

**ГОСУДАРСТВЕННЫЙ КОМИТЕТ СВЯЗИ, ИНФОРМАТИЗАЦИИ И
ТЕЛЕКОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ
РЕСПУБЛИКИ УЗБЕКИСТАН**

ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

На правах рукописи

УДК 004.621.397

КИМ ЕВГЕНИЯ АЛЕКСАНДРОВНА

Исследование методов сжатия мультимедийного потока ТВ программ

5А330202 – Информатика и Мультимедийные технологии

Диссертация на соискание академической степени магистра

Научный руководитель

к.т.н., доцент

Гаврилов И. А.

Ташкент-2014

ОГЛАВЛЕНИЕ

Введение	3
Глава I. Обзор и анализ методов сжатия мультимедийного потока ТВ – программ	9
1. Общие положения мультимедийного потока	9
2. Классификация типов избыточной информации мультимедийных данных ТВ программ и методы их устранения	11
3. Анализ методов сжатия видеоданных ТВ программ.....	17
4. Анализ методов сжатия аудиоданных ТВ программ.....	39
5. Анализ методов сжатия текстовой и служебной информации	57
Выводы по I главе	60
Глава II. Аналитический обзор мультимедийных кодеков и оценка их характеристик	63
1. Общие положения мультимедийных кодеков	63
2. Программные видеокодеки.....	64
3. Аппаратные устройства кодирования мультимедийного потока	73
Выводы по II главе	81
Глава III. Анализ принципов построения и работы кодеков мультимедийного контента ТВ программ и оценка их эффективности.....	83
1. Общие положения работы кодеков	83
2. Анализ построения кодеков стандарта MPEG-4 и его модификаций	84
3. Анализ видео кодеков на основе вейвлет-преобразований	98
4. Сравнительная оценка различных методов кодирования видео сигналов ТВ программ	103
5. Исследование эффективности сжатия аудио данных различными кодеками MP3, AAC, Ogg Vorbis	123
Выводы по III главе	129
Заключение	132
Список литературы.....	138

Введение

Актуальность данного направления исследований заключается в том, что в связи с увеличением потока мультимедийной информации и ограниченной скорости передачи данных ТВ программ, проведенные исследования позволят более эффективно использовать частотный ресурс и организовать передачу ТВ программ по узкополосным каналам Интернета и мобильной связи без заметного ухудшения качества изображений и звука.

В современном мире повсеместно внедряются цифровые технологии в телевизионную отрасль. Это приводит к необходимости передавать все больше информации по каналам связи, поэтому важно искать новые методы сжатия видеoinформации, так как сжатие сокращает объем информации, и количество времени, необходимого для ее передачи по каналу связи установленной ширины пропускания.

В настоящее время в мире телевидение, как средство массовой информации играет очень важную роль в жизни каждого человека, поскольку позволяет донести до него информацию практически из любой части света.

Президент Ислам Каримов в своих выступлениях неоднократно отмечал важность информационных и компьютерных технологий, цифровых и беспроводных широкополосных сетей телекоммуникации, интернета и цифрового телевидения, подчеркивая, что масштабность внедрения современных систем связи служит показателям уровня развития страны и общества [1]. Поэтому существенное внимание этой высокотехнологичной сфере уделяется буквально с первых дней после обретения государственного суверенитета.

В Постановлении Президента Республики Узбекистан «О мерах по дальнейшему внедрению и развитию современных информационно-коммуникационных технологий» принята «Программа дальнейшего внедрения и развития информационно-коммуникационных технологий в Республике Узбекистан на 2012–2014 годы», в которой большое внимание

уделяется развитию инфраструктуры информационно-коммуникационных технологий, в частности, разработке и реализации технических проектов, направленных на обеспечение предоставления населению услуг широкополосного доступа [2, 3]. Поэтому в настоящее время в области обработки видеоданных ведется разработка новых алгоритмов и внедрение новых методов, направленных как на повышение качества телепередач, так и расширение возможностей их передачи, в том числе за счет использования систем сотовой связи 3 поколения, что в условиях ограниченного частотного ресурса представляет серьезную проблему.

Огромным прорывом в телевизионной сфере стала возможность перехода с аналогового телевизионного вещания на качественно новый уровень – наземное эфирное цифровое телевидение. Главным отличием цифрового телевидения от аналогового является качество изображения: более насыщенные и яркие цвета, четкие детали. Второе немаловажное преимущество – это звук в формате 5.1.

Основным программным документом в этой работе служит постановление Президента Ислама Каримова о Государственной программе по техническому и технологическому переходу на цифровое телевидение в Республике Узбекистан от 17 апреля 2012 года. Это постановление предусматривает в первую очередь дальнейшее совершенствование нормативно-правовой базы в области внедрения цифрового телевидения с учетом технико-технологического развития, поэтапный переход к цифровому телевидению: I этап - 2012-2015 годы; II этап - 2016-2017 годы [3].

В соответствии с Государственной программой развитие цифрового телевидения в последние годы переживает в нашей стране поэтапный переход от аналогового к цифровому, становится наиболее интенсивно развивающейся и конкурентной отраслью рынка. В соответствии с концепцией развития телевидения в Узбекистане происходит постепенный переход на цифровой стандарт вещания, при этом на сегодняшний день охват населения цифровым телевидением увеличился до 38% [1]. Транслируются

36 обычных цифровых ТВ программ и 5 цифровых программ с высокой четкостью, одна из которых национальная.

С развитием цифрового телевидения возникают потребности в увеличении количества программ, для предоставления дополнительных сервисных услуг, а также для организации интерактивного телевидения, требующего большого количества свободных каналов. Однако, в условиях ограниченного частотного ресурса, увеличить число передаваемых программ можно только совершенствуя методы сжатия аудио-видео информации.

На сегодняшний день существует довольно много методов и стандартов сжатия видео данных, к которым можно отнести: Motion JPEG (Joint Photographic Experts Group), MPEG-2, MPEG -4 и его модификации, имеющие свои достоинства и недостатки. При этом одни обеспечивают приемлемое качество изображения, но малую степень сжатия, другие, - хорошую степень сжатия, но высокую сложность реализации алгоритма, который в свою очередь сильно влияет на скорость компрессии и декомпрессии.

Большинство кодеков для аудио и видео данных используют сжатие с потерями, для того чтобы существенно уменьшить размер сжатого файла. Существуют также кодеки, сжимающие без потерь, но для большинства из них, малозаметное улучшение качества не оправдывает значительного увеличения объема получаемых файлов. Главным исключением может служить ситуация, когда данные будут подвергаться дальнейшей обработке: в этом случае повторяющиеся потери на кодировании и декодировании окажут серьезное влияние на качество.

В связи с увеличением потока мультимедийной информации и ограниченной скорости передачи данных ТВ программ огромное внимание уделяется исследованию и разработке эффективных методов сжатия мультимедийного потока.

Объектом исследования является поток мультимедийных данных и его сжатие различными типами видеокодеков.

Предметом данного исследования является рассмотрение и совершенствование существующих методов компрессии мультимедийных данных с учетом устранения избыточности данных ТВ изображений без заметного ухудшения качества изображений и звука.

Целью работы является исследование эффективности сжатия мультимедийного потока аудио-, видеоданных ТВ программ различными кодеками и выбор алгоритмов их работы, обеспечивающих хорошее визуальное качество изображений при скоростях цифрового потока не более 2 Мбит/с.

Задачи диссертации заключаются в следующем:

- анализ типов избыточной информации мультимедийных данных ТВ и методов их устранения;
- анализ методов сжатия аудио-, видео данных на основе спектральных преобразований, компенсации движения и психоакустики;
- анализ построения и работы современных кодеков мультимедийного контента;
- экспериментальная оценка эффективности работы мультимедийных видеокодеков;
- оценка эффективности полученных результатов и выбор алгоритма и структуры построения видео кодека.

Методы исследования: для достижения поставленной цели в диссертационной работе использовались методы системного анализа, методы цифровой обработки сигналов, теория дискретных сигналов, теория информации, методы вейвлет-анализа и методы компьютерного моделирования.

Краткий литературный обзор по теме исследования

В настоящее время во всемирной научной литературе большое внимание уделяется сжатию информации. Основные методы и алгоритмы сжатия данных нашли свое отражение в работах таких авторов как Ватолин Д., Ратушняк А., Смирнов М., Юкин В, Сэлмон Д.[21, 38, 39], которые

анализировали эти методы и их применение. Также исследованы алгоритмы архивации данных, впервые изученные аспирантом Массачусетского технологического института Дэвидом Хаффманом, а также предложенные американскими учёными Шенноном и Фано [38, 39].

В настоящее время получают все большее распространение альтернативные методы сжатия изображения и звука на основе вейвлет-преобразований, введенных Гроссманом и Морле в середине 1980-х годов.

В настоящее время для сжатия ТВ программ широкое распространение получили стандарты вещательного ТВ MPEG-4 [29] и стандарт MPEG-4-10, принятый ISO в 2001 [36], позволяет получить большие коэффициенты сжатия. Но на скоростях видеопотока менее 3 Мбит/с в нем проявляются искажения в виде блочного эффекта.

Научная новизна данной работы заключается в результатах исследований новых методов обработки мультимедийного потока ТВ программ, которые позволят эффективно использовать и передавать видеoinформацию по узкополосным каналам связи не более 2 Мбит/с и мобильной связи без заметного ухудшения качества изображений и звука.

Теоретическая и практическая значимость результатов исследования состоят в том, что проведенные исследования позволяют выявить наиболее эффективные методы сжатия мультимедийных данных ТВ программ с наилучшим соотношением коэффициента сжатия и визуального качества восстановленных изображений, что позволит более эффективно использовать каналы связи или накопители для хранения видеофильмов.

Апробация работы. **Основные положения диссертационной работы докладывались и обсуждались на следующих НТК:**

На Республиканской научно-технической конференции молодых ученых, исследователей, магистрантов и студентов на тему «Информационные технологии и проблемы телекоммуникаций», Ташкент, 14-15 марта 2013 г.

На Республиканской научно-технической конференции молодых ученых, исследователей, магистрантов и студентов на тему «Информационные технологии и проблемы телекоммуникаций», Ташкент, 14-15 марта 2014 г.

Диссертационная работа состоит из введения, 3 глав, заключения и списка литературы из 49 наименований. Основной текст содержит 135 страницы и иллюстрируется 57 рисунками и 10 таблицами.

Содержание работы.

Во введении обоснована актуальность темы, научная новизна и практическая ценность исследований, изложены цель и задачи исследования выполненной работы.

В первой главе приведены основные понятия мультимедийного контента ТВ программ. Проведена классификация избыточности мультимедийной информации и проанализированы основные методы сжатия видео, аудио и текстовой информации.

Во второй главе проведен обзор современных мультимедийных кодеков ТВ программ и сравнительному анализу их технических характеристик.

В третьей главе был проведен анализ построения существующих кодеков мультимедийного контента ТВ программ, а также проведены экспериментальные исследования эффективности работы кодеков стандартов H.264, H.265 и кодек Divx и их сравнение с выявлением всех их достоинств и недостатков. На основе полученных экспериментальных данных предложен алгоритм, обеспечивающий наилучшее соотношение объема данных и качества для более эффективного использования частотного ресурса и носителей информации.

В заключении сформулированы теоретические и практические выводы к диссертационной работе.

Все результаты получены автором лично.

Глава I. Обзор и анализ методов сжатия мультимедийного потока ТВ – программ

1. Общие положения мультимедийного потока

В настоящее время мультимедиа-технологии являются бурно развивающейся областью информационных технологий. В этом направлении активно работает значительное число крупных и мелких фирм, технических университетов и студий.

Термин "мультимедиа" образован из слов "мульти" — много, и "медиа" — среда, носитель, средства сообщения, и в первом приближении его можно перевести как "многосредность" [5].

Мультимедиа — это собирательное понятие для различных компьютерных технологий, при которых используется несколько информационных сред, таких, как графика, текст, видео, фотография, движущиеся образы (анимация), звуковые эффекты, высококачественное звуковое сопровождение.

Таким образом, в широком смысле термин "мультимедиа" означает спектр информационных технологий, использующих различные программные и технические средства с целью наиболее эффективного воздействия на пользователя (ставшего одновременно и читателем, и слушателем, и зрителем) [5].

Средства мультимедиа — комплекс аппаратных и программных средств, позволяющих человеку общаться с компьютером, используя самые разные, естественные для себя среды: звук, видео, графику, тексты, анимацию и т. д.

Основными характерными особенностями этой технологии являются:

- объединение многокомпонентной информационной среды (текста, звука, графики, фото, видео) в однородном цифровом представлении;
- обеспечение надежного и долговечного хранения больших объемов информации;

– простота переработки информации.

Многокомпонентную мультимедиа-среду целесообразно разделить на три группы: текстовая информация, аудиоряд, видеоряд.

Рассмотрим основные компоненты мультимедийных объектов.

Текст – это упорядоченный набор предложений, предназначенный для того, чтобы выразить некий смысл. В смысловой цельности текста отражаются те связи и зависимости, которые имеются в самой действительности.

Аудио – общий термин, относящийся к звуковым технологиям. Как правило, под термином аудио понимают звук, записанный на звуковом носителе, а также запись и воспроизведение звука, звукозаписывающая и звуковоспроизводящая аппаратура. Таким образом, аудиальный компонент мультимедийной информации предназначен для передачи звуковых данных.

Аудиоряд может включать речь, музыку, эффекты, файлы, содержащие цифровую запись аудиоданных (голоса, музыкальных произведений или их фрагментов и т.п.), объединяемые обозначением WAVE (волна) [6]. Главной проблемой при использовании этой группы мультисреды является информационная емкость. Для записи одной минуты WAVE-звука высшего качества необходима память порядка 10 Мбайт [6], поэтому стандартный объем CD (до 640 Мбайт) позволяет записать не более часа WAVE. Для решения этой проблемы используются методы компрессии звуковой информации.

Видеоряд по сравнению с аудиорядом характеризуется большим числом элементов. Выделяют статический и динамический видеоряды.

Статический видеоряд включает графику (рисунки, интерьеры, поверхности, символы в графическом режиме) и фото (фотографии и сканированные изображения).

Динамический видеоряд представляет собой последовательность статических элементов (кадров), что характерно для видеофильмов и телевизионных передач.

Цифровой видеопоток имеет четыре основные характеристики: экранное разрешение, частота кадров, битрейт (скорость видеопотока) и качество изображения.

Экранное разрешение (Resolution) – обозначает количество точек (пикселей) по горизонтали и вертикали, участвующих в построении изображения (видеокадра) на экране. Для европейского видеостандарта PAL размер кадра составляет 720x576 пикселей, для североамериканского стандарта NTSC – 720x480, для видео высокой четкости (HD 720p) – 1280x720, а для новомодного стандарта HDTV (Full HD) – 1920x1080 точек.

Частота кадров – величина указывающая, на количество кадров сменяющихся за секунду. Стандартной скоростью воспроизведения видеосигнала считается величина равная 25 кадрам/с. Для кино этот показатель несколько меньше и составляет 24 кадра/с [6].

Битрейт (скорость видеопотока) – показывает количество передаваемых бит видеоинформации за одну секунду времени. Иначе говоря – это скорость видеопотока, измеряемая в мегабитах в секунду (Мбит/с). Чем она выше, тем лучше качество.

Качество изображения – характеристика призванная оценить качество обработанного видео в сравнении с оригиналом и определяющаяся совокупностью значений разрешения, глубины цвета и скорости видеопотока.

2. Классификация типов избыточной информации мультимедийных данных ТВ программ и методы их устранения

Компрессией телевизионного сигнала называется устранение из него избыточной информации для уменьшения скорости цифрового потока и, как следствие, более эффективного использования каналов передачи [8].

Из анализа ТВ изображений известно, что они обладают большим объемом избыточной информации, которую можно разделить на классы [8]:

- кодовая избыточность;
- статистическая или межэлементная;
- психовизуальная;
- структурная;
- временная или межкадровая избыточность.

При этом сжатие информации производится вследствие устранения одного или нескольких указанных типов избыточности, которые рассмотрим подробнее.

Кодовая избыточность обычно возникает в изображениях имеющих довольно много объектов регулярной формы и отражательных свойств поверхности, в результате чего определенные значения яркости встречаются более часто, чем другие (рис.1.1.). А поскольку при двоичном кодировании яркостей пикселей используются коды одинаковой длины, то это и приводит к появлению кодовой избыточности. Для устранения кодовой избыточности часто применяют энтропийное кодирование кодами переменной длины на основе таблиц Хаффмана, где наиболее часто повторяющиеся кодовые комбинации заменяются короткими кодами, а редко встречающиеся - длинными, как в азбуке Морзе. Такой подход позволяет на 20-25% снизить объем передаваемой информации [9].

Статическая или межэлементная избыточность возникает из-за высокой разрешающей способности изображения, которая необходима для отображения мелкоструктурных участков изображения или границ контуров объектов, а на всех гладких участках изображения она расходуется впустую, резко увеличивая объем информации (рис.1.1.). Межэлементная статистическая избыточность основана на довольно сильной корреляционной связи пикселей и обычно устраняется методами на основе спектральных преобразований, которые оценивают распределения энергии сигнала по его спектральным составляющим. В настоящее время известно довольно много

математических функций используемых для спектральных преобразований, но на практике наибольшее распространение для сжатия изображений получили дискретно-косинусное и вейвлет преобразования [10].

Психофизическая избыточность основывается на особенностях нашего зрительного восприятия. Так, глаз меньше замечает изменения цветности, чем яркости. Кроме того, установлено, что при наблюдении человек стремится, в первую очередь, отыскать в изображении его наиболее важные его части (контуры объектов или текстурные области) и образовать из них комбинации, поддающиеся распознаванию. А цвет и яркость элементов при этом играют вспомогательную роль. Поэтому часть информации в изображении может быть исключена без заметного визуального ухудшения его качества [10].

Структурная избыточность основывается на присутствии подобных участков в изображении, как показано на рис.1.2. Для устранения структурной избыточности используют сканирование изображения для нахождения его повторяющихся фрагментов (фракталов), которые заменяются ссылками на уже найденный фрагмент, что существенно снижает объем передаваемой информации.

Временная или межкадровая избыточность проявляется в телевизионных изображениях, если в пределах одного видеосюжета информация в соседних кадрах обычно изменяется мало (рис.1.3). Поэтому, если передавать только изменения изображений относительно опорного кадра, например, в виде межкадровой разности, то можно получить довольно большие коэффициенты сжатия видеопотока. Однако на практике межкадровая разность обычно не применяется, поскольку при небольшом изменении ракурса съемки взаимные координаты пикселей изображений смещаются и значения межкадровой разности возрастает, увеличивая объем межкадровой информации.

На практике применяют более сложную межкадровую обработку на основе компенсации движения фрагментов изображений смежных кадров.

