

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное

автономное образовательное учреждение  
высшего профессионального образования

“Южный федеральный университет”



Кафедра радиоприемных устройств и телевидения

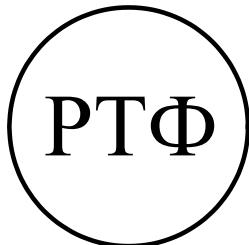
РУКОВОДСТВО  
К ЛАБОРАТОРНОЙ РАБОТЕ

ИССЛЕДОВАНИЕ ПРОЦЕССА  
СЖАТИЯ ВИДЕОСИГНАЛА АППАРАТНЫМИ  
СРЕДСТВАМИ ВИДЕОПРОЦЕССОРА PHILIPS SAA6752HS

по курсам

ОСНОВЫ ЗАПИСИ АУДИО - И ВИДЕОСИГНАЛОВ,  
ЗАПИСЬ АУДИО - И ВИДЕОСИГНАЛОВ,  
ОСНОВЫ ЦИФРОВОГО ТЕЛЕВИДЕНИЯ,  
ЦИФРОВОЕ ТЕЛЕВИДЕНИЕ

Для студентов РТФ  
всех форм обучения



Таганрог 2014

УДК 621.397.13.037.372(076.5) + 621.391.24(076.5)

Составители: Галустов Г.Г., Мелешкин С.Н., Сидько И.В.

Руководство к лабораторной работе “Исследование процесса сжатия видеосигнала аппаратными средствами видеопроцессора PHILIPS SAA6752HS” по курсам: “Основы записи аудио - и видеосигналов”, “Запись аудио - и видеосигналов”, “Основы цифрового телевидения”, “Цифровое телевидение”. – Таганрог: Изд-во ЮФУ, 2014. – 30 с.

В работе использованы инновационные образовательные методы, способствующие развитию индивидуального творческого мышления студентов.

Изучаются особенности аппаратного цифрового сжатия видео- информации с использованием стандарта MPEG-2 в реальном масштабе времени на базе видеопроцессора PHILIPS SAA6752HS.

Предназначено для студентов радиотехнического факультета всех форм обучения, изучающих дисциплины «Видеотехника» и «Цифровое телевидение».

Может быть использовано лицами, самостоятельно изучающими цифровую телевизионную и видеотехнику.

Ил. 17. Библиогр.: 8 назв.

Рецензент Б.К. Лебедев, доктор технических наук, профессор кафедры САПР ЮФУ.

## **1. ЦЕЛЬ РАБОТЫ**

Целью работы являются максимальная интенсификация и активизация самостоятельной работы студента на основе применения методик информационного обеспечения, выполнения комплекса лабораторных исследований в направлении цифрового аппаратного сжатия видеинформации в реальном масштабе времени.

## **2. КРАТКИЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ**

При захвате видео в реальном масштабе времени требуются большие объемы дискового пространства, необходимого для хранения даже небольших фрагментов видеофайлов и высокая скорость работы винчестера. Причем применение современных алгоритмов сжатия не изменяет ситуации кардинально.

Видео "телевизионного" формата 720x576 пикселов 25 кадров в секунду в системе *RGB* требует потока данных примерно в 240 Мбит/с (т. е. 1,8 Гбит/мин).

При этом традиционные алгоритмы сжатия изображений, ориентированные на отдельные кадры, не спасают ситуации, поскольку даже при уменьшении потока в 10 раз он составляет достаточно большие величины.

В результате подавляющее большинство сегодняшних алгоритмов сжатия видео являются алгоритмами с потерей данных. При сжатии используется несколько типов избыточности:

**когерентность областей изображения** – малое изменение цвета изображения в соседних пикселях (свойство, которое эксплуатируют все алгоритмы сжатия изображений с потерями);

**избыточность в цветовых плоскостях** – используется большая важность яркости изображения для восприятия;

**подобие между кадрами** – использование того факта, что на скорости 25 кадров в секунду, как правило, соседние кадры изменяются незначительно.

Первые два пункта знакомы вам по алгоритмам сжатия графики. Использование подобия между кадрами в самом простом и наиболее часто используемом случае означает кодирование не самого нового кадра, а его разности с предыдущим кадром.

В настоящее время для сжатия видеопотока наиболее часто применяется стандарт **MPEG-2**.

## Базовые технологии сжатия видеоданных

Технология сжатия видео в MPEG-2 распадается на две части: уменьшение избыточности видеинформации во временном измерении, основанное на том, что соседние кадры, как правило, отличаются незначительно, и сжатие отдельных изображений. Для того чтобы удовлетворить противоречивым требованиям и увеличить гибкость алгоритма, рассматривается 4 типа кадров:

*I*-кадры – кадры, сжатые независимо от других кадров (*I-Intra pictures*);

*P*-кадры – кадры, сжатые с использованием ссылки на одно изображение (*P-Predicted*);

*B*-кадры – кадры, сжатые с использованием ссылки на два изображения (*B-Bidirection*);

*DC*-кадры – кадры, независимо сжатые с большой потерей качества (используются только при быстром поиске).

*I*-кадры обеспечивают возможность произвольного доступа к любому кадру, являясь своеобразными входными точками в поток данных для декодера. *P*-кадры используют при архивации ссылку на один *I*- или *P*-кадр, повышая тем самым степень сжатия фильма в целом. *B*-кадры, используя ссылки на два кадра, находящиеся впереди и позади, обеспечивают наивысшую степень сжатия. Сами в качестве ссылки использоваться не могут. Последовательность кадров в фильме может быть, например, такой: *IBBPBVPBBPBBIBBPBB...* Или, если мы не экономим на степени сжатия, такой, как на рис. 1.

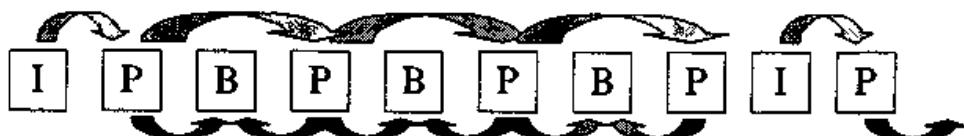


Рис. 1. I-кадры, P-кадры, B-кадры

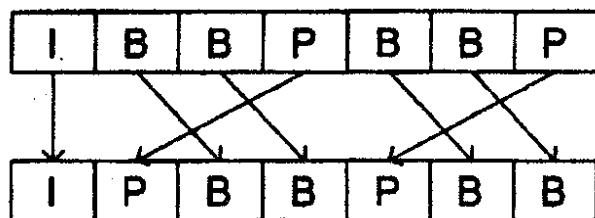


Рис. 2. Последовательность передачи видеокадров

Частота *I*-кадров выбирается в зависимости от требований на время произвольного доступа и надежности потока при передаче через канал с ошибками. Соотношение *P*- и *B*-кадров подбирается исходя из требований к величине компрессии и ограничений декодера.

Как правило, декодирование *B*-кадров требует больше вычислительных мощностей, однако позволяет повысить степень сжатия. Именно варьирование частоты кадров разных типов обеспечивает алгоритму необходимую гибкость и возможность расширения. Понятно, что для того чтобы распаковать *B*-кадр, мы должны уже распаковать те кадры, на которые он ссылается.

Поэтому для последовательности *IBVRVVRVVRVIBVRVB* кадры в фильме будут записаны так:  $0^{**}312645\dots$ , где цифры – номера кадров, а звездочкам соответствуют либо *B*-кадры с номерами -1 и -2, если мы находимся в середине потока, либо пустые кадры, если мы в начале фильма (рис. 2). Одним из основных понятий при сжатии нескольких изображений является понятие *макроблока*. При сжатии кадр из цветового пространства *RGB* переводится в цветовое пространство *YUV*. Каждая из плоскостей сжимаемого изображения (*Y*, *U*, *V*) разделяется на блоки 8x8, с которыми работает ДКП. Причем плоскости *U* и *V*, соответствующие компоненте цветности, берутся с разрешением в 2 раза меньшим (по вертикали и горизонтали), чем исходное изображение.

Таким образом, мы получаем сжатие в 2 раза, пользуясь тем, что глаз человека хуже различает цвет отдельной точки изображения, чем ее яркость. Подробнее об этих преобразованиях смотрите в описании алгоритма *JPEG*. Блоки 8x8 группируются в макроблоки. Макроблок – это группа из четырех соседних блоков в плоскости яркостной компоненты *Y* (матрица пикселов 16x16 элементов) и два соответствующих им по расположению блока из плоскостей цветности *U* и *V*. Таким образом, кадр разбивается на независимые единицы, несущие полную информацию о части изображения. При этом размер изображения должен быть кратен 16. Отдельные макроблоки сжимаются независимо. Алгоритм сжатия отдельных кадров в *MPEG* похож на соответствующий алгоритм для статических изображений – *JPEG*. Если говорить коротко, то сам алгоритм сжатия представляет собой конвейер преобразований. Это дискретное косинусное преобразование исходной матрицы 8x8, квантование матрицы и вытягивание ее в вектор *vll, vl2, v21, v31, v22, ..., v88* (зигзаг-сканирование), сжатие вектора групповым кодированием и сжатие по алгоритму Хаффмана.

**Квантование** является тем звеном обработки сигнала, на котором JPEG вносит потери. Оно определяет точность хранения результатов DCT и коэффициент сжатия. JPEG использует простое линейное квантование.

При этом каждое из значений DCT делится на коэффициент квантования, индивидуальный для каждой пространственной частоты, который берется из заранее определенной таблицы коэффициентов квантования размером 8x8 (естественно, она должна быть одинакова для кодера и декодера).

Эта таблица может быть взята по умолчанию или формируется кодером для конкретной статистики изображения и передается декодеру вместе со сжатыми данными (рис.3).

