывает стереосигнал в формате *joint stereo*. Если кодеры нижележащих уровней работают только в режиме кодирования по интенсивности, когда левый и правый каналы в полосах выше 2 кГц кодируются как один сигнал (но с независимыми

масштабирующими множителями), кодер третьего уровня может работать и в режиме «сумма-разность», обеспечивая более высокую степень сжатия разностного канала.

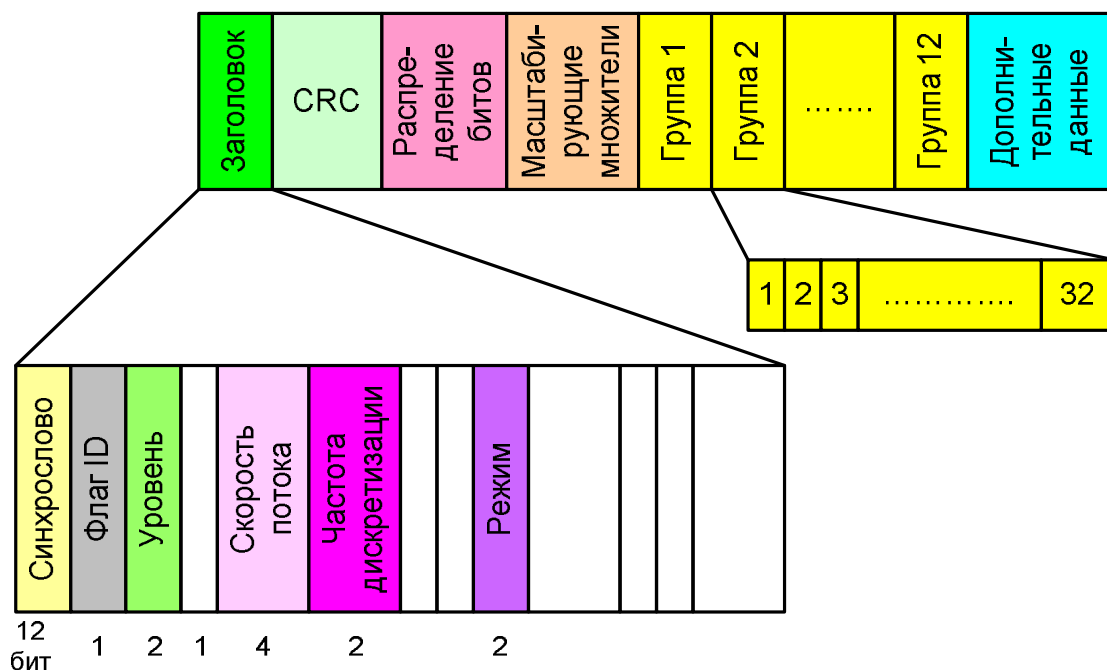


Рисунок 2.15 — Структура потока звукового кодера I и II уровней

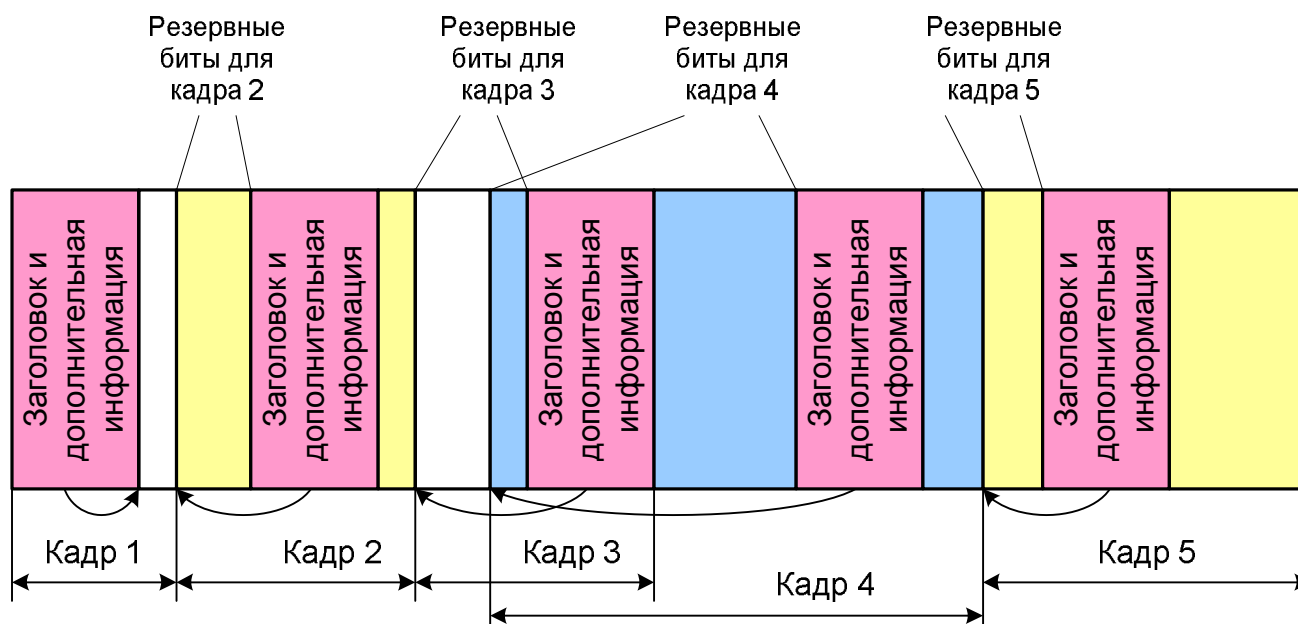


Рисунок 2.16 — Структура потока звукового кодера III уровня

Структура элементарных потоков звукоданных несколько различается для разных уровней. На рисунке 2.16 показана последовательность звукоданных Уровня I. Пакет начинается с заголовка длиной 32 бита, имеющего одинаковую структуру для всех уровней, который включает синхрослово из 12 единиц и более десятка указателей и флагов. Флаг *ID* указывает принадлежность потока («1» — *MPEG-1*, «0» — *MPEG-2*), флаг уровня идентифицирует уровень кодирования («01» = 1, «10» = 2, «11» = 3), четырехбитовый указатель скорости потока ссылается на стандартизованную таблицу, где перечислены возможные скорости потока, табулированы также значения частоты дискретизации, которые сообщает соответствующий указатель (для *MPEG-1* только 48, 44,1 и 32 кГц), флаг режима определяет число каналов звука (не более двух для *MPEG-1*).

Следующее поле после заголовка занимают коды распределения битов, затем в потоке передаются 32 масштабирующих множителя и 384 отсчета звукового кадра, группами по 32 отсчета (по одному отсчету из каждого поддиапазона). Замыкает пакет поле дополнительных данных переменной длины, используемое в стандарте *MPEG-2* для передачи многоканальных сигналов, отсчетов с пониженной частотой дискретизации и других данных, не совместимых с синтаксисом *MPEG-1*.