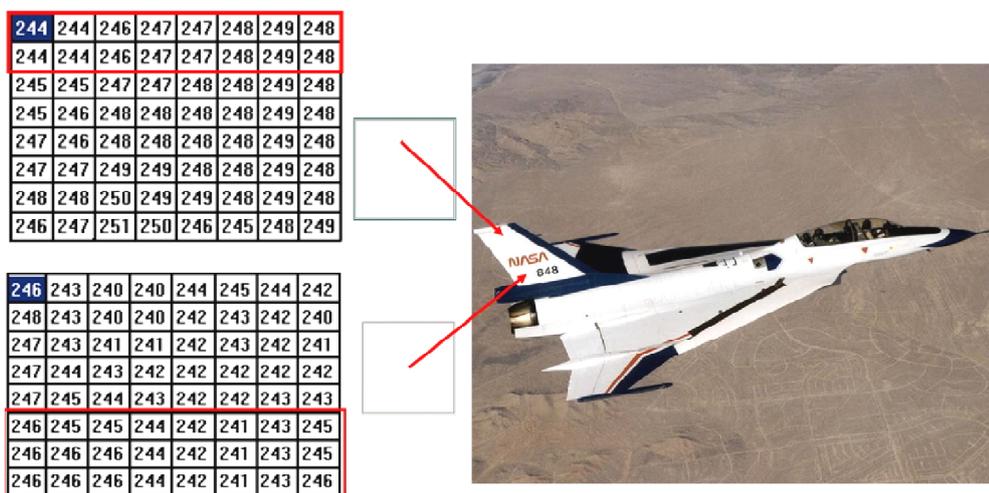


Рис.1.1. Пример изображения с выраженной кодовой и меж элементной избыточностью



Рис.1.2. Пример изображения с выраженной структурной избыточностью



Рис.1.3. Изображения смежных кадров видесюжета и их межкадровой разницы

Суть метода заключается в поиске фрагментов изображения первого кадра в зонах их предполагаемого смещения в следующем кадре. Если такие фрагменты находятся, то вместо них передаются их новые координаты – вектора перемещения. При этом структура видеопотока состоит из опорного кадра, где устраняется только внутрикадровая избыточность и одного или нескольких типов кадров передающих межкадровые различия и векторы перемещений блоков, что реализуется в стандартах сжатия семейства MPEG и многих других кодеках.

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 [10].

На сегодняшний день разработано довольно много различных методов и алгоритмов сжатия видеоинформации, обладающие разной эффективностью сжатия, качественными показателями, сложностью реализуемых алгоритмов и быстродействием. При этом в механизмах обработки изображений можно выделить следующие направления:

- Сжатие на основе спектральных преобразований;
- Фрактальное сжатие;
- Векторное квантование.

Для звуковых сигналов различают статистическую и психоакустическую избыточность первичных цифровых сигналов. Сокращение статистической избыточности базируется на учете свойств самих звуковых сигналов, а психоакустической - на учете свойств слухового восприятия.

Статистическая избыточность обусловлена наличием корреляционной связи между соседними отсчетами временной функции звукового сигнала при его дискретизации. Для ее уменьшения применяют достаточно сложные алгоритмы обработки. При их использовании потери информации нет, однако исходный сигнал оказывается представленным в более компактной форме, что требует меньшего количества бит при его кодировании. Важно, чтобы все эти алгоритмы позволяли бы при обратном преобразовании восстанавливать исходные сигналы без искажений. Наиболее часто для этой цели используют ортогональные преобразования. Оптимальным с этой точки зрения является преобразование Карунена – Лозва [11]. Незначительно по эффективности ему уступает модифицированное дискретное косинусное преобразование (МДКП). Кроме того, между коэффициентами преобразования Фурье и коэффициентами МДКП существует простая связь, что позволяет представлять результаты вычислений в форме, достаточно хорошо согласующейся с работой механизмов слуха. Дополнительно уменьшить скорость цифрового потока позволяет также метод кодирования, учитывающий стати звуковых сигналов, код Хаффмана. Именно в силу этих двух причин в наиболее эффективных алгоритмах компрессии цифровых аудиоданных кодирование подвергаются не сами отсчеты ЗС, а коэффициенты МДКП, и для их кодирования используются кодовые таблицы Хаффмана.

После устранения статистической избыточности скорость цифрового потока при передаче высококачественных ЗС и возможности человека по их обработке отличаются, по крайней мере, на несколько порядков. Это свидетельствует также о существенной *психоакустической избыточности* первичных цифровых ЗС и, следовательно, о возможности ее уменьшения. Наиболее перспективными с этой точки зрения оказались методы, учитывающие такие свойства слуха, как маскировка, предмаскировка и послемаскировка. Если известно, какие части звукового сигнала ухо воспринимает, а какие нет вследствие маскировки, то можно вычлени

затем передать по каналу связи лишь те части сигнала, которые ухо способно воспринять, а неслышимые доли можно отбросить. Кроме того, сигналы можно квантовать с возможно меньшим разрешением по уровню, так, чтобы искажения квантования, изменялись по величине с изменением уровня самого сигнала, еще оставались бы неслышимыми, т.е. маскировались бы исходным сигналом. Однако после устранения психоакустической избыточности точное восстановление формы временной функции ЗС при декодировании оказывается уже невозможным.

3. Анализ методов сжатия видеоданных ТВ программ

Методы сжатия данных используют математические алгоритмы для устранения, группировки и/или усреднения схожих данных, присутствующих в видеосигнале. Существует большое разнообразие алгоритмов сжатия, но только Motion JPEG (Joint Photographic Experts Group), MPEG-1, MPEG-2, MPEG -4 и MPEG - 7 признаны международными стандартами для сжатия видео [12].

Одним из наиболее распространенных способов сжатия видеоданных является применение ортогональных преобразований. Наиболее часто используются методы линейных ортогональных преобразований к которым относятся [12]:

- преобразование Уолша-Адамара;
- преобразование Карунена-Лоэва;
- дискретное косинус-преобразование (ДКП);
- вейвлет-преобразование (ВП).

У каждого из приведенных преобразований есть свои достоинства, недостатки и области применения. Так достоинством преобразования Адамара является простота реализации и низкая вычислительная сложность. Данное разложение дает хорошие результаты для кусочно-постоянных функций, выделяющих постоянные составляющие сигналов, однако, в

реальных изображениях такие сигналы встречаются достаточно редко. Основным недостатком преобразования Карунена-Лоэва является то, что пока не разработан быстрый метод вычисления его векторов, поэтому данный метод пока является сугубо теоретическим.

Таким образом, из перечисленных выше преобразований на практике активно используются только ДКП и ВП, которые рассмотрим более подробно.

Сжатие видеоданных на основе дискретно-косинусного преобразования (ДКП).

ДКП является хорошо изученным и весьма эффективным преобразованием, предложенным В. Ченом в 1981 году и используемое в форматах JPEG, MJPEG, MPEG-1, MPEG-2, MPEG-4 [13]. По сути, этот метод сходен с двумерным дискретным преобразованием Фурье и отличается от него только используемыми базисными функциями. Достоинством ДКП является быстрая сходимость ряда, что обеспечивает меньшую погрешность ошибки преобразования.

Прямое и обратное ДКП описываются уравнениями (1.1, 1.2)

$$F(u, v) = (1/4)C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 p(x, y) \left[\cos \frac{(2x+1)u\pi}{16} \right] \left[\cos \frac{(2y+1)v\pi}{16} \right] \quad (1.1)$$

$$f(x, y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u, v) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (1.2)$$

где v - горизонтальная координата графического блока, u - вертикальная, x - вертикальная координата внутри блока, а y - горизонтальная координата внутри блока, $C(u), C(v) = 1/\sqrt{2}$ для $u, v = 0$ и $C(u), C(v) = 1$ в противном случае.

Данный метод предусматривает разбиение кадра (рис.1.4) на блоки по 64 (8x8) отсчета, называемые сигнальными матрицами (рис.1.5,а). После чего сигнальные матрицы преобразуются в матрицы частотных коэффициентов

(рис.1.5,б) такого же размера, которые можно рассматривать как двумерный спектр изображения в горизонтальном и вертикальном направлениях.

В такой матрице коэффициенты в левом верхнем углу соответствуют низкочастотной составляющей изображения, а в правом нижнем — высокочастотной.



Рис.1.4. Исходное изображение

Особенность спектра ДКП состоит в том, что основная энергия частотных составляющих этого спектра концентрируется в небольшой области около нулевых частот. Амплитуда высокочастотных составляющих или мала, или просто равна нулю, поэтому передаче подлежат только те частотные коэффициенты матрицы ДКП, величины которых превышают принятые пороговые значения. Коэффициенты ниже порогового значения считаются нулевыми и не передаются для чего производится их зигзагообразное считывание (рис.1.5,б) и сжатие статистическим компрессором длинных серий (RLE).

Если при передаче отбрасываются только нулевые коэффициенты, то получается сжатие без потерь качества.

216	213	211	211	213	217	220	221
218	217	218	221	224	226	226	226
217	219	223	227	229	229	228	227
221	223	227	230	230	230	230	230
231	231	233	232	230	230	233	236
229	229	228	226	223	225	230	235
218	220	221	219	217	219	226	232
213	218	221	220	219	221	228	234

а)

1795	-27	6	7	1	1	0	8
-12	0	-7	10	1	1	0	0
-35	-6	7	0	-1	-1	0	0
0	9	11	0	0	1	0	0
0	0	0	-1	0	0	0	0
-12	-1	0	0	0	0	8	8
0	1	0	0	0	0	0	0
0	1	0	-1	0	0	0	1

б)

Рис.1.5. Сигнальная матрица яркостей пикселей исходного изображения размером 8x8 пикселей (а) и матрица коэффициентов после прямого ДКП (б)



Рис.1.6. Исходное и восстановленное изображение при сжатии в 100 раз

Однако при этом коэффициент сжатия не высок и в среднем составляет 10-20 раз в зависимости от детальности изображения. Для управления коэффициентом сжатия применяют деление коэффициентов ДКП на определенные числа (матрицу квантования) с последующим округлением до целого числа, что увеличивает длину цепочек нулевых коэффициентов и соответственно коэффициент сжатия. Однако, это округление данных с одной стороны приводит к увеличению сжатия изображения, с другой стороны, к необратимым потерям информации в результате чего при больших коэффициентах сжатия нарушается плавность изменения яркости на границах блоков, что приводит к возникновению искажений в виде блочного эффекта снижающего разборчивость и качество восстановленного изображения, как показано на рис.1.6.

Данный метод обладает хорошей производительностью, хорошо сочетается с блочным методом компенсации движения и обеспечивает хорошее качество изображений при скоростях видеопотока более 5 Мбит/с. Однако, на меньших скоростях сильно проявляются искажения в виде блочного эффекта, в результате которого изображение приобретает мозаичный вид, что является основным недостатком этого метода сжатия.

На практике, ДКП используется в стандартах сжатия изображений JPEG и MPEG-1, MPEG-2, MPEG-4-10, H.264 и др.

Сжатие изображений на основе вейвлет-преобразований (ВП). В последние годы все большее распространение получают методы сжатия изображения и звука на основе вейвлет-преобразований, введенных Гроссманом и Морле в середине 80-х годов прошлого века. Основным недостатком преобразований Фурье и ДКП, заключается в том, что их гармонические базисные функции плохо работают с непериодическими сигналами, в результате чего безвозвратно теряется часть полезной информации. ВП основано на представлении некоторой функции композицией функций-вейвлетов, где вейвлет - это всплеск или маленькая волна. В настоящее время существует много вейвлет-функций,

отличающихся свойствами и областями применения. Некоторые из наиболее распространенных вейвлетов, представленных на рис.1.7, эффективно используются для обработки сигналов изображения [14].

Преобразуемый сигнал подвергается фильтрации с помощью фильтров нижних и верхних частот, которые делят диапазон частот исходного сигнала на две половины, поэтому НЧ и ВЧ компоненты сигнала, полученные при фильтрации, имеют в два раза более узкую полосу частотных составляющих.

На практике вейвлет-преобразование может рассматриваться как процесс обработки изображения фильтрами нижних и высоких частот с последующим прореживанием полученных коэффициентов (рис.1.8.), поскольку каждое второе значение несет избыточную информацию.

В результате проведенной обработки исходный массив данных bp преобразуется в 2 массива коэффициентов: НЧ (A_n - аппроксимации) и ВЧ (D_n), имеющих в 2 раза более узкую полосу частотных составляющих (рис.1.9). Причем, при обработке изображений ВП обычно производится не однократно, а повторяется 3-5 раз.

Принцип вейвлет преобразования можно рассмотреть на примере обработки изображения, представленного на рис.1.10. Фильтр представляет собой небольшое «окно», в котором значения яркости и цветности пикселей умножаются на заданный набор коэффициентов вейвлет-функции, а полученные значения суммируются, после чего «окно» сдвигается для расчета следующего значения. В результате фильтрации вместо одного изображения размером $m \times n$ вейвлет-преобразование дает четыре изображения размером $(m/2) \times (n/2)$ (рис.1.11). Фильтрация НЧ фильтром по горизонтали и по вертикали дает самое высокоинформативное изображение, которое подвергается дальнейшей фильтрации (число уровней фильтрации обычно составляет от 4 до 6). Изображения, полученные после НЧ и ВЧ фильтрации представляют собой массив числовых коэффициентов, которые также, как и при ДКП, для управления коэффициентом сжатия делятся на определенные числа, квантуются и сжимаются статистическими

компрессорами, после чего попадают в выходной поток.

Таким образом, в результате многократной вейвлет декомпозиции получается совсем малое изображение (в левом верхней части экрана) в массиве аппроксимации, занимающее небольшой объем информации, и большая область массива детализации, заполненная нулями или малыми коэффициентами (на рис.2.11. нулевым значениям соответствует серый цвет со значением 128), которые хорошо сжимаются статистическими компрессорами, что позволяет получить довольно большие коэффициенты сжатия.

Одной из особенностей механизма сжатия изображений и помощью ВП является то, что, как правило, изображение не делится на блоки, а обрабатывается целиком. Это устраняет возникновение блочных искажений и позволяет в 1,5-2 раза увеличить сжатие статического изображения без заметного ухудшения его качества, или соответственно улучшить качество восстановленных изображений, как показано на рис.1.11.

Как видно на рис.1.12, безблочная обработка изображений с применением вейвлетов обеспечивает значительно лучшее качество восстановленных изображений. Однако, именно безблочная обработка изображений не позволяет применять компенсацию движения, как это сделано в стандартах MPEG. Поэтому вейвлет - кодеки по эффективности сжатия видеопотока в 2-3 раза уступают кодам стандартов MPEG. Кроме того, при высокой степени сжатия вейвлет - кодеки ухудшают четкость изображения и могут давать искажения, имеющие вид ряби вблизи резких границ, но такие искажения, в среднем, меньше замечаются глазом, чем «мозаика», создаваемая ДКП.

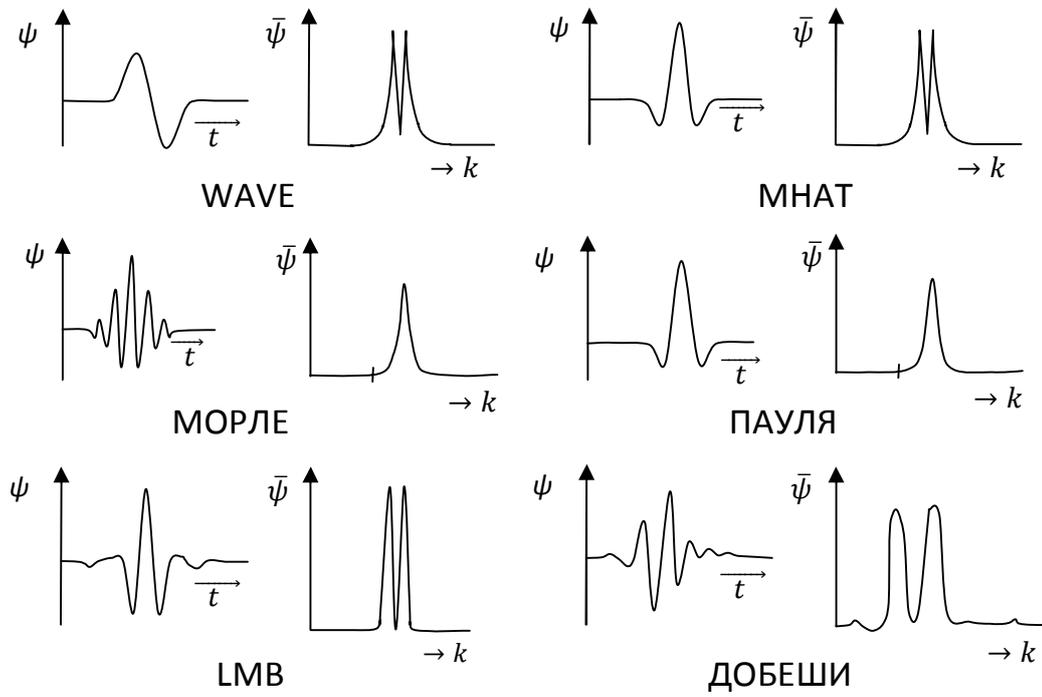


Рис.1.7. Некоторые наиболее распространенные вейвлеты

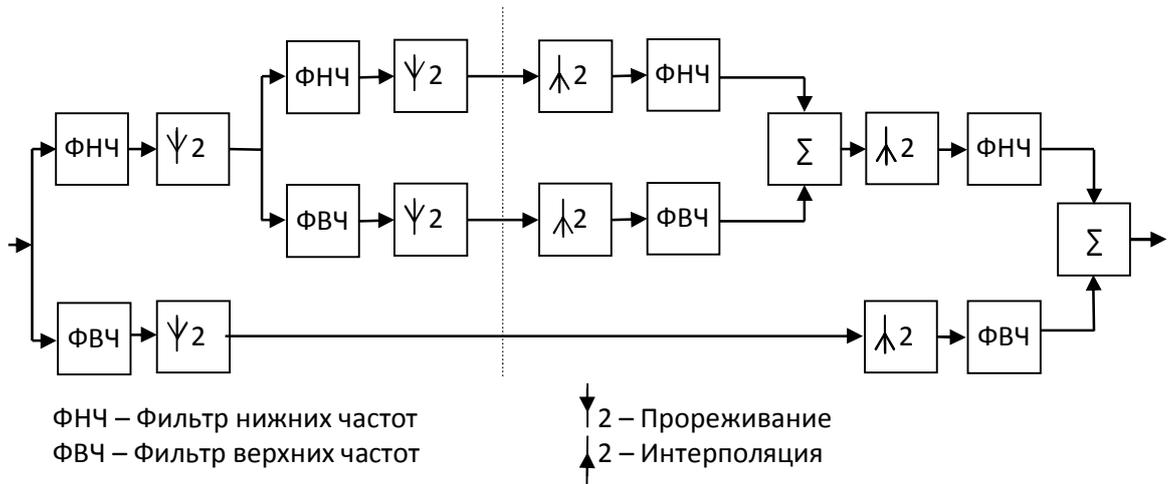


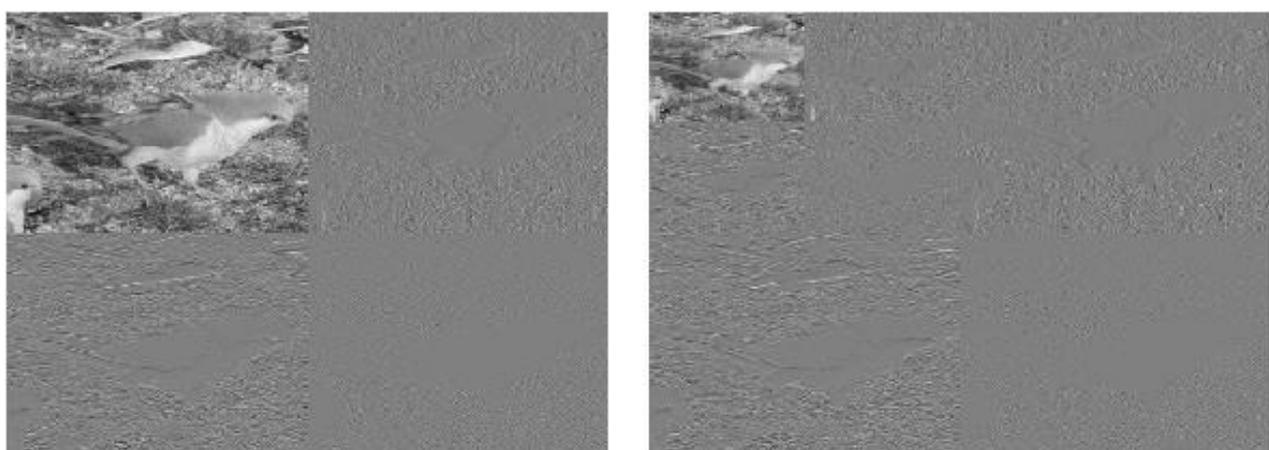
Рис.1.8. Обобщенная структурная схема вейвлет-преобразования



Рис.1.9. Принцип Вейвлет- декомпозиции



Рис.1.10. Исходное изображение и после вейвлет декомпозиции по строкам



а)

б)

Рис.1.11. Изображение после ВП по горизонтали и вертикали, (а), и после двойного преобразования – (б)



Рис.1.12. Сравнительное качество восстановленных изображений на основе ВП и ДКП при сжатии в 100 раз

В настоящее время для сжатия изображений широкое применение находят вейвлет-функции *Коэна-Добеши-Фово* (*CDF22*, *CDF24*, *CDF97*), *Вилласенора -V610*, *Койфмана-BCW3*, а также вейвлеты *TS2/6*, *MIT97*, используемые в стандартах JPEG-2000 и MPEG-4 [15].