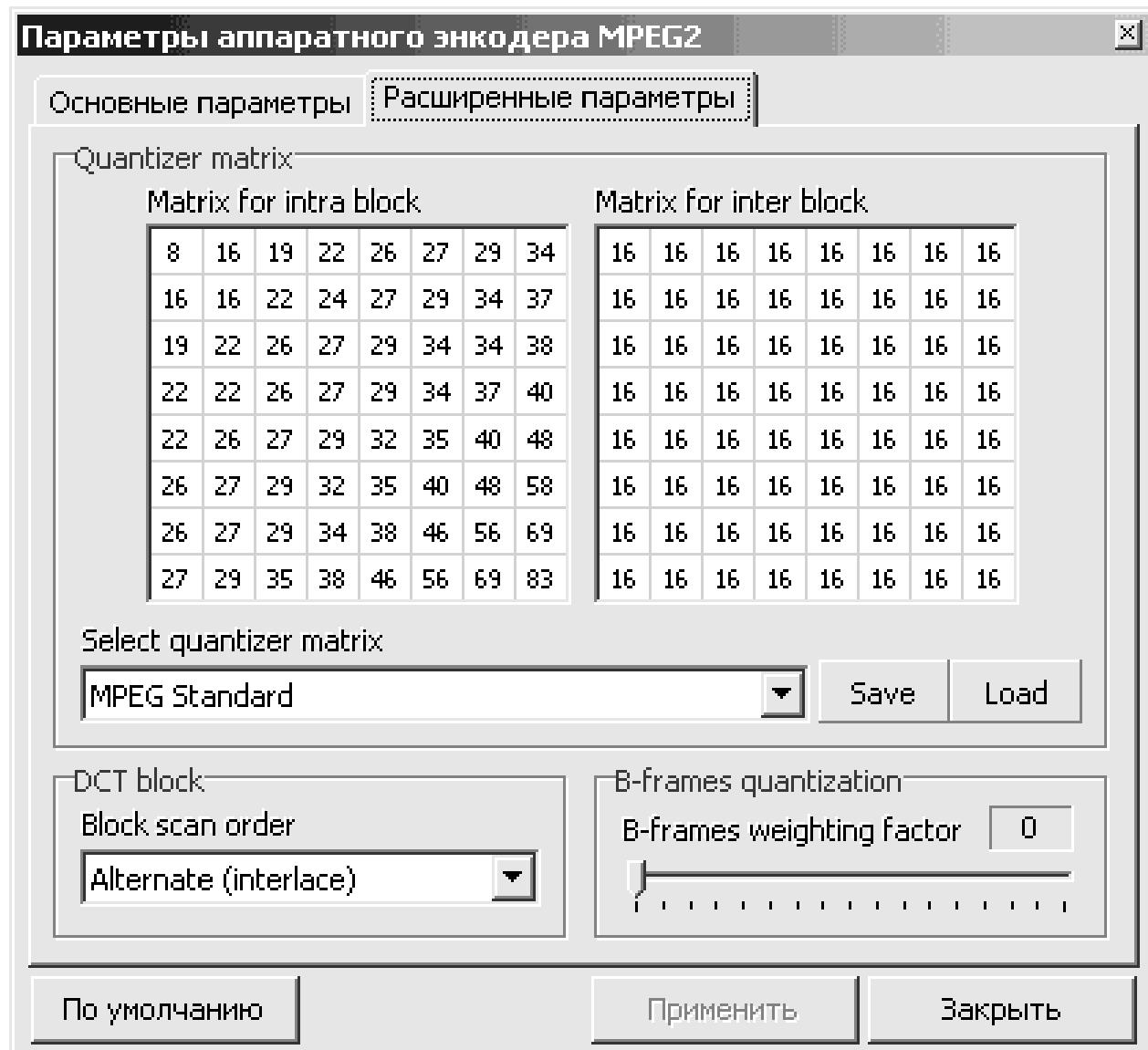


Рис. 3. Стандартная матрица квантования MPEG-2

После квантования коэффициентов DCT происходит сокращение избыточности их последовательности с помощью модифицированного кода Хаффмана, дающего возможность кодирования со средней длиной кодового слова меньше 1 бита на каждый коэффициент и называемого также кодом переменной длины (VLC – variable length code) или арифметического кодирования. Для этого формируется линейная последовательность квантованных коэффициентов DCT - блока путем их считывания в зигзагообразном порядке (рис. 4).

Такая последовательность кодируется методом Хаффмана. Кодированию подвергаются пары чисел {RUN, LENGTH}. При этом общий поток разбивается на участки, начинающиеся с нулевых квантованных значений DCT (их количество образует число RUN) и заканчивающиеся первым встретившимся ненулевым значением (это значение присваивается числу LENGTH). К этим последовательностям применяется кодовая таблица.

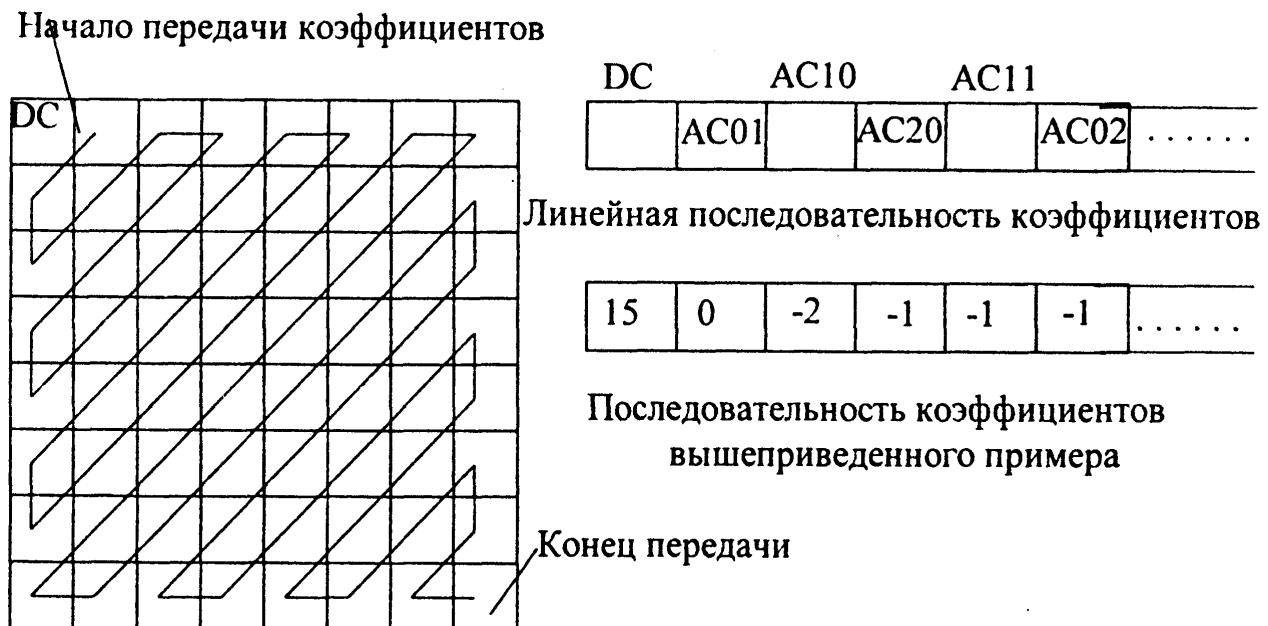


Рис. 4. Формирование линейной последовательности коэффициентов DCT

В рассмотренном примере блок будет закодирован потоком 01000 0001101 111 111 111 0111 111 10, занимающим 30 бит или 3,75 байта. Таким образом, на уровне блока удалось достигнуть коэффициента сжатия около 17 (исходный блок содержал 64 байта).

Вместо постоянной составляющей кодируется ее разность с постоянной составляющей предыдущего блока. Схематически процесс сокращения избыточности представлен на рис. 5.

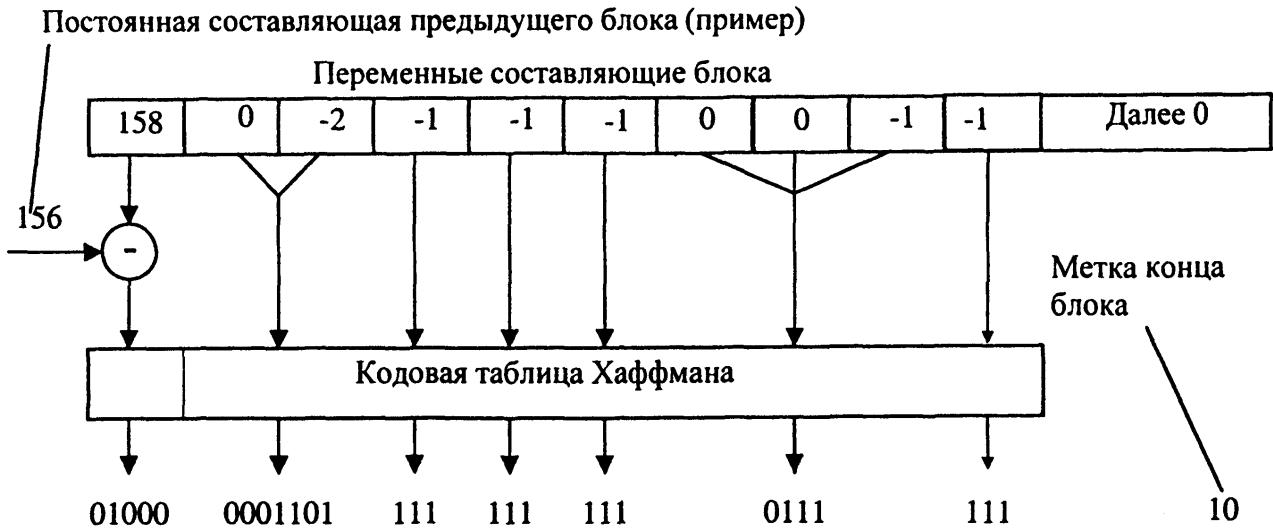


Рис. 5. Сокращение избыточности коэффициентов DCT

### Общая схема алгоритма

В целом весь конвейер преобразований можно представить так:

- Подготовка макроблоков. Для каждого макроблока определяется, каким образом он будет сжат. В  $I$ -кадрах все макроблоки сжимаются независимо. В  $P$ -кадрах блок либо сжимается независимо, либо представляет собой разность с одним из макроблоков в предыдущем опорном кадре, на который ссылается  $P$ -кадр.