Отличие формата Уровня II от Уровня I незначительно и заключается в передаче дополнительной информации о выборе шкалы квантования для каждой из 32 полос, так как не все масштабирующие множители могут передаваться в потоке. Более сложная структура пакета у кодера третьего уровня. Здесь для каждого звукового кадра передаются заголовок, дополнительная информация и основные данные, причем неиспользованная часть области основных данных предшествующих кадров (резервуар битов) может быть заимствована последующим кадром, испытывающим нехватку битов (рисунок 2.16).

3. ФОРМИРОВАНИЕ ЦИФРОВЫХ ПОТОКОВ ВИДЕО И ЗВУКА

Цифровой сигнал непосредственно с выхода кодера MPEG-2 называется «сырым» (необработанным) *элементарным потоком* (ES) и представляет собой бесконечную последовательность видео- или звукоданных в том порядке, как они появляются на выходе в процессе кодирования. Он не содержит необходимой информации для идентификации потока, для его синхронизации с другими потоками и не может непосредственно использоваться для организации вещания или других служб.

Для дальнейшего использования элементарные потоки переформируют в *пакетированные элементарные потоки* (PES), в которых данные разделены на пакеты удобного размера с заголовками, содержащими необходимую информацию о потоках и синхронизации. В зависимости от приложений длина PES-пакета может составлять от единиц до 64 килобайт. Обычно пакет содержит одну единицу воспроизведения данных определенного типа. Единица воспроизведения видеопотока — один видеокادر, звукопотока — один звуковой кадр AES / EBU. Заголовок (рисунок 3.1) начинается со стартового кода длиной 3 байта и однобайтового идентификатора потока, показывающего вид информации (видео, аудио, вспомогательные данные) и параметры кодирования.

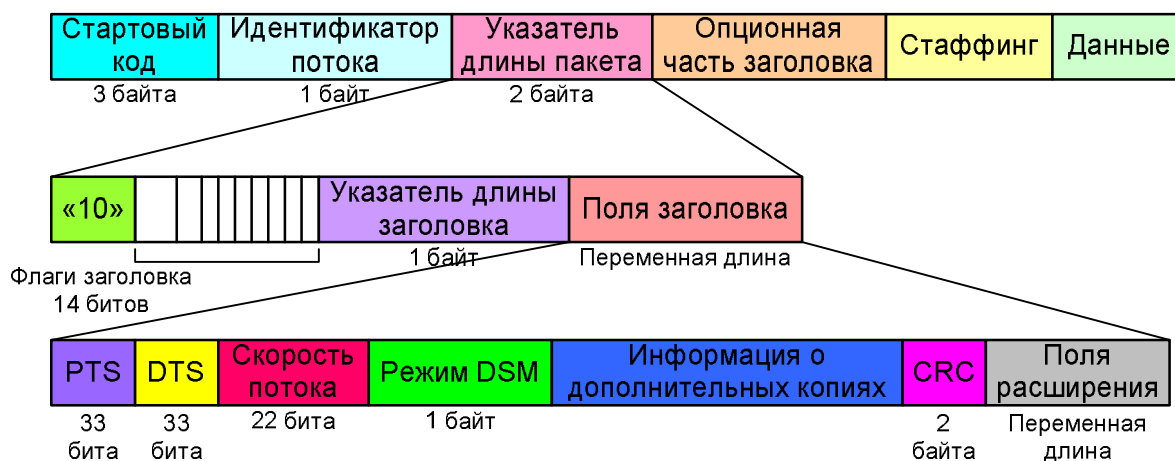


Рисунок 3.1 — Структура PES-пакета

Следующие 2 байта — это указатель длины пакета, он показывает число байтов в пакете. Далее следуют 12 флагов, несущих сведения о свойствах цифрового потока. Они сообщают информацию о скремблировании, о приоритете пакета, условия авторского права, тип временной метки (присутствует ли только *PTS* или *PTS* и *DTS*), наличие или отсутствие определенных полей (поля скорости потока, режима *DSM*, информации о дополнительных копиях, полей расширения и др.). Опционные поля входят в состав заголовка, если соответствующие флаги установлены на «1». Поле режима *DSM* (цифровой видеозаписи) содержит биты, управляющие движением ленты (перемотка, стоп-кадр, замедленное воспроизведение) и другими видеоэффектами; информация о дополнительных копиях необходима для решения вопросов авторского права. Поле длины заголовка указывает размер заголовка в байтах.

Привязка *PES*-сборки к определенному моменту времени 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. Полный цикл счетчика завершается за 26 с небольшим часов.

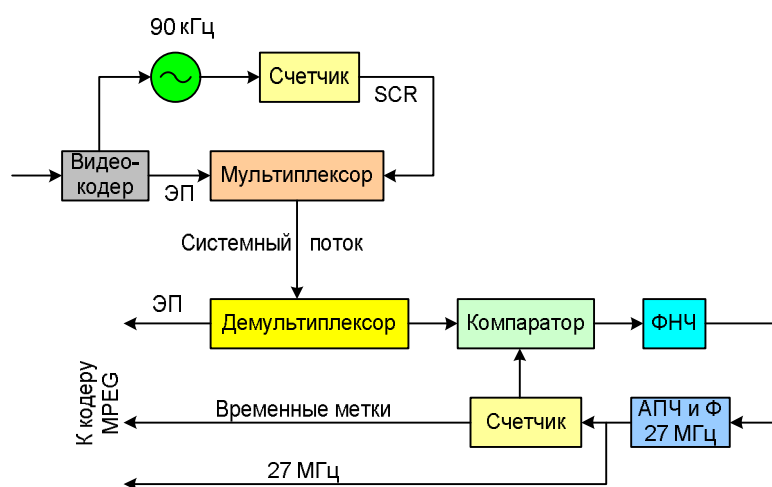


Рисунок 3.2 — Восстановление тактовой частоты в приемнике *MPEG-2* по меткам *SCR*

Основная задача, решаемая системными часами, — обеспечение синхронизации тактовых частот кодера и декодера. Она решается с помощью схемы дискретной автоподстройки частоты в декодере (рисунок 3.2).

Частота местного генератора подсчитывается с помощью многоразрядного счетчика и периодически сравнивается с приходящими от кодера значениями SCR. Выходной сигнал компаратора, пропорциональный разности двух чисел, управляет через ФНЧ частотой генератора и корректирует ее в соответствии со знаком разности. На передающей стороне точность поддержания частоты должна быть не хуже $3 \cdot 10^{-5}$.

При двунаправленном кодировании, как отмечалось, порядок следования и порядок передачи *P*- и *B*-кадров не совпадают, и для синхронизации используются два вида временных меток. *Временная метка декодирования* (*DTS — Decoding Time Stamp*) указывает время, когда видеокادر должен быть декодирован, а *временная метка воспроизведения* (*PTS — Presentation Time Stamp*) — момент времени, когда он должен появиться на выходе декодера. Если, например, принимается последовательность *IPBB* декодер декодирует *I*-кадр и запоминает его, в следующий момент он декодирует *P*-кадр и выдает на выход *I*-кадр, затем вычисляет и выдает кадр *B*₁, затем *B*₂ и только после этого *P*₁ (рисунок 3.3).