Сжатие сигналов изображений на основе фракталов.

Для получения высоких коэффициентов сжатия порядка 200-2000 могут использоваться фрактальные методы сжатия изображений. Основой метода является рассмотрение естественных объектов как «подобных самим себе» и подчиняющихся требованиям фрактальной геометрии, в которой сложные структуры выглядят точно так же, как и простые структуры, т.е. повторяют их. Задачей кодирования является отыскание таких совпадений в цифровых изображениях и описание таких фракталов с дальнейшим эффективным повторением.

Понятия фракталов были предложены математиком *Б. Мандельбротом* в 1975 г. для обозначения нерегулярных, но самоподобных структур, для которых некоторые свойства реального изображения, сохраняются при масштабировании пространства. При фрактальном кодировании используется свойство подобия деталей разного масштаба, встречающиеся в реальных изображениях [16].

Фрактальная компрессия основана на том, что изображение представляется в более компактной форме с помощью коэффициентов системы итерационных функций (IFS). Система итерирующих функций (СИФ) – это совокупность сжимающих аффинных преобразований, которые представляют множительный механизм, многократно искажающий и перемещающий исходное изображение и включают в себя масштабирование, поворот и параллельный перенос. Например, при помощи СИФ из трех функций (рис.1.13) можно построить треугольник Серпинского.

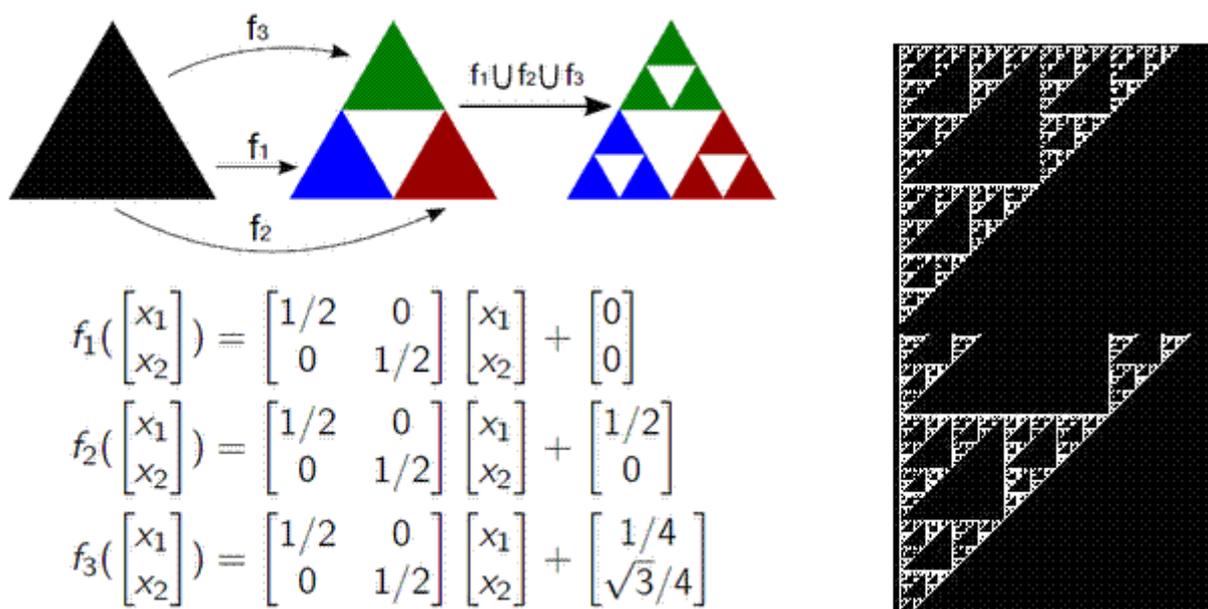


Рис.1.13. Применение СИФ для построения треугольника Серпинского

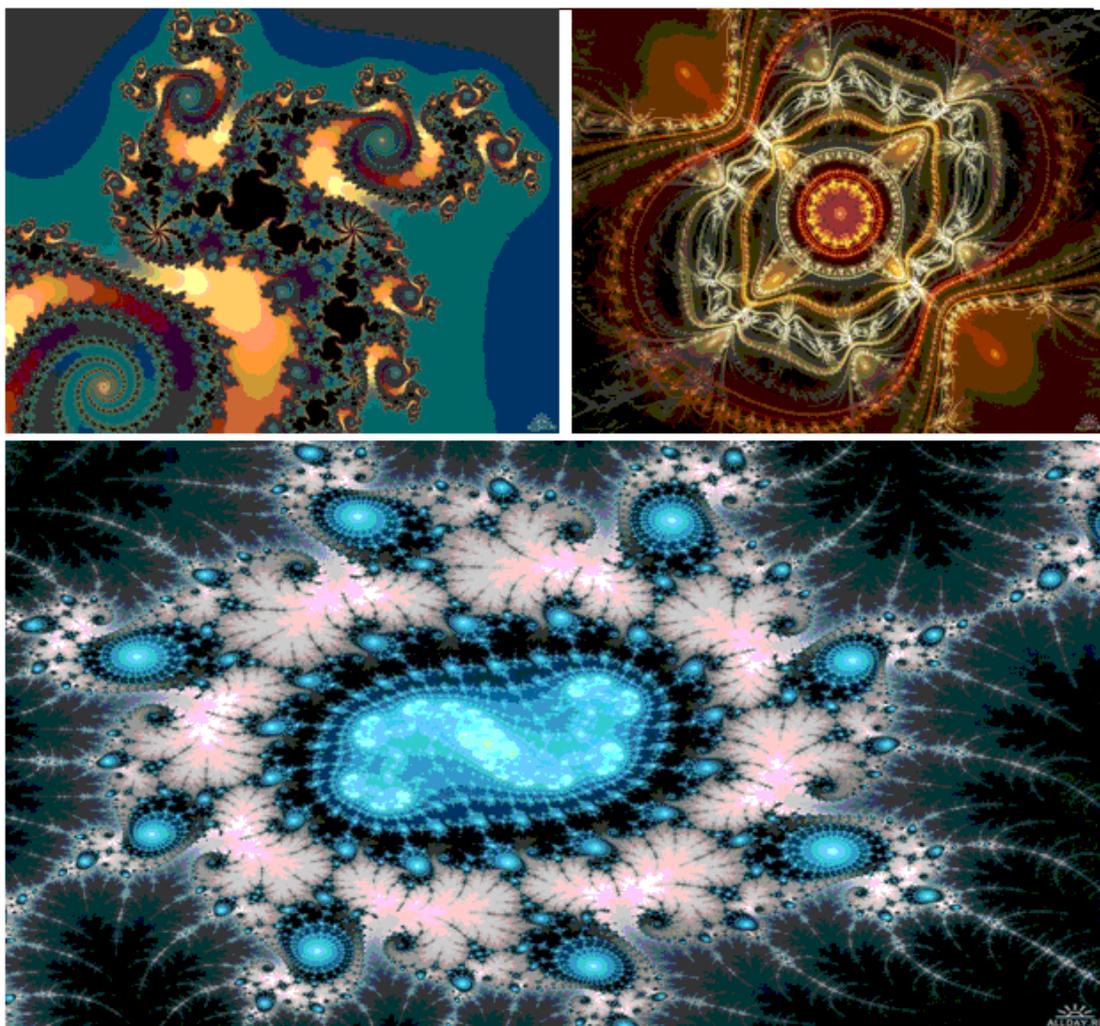


Рис.1.14. Применение фракталов для синтеза сложных изображений

Как видно из рисунка, исходный треугольник три раза множится, уменьшается и переносится. Если так продолжать до бесконечности, получится фрактал Серпинского с площадью 0 и размерностью 1,585. Причем вместо треугольника можно использовать квадрат, круг или другие геометрические примитивы. При этом, для кодирования объекта изображения достаточно сохранить функцию данного фрактала с указанием: в каких координатах, параметрами масштабирования и ориентации нужно синтезировать данный объект. Таким образом, довольно крупный объект можно представить в виде 10-20 байтов его параметров и тем самым существенно снизить объем передаваемой информации.

Фрактальное сжатие хорошо работает на изображениях технических чертежей, текста, карт местности и изображениях компьютерной графики (рис.1.14). При этом коэффициент сжатия таких изображений может достигать **200-2000 раз** [16]. В последнее время фракталы используются в редакторах фрактальной графики для синтеза сложных изображений (рис.1.14), позволяющих представить изображение небольшим набором параметров фрактальных объектов, находящихся в специальной библиотеке. Таким образом коэффициент сжатия видеоданных таких синтезированных изображений получается очень большой и составляет 1000 и более раз без потерь данных и ухудшения визуального качества.

На реальных ТВ изображениях фрактальное сжатие дает хорошие результаты, однако, на поиск фракталов с учетом масштабирования и взаимной ориентации требуются очень большие затраты времени (иногда до 1-5 минут на кадр, что для часовой передачи из 90 000 кадров потребует более 1,5 тысячи часов. Поэтому для телевизионных программ фрактальные методы сжатия пока не применяются, хотя активно ведутся работы по разработке более быстродействующих алгоритмов.

Сжатие сигналов ТВ изображений на основе устранения межкадровой избыточности.

Как было указано выше, в видеопотоке информация в смежных кадрах одного видеосюжета меняется слабо, поэтому, если передавать только изменения изображений относительно некоего опорного или промежуточного кадра, например, в виде межкадровой разности (рис.1.15), то объем информации существенно снижается и можно получить довольно большие коэффициенты сжатия видеопотока.

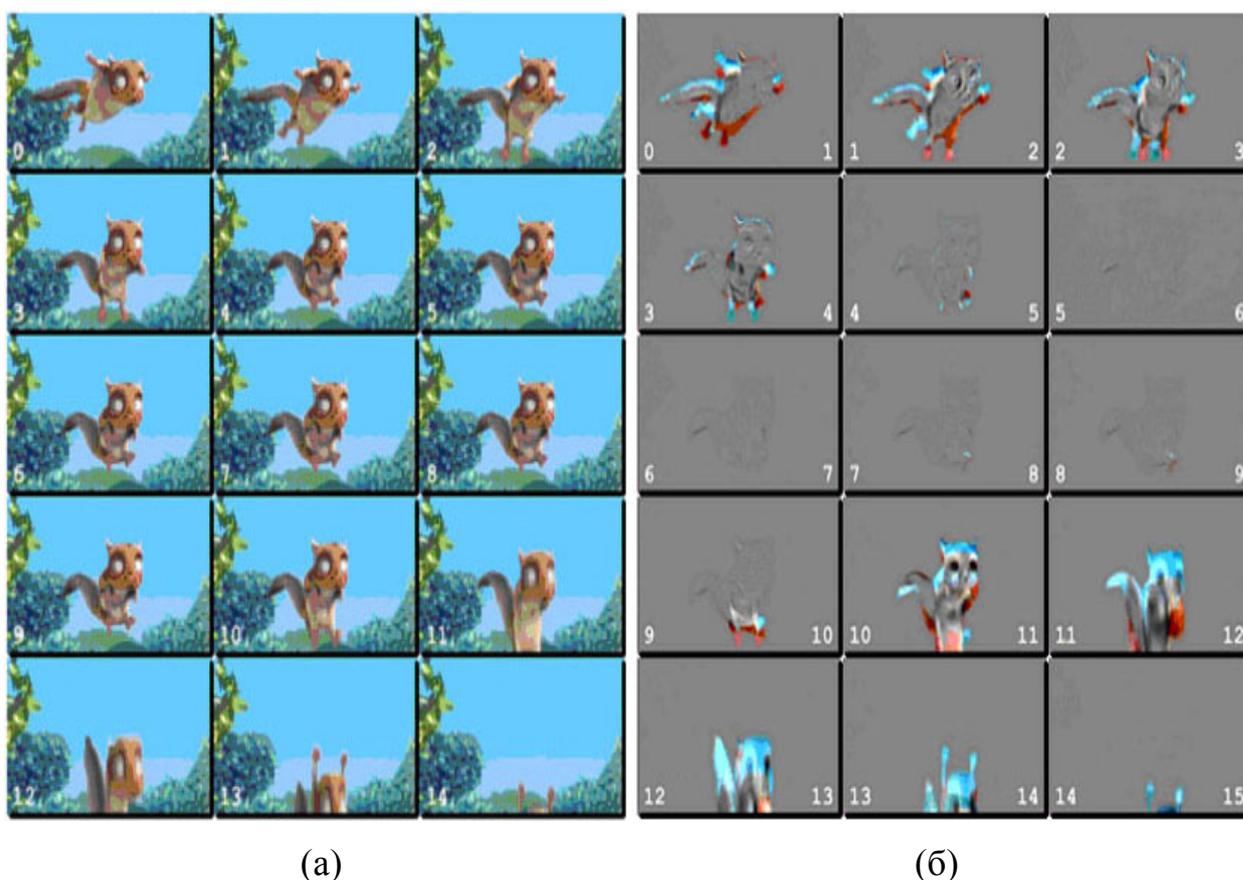


Рис.1.15. Изображения 15 кадров видеопоследовательности (а) и их межкадровая разница (б)

Как видно из приведенного рисунка, только в 5 кадрах межкадровой разницы (4-5, 5-6, 6-7, 7-8) ее информативность близка к 0, а на остальных кадрах наблюдается довольно сильные межкадровые различия, что приводит к резкому увеличению межкадровой информации. В связи с этим простая межкадровая разница обычно не применяется, так как даже при небольшом

изменении ракурса съемки или перемещения объектов передаваемой сцены, вычитаются значения нетождественных пикселей. Это приводит к увеличению объема межкадровой информации в виде появления множества мелких деталей, которые сжимаются плохо, в результате чего выигрыш от ее применения получается небольшой и обычно не превышает 2-3 раза. Поэтому на практике применяют более сложную межкадровую обработку на основе компенсации движения фрагментов изображений на смежных кадрах. Суть метода заключается в поиске фрагментов изображения первого кадра в зонах их предполагаемого смещения в следующем кадре. Если такие фрагменты находятся, то вместо них передаются их новые координаты – вектора перемещения, тем самым, например 255 байтов блока размером 16x16 пикселей можно заменить 1-2 байтами его новых координат. На рис.1.16,а представлены изображения смежных кадров видеосюжета и результат их компенсации движения (рис.1.16,б), где черные области на скомпенсированном изображении показывают полностью идентичные блоки, вместо которых передаются только значения их новых координат (2-4 байта на блок). А светлые фрагменты изображения показывают блоки, для которых результат их межкадровой разницы превышает допустимую погрешность позиционирования и такие блоки далее подвергаются сжатию с помощью ДКП. В этом случае структура видеопотока состоит из опорного кадра, где устраняется только внутрикадровая избыточность и одного или нескольких типов кадров передающих межкадровые различия и векторы перемещений блоков, что реализуется в стандартах сжатия семейства MPEG и многих других кодеках.

При этом компенсация движения посредством нахождения векторов смещения состоит из следующих действий:

- 1) Загружается опорный (ключевой) кадр, а последующий кадр будет скомпенсирован относительно опорного кадра.

2) Производится разделение кадра на прямоугольные блоки (макроблоки), обычно размером 16x16 пикселей (рис.1.17,а), хотя их размер может быть произвольным.

3) Для пикселей первого макроблока ищется его максимальное подобие в некоторой окрестности его возможного смещения в следующем кадре.

Если такое соответствие находится, то в массив метаданных заносятся данные о номере блока и его новых координатах, а сам макроблок не передается. Если такое соответствие не находится, то содержимое блока подвергается обработке с помощью ДКП.



а)



б)

Рис.1.16. Изображения смежных кадров (а) и результат их компенсации движения (б)

Далее весь процесс повторяется над всеми макроблоками первого кадра и так далее.

Наглядная иллюстрация поиска похожих блоков в предыдущем и последующем кадре показана рис.1.17,б.

Рассмотрим более подробно наиболее распространенные методы компенсации движения фрагментов телевизионного изображения.

К наиболее эффективным методам межкадровой обработки ТВ изображений относятся методы компенсации движения, которые можно разбить на классы по ряду признаков и свойств, но наибольший интерес представляет классификация по способу работы (или архитектуры) и по назначению (области применения алгоритма) [35].