- Перевод макроблока в цветовое пространство  $YUV$ . Получение нужного количества матриц  $8 \times 8$ .

- Для  $P$ - и  $B$ -блоков производится вычисление разности с соответствующим макроблоком в опорном кадре.

- ДКП.
- Квантование.
- Зигзаг-сканирование.
- Групповое кодирование.
- Кодирование Хаффмана.

При декодировании весь конвейер повторяется для обратных преобразований, начиная с конца.

### Использование векторов смещений блоков

Простейший способ учитывать подобие соседних кадров – это вычесть каждый блок сжимаемого кадра из соответствующего блока предыдущего. Однако более гибким является алгоритм поиска векторов, на которые сдвинулись блоки текущего кадра по отношению к предыдущему (рис. 6).

Для каждого блока в изображении мы находим блок, близкий по некоторой метрике (например, по сумме квадратов разности пикселов), в предыдущем кадре в некоторой окрестности текущего положения блока. Если минимальное расстояние по выбранной метрике с блоками в предыдущем кадре больше выбранного порога, блок сжимается независимо.



Рис. 6. Поиск векторов смещения блока

На рис. 7 представлена схема кодера MPEG. В основе структурной схемы кодера легко различить те же ступени, которые используют и в алгоритме JPEG, – дискретное косинусное преобразование (DCT), квантование ( $Q$ ) и сокращение избыточности с помощью кодирования Хаффмана. Новым является дифференциальное кодирование с использованием кадровой памяти.

Обычная дифференциально-импульсная кодовая модуляция дает хорошее предсказание только на неподвижных изображениях. На динамических объектах ошибка предсказания может быть велика. Наиболее распространенное движение в изображениях – поступательное. Кроме того, поступательно движущиеся объекты могут быть отслежены глазом человека (и оказываются для него как бы статическими). При этом сохраняется высокая пространственная разрешающая способность и применение более грубого квантования трансформант будет заметно. Поэтому в MPEG используется предсказание с использованием векторов движения (см. рис. 6), которое ставит в соответствие каждомуциальному макроблоку (размером 16x16 пикселов) кодируемого кадра макроблок того же размера предыдущего (опорного) кадра, смещенный таким образом, чтобы между этими макроблоками достигалась максимальная корреляция. Между ними и производится ДИКМ.

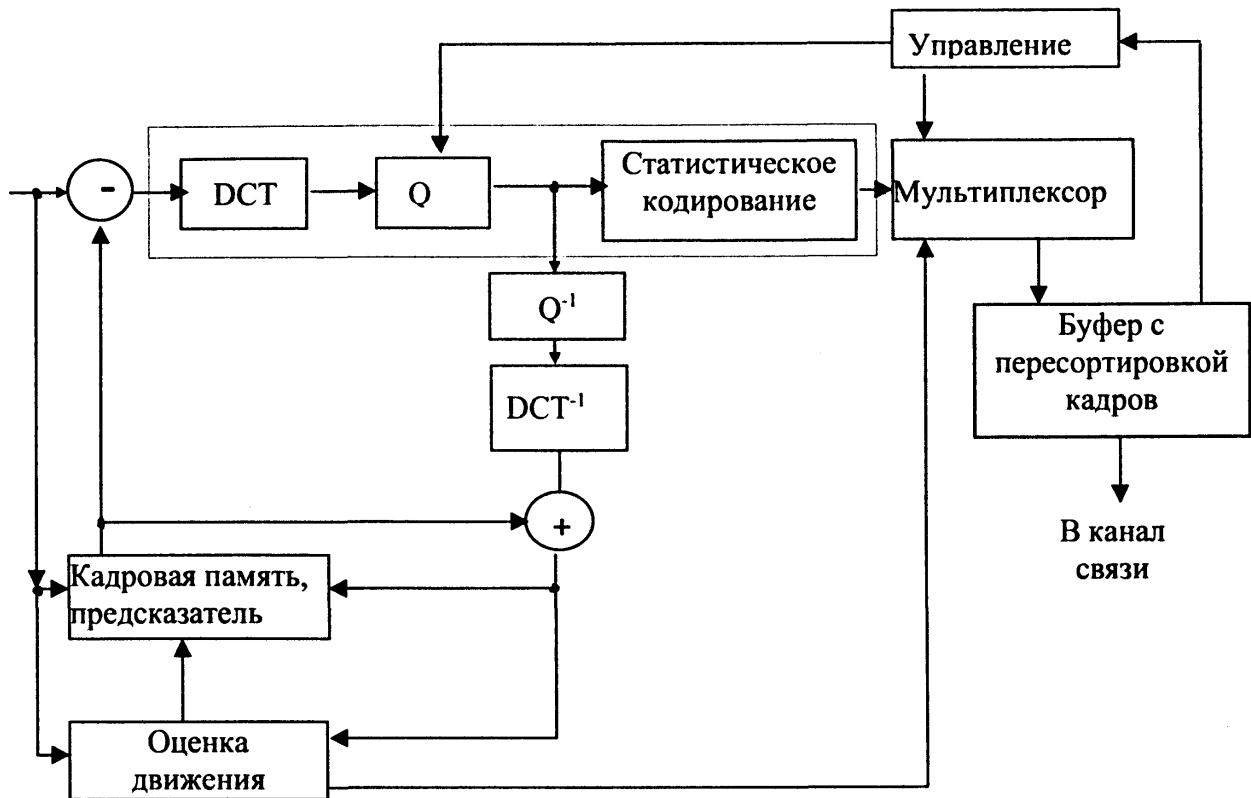


Рис. 7. Структурная схема кодера MPEG

Величина такого смещения, называемая также вектором движения, определяется в блоке оценки движения (motion estimation). Для неподвижного изображения все векторы движения равны нулю.

С целью устранения накопления ошибок квантования для предсказания в кодере «последующего» изображения должно использоваться декодированное «предыдущее» изображение (а не оригинал). Для этого кодер содержит полный декодер (инверсное квантование, инверсное дискретное косинусное преобразование). Таким образом, генерируется декодированное «предыдущее» изображение, на основе которого (и на основе векторов движения) далее производится предсказание.

Предсказанное изображение вычитается из поступающего оригинала (вычитатель в левой части схемы). Полученное таким образом дифференциальное изображение (DFD – displaced frame difference) подвергается кодированию. После декодирования в кодере формируется дифференциальное изображение, к которому должно быть добавлено то, что было подвергнуто операции вычитания перед кодированием, т. е. предсказанное «последующее» изображение (сумматор в нижней части схемы).

После этого будет получено правильно декодированное изображение, пригодное для предсказания дальнейших изображений.

Для обратного предсказания требуется информация о движении, полученные векторы движения также передаются с потоком данных. Для использования корреляции не только с предыдущими, но и с последующими кадрами перед кодированием производится пересортировка кадров.

Следующее существенное дополнение заключается в управлении квантователем для обеспечения постоянства цифрового потока, передаваемого в канал. С этой целью на выходе кодера установлен буфер, принимающий данные с переменным цифровым потоком и выдающий их далее с постоянной скоростью. Его объем при потоке данных 6 Мбит/с выбирается порядка 1...2 Мбит. При возникновении угрозы переполнения буфера он влияет на коэффициенты квантования таким образом, что степень огрубления возрастает и дальнейший поток данных, поступающих в буфер, уменьшается или наоборот. Так поддерживается постоянный поток данных.

Схема декодера представлена на рис. 8.

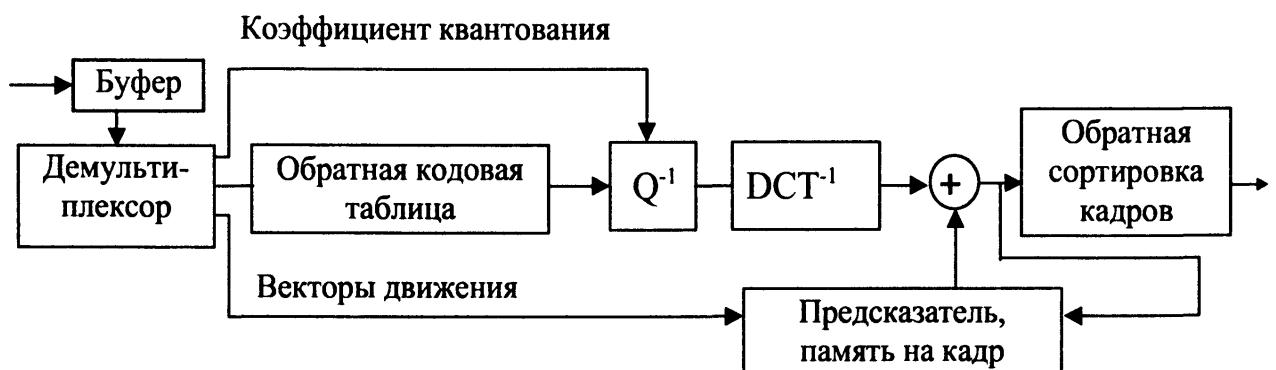


Рис. 8. Структурная схема декодера MPEG

Поступающие постоянным потоком данные записываются во входной буфер ичитываются из него в демультиплексор с необходимой переменной скоростью. Демультиплексор отделяет видеоданные дополнительной информации, в частности от передаваемых в цифровом потоке данных о коэффициенте дискретизации и векторах движения. После операции, обратной сокращению избыточности, происходит обратное квантование, при котором поступающий поток данных взвешивается с текущим значением коэффициента дискретизации (степенью огрубления).

Инверсное дискретное косинусное преобразование преобразует частотные коэффициенты в пространственную область, к которым затем добавляется предсказание из предыдущего кадра, для чего используются передаваемые векторы движения. Наконец, декодированные изображения сортируются в исходную последовательность.

Стандарт MPEG представляет собой несимметричную систему, в которой сложность кодера, обуславливаемая прежде всего наличием устройства выработки *векторов движения (оценки движения)*, компенсируется относительной простотой декодера, который реализуют в виде полупроводниковых кристаллов – массового продукта.

Стандарт MPEG-2 занимается сжатием оцифрованного видео при потоке данных от 3 до 10 Мбит/с. Строки изображения передаются с чередованием и два полукадра составляют кадр. Проблема состоит в чередовании полукадров во входном формате.