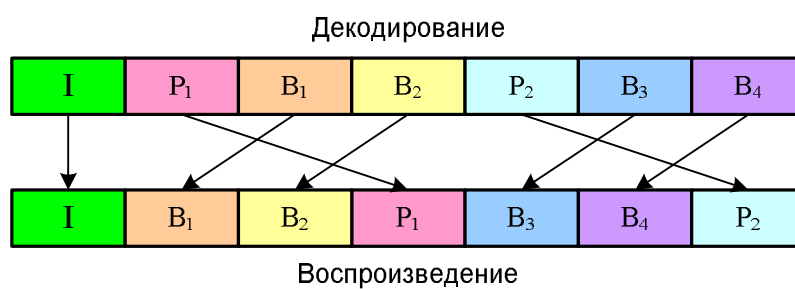


Рисунок 3.3 — Восстановление порядка следования видеок кадров по меткам *DTS* и *PTS*

Временные метки *PTS* и *DTS* формируются по тому же принципу, что и ссылка на системные часы (*SCR*).

Привязка к временным меткам позволяет обеспечить одинаковую задержку кодера и декодера для всех составляющих программы при их независимой передаче с разбивкой на пакеты, синхронизировать элементарные потоки видео и аудио и образовать из них ТВ программу. Декодер может интерполировать временные промежутки между двумя появлениями временной метки, поэтому нет необходимости размещать метку в каждой единице воспроизведения.

При формировании единого потока данных стандарт *MPEG-2* предусматривает две формы такого потока: программный (*Program Stream, PS*) и транспортный (*Transport Stream, TS*).

Программный поток ПП (PS). Структура *PS* включает пачки (*pack*), содержащие один или несколько *PES*-пакетов и заголовков со ссылкой на системные часы (*SCR — System Clock Reference*) (рисунок 3.4). В потоке может быть до 16 видео и до 32 звуковых потоков, но все они считаются компонентами одной программы, так как имеют общую временную базу (единый генератор тактовой частоты).



Рисунок 3.4 — Структура программного потока *MPEG-2*

Транспортный поток ТП (TS) представляет собой более высокий уровень организации данных. В *TS* пакетированные элементарные потоки, принадлежащие разным программам, переносятся в различных транспортных пакетах небольшой длины, снабженных кодозащитой для передачи в каналах с ошибками, в частности, для вещания. Один транспортный поток может переносить несколько программ, не связанных единой временной базой, каждая из нескольких компонентов. Передача оказывается, по сути дела,

асинхронной (*Asynchronous Transfer Mode, ATM*) и потому не может управляться единым синхронизирующим сигналом. Рассмотрим, как решены вопросы синхронизации в транспортном потоке *MPEG-2*.

Пакеты *TS* имеют фиксированную длину 188 байт, из которых 4 байта выделяются на заголовок и 184 байта — на полезные данные. Это могут быть видео- или звукоданные, данные пользователя или пустые байты (1111 1111), называемые *стаффингом*. Каждый пакет переносит данные только одного вида. Размер пакета выбран для совместимости с широко применяемым на сетях связи стандартом асинхронной передачи данных *ATM*. Ячейка *ATM* имеет длину 53 байта, из которых полезных — 47 байт (с учетом 1 байта на уровень адаптации). Один пакет *MPEG-2* длиной 188 байт упаковывается в 4 пакета *ATM*. Структура заголовка показана на рисунке 3.5.

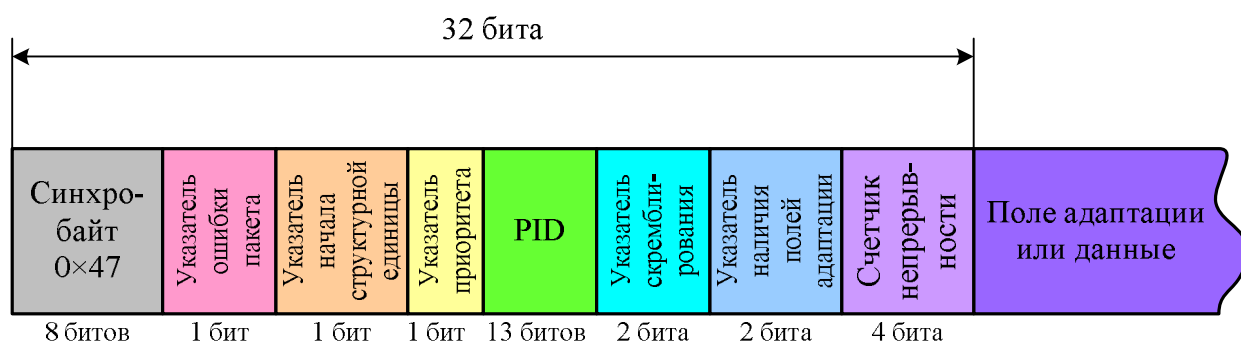


Рисунок 3.5 — Структура заголовка пакета транспортного потока *MPEG-2*

Первый байт — байт синхронизации — содержит зарезервированное кодовое число 0×47, легко опознаваемое демультимплексором. Далее идут три однобитовых флага (ошибки передачи, начала *PES*-пакета и приоритетной передачи) и 13-битовый идентификатор типа пакета *PID* (*Packet IDentifier*), указывающий на принадлежность пакета тому или иному потоку данных. *PID* служит основным признаком, по которому демультимплексор сортирует приходящие пакеты на приемной стороне. Из общего числа 8192 возможных значений *PID* 16 выделены на общесистемные цели, номер 8191 (0 × 1FFF)

закреплен за стаффиновыми байтами, остальные могут назначаться пользователями произвольно для отдельных компонент своих программ.

Передача информации пакетами фиксированной длины и система идентификаторов имеют свои преимущества. Пропускная способность в этом случае используется в максимальной степени и может динамически перераспределяться между программами. Можно добавлять новые *PES* или удалять ненужные, не меняя в целом структуру потока. Для извлечения необходимой информации декодеру не нужно знать детальную структуру всего потока, достаточно указать только *PID*, который всегда находится в заголовке на одном и том же месте.

Еще один байт в заголовке содержит три необходимых указателя:

2-битовый указатель скремблирования TS — свидетельствует о наличии или отсутствии скремблирования *TS*;

2-битовый индикатор наличия полей адаптации в нагрузке пакета.

Поле адаптации занимает часть области полезных данных и служит для ввода управляющих и вспомогательных сигналов, передаваемых не в каждом пакете. В частности, в полях адаптации передаются таблицы *программно-зависимой информации* (*PSI — Program Specific Information*) и сигналы синхронизации. Поле адаптации может также использоваться для передачи данных пользователя, в этом случае оно разбивается на секции.