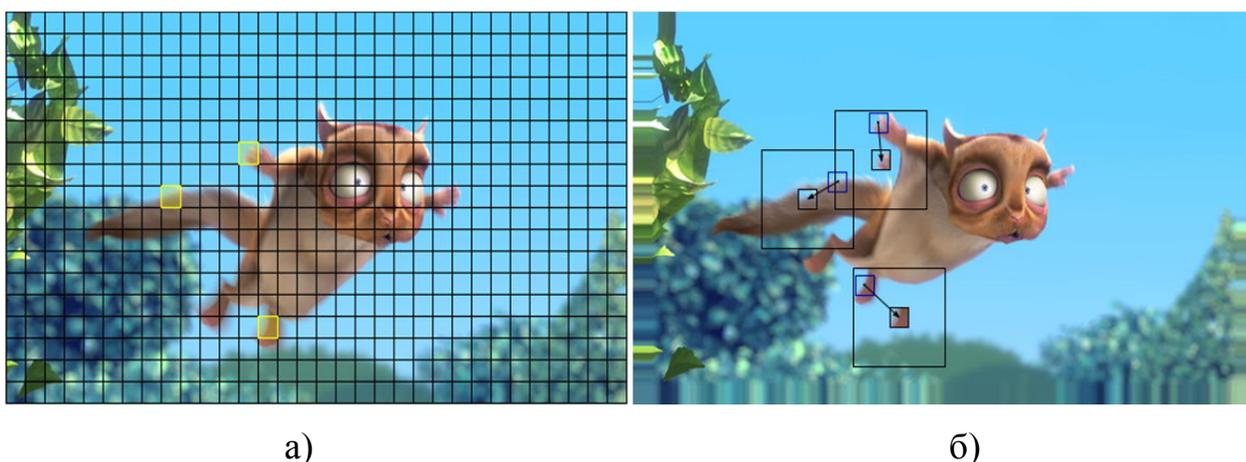


Рис.1.17. Разделение кадра на макроблоки (а) и поиск похожих блоков в соседних кадрах

Классификация по способу работы алгоритма учитывает следующие его архитектурные особенности [35]:

- Компенсация движения может производиться над объектами, блоками либо фиксированного размера, либо с подразбиением, а также над кадром целиком.
- Движение чаще всего ищется в классе параллельных сдвигов с ограниченным максимальным сдвигом, хотя он может быть расширен операциями поворота и масштабирования.

Для определения позиции переместившегося блока используется несколько мер, минимальное значение которых и определяют вероятное новое положение переместившегося фрагмента изображения. К этим мерам относятся [35]:

1. Сумма абсолютных разностей (Sum of Absolute Differences, SAD)

$$SAD = \sum_{p \in Obj} |F_{Orig}(p) - F_{Comp}(p)| \quad (1.3)$$

2. Сумма квадратов разностей (Sum of Squared Differences, SSD):

$$SSD = \sum_{p \in Obj} [F_{Orig}(p) - F_{Comp}(p)]^2 \quad (1.4)$$

где суммирование производится по всем точкам объекта компенсации (например, прямоугольного блока), а F_{Orig} и F_{Comp} - яркости исходного и скомпенсированного кадра, соответственно, в точке с координатами $p=(x,y)$.

При классификации по назначению обычно выделяют две большие группы алгоритмов компенсации движения [35]:

1. Алгоритмы, использующиеся при сжатии видео;
2. Алгоритмы, использующиеся при обработке видео (деинтерлейсинге, изменении частоты кадров).

Разница между этими алгоритмами в том, что алгоритмы первой группы ориентированы на уменьшение попиксельной разницы исходного и скомпенсированного кадров, так как именно от него зависит степень сжатия видео. При этом для них, как правило, не важно, правильно ли определено движение, или просто найдены близкие по яркости области на двух последовательных кадрах. Для алгоритмов же второй группы истинность найденных параметров движения очень важна, поскольку на основе этих параметров производится настройка других параметров алгоритма.

Рассмотрим более подробно наиболее распространенные методы компенсации движения фрагментов телевизионного изображения.

Метод сопоставления блоков. К наиболее эффективным методам межкадровой обработки ТВ изображений относятся методы компенсации движения, которые можно разбить на классы по ряду признаков и свойств, но наибольший интерес представляет классификация по способу работы (или архитектуры) и по назначению (области применения алгоритма).

Классификация по способу работы алгоритма учитывает следующие его архитектурные особенности:

- Компенсация движения может производиться над объектами, блоками либо фиксированного размера, либо с подразбиением, а также над кадром целиком.
- Движение чаще всего ищется в классе параллельных сдвигов с ограниченным максимальным сдвигом, хотя он может быть расширен операциями поворота и масштабирования.

Для определения позиции переместившегося блока используются минимальные значения суммы абсолютных разностей или суммы квадратов разностей значений пикселей в блоках.

В настоящее время в телевидении наибольшее применение находят методы компенсации движения на основе сопоставления блоков, которые используют разбиение изображения на прямоугольные блоки размером 8×8 или 16×16 пикселей. Движение ищется в классе линейных смещений, поэтому описывается такое движение двумерным вектором смещения каждого блока по координатам X и Y .

Для работы данного метода необходимо, чтобы в межкадровом интервале положение объектов в сцене мало изменялось. На практике данное условие выполняется в большинстве обычных видеопоследовательностей, за исключением участков с резкой сменой сцены, поэтому характер движения объектов можно считать почти всюду непрерывным.

Принцип работы данного метода следующий (рис. 1.18) [35]:

1. Текущий кадр разбивается на непересекающиеся блоки одного размера $B(x,y)$.

2. Для каждого блока $B(x,y)$ в небольшой окрестности предполагаемого смещения ищется наиболее «похожий» на него блок $B_{prev}(x+u, y+v)$ в предыдущем кадре. «Похожесть» определяется выбранной метрикой, обычно SAD или SSD.

3. Новые координаты блока, при котором достигается минимум выбранной ошибки позиционирования, считается вектором смещения для данного блока $d=(u,v)^T$.



Рис.1.18. Схема работы алгоритма сопоставления блоков

В качестве функции ошибки компенсации чаще всего используется мера SAD для скомпенсированного блока:

$$SAD(\vec{d}, n) = \sum_{p \in B(x,y)} |F(p, n) - F(p + \vec{d}, n - 1)| \quad (1.5)$$

так как её вычисление проще реализуется на некоторых архитектурах процессоров.

Данный метод хорошо сочетается с размерами сигнальных матриц ДКП и получил широкое распространение в потоковых стандартах сжатия MPEG-1, MPEG-2, MPEG-4-10, H-264. Однако, полный перебор возможных позиций блоков для их точного позиционирования занимает много вычислительных ресурсов и времени. Поэтому для ускорения работы алгоритма часто применяют поиск не по всему полю, а по определенным шаблонам.

Методы, основанные на шаблонах, обеспечивают неплохую скорость работы, однако, часто не обеспечивают нужной точности позиционирования блоков, находя локальный минимум вместо истинного.

Метод параметрических моделей. Данный метод хорошо работает в тех случаях, когда все изменение сюжета в кадре вызвано только движением передающей камеры. Тогда это движение можно описать моделью с четырьмя параметрами [17]:

$$\vec{d}(x, y, n) = \begin{pmatrix} p_1(n) + p_3(n) \cdot x \\ p_2(n) + p_4(n) \cdot y \end{pmatrix} \quad (1.6)$$

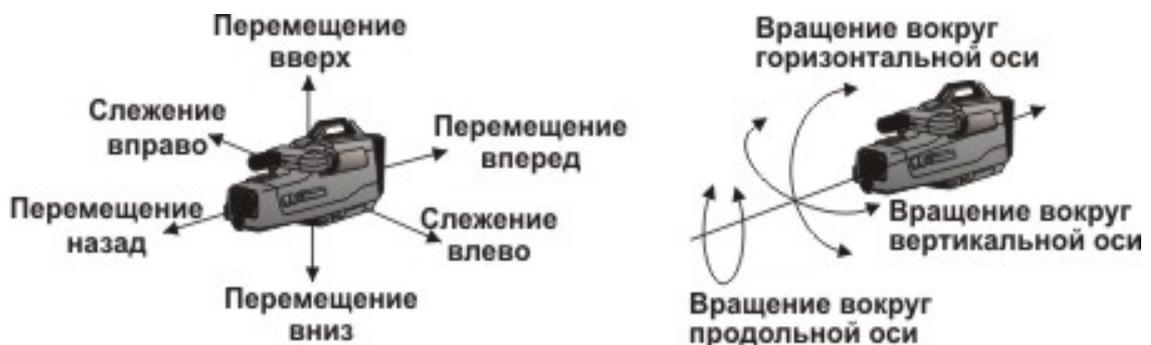


Рис. 1.19. Вариации движения видео-камеры во время съемок видео

Параметры $p_1(n)$ и $p_2(n)$ описывают параллельное перемещение камеры, в то время, как $p_3(n)$ и $p_4(n)$ описывают ее вращение и приближение или удаление. Можно легко обобщить этот метод на общий случай, когда в кадре есть множество независимо движущихся объектов: после разбиения кадра на блоки данная модель может быть применена для каждого из блоков кадра. Для повышения эффективности данного метода, его следует использовать совместно с каким-либо блочным методом.

Теоретически параметрическую модель для описания движения области изображения можно применять, если:

- 1) Области в пространстве соответствует плоский участок поверхности.
- 2) Область «покрывает» разные и находящиеся на разном расстоянии от камеры объекты, но съемки ведутся не смещающейся, относительно этих объектов, камерой.

В иных случаях изображение объектов «движется в кадре» более сложным образом, чем это допускает любая параметрическая модель, поэтому в таких случаях целесообразно применять несколько методов компенсации движения сразу.

Метод объектного подхода. При сжатии ТВ изображений часто возникают ситуации, когда намного удобнее обрабатывать видео объекты кадра целиком, а не разбивать изображение на блоки. Данная методика выделения объектов носит название – сегментация. Объектно-ориентированное представление ТВ изображений входит в область применения алгоритмов сегментации. При сегментации выделяется на изображении области с примерно одинаковыми свойствами внутри – текстурами, и посылается по каналам лишь их описание. Таким образом, степень компрессии возрастет. Подобная методика позволяет увеличить эффективность сжатия изображения без снижения его качества.

Наиболее часто объектно-ориентированный подход используется для более точной компенсации движения, что находит применение в стандарте MPEG-4. При этом, изображение разделяется на блоки не фиксированного размера, а на объекты произвольной формы, выделенные из сюжетов самого изображения, при помощи методов сегментации, которые применяются наряду с другими методами (метод параметрических моделей, метод сопоставления блоков и другие).

Процесс сегментации может происходить независимо от процесса поиска параметров движения, либо и то, и другое может определяться в рамках единого процесса, повторяющегося итеративно. В первом случае основанием для сегментации служит обычно яркостная информация, во втором случае сегментация производится с учетом найденных параметров движения, которые затем уточняются [18]. Иногда сегментация кадра на объекты применяется после определения векторов смещения для отдельных блоков с целью коррекции найденного векторного поля [19].

На рис.1.20 показана обобщенная схема объектно-ориентированного подхода к многокадровой компенсации движения. Работа начинается с загрузки в буфер памяти текущего кадра и векторов движения, найденных кодеком при помощи алгоритма предсказания векторов движения. Если текущий кадр является первым кадром видеопотока, то векторы движения отсутствуют. На основе этих данных происходит дополнение информации при передвижении сегментированных объектов.



Рис.1.20. Обобщенная структурная схема процесса объектно-ориентированной компенсации движения

При этом, сегментация объекта производится, как по яркостным или цветовым признакам, так и по принципу однородности блоков, которые образуют объект. Для поиска или селекции объектов в кадре разработано довольно много методов, отличающихся сложностью и быстродействием используемых алгоритмов, точностью выделения границ контуров и т.д.

В структуре цифрового потока используются опорные кадры, в которых устраняется только внутрикадровая статистическая избыточность на основе

методов спектральных преобразований и промежуточных кадров в которых содержится информация о типе передаваемого кадра, найденных динамических объектов сцены, их координаты размещения в кадре и другая служебная информация.

Следует отметить, что все методы компенсации движения фрагментов сцены, использующие блочную структуру изображений, формируют дополнительный массив метаданных, который присоединяется к сжатым видеоданным опорного кадра и межкадровых различий. Причем, чем точнее осуществляется компенсация движения, тем больше становится объем метаданных, снижая результирующий коэффициент сжатия видеопотока. Поэтому, применяемые различные методы компенсации движения должны минимизировать формируемый объем метаданных, иначе их возрастающий объем может свести к нулю достоинства самих методов компенсации движения.

4. Анализ методов сжатия аудиоданных ТВ программ

Аудио - общий термин, относящийся к звуковым технологиям. Как правило, под термином аудио понимают звук, записанный на звуковом носителе, а также запись и воспроизведение звука, звукозаписывающая и звуковоспроизводящая аппаратура.

У звуковых сигналов корреляционная связь между отсчетами довольно маленькая и поэтому такие сигналы сжимаются слабо, обычно не более 2-4 раза. Это связано с тем, что вещательный звуковой сигнал представляет большую сумму сигналов совершенно различных источников не связанных между собой (различных инструментов оркестра, голоса людей, шум улицы, ветра и.т.д.). Да и, кроме того, динамический диапазон звукового сигнала значительно более широкий, чем у видеосигнала, поэтому одинаковые значения отсчетов в звуковом сигнале практически не встречаются. Поэтому основное сжатие вещательных звуковых сигналов осуществляется на основе

их психоакустической обработки, учитывающей особенности нашего слухового восприятия.

Важной проблемой при цифровом представлении звуковых сигналов является сокращение имеющейся в них статистической и психоакустической избыточности. Сокращение статистической избыточности основывается на учете свойств самих звуковых сигналов. Она обусловлена наличием корреляционной связи между соседними отсчетами звукового сигнала при его дискретизации.

Устранение статистической избыточности не приводит к значительному уменьшению скорости цифрового потока. Методами устранения психофизической избыточности можно обеспечить сжатие цифровых аудиоданных в 10-12 раз без существенных потерь в качестве.

Особенности психоакустической обработки ЗС.

Работа психоакустических моделей основывается на свойствах слухового аппарата человека, представляющего большой набор узкополосных резонаторов, которые могут маскировать определенные компоненты ЗС по уровню, частоте и времени [20].

Маскировка по уровню основывается на неравномерности распределения границы слышимости звука по частотам, представленной на рис.1.20. Наилучшим образом мы слышим частоты в районе 2-4 кГц, в котором примерно находится речевой диапазон, а к низким и высоким частотам чувствительность уха снижается [20].

Таким образом, чем дальше частота слышимого звука от 2-4 кГц, тем выше граница слышимого звука, тем больше информации можно вырезать без заметных потерь в качестве.

Частотная маскировка основана на том, что каждый резонатор слуховой системы человека имеет определенную полосу пропускания. Так, на частоте 1000 Гц порог различимости составляет 30 Гц. Поэтому если в ЗС присутствуют частоты 1000 и 1020 Гц, то они ухом не различимы и поэтому вместо тона 1020 Гц достаточно продлить тон частоты 1000 Гц и при этом

ничего не потерять [20].

Любой слышимый тон изменяет восприятие остальной звуковой картины. При воспроизведении, какого либо тона граница слышимости соседних с ним по частотам звуков изменяется. В этом случае воспроизводимый тон называется маскирующим, а граница слышимости окружающих его тонов поднимается тем выше, чем ближе их частота к частоте маскирующего сигнала. То есть, слушая низкий звук, значительно проще услышать высокий звук, нежели низкий другой частоты. Под воздействием тона частотой 1 кГц и интенсивностью 60 дБ измененная граница слышимости будет выглядеть примерно, как на рис.1.22.

Обычно в музыке одновременно присутствуют самые разнотоновые компоненты. Таким образом, маскирующих тонов может быть несколько. При использовании сразу нескольких маскирующих тонов (частотой 0,25, 1, 4, 8 кГц, как показано на рис.1.23.) граница слышимости остальных сигналов сильно поднимается.

На рис.1.24 видно, что наилучшим образом маскируются высокие частоты. Уже при воспроизведении 8-килогерцового маскирующего тона граница слышимости на 14 килогерцах поднимается. Алгоритмы компрессии этим активно пользуются — при сжатии качество высоких частот обычно ухудшается в первую очередь, что особенно заметно на низких битрейтах.

Временная маскировка основывается на том, что очень короткие звуковые сигналы человеческое ухо не замечает. Кроме того, после резкого прекращения действия маскирующего тона, в течение короткого времени (около сотни миллисекунд, в зависимости от частоты и амплитуды сигнала) граница слышимости уха изменяется, причем нелинейно. На рис.1.24 представлен график, иллюстрирующий временную маскировку [34].

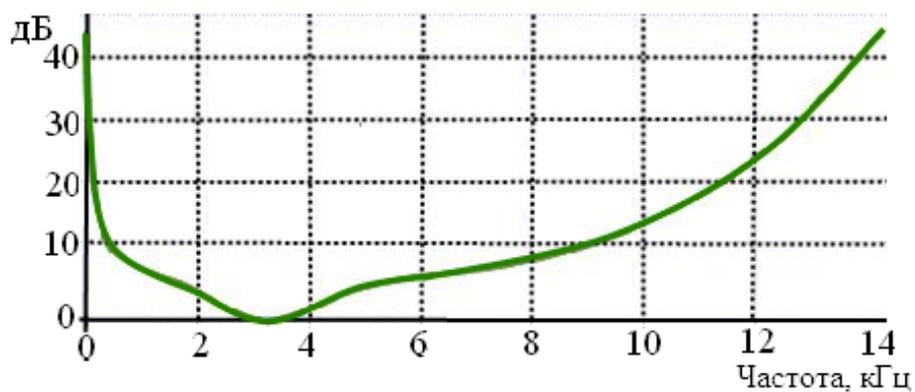


Рис.1.21. Граница слышимости в тишине

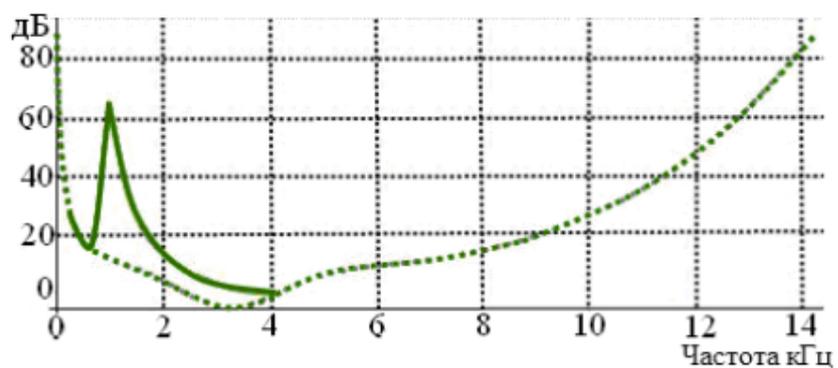


Рис.1.22. Граница слышимости под воздействием тона частотой 1 кГц и интенсивностью 60 дБ

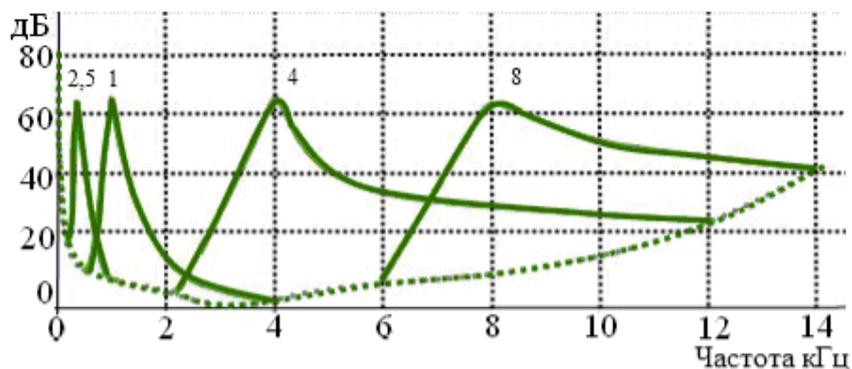


Рис.1.23. Граница слышимости под воздействием сразу нескольких маскирующих тонов (частотой 0,25, 1, 4, 8 кГц)

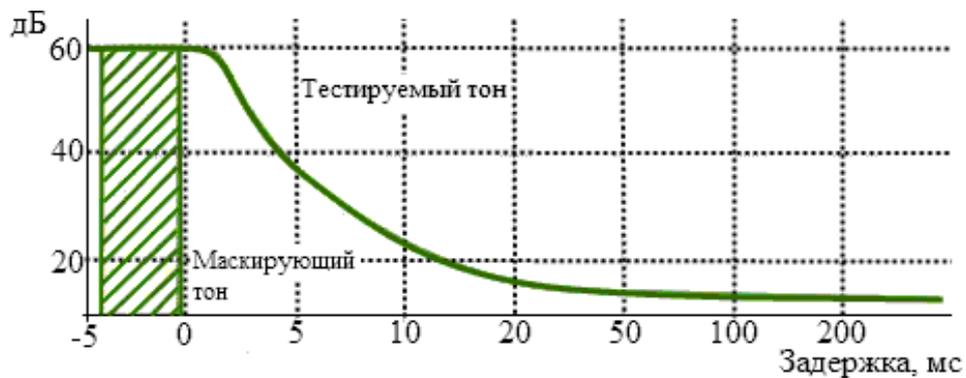


Рис.1.24. Временная маскировка

Как видно, объем, заключенный под поверхностью слышимых звуков, составляет немалую часть от общего количества звуков. Удаление замаскированных тонов дает самое значимое сжатие, однако оно же и приводит к наиболее заметной потере качества.