Так как кадр формируется из исходных полукадров через строку, то при сжатии происходит размытие движущегося объекта.

Виновно в этом эффекте ДКП, и как-то исправить ситуацию, не уменьшив степени сжатия видео или не потеряв в визуальном качестве, нельзя. Достаточно распространенным является применение "*дe-интерлейсинга*" (от англ. *deinterlacing* – удаление чередования строк).

Эта операция позволяет удалить чередование, смешая четные строки в одном направлении, а нечетные – в другом, пропорционально относительному движению объекта в данной области экрана. В результате мы получаем визуально более качественное изображение, но несколько более длительную предобработку перед сжатием.

Поток, разрешение: 3-15 Мбит/с, универсальный.

Плюсы: поддержка серьезных звуковых стандартов *Dolby Digital 5.1*, *DTS*, высокая универсальность, простота аппаратной реализации.

Минусы: недостаточная на сегодня степень сжатия, недостаточная гибкость.

Объем данных измеряется в битах, но может быть и рациональным числом, а не только целым.

**Битовая скорость.** В общем случае термин "*битовая скорость*" (по-английски *bitrate*) обозначает и bpb (*bit-per-bit* – бит на бит) и bpc (*bit-per-character* – бит на букву). При сжатии звука в *MPEG*, этот термин используется для скорости, с которой сжатый поток считывается декодером. Эта скорость зависит от его источника или носителя (диск, канал связи, оперативная память компьютера и т.п.).

Если битовая скорость аудиофайла *MPEG* определена, скажем, в 128 Кбит/с, то это означает, что кодер конвертирует каждую секунду исходного звукового потока в 128 Кбит сжатых данных, а декодер соответственно, переводит каждые 128 Кбит сжатого файла в одну секунду звучания. Низкая битовая скорость означает малый размер сжатого файла.

**Коэффициент сжатия** – важная величина, которая постоянно используется для определения эффективности метода сжатия. Она равна частному:

$$\text{Коэффициент сжатия} = \frac{\text{размер выходного файла}}{\text{размер входного файла}}.$$

Коэффициент 0,6 означает, что сжатые данные занимают 60 % от исходного.

**Блок** – конечная последовательность цифровой информации.

**Поток** – последовательность с неизвестными границами: данные поступают маленькими блоками, и нужно обрабатывать их сразу, не накапливая. Блок – последовательность с произвольным доступом, а поток – с последовательным. Используют и такие пары терминов: компрессия/декомпрессия, кодирование/декодирование, упаковка/распаковка.

**Сжатие без потерь** (*lossless compression*).

**Сжатие с потерями** (*lossy compression*) – это два разных процесса: выделение сохраняемой части информации с помощью модели, зависящей от цели сжатия и особенностей источника и приемника информации, и, собственно, сжатие, без потерь. Приемлемость сжатия изображения и звука со значительными потерями обусловлена особенностями восприятия такой информации органами чувств человека. Если же предполагается компьютерная обработка изображения или звука, то требования к потерям гораздо более жесткие.

**Эффективность сжатия** учитывает не только **степень сжатия** (отношение длины несжатых данных к длине соответствующих им сжатых данных), но и скорости сжатия и разжатия. Часто пользуются обратной к степени сжатия величиной – **коэффициентом сжатия**, определяемым как отношение длины сжатых данных к длине соответствующих им несжатых. Еще две важные характеристики алгоритма сжатия – объемы памяти, необходимые для сжатия и для разжатия (для хранения данных, создаваемых и/или используемых алгоритмом). Для сжатия видеопотока используется видеокодек.

**Видеокодеком** называется набор инструкций, описывающий правила и алгоритмы кодирования и декодирования цифрового видеосигнала. Видеокодек может быть реализован программно или аппаратно.

Для захвата видеопотока в реальном масштабе времени с помощью программного кодека необходимы мощные вычислительные ресурсы компьютера, и любой сбой в программе приведет к необратимому искажению сигнала и рассинхронизации звука с изображением. Достоинством программных кодеков является их гибкость и возможность обновления.

Структура программного сжатия представлена на рис. 9. Аналогово-цифровой преобразователь (АЦП, его еще называют декодером, так как он выполняет функции декодирования видеосигналов различных стандартов) чаще всего совмещен с PCI-контроллером для передачи несжатого видео напрямую в память компьютера через DMA-канал (канал прямого доступа к памяти). В этом случае PCI-декодер имеет небольшой буфер, никогда не вмещающий целый кадр, что зачастую создает трудности в передаче большого количества потоков через шину и часто приводит к своеобразным помехам на изображении (полосы).

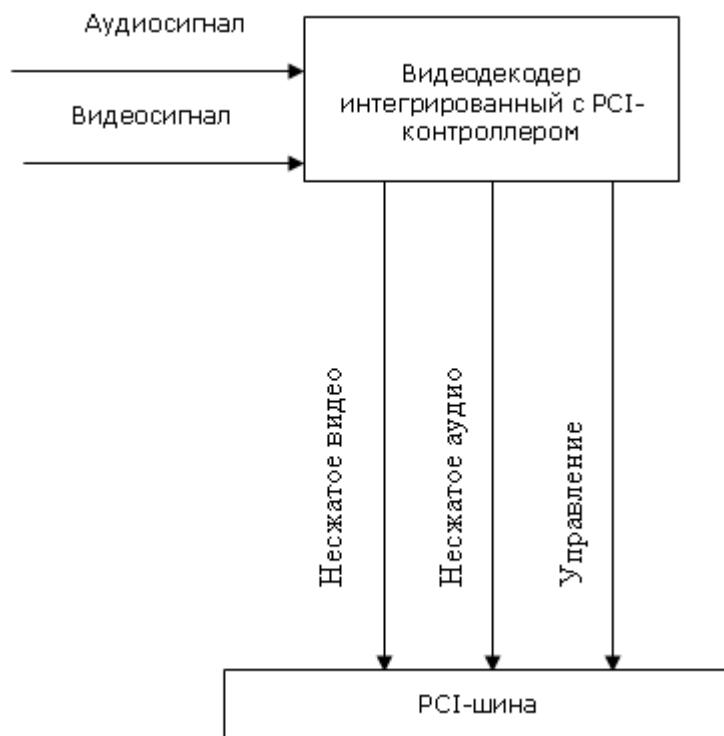


Рис. 9. Структурная схема плат с программным сжатием

Структура аппаратного сжатия представлена на рис. 10. В аппаратных платах такая схема используется чрезвычайно редко. Видеодекодер представляет собой отдельную микросхему, так как он передает оцифрованное видео не напрямую центральному процессору, а специализированной микросхеме сжатия. Уже эти специальные компрессоры выполняют несколько задач помимо компрессии, в том числе и функции PCI-контроллеров. Они масштабируют видео для вывода его напрямую на экран. Суммарный поток через шину никогда не превысит разрешения экрана.

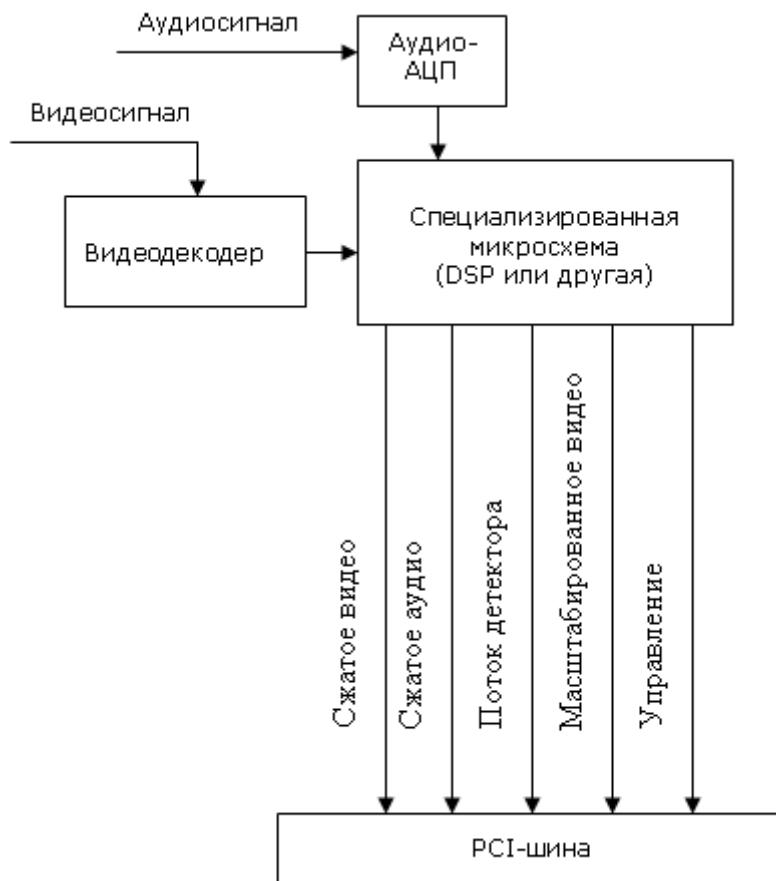


Рис. 10. Структурная схема платы с аппаратным сжатием

Сжатое же видео в желаемом размере и формате поступает через такие же DMA-каналы напрямую на запись и/или в память для передачи в сеть. Таким же образом передается и остальная информация – потоки детекторов, аудио, настройки.

В результате центральный процессор компьютера остается свободным для выполнения задач по работе с архивами, обработке данных детекторов движения и интерфейса, для других необходимых работ.