Значение указателя наличия полей адаптации 0×1 означает, что поле адаптации отсутствует, и в пакете передаются данные видео или звука; 0×2 присваивается пакету, в котором часть области полезных данных занимает поле адаптации, остальное занято стаффином; индикатор, установленный на 0×3 , свидетельствует об одновременном наличии и поля адаптации, и полезных данных. Отметим, что заголовок пакета и поле адаптации не скремблируются.

4-битный счетчик непрерывности пакетов увеличивает свое значение на единицу при поступлении каждого следующего пакета с данным *PID* и

обнуляется после каждого 15-го пакета. Он позволяет декодеру обнаруживать потерю пакета и принимать меры по его замене.

Область полезных данных транспортного пакета значительно меньше, чем обычная длина *PES*-пакета, поэтому последний для укладки в пакеты *TS* должен разрезаться. Стандарт определяет, что первый байт *PES*-пакета должен обязательно совпадать с первым байтом области полезных данных, а конец *PES*-пакета — с концом одного из последующих транспортных пакетов. Незаполненные начальные байты этого последнего пакета дополняются пустыми байтами.

Если в пакете присутствует поле адаптации, оно имеет структуру, показанную на рисунке 3.6.

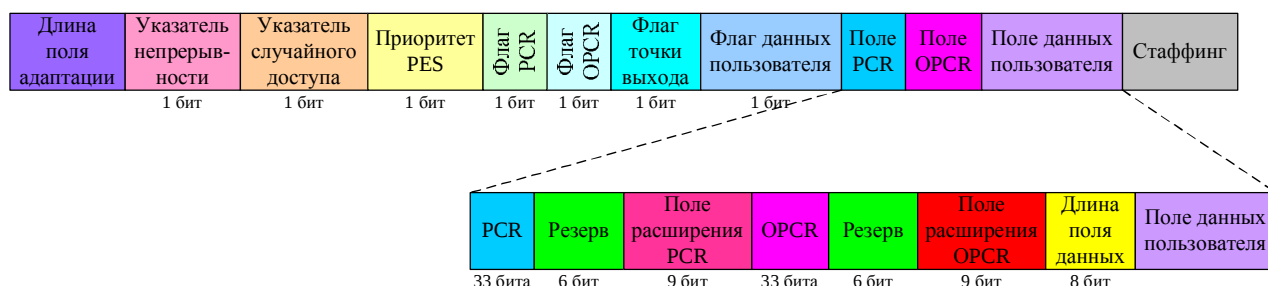


Рисунок 3.6 — Структура поля адаптации транспортного потока

Первый байт указывает длину поля, затем следуют три однобитовых указателя — непрерывности, случайного доступа и приоритета элементарного потока. Первый указатель свидетельствует о непрерывности счета времени во временных метках и необходим на приеме для мониторинга *TS*. Он устанавливается в «1», если на передающей стороне изменена база отсчета времени, например, гладко введена другая программа с иным *PCR*.

Если в данном поле адаптации передается *PCR*, об этом свидетельствует однобитовый флаг *PCR*. Метка *PCR* занимает 48 битов, в том числе 33 бита собственно *PCR*, 6 битов зарезервировано для будущих применений и 9 битов занимает поле расширения *PCR*. В поле расширения

ведется счет импульсов тактовой частоты 27 МГц. Как только их число достигает 300, поле обнуляется, и значение *PCR* увеличивается на единицу. Так обеспечивается совместимость с *MPEG-1*, где отсчет времени ведется в периодах тактового генератора 90 кГц. Такой же формат имеет метка *Исходной PCR (OPCR — Original PCR)*, передаваемой вместе с программой, заимствованной из другого транспортного потока. В отличие от *PCR*, значение *OPCR* не изменяется в процессе передачи. Оно может использоваться, например, при записи и воспроизведении программы.

Еще один флаг, *splice_countdown*, указывает число пакетов с тем же *PID* в транспортном потоке, оставшихся до точки гладкого входа в поток (например, для ввода рекламы или местных программ).

Оставшуюся часть поля адаптации могут занимать либо служебные данные, либо данные пользователя, либо расширения поля адаптации, на что указывают соответствующие флаги. Формат данных пользователя включает однобайтовый указатель длины поля и собственно поле данных, которые в этом случае вводятся в так называемые *секции транспортного потока*.

3.1. Мультиплексирование, ремультиплексирование и статистическое мультиплексирование

Принятая в стандарте *MPEG-2* схема построения ТП может быть описана как двухступенчатая. На первом этапе формируется программный поток (ПП) путем мультиплексирования одного или нескольких *PES* с общей временной базой и индивидуальными *PID* (рисунок 3.7,а).

На втором шаге ПП нескольких программ и управляющие потоки объединяются методом асинхронного пакетного мультиплексирования в единый транспортный поток (рисунок 3.7,б). Транспортный поток может быть построен и непосредственно из *PES* или других ТП, если при этом сохраняются общие правила синтаксиса *MPEG*. Такая иерархическая структура обеспечивает большую гибкость в построении систем вещания. Можно, например, объявить один *PES* принадлежащим более чем одной программе и организовать так называемые *виртуальные каналы*, можно

использовать несколько потоков одного вида для масштабирования, возможна организация межрегионального вещания с выбором отдельного языка для каждого региона и т.д.

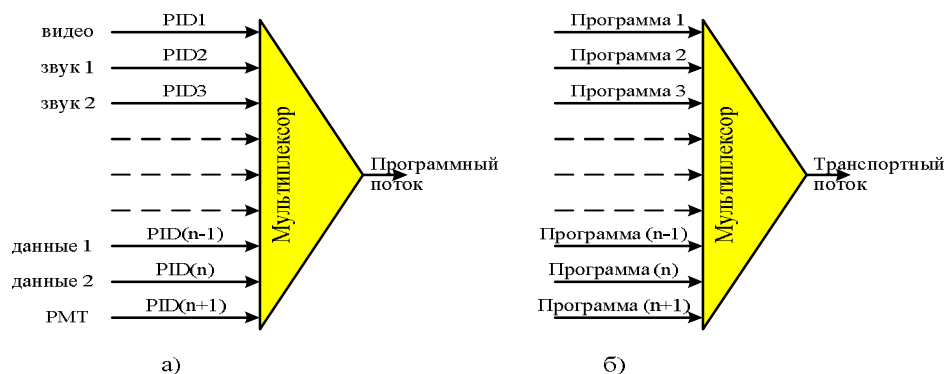


Рисунок 3.7 — Объединение *PES* в программный поток (а) и программных потоков в транспортный поток (б)

Гибкость мультиплексирования ТП представляет одну интересную возможность построения многопрограммной системы вещания, связанную с вводом в ТП, формируемый мультиплексором из программ, кодированных локальными кодерами, «постороннего» транспортного потока, поступившего по линии связи и имеющего другую временную базу (режим *ремultipлексирования*). Современные мультиплексоры легко справляются с такой задачей, позволяя при этом выбрать из компонентов пришедшего ТП желаемые, при необходимости изменить *PID* этих компонент, режим скремблирования и другие параметры.