Методы сжатия в стандарте MP3.

MP3 (более точно, англ. MPEG-1/2/2.5 Layer 3 (но не MPEG-3) — третий формат кодирования звуковой дорожки MPEG) — лицензируемый формат файла для хранения аудио-информации [33].

На данный момент MP3 является самым известным и популярным из распространённых форматов цифрового кодирования звуковой информации с потерями. Он широко используется в файлообменных сетях для оценочной передачи музыкальных произведений. Формат может проигрываться практически в любой популярной операционной системе, на практически любом портативном аудио-плеере, а также поддерживается всеми современными моделями музыкальных центров и DVD-плееров.

В формате MP3 используется алгоритм сжатия с потерями, разработанный для существенного уменьшения размера данных, необходимых для воспроизведения записи и обеспечения качества воспроизведения очень близкого к оригинальному. При этом при среднем битрейте 128 кбит/с обеспечивается сжатие около 10 раз. Принцип сжатия заключается в снижении точности некоторых частей звукового потока, что практически неразличимо для слуха большинства людей. Данный метод называют кодированием восприятия. При этом на первом этапе строится диаграмма звука в виде последовательности коротких промежутков времени, затем на ней удаляется информация не различимая человеческим ухом, а оставшаяся информация сохраняется в компактном виде. Данный подход похож на метод сжатия, используемый при сжатии картинок в формат JPEG.

Рассмотрим работу кодека по обобщенной структурной схеме, представленной на рис.1.25.

целого) и затем кодируются. Масштабирование в каждом субполосном канале представляет собой выбор из 18 отсчетов ЗС максимального значения, которое называется масштабным коэффициентом SCF (Scale Factor) выборки субполосного ЗС. В кодеке имеем 32 масштабных коэффициента, причем каждый блок из 36 отсчетов ЗС делится на три подблока, называемые гранулами. В каждой грануле, включающей 18 отсчетов ЗС, также определяется максимальный отсчет, его значение является масштабным коэффициентом SCF гранулы. Всего субполос 32, поэтому для Layer-2 общее количество SCF равно $3 \times 32 = 96$. Передаваемые декодеру значения SCF заданы в стандарте таблицей. Поэтому максимальное значение отсчета в грануле сравнивается с набором табличных значений SCF. Из множества последних выбирается ближайшее большее. Оно и принимается за SCF гранулы.

Квантование и кодирование контролирует модифицированная психоакустическая модель. Для осуществления этого контроля квантования используются оценки границ (пороговых значений) маскирования. Распределение бит в субполосах рассчитывается на основе соотношений сигнал/маска всех субполос, при котором эта маска (шум) ещё маскируется полезным сигналом. Максимальный уровень сигнала и минимальная граница маскирования выводятся из быстрого преобразования Фурье (FFT transform) дискретизированного входного сигнала.

Для каждой скорости передачи (или группы скоростей) существует своя таблица. Значения f_d и скорость передачи цифровых данных ν определяют также и максимальный номер субполосы n , отсчеты ЗС которой ещё кодируются и передаются на приёмную сторону системы.

В результате, после многочисленных вычислений на выходе блока «Форматирование цифровых потоков и помехоустойчивое кодирование» формируется звуковой кадр. Далее кодирования коэффициентов для обеспечения сжатия аудио потока обеспечивается кодером длинных серий (RLE) в котором повторяющиеся кодовые комбинации заменяются числом их

повторений и дожимаются энтропийным кодером на основе таблиц Хаффмана. Энтропийное кодирование, учитывает статистические особенности ЗС, и наиболее часто повторяющиеся кодовые комбинации передаются короткими кодовыми комбинациями, а редко встречающимися – длинными. Этот способ кодирования на 20-25% повышает эффективность сжатия кодека.

Методы сжатия звуковых сигналов в стандарте MPEG-2AAC.

Алгоритм AAC (Advanced Audio Coding - передовое кодирование звука) (рис.1.26) разработан в рамках стандарта MPEG-2 ISO/IEC 13818-7 и позже включен в стандарт MPEG-4. Он базируется на учете опыта, накопленного при разработке алгоритма компрессии MPEG-1/2/2.5 Layer 3 (третий формат кодирования звуковой дорожки MPEG) стандартов ISO/IEC11172-3 и 13818-3, который поддерживает все известные звуковые форматы: моно (1/0), обычное стерео (2/0), разновидности систем Dolby (Dolby Stereo 3/1; Dolby Surround 3/2; Dolby Pro Logic I и II, 3/2) и пятиканальный Surround-формат 5.1.

Полное название – MPEG-2 AAC (Advanced Audio Coding – передовое кодирование звука) был разработан в стенах института Fraunhofer, при активном участии компаний AT&T, Sony, NEC и Dolby в начале 1998 года [20]. Этот формат изначально позиционировался разработчиками как преемник MP3, так как обладал по сравнению с последним рядом несомненных достоинств. Как и в MP3 в основе алгоритма AAC лежит психоакустическая модель кодирования, то есть при сжатии какая-то часть звукового спектра удаляется. При этом алгоритм AAC содержит большое количество усовершенствований, направленных именно на улучшение качества выходного аудиосигнала. В MPEG-2 AAC используются другие алгоритмы преобразований, улучшенные обработчики шумов и новый банк фильтров. Из специальных возможностей можно назвать, так называемые "водяные знаки" (watermarks) - информацию об авторских правах, которую AAC позволяет хранить в теле аудиокomпозиции, причем удалить эту

информацию не разрушив целостность аудиоданных невозможно. При этом MPEG-2 AAC обладает высоким качеством звучания и хорошей степенью компрессии аудиоконпозиций. Так, например, аудиоконпозиция в формате AAC с bitrate 96 kbs обеспечивает качество звучания, аналогичное потоку MPEG-1 Layer III bitrate 128 kbs. При сравнении же файлов AAC с bitrate 128 kbs, качество звучания ощутимо превосходит MPEG-1 Layer III с такой же степенью сжатия [34]. Однако, после окончания работ над MPEG-2 AAC некоторые из компаний соучредителей забрали причитающиеся им исходные коды стандарта и на их базе создали собственные форматы, коммерческие и не совместимые друг с другом. Таким образом и появились несовместимые форматы, которые называют "семейством форматов аудиосжатия AAC". На данный момент существуют четыре разновидности формата AAC [20]. Homeboy AAC, AT&T a2b AAC, Liquifier PRO AAC, Astrid/Quartex AAC. Все эти модификации несовместимы между собой, имеют собственные кодеры/декодеры и неодинаковы по качеству. Так, последние две модификации по ряду параметров превосходят первую пару. В целях ознакомления и сравнения этих модификаций между собой рассмотрим все представленные форматы этого семейства. Утверждается, что AAC, стандартизованная ISO как часть спецификации MPEG-2, обеспечивает более высокое качество звука, нежели MP3 (MPEG1 Layer-3), а при сопоставимом качестве требует на 30% меньше дискового пространства или величины сжатого потока. AAC-поток способен нести 48 каналов аудио при изначальной оцифровке до 96 кГц, 15 встроенных потоков данных и "говорить" на разных языках. AAC выбран для цифрового радиовещания в АМ-диапазоне (<30 МГц). Высококачественное радиовещание достигается при 320 кбит/с для 5+1-канальной программы. Более того, AAC (с некоторыми модификациями) является единственной высококачественной аудиотехнологией, используемой в рамках стандарта MPEG-4 - глобального мультимедийного формата будущего. Среди создателей AAC - AT&T, Dolby Laboratories, Fraunhofer IIS, Lucent Technologies, Sony Corporation и другие. Так же как в

MP3, сжатие в AAC основано на психоакустических особенностях восприятия звука. Вместе с тем между ними имеются существенные различия: Вместо гибридного (каскадного) банка фильтров AAC использует модифицированное дискретное косинусное преобразование в частотную область (MDCT) с импульсным откликом 5,3 мс (18,6 мс для MP3) при дискретизации 48 кГц, что уменьшает артефакты сжатия в момент взрывного нарастания амплитуды. Наряду с увеличением разрешения по частоте (1024 линии вместо 576 для MP3) MDCT повышает эффективность кодирования. Для улучшения кодирования чистых тонов применяется обратное адаптивное предсказание. Информация, которая должна быть донесена до слушателя, подвергается энтропийному кодированию, чтобы устранить избыточность, насколько это возможно. Минимизация переходных шумов (Temporal Noise Shaping - TNS) сглаживает распределение шума квантизации во времени посредством предсказания в частотной области. В частности, благодаря TNS улучшается качество воспроизведения голосовых сигналов, особенно - на низких потоках. Как MP3, так и AAC имеют критическое значение битрейта, выше которого качество звука улучшается очень медленно (по мере дальнейшего увеличения потока), зато ниже - ухудшается очень быстро. Для MP3 критический битрейт составляет около 1,33 бита на дискрету (128 кбит/с для стереосигнала 48 кГц), а для AAC - 1 бит на дискрету (96 кбит/с для стерео 48 кГц), т.е. данное значение показывает, какое количество раз в единицу времени АЦП производит замер той или иной характеристики аналогового сигнала. В плане критического битрейта многоканальное кодирование еще более эффективно. Технология AAC позаимствовала у AC3 возможность кодирования нескольких параллельных каналов, а у MP3 - принципы психоакустической модели [20]. Кроме того, разработчики AAC внесли параметр настройки уровня "громкости" для каждого блока данных. Этот параметр обеспечивает динамический диапазон AAC, эквивалентный 24 битам. Оригинальные нововведения AAC привели к увеличению

продолжительности кодирования чуть ли не вдвое, но скорость декодирования практически не изменилась.

Кроме того существует упрощенная версия AAC (MPEG-2 AAC low complexity profile) в которой нет предсказания, а TNS ограничено двенадцатью коэффициентами, но с рабочей полосой до 18 кГц.

Таким образом сочетая в себе алгоритмы кодирования MPEG2-AAC имеет более широкие, чем MP3, возможности, например, возможность кодирования нескольких звуковых каналов с частотой дискретизации до 96кГц. Более высокое, чем у MP3, соотношение “качество/размер” делает его весьма удобным как для создания музыкальной коллекции, так и для кодирования многоканальных звуковых дорожек. Качество звучания файлов, сжатых при помощи AAC, оценивается как отличное. Тем не менее в использовании AAC есть свои трудности: алгоритмы кодирования, используемые в данном формате, достаточно сложны, поэтому для создания AAC-файла требуется значительное количество времени и системных ресурсов [34].

Алгоритм MPEG-2 ISO/IEC 13818-7 AAC обеспечивает наиболее высокое качество кодирования звукового сигнала (рис.1.26.), например, при формате 5.1. Искажения, вызванные компрессией, лежат ниже порогов их слуховой заметности уже при суммарной скорости цифрового потока 320...384 кбит/с. А при скорости 64 кбит/с. возможно использовать 48 основных каналов, 16 низкочастотных каналов для звуковых эффектов, 16 многоязыковых каналов и 16 каналов данных. До 16 программ может быть, описано используя любое количество элементов звуковых и других данных.

Во внутреннем цикле выполняется квантование и кодирование коэффициентов МДКП, во внешнем – оценка фактических величин искажений квантования в каждой субполосе кодирования и необходимая коррекция искажений квантования. При этом, если в какой-либо одной или нескольких субполос кодирования фактическое значение искажений квантования превосходит допустимое значение, рассчитанное

психоакустической моделью, то выполняется процедура коррекции коэффициентов МДКП, которая заключается в следующем: вначале включается режим предыскажений, затем, если это не помогает, коэффициенты МДКП в той субполосе кодирования, где не выполняются требования психоакустической модели, умножаются на корректирующий множитель. Алгоритм имеет несколько их значений, перебираемых последовательно начиная с наименьшего значения. После каждого цикла коррекции вновь выполняется вся процедура квантования и кодирования коэффициентов МДКП во внутреннем цикле, и так продолжается до тех пор, пока требования психоакустической модели не будут выполнены.

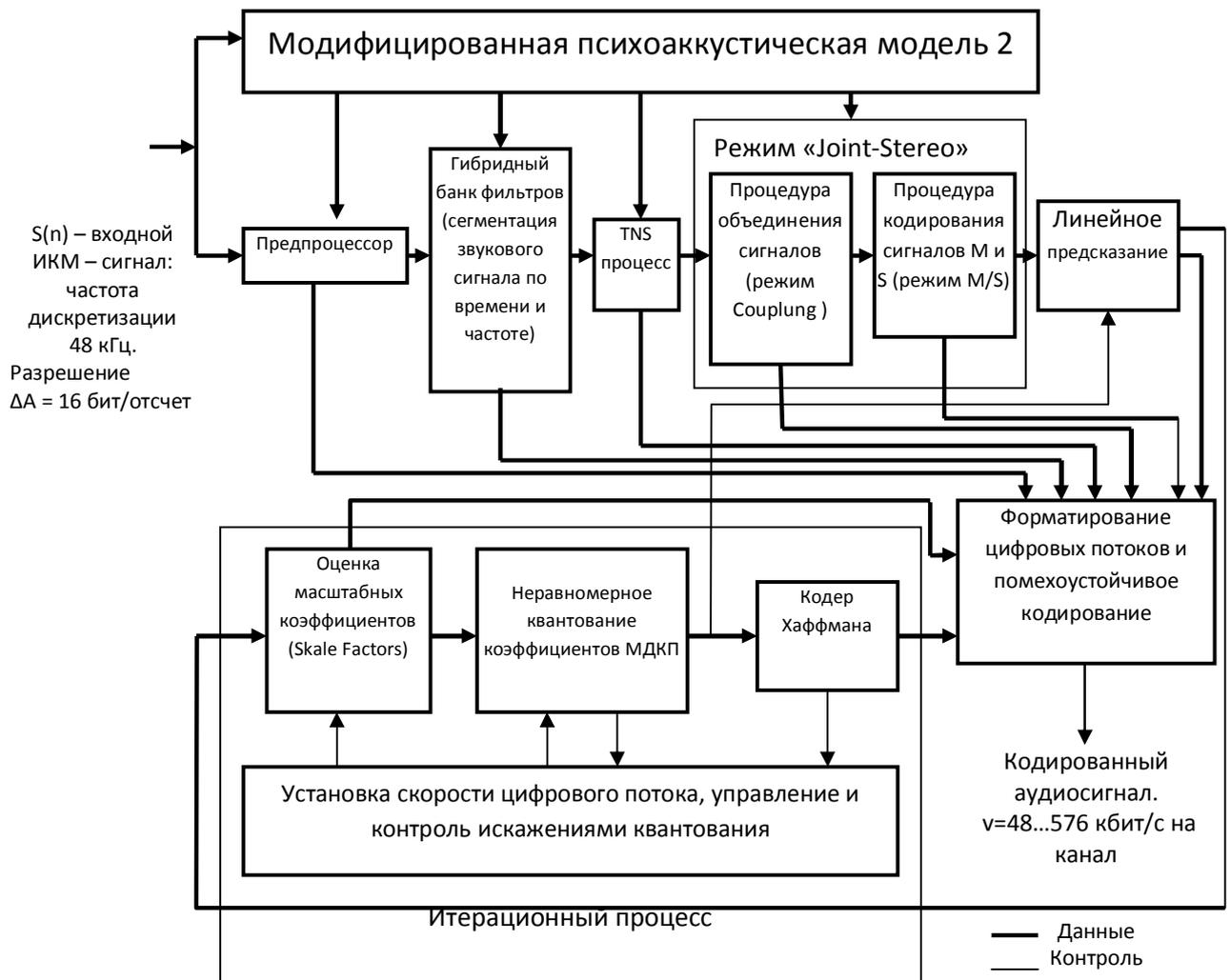


Рис.1.26. Структурная схема кодера AAC (MPEG-2 ISO/IEC 13818-7 и MPEG-4 ISO/IEC 14496-3)

В алгоритме AAC для повышения качества компрессии цифровых аудиоданных применены следующие средства:

- специальные процедуры управления структурой искажений квантования (так называемая техника Temporal Noise Shaping (TNS));
- изменены процедуры объединения субполосных сигналов при их кодировании (Coupling);
- при обработке стерео сигналов в отдельных субполосах предусмотрена возможность кодировать не левый (L) и правый (R) сигналы стереопары, а их сумму и разность:

$$M = (L + R)/\sqrt{2}, S = (L - R)/\sqrt{2}. \quad (1.5)$$

Следует отметить, что при линейном предсказании учитывается не только корреляция между отсчетами сигнала, но и форма спектра шумов квантования и его изменение во времени. При этом, как и в Layer 3, основу кодера MPEG-4 AAC составляет модифицированная психоакустическая модель 2, однако при расчете глобального порога маскировки в нее введены дополнительные процедуры и уточнения.

Технология TNS реализована в алгоритме AAC на базе цифровых фильтров. С их помощью энергия искажений квантования перераспределяется в пределах сигнала выборки в ту ее часть, где энергия полезного сигнала более значительна, что позволяет снизить слуховую заметность шумов квантования. После выполнения этой процедуры огибающие полезного сигнала и искажений квантования оказываются близкими по форме во временной и частотной областях, что значительно повышает эффективность сжатия аудио потока. Так тестовые прослушивания показали, что алгоритм компрессии AAC обеспечивает так называемое «прозрачное кодирование» при скорости цифрового потока около 64 кбит/с на канал.