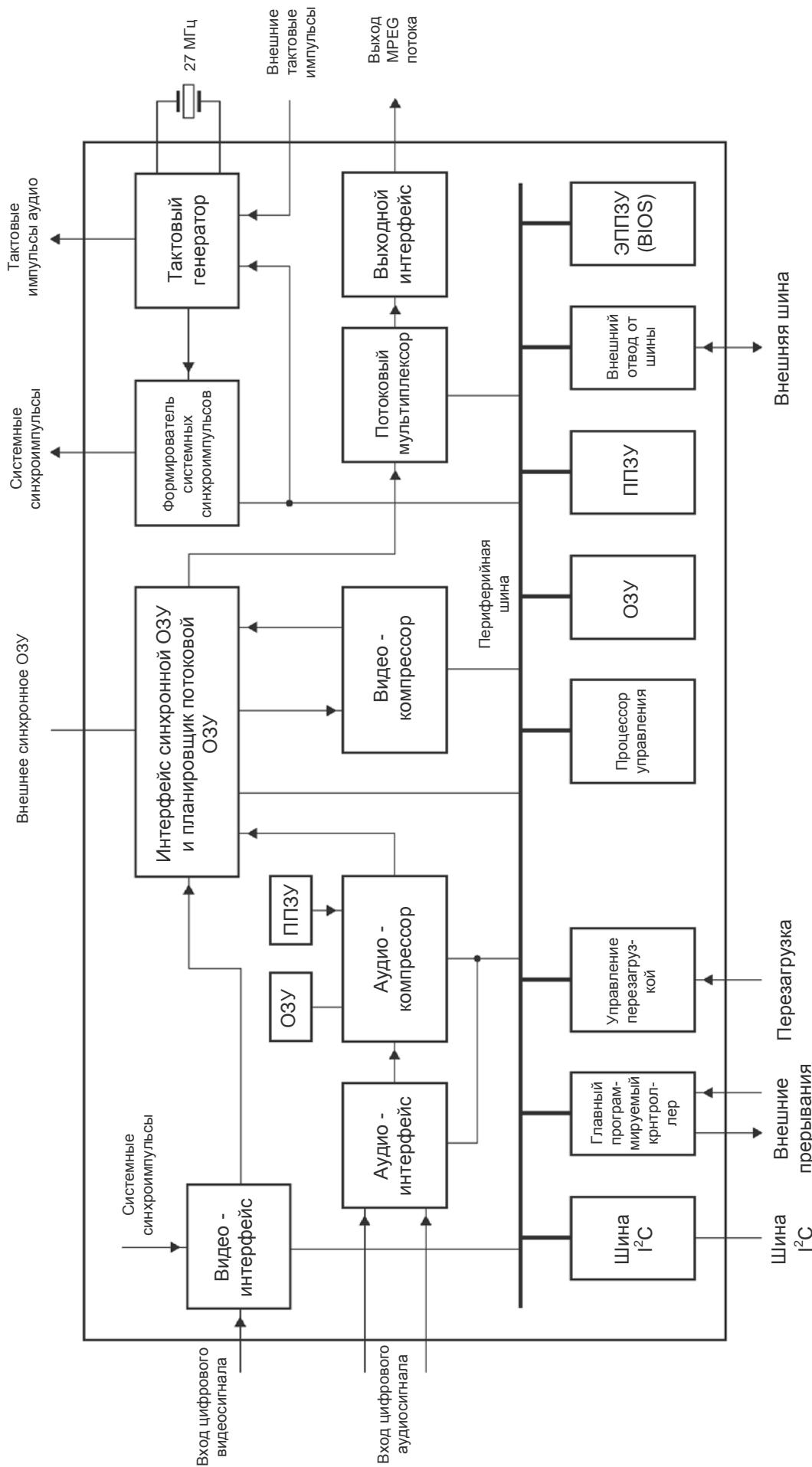


Рис. 11. Внутренняя структура видеопроцессора PHILIPS SAA6752HS

В данной лабораторной работе для аппаратного сжатия видео-потока в реальном масштабе времени используется видеопроцессор PHILIPS SAA6752HS, входящий в состав PC-TV-тюнера Behold TV M6 Extra. Внутренняя структура видеопроцессора PHILIPS SAA6752HS показана на рис. 11.

Запись видео совместно со звуком имеет ряд особенностей. Существуют несколько моментов, которые следует помнить. В захвате видео - и аудиоряда с помощью ТВ - тюнеров, как правило, всегда принимают участие два независимых устройства – устройство оцифровки видео (собственно, сама плата ТВ - тюнера) и устройство оцифровки звука (звуковая плата компьютера).

Эти устройства работают на опорной частоте своих собственных кварцевых генераторов, каждый из которых может иметь некоторую погрешность. Вследствие этого при оцифровке данных происходит рассинхронизация по времени видео - и аудио потоков, что в конечном итоге приводит к отставанию или опережению звука относительно изображения в захваченном файле.

Однако микросхема Philips SAA7135, входящая в состав PC-TV-тюнера Behold TV M6 Extra, помимо видеодекодера, содержит встроенный, независимый от звуковой платы компьютера аудиодекодер и умеет аппаратно связывать аудио - и видео потоки, благодаря чему программа Behold TV может гарантированно оцифровывать звук синхронно с видеорядом. Звуковая плата компьютера в данном режиме не используется. Если тюнер построен на чипсете Philips SAA7135, всегда используйте эту возможность. Аудиодекодер чипсета SAA7135 может оцифровывать звук с частотой 32 кГц и 48кГц (а также частотой 44,1 кГц, но только для источника, подключенного к аналоговому входу чипсета).

- По команде *Параметры MPEG2 энкодера* доступен диалог для полной конфигурации энкодера:

- *Video bitrate* – скорость записи потока видеоданных в файл. Чем выше битрейт, тем выше качество изображения при большем размере конечного файла. Битрейт может быть как постоянным (CBR), так и переменным (VBR). Постоянный битрейт используется в тех случаях, когда необходимо получить чётко заданную, постоянную скорость потока данных, например, для использования в системах передачи данных. В сложных, насыщенных мелкими деталями или динамичных сценах битрейт сохраняется постоянным за счёт ухудшения качества изображения.

Чтобы этого не происходило, при сжатии видеоданных может быть использован переменный битрейт, который будет динамически увеличиваться на сложных сценах, сохраняя постоянным качество изображения.

- *Bitrate mode*
- *Constant (CBR)* – скорость видеопотока является постоянной величиной, которая задаётся настройкой Bitrate.
- *Variable (VBR average)* – значение битрейта может изменяться в диапазоне, установленном Bitrate  $+/- 40\%$  от Bitrate.
- *Variable (VBR peak)* – в процессе записи битрейт будет изменяться от базового Bitrate, строго в пределах пикового значения Bitrate peak.

Если записанный файл в дальнейшем планируется использовать для подготовки DVD и просмотра на стационарном проигрывателе, то не рекомендуется устанавливать пиковый битрейт больше 9200, так как некоторые стационарные проигрыватели не справляются с декодированием высоких битрейтов. Также в стандарте MPEG2 DVD заложен максимальный битрейт, равный 9600. При максимальном MuxRate, равный 10080. Рекомендуемые значения битрейта – Bitrate (kBit/s) = 6500, Bitrate peak (kBit/s) = 9000.

- *Video GOP structure* – структура видеоданных MPEG-потока включает в себя последовательности так называемых групповых кадров (Group of Pictures – GOP) – структур, которые, в свою очередь, состоят из последовательности кадров разного типа (I, P, B). Каждый кадр в структуре GOP кодируется (сжимается) в соответствии с определёнными правилами. Первый кадр в структуре GOP, как правило, I-кадр (Intracoded), который сжимается без учёта содержимого соседних (предыдущего и последующего) кадров. P-кадры (Predicted – кодирование с предсказанием) кодируются с учётом предыдущего I- или такого же P-кадра. B-кадры кодируются с учётом двунаправленного (Bidirectional) предсказания на основе информации в предыдущих и последующих I- или P-кадрах.

- *GOP size* – устанавливает общую длину последовательности GOP.

- *SubGOP* – определяет периодичность расположения P и B кадров в структуре GOP. При  $\text{SubGOP} = 1$  вся последовательность кадров, за исключением первого I-кадра, будет состоять из P-кадров.

Рекомендуемое значение GOP (Recommended GOP) будет меняться в зависимости от установленного битрейта (Bitrate kBit/s).

При кодировании взаимосвязь может учитываться как между кадрами только одной GOP, так и между кадрами из соседних групп. Во втором случае последовательность GOP начинается не с I-, а с B-кадра. Такая последовательность обеспечивает более высокое качество кодирования и называется открытой, так как данные из неё используются для кодирования B-кадра в начале следующей группы. Открытая последовательность считается нередактируемой (Non-editable GOP). Если установить флаг *Force close GOP*, энкодер будет ставить метки "Closed" на открытые последовательности, а B-кадры будут рассчитываться в обратном порядке (Backward predicted closed GOP). Последовательности GOP, которые изначально закрыты, всегда начинаются с I-кадров (Real closed GOP) и флаг *Force close GOP* не оказывает на них никакого влияния.

- *Static denoise* – метод уменьшения шума, основанный на анализе отдельно взятого кадра видеоряда.
- *Off* – статическое шумоподавление не используется.
- *Median* – алгоритм, эффективно устраняющий ярко выраженные пиковые шумы, сохраняя при этом резкость отдельных яркостных и тоновых переходов.
- *Average* – алгоритм, основанный на использовании фильтра нижних частот. Значение Denoise устанавливает частоту среза фильтра. При большом значении среза, снижается четкость.
- *Motion compensated denoise* – метод уменьшения шума, основанный на анализе последовательности кадров видеоряда с использованием векторов движения. Этот метод даёт наилучшее качество без снижения резкости изображения в целом, однако при больших значениях коэффициента Denoise возможно появление артефактов в виде "горячего воздуха".
- *Denoise* – устанавливает коэффициент для определения наличия шума по вектору движения. Рекомендуемое значение 6...8.
- *Multiplexer settings* – параметры ISO/IEC13818-1 (MPEG2 Program Stream) мультиплексора.
- *SCR clock* – System Clock Reference - базовые временные метки в MPEG-потоке.
- *Locked to video* – тактовый генератор SCR привязывается к скорости потока кодируемого видео. Этот режим рекомендуется выбирать при прямой записи видео в файл, так как он обеспечивает максимально эффективную синхронизацию видео и аудиопотоков.

- *Fixed clock* – рекомендуется для просмотра кодированного видео налету через программный декодер. В этом случае, если теряется синхронизация видео (пропадает сигнал), не будет заметных подтормаживаний изображения.