Еще одна возможность гибкого изменения параметров сжатого сигнала — *статистическое мультиплексирование*, заключающееся в динамическом перераспределении ресурса битов между кодерами в многопрограммном цифровом потоке с учетом конкретных особенностей изображения в каждой программе таким образом, чтобы качество всех программ оставалось, возможно, более высоким. Более сложный контроллер может использовать технологию «заглядывания вперед» для оценки сложности представленных кадров и эта информация может служить для подстройки управления скоростью.

Система работает следующим образом. Управляющий контроллер статистического мультиплексора получает от всех кодеров информацию о сложности обрабатываемого в данный момент изображения, оценивает ее и выделяет каждому кодеру ресурс битов, пропорциональный потребности. При суммировании скоростей нескольких каналов среднее значение скорости растет пропорционально сумме средних значений составляющих, а среднеквадратичное отклонение изменяется пропорционально квадратному корню из числа каналов. Если в пакете программ присутствуют, например, фильмовые, детские, спортивные, развлекательные программы, в разные моменты времени в зависимости от наличия, детальности изображения, цветовой насыщенности они требуют разной скорости выходного потока для сохранения примерно одинакового субъективного качества изображения и, что еще важнее, пики скорости наступают в разные моменты времени, так что суммарная скорость может быть заметно снижена. Исследования показали, что без ухудшения качества можно выиграть примерно 30% пропускной способности канала. Применительно к спутниковому ТВ вещанию это означает, что вместо 6 программ в спутниковом стволе можно передать 8, вместо 8 до 10...11.

3.2. Таблицы программно-зависимой информации

Вся информация, необходимая декодеру для обработки принятого цифрового потока и выделения нужных компонент программы, сосредоточена в управляющей информации (ее еще называют *метаданными*), передаваемой в составе транспортного потока. Она организована в виде нескольких таблиц, содержащих сведения о составе программ и идентификаторах их компонент и называемых таблицами программно-зависимой информации *PSI (Program Specific Information)*.

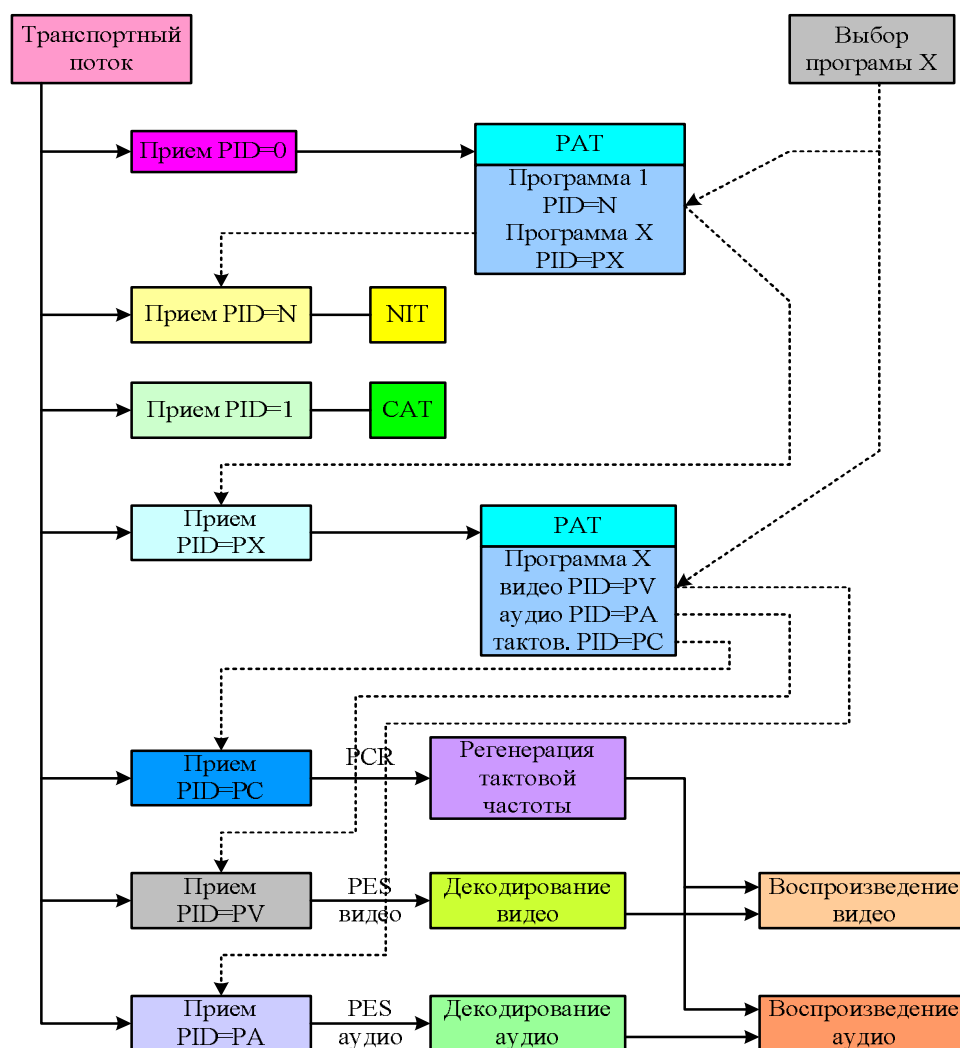


Рисунок 3.8 — Алгоритм действия декодера при прочтении таблиц *PSI*

Алгоритм работы декодера при прочтении таблиц *PSI* поясняется рисунком 3.8. Первая таблица, пакеты которой находит кодер в потоке — *Таблица объединения программ (PAT — Program Association Table)*. Таблица *PAT* имеет по умолчанию *PID = 0* и включает информацию о программах, передаваемых в данном потоке, и идентификаторы, относящиеся к этим программам. Каждый такой *PID*, в свою очередь, определяет для выбранной программы *Таблицу состава программы (PMT — Program Map Table)*, в которой перечислены все компоненты, входящие в данную программу, с их идентификаторами. Декодеру остается отобрать из общего потока пакеты с нужными идентификаторами и декодировать их, восстановив изображение и звуковое сопровождение. Если программа платная, декодеру придется

анализировать еще и содержание *Таблицы условного доступа (CAT — Conditional Access Table)* с $PID = 1$, в которой указаны идентификаторы пакетов с сообщениями системы условного доступа. Иногда в набор таблиц *PSI* включают еще необязательную *Таблицу сетевой информации (NIT — Network Information Table)*, которая определяет все транспортные потоки, относящиеся к данной сети. Содержимое таблиц вводится в секции — области поля адаптации определенной длины, снабжаемые указателями. Размер секции для передачи служебной информации не должен превышать 1024 байта. Возможна передача нескольких коротких секций в поле адаптации одного пакета транспортного потока или одной длинной секции в нескольких пакетах.