4.4. Аудио кодирование в стандарте MPEG-4

Стандарт MPEG-4, разработанный для мультимедийных приложений, объединяет опыт, накопленный группой MPEG при разработке всех рассмотренных алгоритмов компрессии цифровых данных. Здесь кроме алгоритма AAC впервые при компрессии цифровых данных высококачественных звуковых сигналов предложено использовать параметрическое кодирование, при котором реальный звуковой сигнал представляется в виде модели, содержащей совокупность тональных и шумоподобных сигналов [35]. При кодировании ЗС определяются ряд параметров, описывающих это сигнал, которые затем передаются декодеру. А декодер, на основе этих параметров при помощи синтезатора восстанавливает исходный ЗС. При этом, чем точнее будут описаны параметры ЗС, тем ближе его звучание будет соответствовать оригиналу. На рис.1.27 и 1.28 представлены упрощенные структурные схемы параметрического кодирования и декодирования соответственно. Рассмотрим подробнее работу параметрического кодека.

Сигнал выборки $S(n)$ в блоке анализа-синтеза разделяется на тональные и шумоподобные составляющие, далее для базовой параметрической модели сигнала оцениваются значения текущих частот, фаз и амплитуд тональных сигналов и уровней энергии шумоподобных сигналов в определенных полосах частот. Значения перечисленных параметров квантуются и кодируются минимально возможным количеством бит, которое определяется с помощью психоакустической модели и затем через канал связи передаются декодеру.

На приемной стороне системы, на основе принятых параметров сигналов декодер управляет тональным генератором импульсов и генератором шума в результате чего производится синтез исходного звукового сигнала. При этом синтез звонких звуков происходит с помощью генератора тональных импульсов, который открывает на определенное время

генерацию основного тона. Возбуждение глухих звуков имитируется шумовым генератором.

До последнего времени параметрическое моделирование использовалось только при компрессии цифровых данных речевых сигналов в вокодерах (кодировщиках голоса), более простых по своей структуре, чем музыкальный сигнал. Однако в последние годы благодаря успехам вычислительной техники математического моделирования, психофизики и электроники, параметрическое представление все чаще начинает применяться и при кодировании высококачественных звуковых сигналов, обеспечивая более высокий уровень компрессии цифровых данных. Параметрическое кодирование, обладая очень сложными процедурами оценки параметров и требующее при реализации существенно больших вычислительных затрат, позволяет получить скорость цифрового потока порядка 16...24 кбит/с при достаточно хорошем качестве звучания.

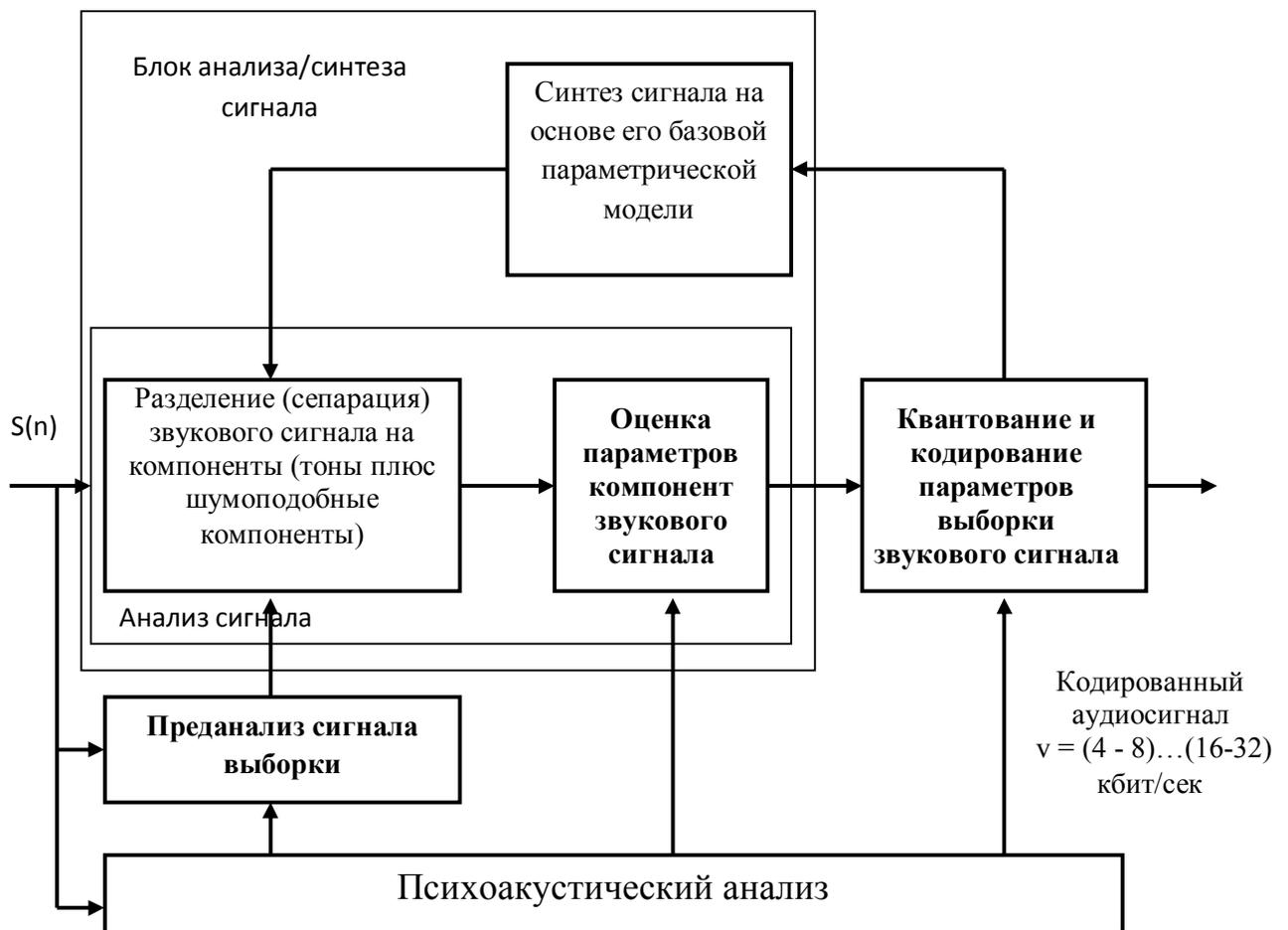


Рис.1.27. Упрощенная схема кодера параметрического кодирования ЗС

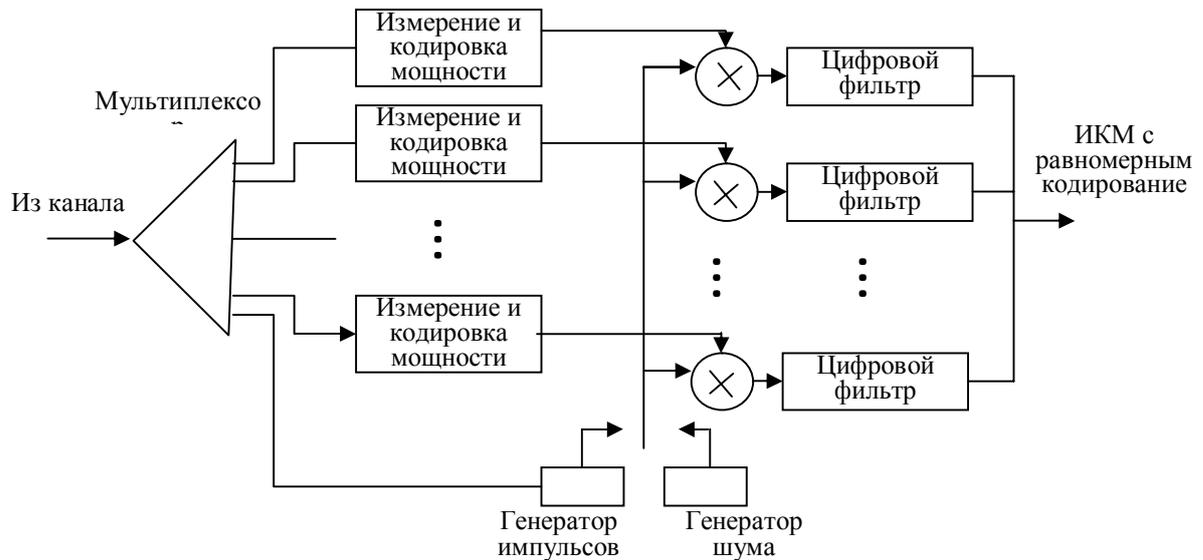


Рис.1.28. Структурная схема параметрического декодера

Рассмотрим подробнее структуру аудио кодирования стандарта MPEG-4 ISO/IEC 14496-3 по обобщенной структурной схеме, представленной на рис.1.29.

Кодер включает в себя следующие алгоритмы компрессии:

- MPEG-4 AAC (Advanced Audio Coding), который предназначен для кодирования музыкальных фрагментов звукового сигнала, имеющих сложную динамическую и временную структуры;
- MPEG-4 AAC+SBR (Spectral Band Replication, буквально «копирование спектральных полос») позволяет дополнительно понизить скорость цифрового потока при кодировании музыкальных фрагментов звукового сигнала, имеющих сравнительно сложную динамическую и временную структуры;
- MPEG-4 CELP (Code Excited Linear Prediction), служащий для кодирования речи;
- MPEG-4 HVXC+HILN (Harmonic Vector Excitation плюс Harmonic and Individual Lines plus Noise), предназначенный для параметрического кодирования сложного сигнала (речь + музыка), а также для кодирования музыкальных фрагментов звукового сигнал со сравнительно несложной структурой.

Выбор алгоритма компрессии происходит автоматически на этапе предварительного анализа выборки звукового сигнала.

В алгоритме MPEG-4 AAC по сравнению с MPEG-2 AAC применен более широкий набор возможных частот дискретизации ЗС: 8, 11.025, 16, 22.05, 24, 32, 44.1, 48, 64, 88.2, 96 кГц, а также вместо синусоидальных оконных функций используется оконная функция Кайзера-Бесселя с двумя размерами окна: длинное окно, включающее 2048 отсчетов ЗС и короткое окно с 256 отсчетами ЗС. Это обеспечивает лучшее соответствие разрешающей способности слуха по частоте и по времени по сравнению с Layer 3. При этом в обоих случаях используется 50%-ное перекрытие выборок отсчетов звукового сигнала.

Как и в Layer 3 (MP-3) кодированию подвергаются не сами семплы, а коэффициенты модифицированного дискретного косинусного преобразования (МДКП). Однако в отличие от Layer 3 здесь изменена форма кривой компрессии при неравномерном квантовании, применены иные коды Хаффмана для кодирования коэффициентов МДКП. Так же управление величиной искажений квантования выполняется с помощью двух итерационных циклов – внутреннего и внешнего.

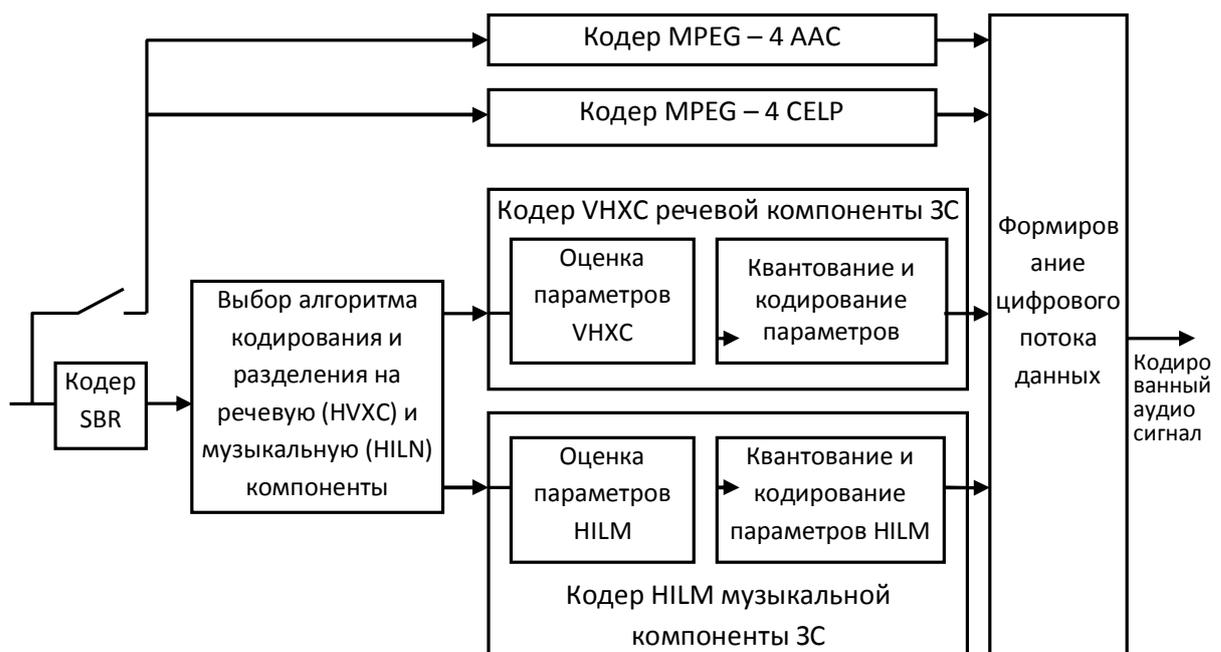


Рис.1.29. Обобщенная структурная схема кодера стандарта MPEG-4 ISO/IEC 14496-3

Кодирование звуковой информации в стандарте MPEG-4 может осуществляться разными способами, отличающихся сложностью и быстродействием применяемых алгоритмов, объемом передаваемых данных и качеством звучания восстановленных фонограмм. Поэтому отметим основные особенности применяемых механизмов кодирования ЗС:

1. Кодирование всех видов звука с обеспечением высокого и среднего качества выполняется методом, основанным на алгоритме AAC стандарта MPEG-2 ISO/IEC 13818-7. При этом обеспечивается передача до 8 каналов звука при скорости передачи двоичных символов 16...64 кбит/с на канал.

2. Кодирование музыки и других звуков с более низкими скоростями выходного потока производится по методу TwinVQ (Transform-domain Weighted Interleave Vector Quantization - Взвешенное векторное квантование с перемежением в области преобразования). Как и в методе AAC, в этом методе выполняются разложение на частотные поддиапазоны и МДКП в каждом поддиапазоне. Отличие TwinVQ от AAC состоит в векторном квантовании спектральных составляющих звукового сигнала. В результате скорость передачи двоичных символов в выходном потоке данных составляет от 6 до 24 кбит/с.

3. Для передачи речи с высоким и средним качеством используется метод кодирования CELP (Code Excited Linear Predictive - кодирование возбуждений с линейным предсказанием), который обеспечивает скорости передачи 6...24 кбит/с при частотах дискретизации 8 кГц или 16 кГц.

4. Параметрическое кодирование речи по методу HVXC (Harmonic Vector eXcitation Coding - кодирование возбуждений гармоническими векторами), которое обеспечивает сжатие при сохранении разборчивости до скоростей 2...4 кбит/с и даже до 1,2 кбит/с при частоте дискретизации 8 кГц.

5. Самые низкие скорости передачи 0,2...1,2 кбит/с достигаются для искусственно синтезированной речи. Для этого в MPEG-4 имеется интерфейс преобразования текста в речь TTSI (Text-to-Speech Interface), который позволяет передавать описание речи в виде последовательности фонем

(звукосочетаний) с указаниями особенностей их произнесения (ударения, длительности и т.д.). По этим данным в декодере синтезируется речь, которая синхронизируется с анимацией изображения лица.

6. Музыка также может передаваться в виде описания и синтезироваться в декодере. Для описания звучания музыкальных инструментов в MPEG-4 дан специальный язык SAOL (Structured Audio Orchestra Language - язык структурированного звукового оркестра). Каждый инструмент представляется как небольшой набор средств обработки сигналов, позволяющий создавать специфический для данного инструмента звук. Описания музыкальных инструментов загружаются в декодер из принимаемого потока данных и могут храниться в нем для последующего использования. Для описания собственно музыки передается описание оркестровки, т.е. команды и данные, по которым осуществляется синтез звуков, соответствующих разным инструментам.

5. Анализ методов сжатия текстовой и служебной информации

В основе всех методов сжатия лежит простая идея: если представить часто используемые элементы короткими кодами, а редко используемые – длинными кодами, то для хранения блока данных требуется меньший объем памяти, чем, если все элементы представлять кодами одинаковой длины [21].

Сжатие текстовой информации общего вида. Принципиальная возможность сжатия текстовой информации связана с тем, что составляющие текста - буквы и словоформы - различаются по частоте встречаемости в тексте, в то время как их длины слабо связаны с частотой. Все алгоритмы сжатия можно классифицировать по используемому методу кодирования и характеру использования статистики и грамматики текста [37].

Методы кодирования можно разделить на четыре класса в зависимости от того, какие группы символов кодируются, и какие коды используются. Рассмотрим наиболее распространенные из них.

Метод Хаффмана. Сжатие Хаффмана - статистический метод сжатия, который уменьшает среднюю длину кодового слова для символов алфавита. Код Хаффмана является примером кода, оптимального в случае, когда все вероятности появления символов в сообщении - целые отрицательные степени двойки [38].

Для заданного распределения частот символов может существовать несколько возможных кодов Хаффмана. Возможно определить 'каноническое' дерево Хаффмана, выбрав одно из возможных деревьев. Такое каноническое дерево может быть очень компактно, передавая только длину в битах для каждого кодового слова. Такой метод используется в большинстве архиваторов (pkzip, lha, zoo, arj, ...).

Кодом Хаффмана можно добиться оптимального кодирования. Код Хаффмана является неравномерным и префиксным. Неравномерность означает, что те символы, которые встречаются в сообщении чаще, кодируются более короткими кодами, а символы, которые встречаются редко – более длинными. Префиксность говорит о том, что ни один код не является началом другого кода, что позволяет достичь однозначности при декодировании. Сжатие методом Хаффмана выполняется за два прохода. На первом проходе читаются все входные данные и подсчитываются частоты встречаемости всех символов. Затем по этим данным строится дерево Хаффмана, по которому вычисляются коды символов. На втором проходе, входные данные читаются еще раз и перекодируются на основе новой кодовой таблицы.

Метод RLE. RLE (Run Length Encoding - кодирование серий последовательностей) – наиболее известный простой подход и алгоритм сжатия информации обратимым путем [39]. Суть методов данного подхода состоит в замене цепочек или серий повторяющихся байтов или их последовательностей на один кодирующий байт и счетчик числа их повторений. Проблема всех аналогичных методов заключается лишь в определении способа, при помощи которого распаковывающий алгоритм мог

бы отличить в результирующем потоке байтов кодированную серию от других - некодированных последовательностей байтов. Решение проблемы достигается обычно постановкой меток вначале кодированных цепочек. Такими метками могут быть, например, характерные значения битов в первом байте кодированной серии, значения первого байта кодированной серии и т.п. Недостатком метода RLE является достаточно низкая степень сжатия или стоимость кодирования файлов с малым числом серий и, что еще хуже - с малым числом повторяющихся байтов в сериях.

LZW. Алгоритм Лемпеля — Зива — Велча (Lempel-Ziv-Welch, LZW) — это универсальный алгоритм сжатия данных без потерь. Алгоритм сжатия данных, основанный на поиске и замене в исходном файле одинаковых последовательностей данных для их исключения и уменьшения размера «архива». В отличие от предыдущих рассмотренных методов сжатия, более «интеллектуально» просматривает сжимаемое содержимое, а все ради большей степени сжатия данных. Данный тип сжатия не вносит искажений в исходный графический файл и подходит для обработки растровых данных любого типа. Наилучшие результаты получаются при компрессии изображений с большими областями одинакового цвета или изображений с повторяющимися одинаковыми структурами. Этот метод демонстрирует самые высокие степени сжатия среди алгоритмов сжатия без потерь данных и используется в форматах TIFF, PDF, GIF, PostScript (в инкапсулированных объектах) и других [40].