- *NAV PACK time slots* –настройка, позволяющая при записи резервировать в MPEG-потоке некоторое количество временных слотов для использования в дальнейшем навигационными пакетами (NAV PACK). Навигационные пакеты применяются для доступа к данным DVD. Минимальная единица данных DVD потока называется VOBU и содержит от 0,4 до 1,2 с. видео. В зависимости от длины GOP размер VOBU составляет от 1 до 4 GOP. NAV PACK записывается перед началом каждого VOBU и используется при подготовке DVD для указания ссылок.

- *Program end code* – указывает, записывать в файл (в соответствии с ISO/IEC 13818-1) или нет код окончания потока (00 00 01 B9).

- *Quantizer matrix* – В процессе MPEG-кодирования кадр разбивается на блоки размером 8x8 пикселов и каждый блок подвергается дискретному косинусному преобразованию (DCT), в результате чего формируется массив данных 8x8 значений, представляющий собой матрицу коэффициентов DCT. Величина и положение коэффициентов в матрице определяют частотные характеристики исходного изображения. Коэффициенты ближе к правому нижнему углу матрицы характеризуют количество, чёткость и прорисовку мелких деталей изображения. Коэффициенты ближе к левому верхнему углу матрицы отвечают за то, как будут выглядеть крупные элементы сцены. Самый крайний коэффициент в левом верхнем углу определяет среднюю яркость квадрата 8x8 пикселов. Имея матрицу коэффициентов DCT можно восстановить исходное изображение (квадрат 8x8 пикселов).

Если мелкие детали полностью отсутствуют в изображении, то значения коэффициентов в правом нижнем углу матрицы будут близки к нулю. Этими значениями можно пренебречь, полностью их обнулив. Чем больше в матрице будет одинаковых коэффициентов (например, нулей) тем сильнее её можно сжать.

Значения других коэффициенты DCT также можно подкорректировать, установив тем самым компромисс между качеством изображения и степенью его сжатия (размером файла). Процесс нормирования и коррекции коэффициентов DCT называется квантованием.

Во сколько раз будет уменьшен тот или иной коэффициент DCT определяется матрицей квантования. Каждый элемент матрицы квантования является делителем для коэффициента, находящегося в той же позиции в матрице DCT. Размер коэффициента-делителя может иметь значение от 8 до 255. В силу разных требований, предъявляемых к качеству I-, B- и P-кадров для их квантования используются две разные матрицы.

- *Matrix for intra block* – матрица коэффициентов для квантования I фреймов.

- *Matrix for inter block* – матрица коэффициентов для квантования B- и P-фреймов. Коэффициенты этой матрицы имеют, как правило, большие значения (большие коэффициенты деления) и не так частотно зависимы.

- *Select quantizer matrix* – предустановленный набор матриц квантования:

- *MPEG standard* – значение коэффициентов матриц устанавливаются стандартом MPEG. Так как матрицы являются стандартными, значения их коэффициентов не записываются в MPEG-файл и при его проигрывании подставляются декодером по умолчанию автоматически.

- *Low interframe compression, Philips default и High interframe compression* матрицы отличаются от MPEG standard только матрицей для квантования B и P кадров (интер-блоков). В порядке перечисления эти матрицы обеспечивают более высокую компрессию B и P кадров.

- *CG/Animation* – матрица квантования I-кадров отличается от стандартной равномерностью распределения коэффициентов во всей области частот. Используется для сжатия контента в виде анимации или компьютерной графики.

- *User defined* – матрица квантования, определяемая пользователем. Коэффициенты матрицы редактируются непосредственно в её ячейках. По стандарту значение коэффициента в позиции (0, 0) изменению не подлежит и всегда должно оставаться равным 8. Значения матриц квантования сохраняются в пресетах для записи.

- *DCT block* – полученные после квантования коэффициенты матрицы DCT дополнительно сжимаются с помощью алгоритма RLE.

Этим методом наиболее плотно упаковываются числовые ряды с периодически повторяющимися последовательностями одинаковых чисел.

Матрица коэффициентов DCT после квантования как раз представляет собой массив таких чисел, значения которых убывают, заканчиваясь серией нулей, если просматривать матрицу начиная с левого верхнего угла к нижнему, двигаясь по диагонали (зигзагом).

Для чересстрочных (интерлейсных) кадров более эффективным является несколько модифицированный метод альтернативного сканирования. *Block scan order* – указывает метод сканирования матрицы коэффициентов DCT:

- *Alternate (interlace)* – рекомендуется использовать при сжатии интерлейсного видео (студийные программы, сериалы и пр.).
- *Zigzag (progressive)* – рекомендуется использовать при сжатии прогрессивного видео (мульти фильмы, старые фильмы). Также метод *Zigzag* принудительно включается при установке размера кадра как 352x288 (352x240).
- *B-frames weighting factor* – дополнительный весовой коэффициент, который усиливает квантование В-кадров, уменьшая тем самым битрейт при сжатии В-кадров.
- При общем низком битрейте потока (CBR 3000) увеличение этого коэффициента, например до 32, позволяет улучшить качество изображения за счёт перераспределения битрейта между I- и В-кадрами.
- Если используется только IP-последовательность GOP, то при квантовании весовой коэффициент применяется для каждого второго P-кадра, которые, в данном случае, выступают в качестве виртуальных В-кадров.

### **3. ОПИСАНИЕ ЛАБОРАТОРНОГО МАКЕТА**

В данной лабораторной работе производится исследование методов аппаратного цифрового сжатия видеинформации в реальном масштабе времени, использующих стандарт MPEG-2. Для этого используется видеопроцессор PHILIPS SAA6752HS, входящий в состав PC-TV-тюнера Behold TV M6 Extra.

TV-тюнер Behold TV M6 Extra работает в составе персонального компьютера. Программное обеспечение, необходимое для выполнения лабораторного практикума, входит в комплект поставки TV-тюнера и может использоваться на законных основаниях

## 4. ДОМАШНЕЕ ЗАДАНИЕ

1. Изучить особенности аппаратного цифрового сжатия видео-информации с использованием стандарта MPEG-2 в реальном масштабе времени на базе видеопроцессора PHILIPS SAA6752HS.
2. Изучить параметры, доступные пользователю MPEG-2 энкодера на уровне, необходимом для выполнения лабораторной работы.
3. Ознакомиться с методикой проведения лабораторной работы, проанализировать каждый пункт лабораторного задания.
4. Подготовиться к ответам на контрольные вопросы.
5. Подготовить отчет по лабораторной работе.

## 5. ЛАБОРАТОРНОЕ ЗАДАНИЕ

1. Включить компьютер, загрузить операционную систему и запустить приложение «Behold TV».
2. Открыть панель записи, которая служит для управления режимами записи, а также для отображения статистики в процессе видео/аудио или аудиозаписи (рис. 12).

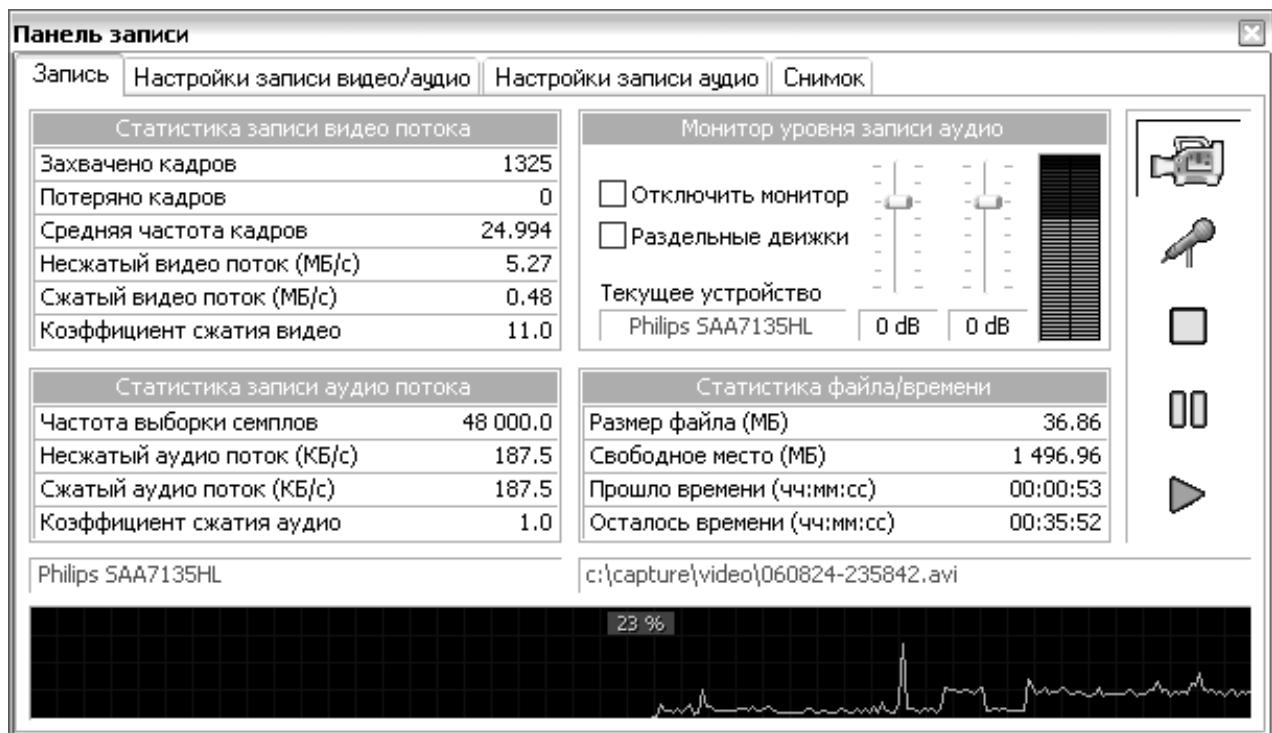


Рис. 12. Панель записи

Далее открыть закладку настроек видео/аудиозаписи (рис. 13).