На рисунке 3.9 показана иерархия *PSI* таблиц и взаимосвязи между ними. Рассмотрим структуру таблиц подробнее.

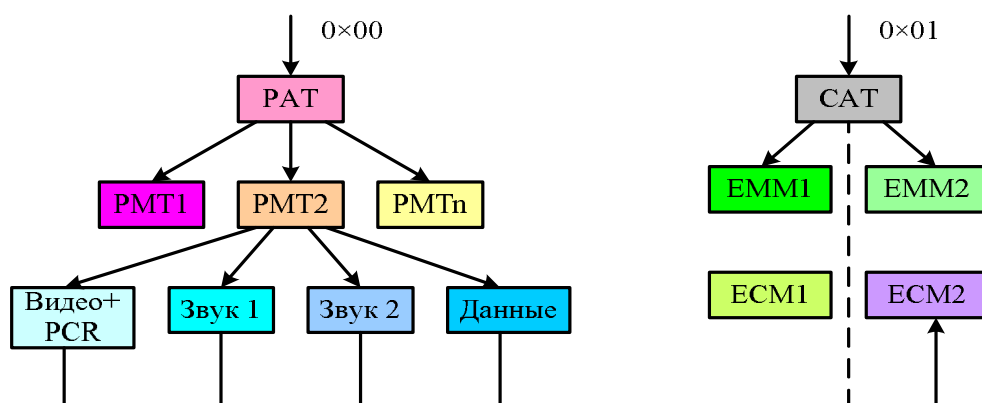


Рисунок 3.9 — Структура таблиц программно-зависимой информации *PSI*

Общий формат таблицы *PAT* показан на рисунке 3.10,а. Таблица содержит заголовок длиной 8 байт и поле данных. Структура заголовка более детально развернута на рисунке 3.10,б. Первым идет *идентификатор таблицы table_id*. Это однобайтовое число обязательно входит в состав любой таблицы и определяет ее тип. Может возникнуть вопрос: разве PID не определяет таблицу полностью, и зачем нужен еще один идентификатор?

Рисунок 3.10 — Формат таблицы *PAT*:

а — общая структура; б — структура заголовка.

PID является более общим указателем, чем *table_id*. Например, две таблицы TDT и TOT (о них мы узнаем в шестой главе) имеют одинаковый *PID*, но разные *table_id*. Следующий существенный указатель — *длина секции* в байтах. Два старших бита из 12 установлены на «0», так что длина секции не может превышать 1024 байта. *Идентификатор транспортного потока* *transport_stream_id* размером 2 байта указывает условный номер в данной сети транспортного потока, в котором передается анализируемая таблица. Указатель *номер версии* изменяется на единицу каждый раз, когда в таблицу вносятся изменения. Если таблица разбита на несколько секций, однобайтовый указатель *номер секции* сообщает номер передаваемой секции. *Номер последней секции* необходим для подтверждения того, что вся таблица принята декодером. В поле данных таблицы *PAT* содержатся сведения о программах, передаваемых в транспортном потоке, с их номерами *PID*. *Номер программы* занимает 2 байта, затем следуют 3-битовый промежуток и 13-битовое значение *PID*.

Таблица PMT создается отдельно для каждой программы, передаваемой в потоке. Общая структура таблицы показана на рисунке 3.11,а, детальная структура заголовка — на рисунке 3.11,б. Заголовок длиной 12 байтов содержит идентификатор таблицы *table_id*, всегда равный 0×02, номер программы, сведения о версии таблицы, номере секции и номере последней секции, а также *PID* того PES в потоке, который переносит значения PCR. В поле данных описывается одна из передаваемых программ

со всеми ее компонентами, а завершается секция контрольной суммой. Описание программы содержит подробный перечень всех элементарных потоков, составляющих программу, с их основными параметрами. *Stream_id* (*stream_id*) указывает на содержимое данного потока (0×01 — MPEG-1 видео, 0×02 — MPEG-2 видео, 0×03 — MPEG-1 звук, 0×04 — MPEG-2 звук, 0×05 — нестандартные секции, и т.д.), *elementary_PID* сообщает значение PID пакетированного потока, несущего данный элементарный поток, *ES_info_length* указывает длину дескриптора элементарного потока. По этим значениям декодер выделяет нужные элементарные потоки из общей цифровой последовательности. На рисунке 3.11 показана одна из таблиц PMT того же цифрового потока, иллюстрирующая общие принципы.



Рисунок 3.11 — Формат таблицы PMT:

а — общая структура; б — структура заголовка.

Основными компонентами таблицы CAT являются *table_id* и дескриптор системы условного доступа — указатель, сообщающий декодеру условное обозначение используемой в потоке системы условного доступа и номер PID потока управляющих сообщений о правах доступа. Дескриптор условного доступа может присутствовать и в PMT таблице, в этом случае он указывает на PID потока сообщений, необходимого для дешифровки скремблированной программы.

Частота повторения пакетов PAT и PMT таблиц должна быть не менее 10 Гц, периодичность сообщений условного доступа определяется конкретной системой условного доступа.

Рассмотренные три таблицы составляют необходимый минимум, без которого декодер *MPEG-2* не сможет декодировать цифровой поток. Для многопрограммного вещания нужны дополнительные данные, описывающие организацию букетов программ, состав вещательной сети, содержание программ и т.д.

3.3. Транспортировка пакетов *MPEG-2* в составе ячеек *ATM*

Механизм транспортировки данных из одной точки в другую в большинстве систем связи включает идентификацию источника и получателя пункта назначения. Механизм идентификации особенно важен в системах с коммутацией пакетов, в которых пакеты данных пересылаются из одного пункта коммутации в другой до тех пор, пока не достигнут конечного пункта назначения. В комплексных системах связи пакеты данных, относящиеся к одному и тому же сообщению, могут доставляться от источника к пункту назначения различными путями в зависимости от величины трафика системы. Пакеты, предназначенные для применения на уровне звена данных и физическом уровне, обычно содержат следующие шесть элементов: флаг начала пакета и синхροинформацию; информацию для управления кадром с указанием типа кадра; идентификатор пункта передачи в виде адреса источника; идентификатор пункта назначения в виде адреса получателя; полезная нагрузка; символы для обнаружения ошибок или циклической проверки на четность (CRC).

Одной из важных технологий, применяющейся для передачи данных по широкополосным цифровым сетям с интеграцией служб (*B-ISDN*), является *асинхронный способ передачи (Asynchronous Transfer Mode, ATM)*. Первоначально *ATM*-технология была разработана для решения задач мультиплексирования и транспортировки применительно к телефонной связи с целью объединения различных видов трафика в одной и той же системе. В структуре *ATM* информация передается короткими пакетами фиксированной длины, которые называются *ячейками*. Длина ячейки составляет 53 байта (октета по терминологии МСЭ-Т). Первые 5 байтов (заголовков) содержат

информацию о мультиплексировании, а остальные 48 байтов (полезная загрузка) содержат данные пользователя.