Процесс сжатия выглядит достаточно просто. Считываются последовательно символы входного потока и проверяется, есть ли в созданной нами таблице строк такая строка. Если строка есть, то считывается следующий символ, а если строки нет, то заносится в поток код для предыдущей найденной строки, заносим строку в таблицу и начинается поиск снова [21].

Особенность LZW заключается в том, что для декомпрессии нам не надо сохранять таблицу строк в файл для распаковки. Алгоритм построен таким

образом, что мы в состоянии восстановить таблицу строк, пользуясь только потоком кодов.

Для каждого кода надо добавлять в таблицу строку, состоящую из уже присутствующей там строки и символа, с которого начинается следующая строка в потоке.

Метод сортировки блока данных (BWT). Метод сортировки блока данных (BWT - сокращение от Burrows Wheeler Transform - по имени авторов).

Это особый вид или группа преобразований, в основе которых лежит сортировка. Такому преобразованию можно подвергать почти любые данные. Сортировка производится над блоками, поэтому данные предварительно разбиваются на части. Основным параметром является размер блока, который подвергается сортировке. Для распаковки данных необходимо проделать почти те же действия, что и при упаковке. Поэтому скорость и требования к оперативной памяти почти одинаковы. Архиваторы, которые используют данный метод, обычно показывают высокую скорость и степень сжатия для текстовых данных [41].

Выводы по I главе

В результате проведенного анализа существующих и перспективных методов и алгоритмов сжатия ТВ изображений было выявлено следующее:

1) В настоящее время мультимедийные технологии являются бурно развивающейся сферой информационных технологий. Следовательно, было определено основное понятие этого направления – «мультимедиа» и основные составляющие компоненты, входящие в состав мультимедиа – текст, аудио, видео. Установлены основные характерные особенности этой технологии. Более подробно были выделены основные характеристики видеопотока ТВ программ.

2) Из анализа видов информации установлено, что любая информация обладает большим объемом избыточности, за счет которой осуществляется компрессия телевизионного сигнала для уменьшения скорости цифрового потока для более эффективного использования каналов передачи.

3) Установлено, что видео и аудио данные облают следующими видами избыточности: кодовая, статистическая или межэлементная, структурная, временная или межкадровая, психовизуальная и психоакустическая, за счет устранения которых и обеспечивается первичное сжатие данных.

4) В результате анализа методов сжатия видеоданных ТВ программ установлено, что сжатие объемов видеоданных производится на основе ортогональных преобразований и межкадровой обработки с компенсацией движения. При этом наибольшее применение среди ортогональных преобразований находит дискретно косинусное и вейвлет преобразования (ДКП и ВП). Установлено, что ДКП обладает хорошей производительностью и хорошо сочетается с блочным методом компенсации движения, и обеспечивает хорошее качество изображения при скоростях видеопотока более 5 Мбит/с, однако на меньших скоростях возникают искажения изображения в виде блочного эффекта, в результате чего появляется мозаичный вид.

5) При использовании ВП изображения обычно не разбиваются на блоки, а обрабатывается целиком, что позволяет в 1,5-2 раза увеличить сжатие статического изображения без заметного ухудшения его качества, и устраняет возникновение блочных искажений, однако без блочная обработка изображений не позволяет использовать методы блочной компенсации движение. Поэтому вейвлет – кодеки уступают по эффективности сжатия видеопотока кодекам стандартов MPEG.

6) Для получения высоких коэффициентов сжатия порядка 200-2000 могут использоваться фрактальные методы сжатия изображений, которые основаны на поиске самоподобных видеообъектов, нахождение которых позволяет передавать ссылки на первоначально найденный фрактал с

указанием в каких координатах, ориентации и масштабе его нужно вставить в синтезируемое изображение. Фрактальное сжатие хорошо работает на изображениях технических чертежей, текста, карт местности и т.д. На реальных ТВ изображениях фрактальное сжатие дает хорошие результаты, но на поиск фракталов с учетом масштабирования и взаимной ориентации требуются большие временные затраты, что не позволяет их использовать в вещательном телевидении, где сжатие видеопотока должно происходить в режиме реального времени.

7) Для звуковых сигналов важной проблемой при цифровом представлении является сокращение имеющейся в них статической и психоакустической избыточности. Устранение статической избыточности не приводит к значительному уменьшению скорости цифрового потока, а благодаря психофизической избыточности можно обеспечить сжатие аудиоданных в 10-12 раз без существенных потерь в качестве. Основное устранение избыточности в аудиоданных достигается с помощью различных видов маскировок, что не приводит к потере в качестве.

8) Для сжатия текстовой и служебной информации необходимо использовать методы сжатия без потерь данных, поскольку нарушение структуры информации не позволит правильно восстановить исходные изображения, звуковое сопровождение и текстовую информацию. В настоящее время для этой цели активно используются алгоритмы RLE, Хаффмана, LZW, BWT, являющиеся основой архиваторов ARJ, ZIP, RAR.

Глава II. Аналитический обзор мультимедийных кодеков и оценка их характеристик

1. Общие положения мультимедийных кодеков

С развитием телевизионных технологий и появлением широкоформатного телевидения все большее значение приобретает компрессия видеосигнала при его хранении и передаче по сети. Кодирование и последующее декодирование сигнала уменьшает нагрузку на каналы передачи данных и позволяет обойтись меньшим объемом архива. Компрессия сигнала производится при помощи кодеков.

Кодеки позволяют кодировать данные для последующей передачи, хранения или шифрования, а также раскодируют их для просмотра (воспроизведения). Использование кодеков уменьшает исходный размер мультимедиа файлов, сохраняя максимально хорошее качество.

Кодеки сами по себе не воспроизводят мультимедийные файлы; они устанавливаются для того, чтобы помочь воспроизводить такие файлы программам-проигрывателям [22].

Наиболее известными являются кодеки, основанные на стандартах H.264, MPEG – 4, перспективным стандартом является H.265.

Большинство существующих кодеков для звуковых и визуальных данных используют сжатие с потерями, для получения приемлемого размера сжатого файла. Существуют также кодеки, сжимающие без потерь (англ. lossless codecs). Но для большинства применений выгоднее кодеки с потерями информации, так как малозаметное ухудшение качества оправдывается значительным уменьшением объема данных.

По типу исполнения кодеры делятся на аппаратные и программные. В первом случае кодирование выполняет специализированный чип, а во втором – программное приложение, опирающееся на ресурсы центрального процессора компьютера, используемые для просмотра видео на

персональных компьютерах, представляют собой наборы программных средств.

По типу обрабатываемой информации существуют два типа кодеков: аудио - и видеокодеки.

Аудиокодек — компьютерная программа или аппаратное средство, предназначенное для кодирования или декодирования аудиоданных.

Видеокодек — программа/алгоритм сжатия видеоданных (видеофайла, видеопотока) и восстановления сжатых данных. Также возможно кодирование кроме видео и аудиоинформации, добавления субтитров, векторных эффектов и т. п.

2. Программные видеокодеки

Программный видеокодек – это специализированная компьютерная программа, кодек, которая производит компрессию или декомпрессию цифровых видео данных в соответствии с файловым видео форматом или потоковым видео форматом.

Мультимедийные файлы имеют довольно большие размеры, особенно это касается видеофайлов хорошего качества. Кодеки призваны решить проблему передачи и хранения файлов больших размеров – при кодировании уменьшается исходный размер мультимедийных файлов, при этом исходное качество сохраняется [22].

Задача программного видеокодека как компрессора заключается в предоставлении видеосигнала с минимально возможным размером и наилучшим уровнем качества. Благодаря сжатию уменьшается объем пространства, требуемого для хранения видеоданных и появляется возможность снизить полосу пропускания канала, по которому передаются видеоданные.

Видеокодеки являются вспомогательными программами, и сами по себе они не воспроизводят медиафайлы. Так как кодеки являются служебными

программами, после их установки не появятся никаких ярлыков. После установки пакетов кодеков они будут работать автоматически при необходимости воспроизведения тех или иных форматов медиа-файлов.

Поскольку существует огромное количество различных видео форматов, следовательно, и кодеков для их работы также немало. Рассмотрим наиболее популярные видеокодеки.

Grass Valley ProCoder. Недорогой программный кодер ProCoder компании Grass Valley (Canopus) (рис.2.1) предназначен для преобразования видео различных форматов. Он может использоваться и как отдельное приложение, и в качестве подключаемого модуля для монтажной программы EDIUS (до версии 4 включительно) [25].

ProCoder позволяет осуществлять пакетное кодирование с соблюдением очередности, а также одновременное кодирование одного файла в несколько разных форматов и наоборот. Фильтрация видеосигнала обеспечивает максимальное качество при передаче. ПО также включает дополнительные функциональные возможности, такие как цветокоррекция, подавление шумов и переходы типа Fade In/Fade Out.

Версия 3 этого кодера включает поддержку кодека H.264, который используется при создании видеофайлов для мобильных устройств и декодирования формата AVCHD с камер MPEG-4 нового поколения. ProCoder 3 может разбить обработку сигнала на несколько задач, что позволяет быстрее выполнить процесс преобразования. Ускорен и процесс кодирования в MPEG-2, что значительно сокращает время при работе с видео формата HDV и создании DVD-дисков [25].

Видеофайлы, которые были преобразованы в чересстрочный формат с использованием понижающего преобразования 3:2, могут быть восстановлены в изначальный формат с прогрессивной разверткой. Обработка звука включает поддержку Dolby Digital (до 5.1). В версии 3 также улучшена работа с временным кодом во время преобразования.



Рис.2.1. Интерфейсы ProCoder 3

ProCoder 3 может быть использован для конвертирования и создания файлов для дальнейшего их просмотра на мобильных устройствах. В него включен кодек Grass Valley (Canopus) HQ, который позволяет кодировать/декодировать файлы Grass Valley (Canopus) HQ вне программного обеспечения Grass Valley EDIUS. Этот кодек поддерживает технологии Panasonic, в частности, обеспечивает возможность декодирования материала, снятого на видеокамеры AVCHD.

Поддержка Grid Encoding для MPEG-2 (Transport Stream) обеспечивает более высокую скорость при кодировании на современных многоядерных и многопроцессорных рабочих станциях. Данная функция автоматически

активируется при выборе форматов кодирования Preview DVD Target или HDV Target.

ProCoder 3 поддерживает кодирование аудиосигнала Dolby Digital, а также дает возможность пользователю добавлять, менять местами или удалять отдельные аудиоканалы.

Ключевые особенности ProCoder 3:

- высокоскоростное кодирование с высоким качеством и преобразование видео в форматы MPEG-1, MPEG-2, Windows Media, QuickTime и др.;
- функция Watch Folder;
- менеджер постановки задач в очередь с функцией соблюдения приоритетности выполнения;
- объединение нескольких файлов в один, одновременный вывод файла в нескольких форматах;
- преобразование формата кадра;
- преобразование NTSC/PAL;
- интерполяция и преобразование частоты кадров, создание и сохранение временного кода;
- автоматическое адаптивное преобразование чересстрочной развертки в прогрессивную, понижающее преобразование 3:2;
- функции инверсного телекино, многопроходного кодирования с переменной скоростью потока (VBR);
- поддержка HD для MPEG и Windows Media: создание с разметкой глав для файла VOB и DVD-Video image;
- поддержка цветовых пространств YUV и RGB.

Поддерживаемые входные и выходные форматы:

- входные – Grass Valley DV/HQ, Microsoft DV, H.264, AVCHD, Dolby Digital, видео для кодеков Windows, кодеки DirectShow и QuickTime; MPEG-1, MPEG-2, Macromedia Flash (только видео), Windows Media,

DivX, MP3- и PCM WAV-файлы, AVISynth скрипт-файлы, статичные изображения, включая BMP, TGA, TIF, PNG и JPG;

- выходные – Grass Valley DV/HQ, Microsoft DV, H.264, Dolby Digital, iPod (MPEG-4), MediaEdge (MPEG-2, MPEG-4), большинство видео для кодеков Windows, Direct - Show и QuickTime; MPEG-1, MPEG-2 (при использовании кодека Grass Valley Canopus), VOB (DVD-Video image), ACM-кодеки, Macromedia Flash-видео, Windows Media, RealVideo, MP3- и PCM WAV-файлы.

DivX 10.2.0. DivX - популярнейший на сегодняшний день видео-кодек.

DivX - технология видеозаписи, позволяющая создавать и просматривать медиа файлы с высокой степенью сжатия [42].

Данный пакет содержит:

- DivX Plus Player - бесплатная версия проигрывателя для воспроизведения файлов в форматах DivX, AVI, MKV, MP4, MOV с функцией DivX To Go (быстрый перенос видео с компьютера на различные устройства, поддерживающие DivX (DVD-плееры, телевизоры, игровые консоли и пр.)).
- DivX Plus Codec Pack - бесплатный набор кодеков для воспроизведения DivX и MKV файлов в любом проигрывателе.
- DivX Plus Web Player - программа для проигрывания потоковых DivX, AVI и MKV файлов, интегрируется в популярные браузеры
- DivX Plus Converter - конвертер видео (15-ти дневная пробная версия)
- DivX 9 включает ряд усовершенствований, рассчитанных на работу с видео высокой четкости в формате H.264 и контейнеров MKV.
- DivX 10.1 включает ряд усовершенствований, рассчитанных на работу с видео ультра высокой четкости 4k в формате H.265.

DivX - новая версия пакета программ для воспроизведения видео, в котором появились опции, привычные по DVD и отсутствующие до сих пор в DivX - интерактивные меню, поддержка нескольких субтитров и нескольких

аудиодорожек, сцены и т.п. Кроме этого, улучшено качество и степень сжатия видео.

DivX Plus поддерживает воспроизведение видео высокой четкости в формате MKV/H.264. При помощи DivX Plus Web Player можно передавать потоковое HD-видео и в браузере. Используя новые функции DivX Plus Converter, можно преобразовывать HD-содержимое 1080p в DivX Plus HD, основанном на контейнере файлов MKV и кодеке H.264.

Достоинством DivX Plus является механизм транскодирования видео, заключается в поддержке аппаратного ускорения операций на всех моделях процессоров Intel из линеек Sandy Bridge и Ivy Bridge.

Это аппаратное ускорение обеспечивает до 10-кратного увеличения скорости во время перекодирования по сравнению с предыдущей версией DivX Converter 8.

Дополнительное удобство создают специальные профили вывода перекодированных материалов для аппаратов iPhone и iPad – они позволяют быстро преобразовать исходные видеозаписи в формат MP4.

Также проигрыватель DivX Player теперь содержит полноценный медиа-сервер для вещания в домашней локальной сети, где поддерживается стандарт DLNA (Digital Living Network Alliance – «Альянс цифровых сетей для дома»).

Остались два платных компонента, плагин для поддержки формата MPEG-2 в модуле Converter и DFX Audio Enhancer 10.

Файловый медиатранскодер ProMedia Carbon. ProMedia Carbon (ранее Carbon Coder), разработанный компанией Rhozet (в настоящее время она входит в Harmonic), является универсальным программным транскодером медиафайлов [26].

ProMedia Carbon может работать как независимое приложение или в качестве одного из узлов автоматизированной фермы визуализации (render farm) под управлением Carbon Server или системы Harmonic WorkFlow System (WFS). Он поддерживает преобразование SD/HD-форматов,

PAL/NTSC-конверсию, вставку логотипов, преобразование цветовых пространств, цветокоррекцию, извлечение субтитров и многое другое. ProMedia Carbon имеет открытый API-интерфейс, позволяющий создавать рабочие процессы для конкретных задач и оборудования.

Отличительной особенностью ProMedia Carbon является его способность конвертировать материал практически из любого медиаформата в любой используемый сегодня в мире видеостандарт. Он работает со всеми основными форматами SD- и HD-видео, постоянно обновляется и дополняется поддержкой новых форматов.

Поддерживаемые видеокодеки:

- MPEG-1 – DV25, DV50, DV100, DVCPPro;
- MPEG-2 – DPS;
- MPEG-4 – DivX;
- H.263 – JPEG 2000;
- H.264 – DNxHD, AVCIntra;
- VC-1 – секвенции изображений;
- Flash – Windows Media, RealVideo.
- Поддерживаемые аудиокодеки:
- Dolby Digital – AAC;
- Dolby Digital Plus – AMR-NB;
- Dolby E – Windows Media Audio;
- PCM – MPEG-1 Layer II, MP3;
- RealAudio.

Интуитивно - понятный пользовательский интерфейс ProMedia Carbon обеспечивает полный контроль над процессом транскодирования.

Пользователь может легко сконфигурировать систему для конвертирования файлов в любое количество форматов одновременно, вставлять логотип и временной код, а для мобильного формата применить специальное масштабирование.

ProMedia Carbon может быть запущен в автоматическом режиме с поддержкой пакетной обработки, мониторинга директорий и автоматической передачи данных по FTP. Технология транскодирования ProMedia Carbon позволяет повысить производительность процесса путем простого определения исходного формата и его автоматического преобразования в требуемые форматы [26].

Для выполнения объемных задач несколько узлов ProMedia Carbon могут быть объединены в render farm под управлением Rhozet Workflow System (WFS), которая будет распределять задачи, задавать приоритеты их выполнения, балансировать нагрузку, контролировать передачу данных по FTP, выполнять мониторинг системы и оповещать о выполнении задач.

ProMedia Carbon можно управлять с помощью набора средств разработчика (SDK) на базе XML, входящего в комплект ПО. SDK обеспечивает все возможности для управления процессом транскодирования, включая выбор размещения исходных и конечных файлов, параметров преобразования, фильтрации, оповещения, вставку рекламы, наложение титров и др.

ProMedia Carbon используется для конвертирования формата, заданного сетевым сервером распределения контента, в формат вещания. Материал с сервера может либо сразу преобразовываться в формат вещания, либо сначала в формат для видеомонтажа, а затем созданная программа конвертируется в вещательный формат. ProMedia Carbon применяется для телевещания, интернет-вещания, транскодирования в рабочих группах, в видеопроизводстве.

Основные функции:

- преобразование размера, частот, форматов кадров;
- преобразование форматов PAL/NTSC;
- преобразование форматов SD и HD;
- преобразование цветовых пространств;
- масштабирование;

- извлечение ключевых кадров;
- преобразование прогрессивной развертки в чересстрочную и обратно;
- вывод в нескольких форматах одновременно;
- телекино/инверсное телекино;
- пакетная обработка.

Видеофильтры:

- переходы Fade In/Out, вращение, медиана, размытие;
- коррекция баланса белого/черного, цвета, гаммы, резкости;
- зона безопасности NTSC-формата;
- временное шумоподавление;
- деблокирование;
- временной фильтр с компенсацией движения.

– Аудиофильтры:

- нормализация по стандарту ITU 1770;
- регулировка громкости;
- НЧ-фильтр;
- компрессор динамического диапазона,
- эффекты Fade In/Out.