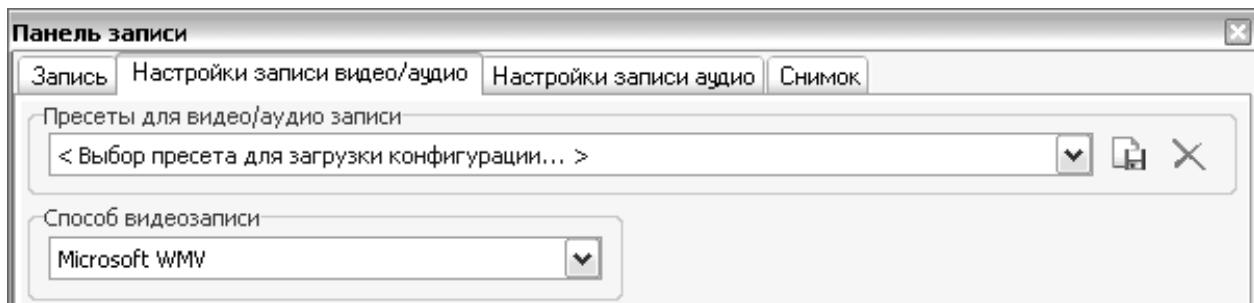


Рис. 13. Закладка настроек видео/аудиозаписи

3. Выбрать параметры записи как показано на рис. 14 и захватить одну минуту несжатого видео.

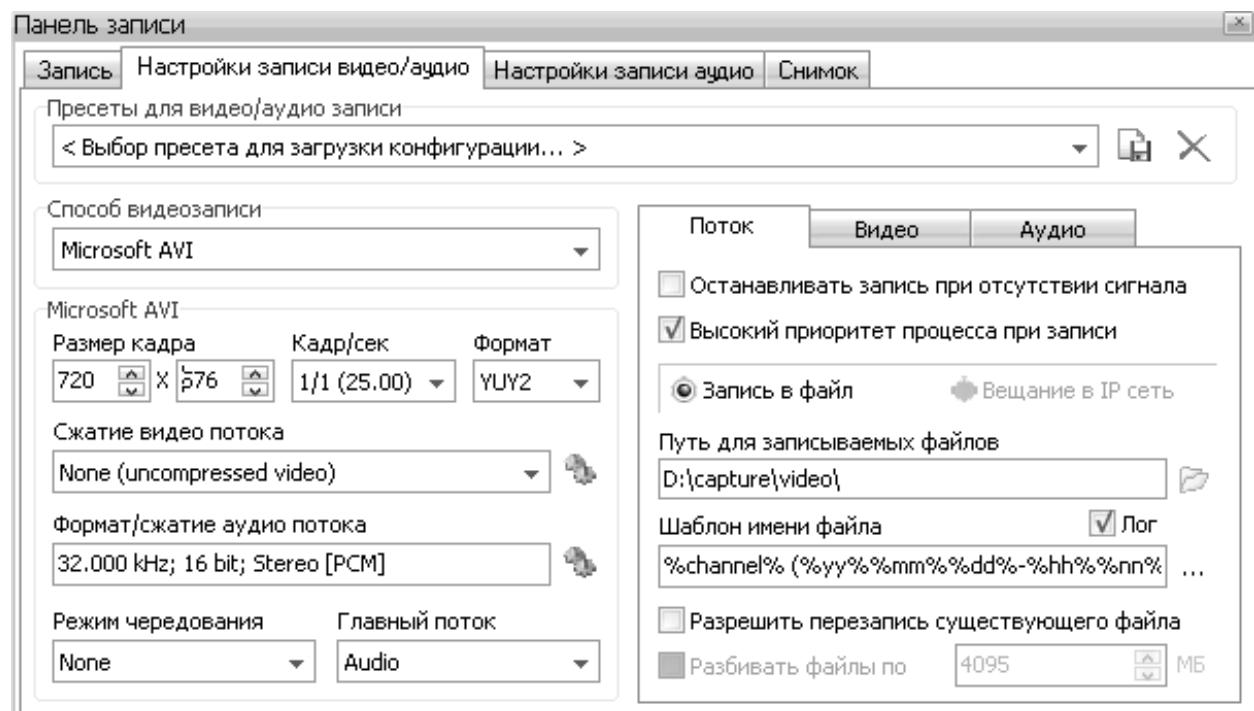


Рис. 14. Настройка видео/аудиозаписи – (*Uncompressed*)

4. Выбрать параметры записи как показано на рис. 15 и захватить одну минуту видео сжатого аппаратно в MPEG-2 DVD. Проанализировать качественные показатели изображения, вычислить коэффициент компрессии по отношению к несжатому файлу, сделать выводы.

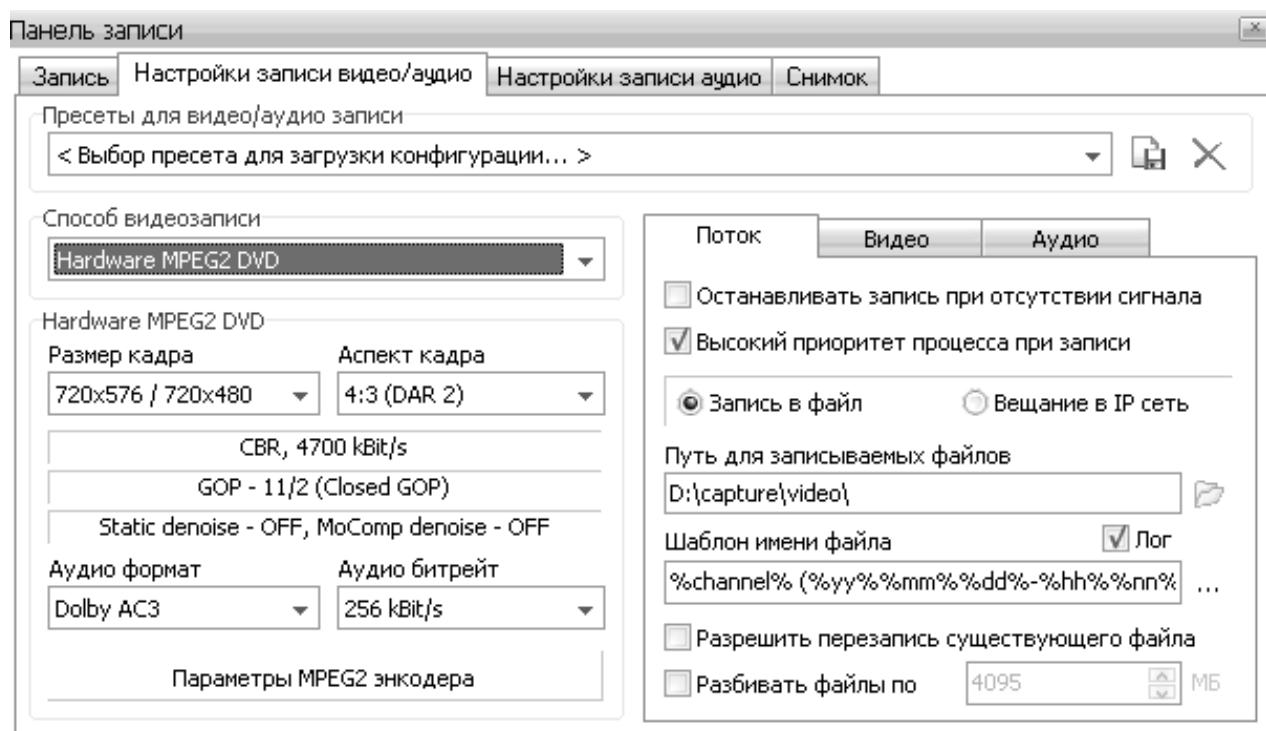


Рис. 15. Настройка видео/аудиозаписи для аппаратного сжатия видео в MPEG-2 DVD

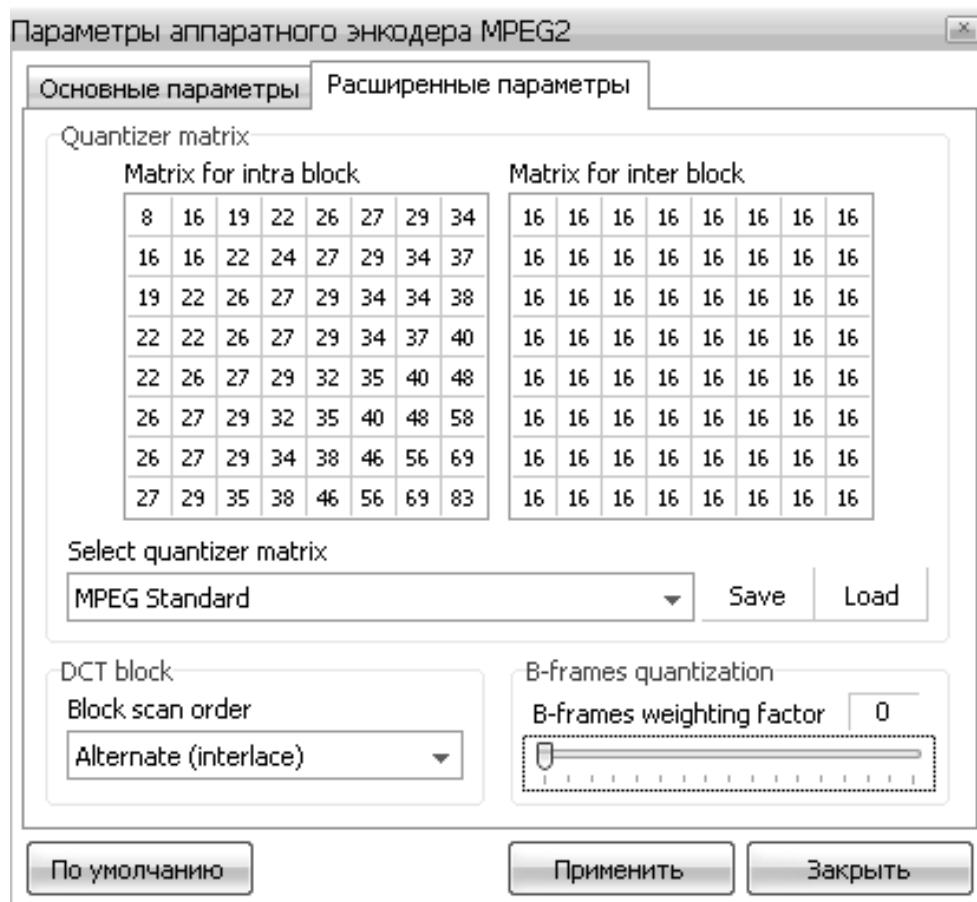


Рис 16. Расширенные параметры аппаратного энкодера MPEG-2

5. Проделать п. 4 для матрицы квантования MPEG-2стандарт как показано на рис. 16. Проанализировать качественные показатели изображения и вычислить коэффициент компрессии по отношению к несжатому файлу, сделать выводы.