Заголовок ячейки *ATM* обеспечивает выполнение сетевых функций и состоит из шести полей:

— *GFC — Generic Flow Control* — 4-битовое поле общего контроля потока, служащее для контроля прохождения трафика через сетевой интерфейс пользователя.

— *VPI — Virtual Path Identifier* — 8-битовое поле идентификатора виртуального пути (указателя фактического маршрута передачи).

— *VCI — Virtual Channel Identifier* — 16-битовое поле идентификатора виртуального (фактически используемого) канала. Функции *VCI* и *VPI* совместно обеспечивают информацию о маршруте передачи информации с преобразованием последовательностей из одного вида в другой. Информация, относящаяся к источнику и конечному пункту назначения, передается не в *ATM* заголовке, а представляется последовательностью этапов маршрута при выделении *ATM* канала.

— *PTI — Payload Type Indicator* — 3-битовое поле индикации типа полезной нагрузки.

— *CLP — Cell-Loss Priority* — 1-битовое поле приоритета потери ячейки, т.е. флаг, определяющий прерывание передачи последовательности при перегрузке сети.

— *HEC — Header Error Control* — 8-битовое поле контроля над ошибками в заголовке, служащее для исправления ошибок в *ATM* заголовке.

Метод *ATM* разработан для применения в сети с коммутацией пакетов, когда различные ячейки одного и того же сообщения могут следовать по разнообразным путям между источником и приемником. Чтобы гарантировать правильную сквозную доставку информации и восстановление полезной нагрузки в порядке ее передачи, были стандартизированы уровни эталонной модели протоколов *B-ISDN*. Применительно к *ATM* важное

значение имеют три первых уровня: физический, уровень *ATM* и уровень адаптации *ATM* (*ATM Adaptation Layer, AAL*).

Уровень адаптации *ATM* обеспечивает выполнение услуг в интересах верхнего уровня, при этом он разбивается на два подуровня: конвергенции (*Convergence Sublayer, CS*); сегментации и сборки (*Segmentation and Reassembly Sublayer, SAR*). Стандартизированы 5 уровней адаптации *ATM* типа 1 — 5, которые обеспечивают маршрутизацию сообщения, ретрансляцию кадров, эмуляцию каналов, идентификацию узлов и информацию о времени, необходимые для правильного восстановления сообщений на приемном конце. Однако, в условиях единственного пути сообщения, характерного для цифрового наземного вещания, байты, относящиеся к идентификации пути сообщения, представляют собой избыточные данные заголовка.

Цифровое наземное телевизионное вещание является системой исключительно для передачи из одной точки на множество точек, и система транспортировки *MPEG-2* предназначена для применения в функции эффективного контейнера сигналов изображения, звука и данных, которые будут подаваться от источника к потребителю. Поскольку участок тракта между источником и потребителем является по существу фиксированным (если не учитывать проблемы многолучевой интерференции) и пропускная способность наземного канала при передаче данных строго ограничена, нет никакой необходимости передавать информацию об источнике, пункте назначения или узловом пункте сети в каждом пакете.

При разработке стандартов *MPEG-2* учитывался вариант сопряжения транспортных потоков цифрового телевидения с сетями *ATM*. Известны два метода последовательного ввода пакетов транспортного потока *MPEG-2* в ячейки системы *ATM*, рассмотренные в Рекомендации *ITU-T J.82*. В первом из них используются уровень адаптации *ATM AAL* типа 1 и введение отдельного транспортного пакета *MPEG-2* в четыре ячейки системы *ATM*.

Подуровень сегментации и сборки ВАН принимает от подуровня конвергенции СВ блоки данных полезной нагрузки длиной 47 октетов и добавляет к ним 1 октет заголовка. Сформированный таким образом блок данных в 48 октетов называется *протокольным блоком данных (Protocol Data Unit, PDU)*, или *SAR-PDU*. Транспортный пакет стандарта *MPEG-2* состоит из 188 байтов, включающих 4 байта заголовка и 184 байта полезной загрузки.

Этот 188-байтовый пакет может быть размещен в пределах четырех ячеек *ATM* как четыре полезные нагрузки по 47 байтов блоков *SAR-PDU* ($4 \times 47 = 188$), оставляя пространство для одного байта адаптации *ATM AAL* типа 1 на одну ячейку *ATM*. Пример такого размещения показан на рисунке 3.13.

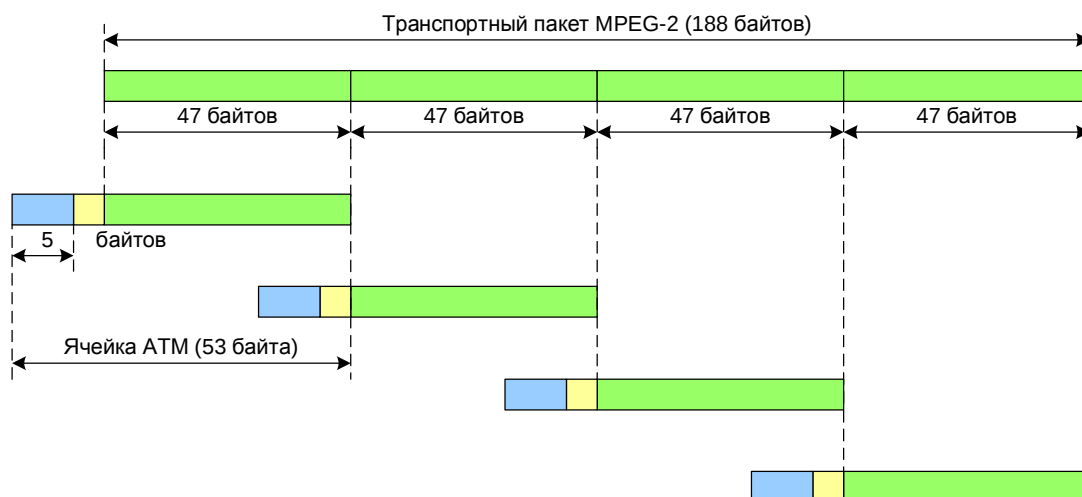


Рисунок 3.13 — Мультиплексирование транспортного пакета *MPEG-2* в ячейки *ATM*

Транспортная подсистема *MPEG-2* с ее 4-байтовым заголовком транспортного пакета из 188 байтов менее перегружена, чем система *ATM*.

Излишнее наполнение заголовка должно приниматься во внимание среди прочих ограничений систем цифрового ТВ вещания.

Во втором методе ввода используется уровень адаптации *ATM AAL* типа 5, для которого блок данных полезной нагрузки равен 48 октетам. Уровень адаптации типа 5 обладает меньшей избыточностью, так как использует служебную информацию, переносимую в заголовке ячейки *ATM*. На этом уровне осуществляется так называемое $1/N$ - перераспределение, при

котором N пакетов транспортного потока загружаются в ATM -ячейки на уровне адаптации для звена служебных данных. Значение N определяется сигнализацией о заполнении ячеек. В некоторых случаях сигнализация не передается, и объем загрузки ячеек не превышает 376 октетов данных.

Если в двух последовательно поступающих пакетах транспортного потока не содержатся эталонные метки времени программы (программный тактовый сигнал — *Program Clock Reference, PCR*), то они могут быть непосредственно введены в восемь ATM -ячеек ($N=2$).

Структурная схема ввода двух транспортных пакетов в 8 ячеек ATM показана на рисунке 3.14. Общий объем данных полезной нагрузки в составе 8 ячеек ATM уровня адаптации AAL типа 5 составляет 384 октета. Из них 376 октетов служат для ввода данных двух транспортных пакетов, а 8 октетов хвостовика протокольного блока данных общей части подуровня конвергенции (*Common Part Convergence Sublayer, CPSC-PDU*) могут использоваться, например, для контроля ошибок.

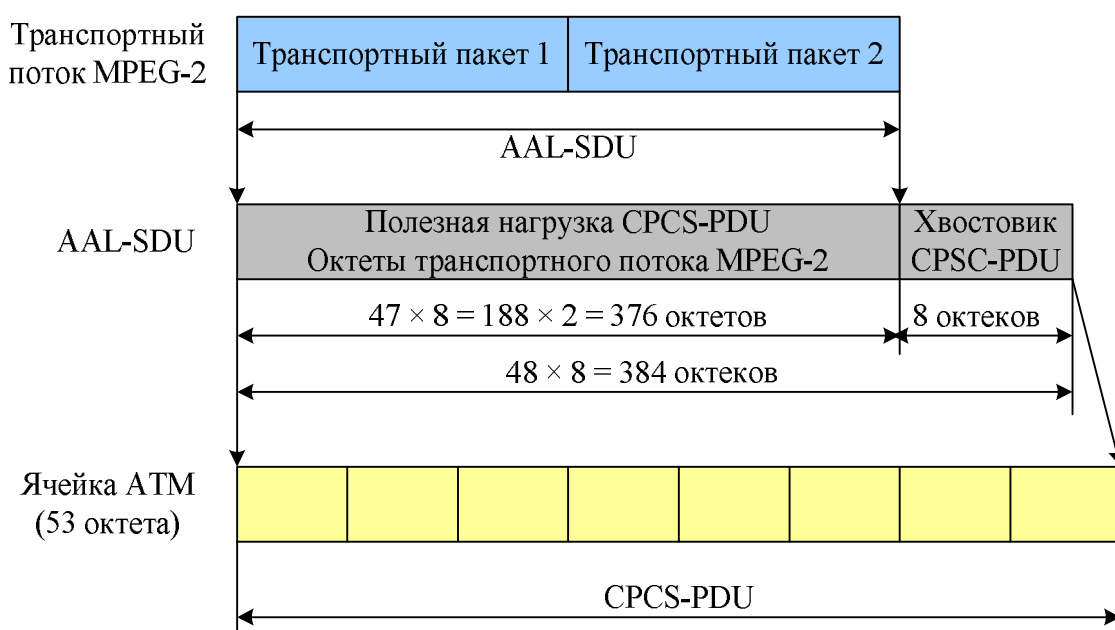


Рисунок 3.14 — Ввод пакетов ($N = 2$) в ячейки ATM (AAL типа 5)

3.4. Принципы обеспечения синхронизации цифровых потоков в стандарте *MPEG-2*

Общий отсчет времени задается ссылками на системные часы (*SCR*), вводимыми в заголовок системного потока, по ним декодер подстраивает свои внутренние часы и «привязывает» потоки видео и звука к абсолютным отсчетам времени. Для выравнивания задержки во времени отдельных пакетов в заголовок *PES*-пакета вводятся временные метки воспроизведения (*PTS*) и временные метки декодирования (*DTS*), которые указывают, в какие моменты этого абсолютного времени декодер должен обработать и вывести на экран соответствующие изображения или звуковые последовательности. Эти же средства используются и в программном потоке *MPEG-2*, где все компоненты потока также имеют общую временную базу.

Транспортный поток содержит программные компоненты с разной предысторией, в том числе и с несколько различающимися тактовыми частотами, поэтому невозможно или весьма трудоемко привести все сигналы к единой временной базе. Для управления такими потоками вводится еще один механизм синхронизации, называемый *ссылкой на программные часы* (*PCR — Program Clock Reference*). *PCR*, как и другие временные метки, также представляет собой 33-битовое число, отсчитываемое в периодах частоты 90 кГц, получаемой делением на 300 частоты тактового генератора 27 МГц. Оно показывает ожидаемое время завершения считывания в декодере поля *PCR* из транспортного потока, после чего декодер может приступить к сравнению пришедшего и местного отсчетов и выработке корректирующего сигнала. Отличие от *SCR* в том, что *PCR* вводится в поток на программном, а не системном уровне, в одном потоке могут передаваться несколько различных *PCR*, по числу программ. Декодер при переключении на каждую новую программу заново синхронизирует внутренний генератор частоты 27 МГц. Стандарт предписывает повторение метки *PCR* не реже чем 1 раз в 0,7 с. В промежутках декодер вычисляет значения меток путем интерполяции.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Цифровое телевидение / М. И. Кривошеев, Л. С. Виленчик и др.; ред. М. И. Кривошеев. — М.: Связь, 1980. — 263 с.
2. Птачек М. Цифровое телевидение. Теория и техника / М. Птачек; Пер. с чеш. В. В. Исаченко; под ред. Л.С. Виленчика. — М.: Радио и связь, 1990. — 528 с.
3. Мамчев Г.В. Основы цифрового телевидения. Учебное пособие./Г.В. Мамчев. — Новосибирск: СибГУТИ, 2003. — 248 с.
4. Телевидение: Учебник для ВУЗов / В.Е. Джакония, А.А. Гоголь, Я.В. Друзин и др.; Под ред. В.Е. Джаконии. — М.: Горячая линия — Телеком, 2007.— 616 с.
5. Смирнов А.В. Основы цифрового телевидения: Учебное пособие. — М.: Горячая линия — Телеком, 2001. — 224 с.
6. Зубарев Ю.Б. Цифровое телевизионное вещание. Основы, методы, системы./ Ю.Б. Зубарев, М.И. Кривошеев, И.Н. Красносельский. — М.: Научно-исследовательский институт радио (НИИР), 2001. — 568 с.
7. Локшин Б.А. Цифровое вещание: От студии к телезрителю./ Б.А. Локшин. — М.: Компания Сайрус системс, 2001. — 446 с.
8. Смирнов А.В. Цифровое телевидение. От теории к практике./ А.В. Смирнов, А.Е. Пескин. — М.: Горячая Линия — Телеком, 2005, — 349 с.