6. Творчески поэкспериментировать с параметрами *B-frames weighting factor* и *Select quantizer matrix > User defined*, выполнив п. 4 для разных значений указанных параметров как показано на рис. 17. Проанализировать качественные показатели изображения и вычислить коэффициенты компрессии по отношению к несжатому файлу, сделать выводы по каждому пользовательскому пресету.

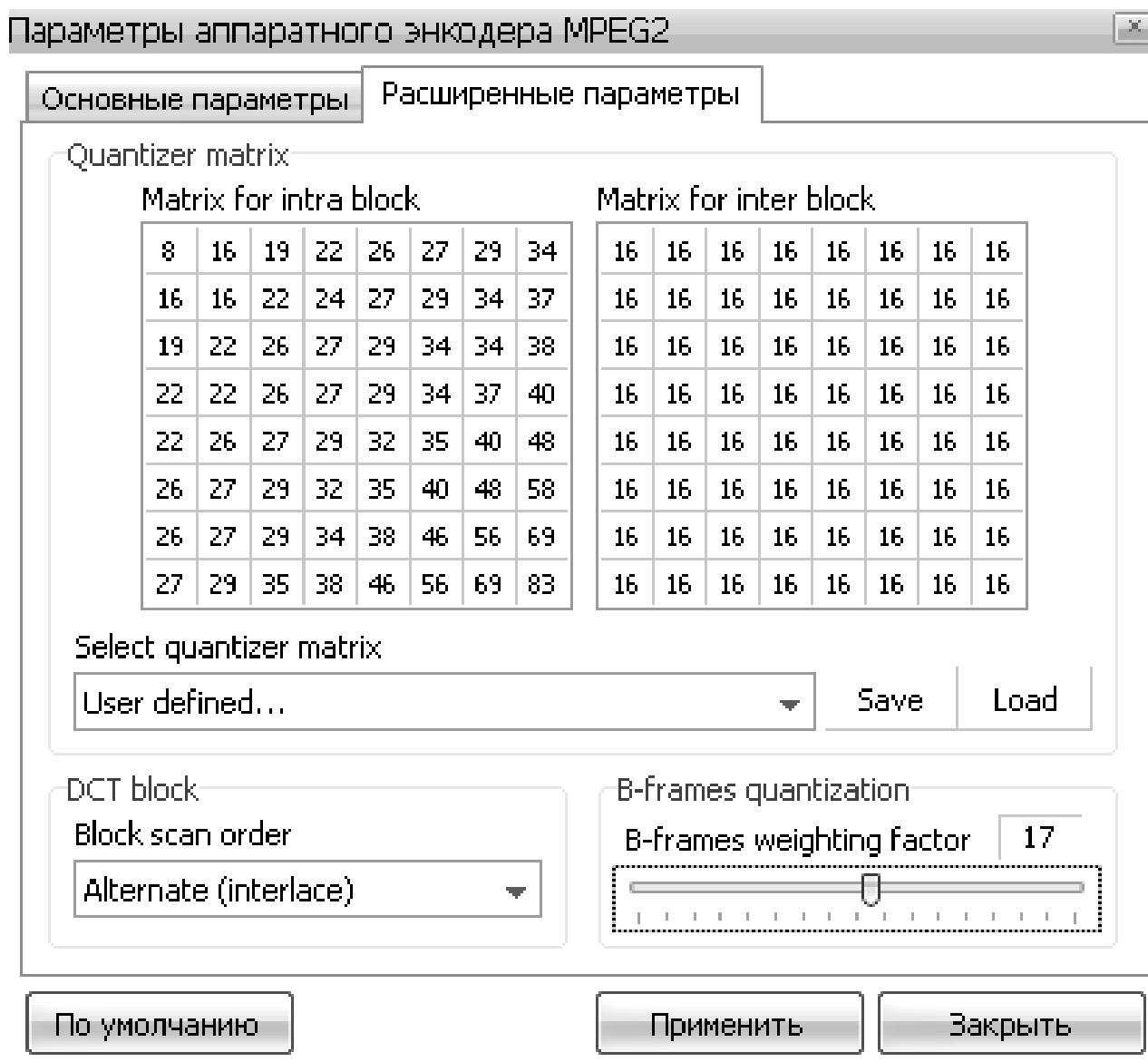


Рис. 17. Применение пользовательских настроек кодирования

## 6. СОДЕРЖАНИЕ ОТЧЕТА

1. Цель работы.
2. Результаты выполнения лабораторного задания в виде директории с файлами.
3. Выводы по результатам выполнения каждого пункта лабораторной работы.
4. Результаты выполнения лабораторного задания в виде сводной таблицы с размерами файлов, рассчитанными коэффициентами компрессии, и качественными субъективными показателями (по пятибалльной шкале) результирующих изображений, кодированных с различными настройками энкодера.

## КОНТРОЛЬНЫЕ ВОПРОСЫ

1. В чем разница между алгоритмами с потерей информации и без потери информации?
2. Какие типы избыточности используются при сжатии видеоизображения?
3. Какими параметрами характеризуется цифровой видеопоток?
4. Какие требования предъявляются к алгоритмам сжатия видео?
5. По каким параметрам сравнивают алгоритмы сжатия видео?
6. Что такое *I*-кадры, *P*-кадры, *B*-кадры, *DC*-кадры?
7. Для чего изменяют последовательность передачи *I*-, *P*- и *B*-кадров?
8. Что означает аппаратное и программное сжатие?
9. Что такое аудиовизуальная синхронизация, и почему выполнение ее требований значительно снижает степень сжатия?
10. Какие достоинства и недостатки у стандарта *MPEG-2*?
11. Какое влияние на работу сжимающего кодека оказывают динамичные быстро меняющиеся сцены?
12. Что такое макроблок и как он используется в алгоритмах сжатия видео?
13. Что такое коэффициент сжатия и как он вычисляется?
14. Что такое битрейт, как он влияет на качество сжатого видео и размер конечного файла?
15. За счет чего достигается сжатие в стандарте *MPEG-2*?
16. Почему сжатый фильм можно монтировать только по ключевым кадрам?

17. Почему возникают искажения изображения, если ключевые кадры следуют с большими интервалами?
18. Какую функцию выполняют коэффициенты квантования в матрицах квантования?
19. С какой целью используют переменный битрейт при кодировании видео?
20. В чем преимущество аппаратного сжатия перед программным?
21. Что такое видеокодек и как он реализуется на практике?

## **БИБЛИОГРАФИЧЕСКИЙ СПИСОК**

### **Основной**

1. Смирнов А.В. Основы цифрового телевидения: Учебное пособие.– М.: Горячая линия - Телеком, 2001. – 224 с.
2. Артюшенко В.М., Шелухин О.И., Афонин М.Ю. Цифровое сжатие видеинформации и звука / Под ред. В.М. Артюшенко. – М., 2003. – 426 с.
3. Ватолин Д., Ратушняк А., Смирнов М., Юркин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: Диалог-МИФИ, 2003. – 384 с.
4. Цифровое преобразование изображений: Учебное пособие для вузов / Р.Е. Быков, Р. Фрайер, К.В. Иванов, А.А. Манцветов; Под ред. проф. Р.Е. Быкова. – М.: Горячая линия – Телеком, 2003. – 228 с.

### **Дополнительный**

1. Птачек М. Цифровое телевидение: Теория и техника / Пер. с чеш.; Под ред. Л.С. Виленчика. – М.: Радио и связь, 1990. – 528 с.
2. Тимофеев Б.С. Цифровое телевидение: Учебное пособие. – СПб.: Изд-во СПбГУ-АП, 1998. – 49 с.
3. Брайс Р. Руководство по цифровому телевидению. – М.: ДМК Пресс, 2002. – 288 с.
4. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. – М.: Физматлит, 2001. – 784 с.

**СОДЕРЖАНИЕ**

1. ЦЕЛЬ РАБОТЫ .....	3
2. КРАТКИЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ .....	3
3. ОПИСАНИЕ ЛАБОРАТОРНОГО МАКЕТА.....	22
4. ДОМАШНЕЕ ЗАДАНИЕ .....	23
5. ЛАБОРАТОРНОЕ ЗАДАНИЕ .....	23
6. СОДЕРЖАНИЕ ОТЧЕТА .....	27
КОНТРОЛЬНЫЕ ВОПРОСЫ .....	27
БИБЛИОГРАФИЧЕСКИЙ СПИСОК .....	28

**Галустов Геннадий Григорьевич  
Мелешкин Сергей Николаевич  
Сидько Иван Владимирович**

# **Руководство к лабораторной работе**

## **Исследование процесса сжатия видеосигнала**

### **аппаратными средствами**

### **видеопроцессора PHILIPS SAA6752HS**

# Для студентов радиотехнического факультета всех форм обучения

# Редактор Корректор

Чиканенко Л.В.  
Селезнева Н.И.

Компьютерная верстка Мелешкин С. Н.  
Ответственный за выпуск Мелешкин С. Н.

ЛР~020565 от 27.06.1997 г. Подписано к печати г.  
Формат 60 x 84 1/16. Печать офсетная.  
Бумага офсетная. Усл. п. л. – 1,9. Уч. - изд. л. – 1,8.  
Заказ № Тираж 100 экз.

"C"

Издательство ЮФУ  
ГСП 17А, Таганрог, 28, Некрасовский, 44  
Типография ЮФУ  
ГСП 17А, Таганрог, 28, Некрасовский, 44