Поддерживаемые медиаконтейнеры: AVI, LXF, GXF, Quicktime, WMV, WMA, ASF, HDV, VOB, MXF (OP1a, OPAtom), 3GPP, MPEG-2 PS, MPEG-2 TS, 3G2, WAV, Broadcast WAV, Streaming Media (H.264, VC-1), HTTP Streaming.

Поддерживаемые системы: ATSC, DVB, CableLabs, Quantel sQ, Panasonic P2, Avid MediaStream, Sony XDCam, Apple Final Cut Pro, Leitch VR и Nexio, Adobe Premiere Pro, Grass Valley Profile и K2, Grass Valley Edius, Omneon Spectrum.

3. Аппаратные устройства кодирования мультимедийного потока

Кодек на аппаратном уровне обозначает отдельную микросхему, которая кодирует и декодирует аналоговый аудио - и/или видео сигнал в цифровой сигнал и, наоборот, при помощи аналогово-цифрового и цифро-аналогового преобразователей.

Кодеры StreamZ. Кодеры StreamZ был разработан канадской компанией Digital Rapids и используются ведущими вещательными сетями, телекомпаниями, киностудиями и дистрибьюторами онлайн-контента, государственными организациями и системными интеграторами в решениях для IPTV, «видео по запросу», Web-вещания, мобильного видео, дистанционного обучения и во многих других областях [27].

Программно-аппаратные кодеры StreamZ выполняют кодирование, архивирование и потоковое вещание. Кроме базовых функций в них реализована аппаратная поддержка преобразования чересстрочной развертки в прогрессивную и шумоподавления. Кодеры StreamZ характеризуются высокой надежностью, что важно для организации прямых трансляций и вещания.

StreamZHD – это системы, осуществляющие захват, кодирование, транскодирование и потоковое вещание. Они поддерживают различные форматы, поэтому можно захватывать видео в форматах 1080i, 1080p, 720p или SD и использовать его для вещания, а также конвертировать видео из HD в SD, SD в HD и выполнять преобразования разных HD-форматов.

Кодирование можно проводить в несколько форматов, с разными разрешениями и скоростями потоков одновременно в реальном времени, что оптимально для обеспечения контентом аудиторий с разной пропускной способностью канала.

Оригинальный портативный кодер Touch-Stream сочетает в себе профессиональное качество (рис.2.2), универсальность и интуитивно-понятный интерфейс, управляемый с сенсорного дисплея. Этот кодер

является полностью автономным и позволяет организовать потоковое вещание на выезде – с передвижной платформы или просто вне студии.



Рис.2.2. Кодер StreamZ – TouchStream

Среди кодировщиков Digital Rapids есть устройства, работающие с аналоговыми и с цифровыми видеоинтерфейсами, одноканальные и двухканальные, для SD- и HD-видео. Все кодировщики могут быть укомплектованы любым набором кодеков и позволяют получать поток с параметрами, заданными пользователем или выбранными из набора предустановленных характеристик. Гибкая архитектура устройств Digital Rapids позволяет в случае необходимости легко расширить список поддерживаемых форматов. Кодировщики StreamZ оптимально подходят для создания комплекса IPTV, проведения презентаций и шоу, дистанционного обучения и множества других профессиональных применений [27].

Кодеры Ericsson. Компания Ericsson, приобретая несколько лет назад активы Tandberg Television, интегрировала выпускавшееся ею оборудование в свою линейку устройств [43]. И теперь кодировщики, носившие некогда имя Tandberg, выпускаются под маркой Ericsson, получая дальнейшее развитие.

В данном классе устройств выпускаются как модульные системы на базе платформы Ericsson Video Processor Chassis (VPC), так и самостоятельные устройства.

Кодеры на базе VPC. VPC (Video Processor Chassis – шасси процессора видео) представляет собой многофункциональную, с высокой плотностью монтажа платформу обработки видео, отвечающую современным потребностям вещательных компаний.

Она компактна (1RU), но позволяет установить до шести модулей при наличии одного блока питания или до четырех модулей, если в корпусе имеются два блока питания (основной и резервный).

Поддерживается широкий спектр опций кодирования, включая MPEG-2 SD и MPEG-4 AVC HD.

В слоты процессора можно установить ряд модулей. Первый из них – это кодер EN8190 MPEG-4 AVC HD (рис.2.3). Модуль не только имеет новую конструкцию, но и создан с поддержкой iRDO (Interpolating Rate Distortion Optimization – оптимизация искажений в зависимости от скорости потока), что позволяет максимально эффективно использовать кодек MPEG-4 AVC. В результате кодер дает возможность передавать 5...6 HD-каналов в одном частотном канале DVB-T2, 7...8 HD-каналов в одном канале DVB-S2 и 8...10 HD-каналов в одном канале DVB-C [43].

Плата имеет входы 3G-SDI и AES-EBU, ряд функций для работы с Dolby Digital, поддерживает работу со скрытыми титрами и обладает рядом других возможностей, который можно расширять путем обновления ПО.

Другой модуль – EN8180 MPEG-2 HD – предназначен для кодирования в MPEG-2 HD. Его функции и возможности во многом такие же, как и у EN8190, разумеется, там, где это применимо к используемому формату сжатия. Поддерживается и статистическое мультиплексирование. Входы видео – HD-SDI, аудио – AES-EBU и HD-SDI.

Третий из модулей – это EN8130 MPEG-4 AVC SD. Он отличается от EN8190 в основном тем, что работает с сигналами стандартного разрешения и не содержит ядра iRDO. По сравнению с предыдущими версиями эффективность сжатия увеличена на 15%.



Рис.2.3. Плата кодера EN8190

Кодер EN8100 MPEG-2 SD (четвертый в ряду, рис.2.4), как нетрудно догадаться по названию, формирует поток MPEG-2 SD и тоже создан на базе новейших достижений в сфере технологий компрессии.

И, наконец, пятый модуль EN7100 MPEG-2 SD тоже работает в стандартном разрешении и оптимален для любых вариантов реализации цифрового ТВ.



Рис.2.4. Кодер EN8100

Самостоятельные кодеры. HD-кодеры MPEG-2 моделей E5780/ E5782 способны работать в режимах высокого и стандартного разрешения. Они эффективны и просты в эксплуатации.

Максимальная скорость потока для E5782 достигает 90 Мбит/с, что делает его пригодным даже для приложений цифрового кино и других, требующих компрессии почти без потерь. E5780 выполняет кодирование по профилю 4:2:0, а E5782 – еще и по 4:2:2.

Устройства легко адаптировать к текущим задачам. В корпусе 2RU есть четыре слота расширения. Органы управления – такие же, как у E5720.

Из функций следует упомянуть формирование таблиц PSIP и PSI, наличие интерфейсов для вставки PSIP/SI, улучшенное иерархическое предсказание движения, восьмиуровневое шумопонижение, детектирование кинорежима (протяжка 3:2), поддержку ввода скрытых титров по RS-232 и HD-SDI и т.д.

Кодер MPEG-4 модель EN8030 – кодер второго поколения MPEG-4 AVC SD.

Он оптимизирован для передачи ТВ через каналы DSL, непосредственного спутникового вещания (DTH), подготовки контента для услуги «видео по запросу», а также для профессиональных приложений. Прибор характеризуется очень малой задержкой, простотой настройки и эксплуатации, надежностью и производительностью.

Удобно, что из одного исходного сигнала кодер формирует потоки MPEG-2 и MPEG-4 AVC (опция).

Специально для IPTV прибору придана такая функция, как PIP (картинка в картинке), а прямое IP-широковещание с кодера позволяет применять его в любых приложениях IP и «ТВ через xDSL». Входы – композитный и SDI. Предусмотрена предобработка видео, включая шумопонижение, изменение разрешения от 1/4 до полного D1 и др. Скорость потока – 0,25...10 Мбит/с

Вершиной технической мысли инженеров Ericsson является кодер EN8092 MPEG-4 AVC HD второго поколения. В нем применено ядро «умной» компрессии (Intelligent Compression Engine) на основе технологии предобработки Clarus, обеспечивающей шумоподавление, удаление

артефактов и улучшенную фильтрацию. В сочетании со статистическим мультиплексированием Reflex эффективность сжатия увеличена на 50% по сравнению с предыдущими устройствами. Хотя кодер одноканальный, он может формировать сразу два потока, но разного разрешения – SD и HD. Режимы кодирования – VBR и CBR, предусмотрены многопроходное кодирование со статистическим мультиплексированием Reflex, прямое IP-широковещание и одновременная генерация потока PiP.

Высота корпуса – 2RU, входы видео – SD/HD-SDI и аналоговый композитный. Скорость потока – 0,25...10 Мбит/с, выход ASI (MPEG-2 TS).

Кодер HD-DVB/C-ASI/IP. IP ставилась задача создания универсального прибора с возможностями, удовлетворяющими большинство пользователей, и доступной ценой. Он оснащен входами HD/SD-SDI, композитным PAL и HDMI, а также имеет широкие возможности для работы со звуком (вложенным цифровым и аналоговым), уровень которого можно корректировать. Выходы у кодера – ASI и IP [28].

Учтено и требование времени – возможность выдавать сигнал для вещания на мобильные устройства (телефоны и планшеты) и в Интернет. Встроенный понижающий конвертер позволяет в широких пределах уменьшать размер изображения для трансляции сигнала на мобильные устройства.

Полезной особенностью аппарата является возможность параллельного вещания с высоким (1920×1080) и пониженным (320×240 или 320×180) разрешениями.

Большое внимание уделено удобству эксплуатации – важными параметрами и настройками можно управлять по протоколу SNMP через Интернет, по сети можно и обновить его прошивку.

Кодер работает в режиме реального времени и может быть использован для преобразования цифрового сигнала, когда необходимо уменьшить объем передаваемых данных без потери качества. Режим HD/MPEG-4 обеспечивает высококачественное HD-изображение на экране телевизора даже при

относительно небольших скоростях потока. В режиме работы MPEG-2 на выходе кодера независимо от формата входного сигнала всегда будет сигнал SD –720×576i.

Скорость выходного потока для SD-сигнала можно настраивать в пределах 2... 15 Мбит/с, а для HD-сигнала – 4...20 Мбит/с.

Поддерживаемые форматы на входе – HDMI 576×50i, 576×50p; SDI 1920×1080i с вложенным звуком.

Группа и стереопара входного сигнала выбирается программно. На выходе кодера формируется транспортный поток ASI (EN 50083-9:1998), а также IP-поток, который может использоваться для мониторинга, передачи и IPTV.

Высокая степень интеграции и схемотехнические решения обеспечивают надежность устройства.

Технические характеристики:

- входы – композитный, HDMI, SD- или HD-SDI;
- стандарты сжатия – MPEG-4 AVC;
- MPEG-2: профили – Main, High, уровни – 1.2...4.2;
- стандарт аудио – MPEG-1 Layer 2;
- скорость потока ASI на выходе – 500 кбит/с...15 Мбит/с;
- выходы (два равноправных) – ASI (EN 50083-9:1998), разъемы BNC;
- выход IP – протоколы UDP, RTP, RTP+;
- управление – Ethernet 100 Мбит/с;
- корпус –1U;
- питание – 220 В, 50 Гц;
- масса – 3 кг.

Двухканальный кодер MPEG-2/4. Устройство представляет собой два кодера, на каждый из которых можно подавать сигналы SD/HD, SDI/CVBS и обрабатывать их в соответствии с алгоритмами цифрового сжатия MPEG-4 или MPEG-2 (одновременно для двух каналов). Для HD-сигнала – с профилем High уровня 4.0 при формате входных сигналов 1440×1080i50,

1920×1080i50; а для SD-сигналов – с профилем Main уровня 3.2 при формате входных сигналов 720×576i50.

Звуковой сигнал кодируется в соответствии со стандартом MPEG-1 Layer 2. Скорость выходного потока для сигнала SD – 2...15 Мбит/с, а для HD – 4... 20 Мбит/с. В стандарте MPEG-2 на выходе кодера всегда будет сигнал 720×576i50.

Кодер позволяет работать с сигналами SDI с вложенным звуком, а также способен внедрять аналоговый звук в выходные сигналы. Группа и стереопара входного сигнала выбираются с помощью программы управления DVB Toolkit, входящей в комплект поставки.

Выходные сигналы кодирующего устройства – мультиплексированный транспортный поток ASI (EN 50083-9:1998), содержащий два канала, и поток IP.

Технические характеристики:

- два видеовхода – SD/HD-SDI, CVBS;
- стандарт сжатия – MPEG-4 (ISO/IEC Part 10), MPEG-2;
- для SD – профиль Main, уровень 3.2, для HD – профиль High, уровень 4.0;
- формат SD-сигнала на входе – 720×576i50;
- формат HD-сигнала на входе – 1920×1080i50, 1440×1080i50;
- стандарт аудио – MPEG-1 Layer 2;
- скорость потока на выходе – 2... 15 Мбит/с для SD, 6...20 Мбит/с для HD;
- выходы (два равноправных) – ASI (EN 50083-9:1998, разъемы BNC);
- выход IP – протоколы UDP, RTP, RTP+;
- управление – Ethernet 100 Мбит/с;
- корпус –1U;
- питание – 220 В, 50 Гц;
- масса – 3 кг.

Аппаратный транскодер H.264 и MPEG-2. Данный транскодер позволяет решать проблемы, связанные с параллельным существованием двух форматов кодирования – MPEG-2 и H.264, и двух вариантов разрешения – HD и SD. Устройство выделяет из входного ASI-потока один из каналов и производит для него:

- транскодирование H.264 в MPEG-2 с выводом сигнала на выход ASI (или IP);
- транскодирование MPEG-2 в H.264;
- конвертацию HD-сигнала MPEG-2 в сигнал SD MPEG-2.

Предусмотрены возможности для изменения скорости потока, понижающего преобразования HD в SD, а также для ремультимплексирования входного потока – выбора нужных каналов и добавления транскодированного. Сформированный поток можно подать на IP-выход. Управление устройством осуществляется по Ethernet.

Выводы по II главе

В результате проделанной работы был проведен анализ и сравнения мультимедийных кодеков, используемых при кодировании и декодировании телевизионного сигнала, для уменьшения объема информации и нагрузки на каналы передачи данных. Как было установлено в результате проведенного анализа, на сегодняшний день существует довольно большое количество разнообразных кодеков используемых при сжатии мультимедийного потока.

Было выявлено, что существующие кодеки, применяемые для звуковых и визуальных данных, используют сжатие с потерями, для получения приемлемого размера файла и сжатие без потерь. Но для большинства применений выгоднее кодеки с потерями информации, так как малозаметное ухудшение качества оправдывается значительным уменьшением объема данных.

Также в результате анализа было выявлена классификация кодеков:

- по типу исполнения кодеры делятся на аппаратные и программные.

Программный видеокodeк – это специализированная компьютерная программа, кодек, которая производит компрессию или декомпрессию цифровых видео данных в соответствии с файловым видео форматом или потоковым видео форматом.

Кодек на аппаратном уровне обозначает отдельную микросхему, которая кодирует и декодирует аналоговый аудио - и/или видео сигнал в цифровой сигнал и, наоборот, при помощи аналогово-цифрового и цифро-аналогового преобразователей.

- по типу обрабатываемой информации существуют аудио - и видеокodeки.

Аудиокodeк — компьютерная программа или аппаратное средство, предназначенное для кодирования или декодирования аудиоданных.

Видеокodeк — программа/алгоритм сжатия видеоданных (видеофайла, видеопотока) и восстановления сжатых данных. Также возможно кодирование кроме видео и аудиоинформации, добавления субтитров, векторных эффектов и т. п.

Анализируя построение кодеков также было выявлено бурное развитие кодеков на основе вейвлет преобразований и дискретно-косинусном преобразовании.

В качестве примера, были рассмотрены несколько кодеков, работающих с мультимедийной информацией, приведены их технические характеристики.

Глава III. Анализ принципов построения и работы кодеков мультимедийного контента ТВ программ и оценка их эффективности

1. Общие положения работы кодеков

Кодек (codec — кодировщик/декодировщик) — совокупность кодера и декодера, устройств, соответственно, кодирующих и декодирующих определённым алгоритмом сжатия аудио и видеоданные [29]. Кодеки позволяют кодировать данные для последующей передачи, хранения или шифрования, а также раскодируют их для просмотра (воспроизведения). Использование кодеков уменьшает исходный размер мультимедиа файлов, сохраняя максимально хорошее качество. Кодирование сигнала требуется для существенного сокращения размера данных, часто при незначительных потерях качества. Восстановление происходит с определёнными потерями, так называемыми «артефактами компрессии». Правда, существуют кодеки, которые позволяют производить сжатие без потери качества, такие кодеки называются «lossless codecs», но часто это неоправданно.

Кодек как устройство может быть не только аппаратным, но и программным — например, кодеки, используемые для просмотра видео на персональных компьютерах, представляют собой наборы программных средств.

MPEG-2, MPEG-4/SP, MPEG-4/ASP, MPEG-4/AVC, H.263/264, H.265, MotionJPEG, Motion Wavelet — всё это названия кодеков, в той или иной мере используемых в отрасли безопасности по прямому назначению: сжимать видеосигнал для повышения эффективности его передачи и записи, а затем разжимать компрессированный поток для отображения видеоданных либо их дальнейшей обработки.

2. Анализ построения кодеков стандарта MPEG-4 и его модификаций

Стандарт MPEG-4 начал разрабатываться в начале 90-х годов для передачи мультимедийной информации (звук, малокадровое телевидение, графика, текст и т.д.) по узкополосным каналам Интернет со скоростями цифрового потока более 64 кбит/с. Изначально MPEG-4 не являлся телевизионным стандартом, а обрабатывал изображения разрешением в четверть формата CIF (352x288) при 15 кадрах/с. Однако новый стандарт очень быстро вышел за эти рамки и в октябре 1998 года была выпущена первая версия MPEG-4, а в декабре 1999 г. - вторая, получившая официальный статус международного стандарта ISO/IEC 14496.

В настоящее время стандарт содержит 22 части, из которых для 10 часть описывает правила кодирования сигналов цифрового телевидения. Такая часть называется MPEG-4-10 или H264 [44].

В отличие от своих предшественников, MPEG-4 не является просто технологией сжатия, хранения и передачи видео или аудио информации. MPEG-4 предназначен для отображения цифровых медиа-данных для трех областей:

- Цифровое телевидение;
- Интерактивные графические приложения;
- Интерактивное мультимедиа World Wide Web.

Если стандарты MPEG-1 и MPEG-2 описывают работу с уже готовыми видеокадрами, то MPEG-4 фактически задает правила организации объектно-ориентированной среды. Он имеет дело не просто с цифровыми потоками и массивами данных, а с медиа - объектами, и его основой является формирование на приемном конце и в реальном масштабе времени итогового изображения и звука из имеющихся в наличии отдельных объектов, основываясь на так называемом "схематическом описании события".

Любая видео сцена разделяется на объекты, и описываются отдельными элементарными потоками (ЭП). Объекты могут быть

натуральными — записанными с видеокамеры или микрофона, и синтетическими — синтезированными в компьютере. Такой подход имеет ряд преимуществ:

- более экономно расходуются биты для описания сцены;
- отдельные объекты легко использовать в других сценах;
- упрощается построение масштабируемых объектов;
- появляются широкие возможности взаимодействия пользователя с выбранным объектом, например, вывод дополнительной информации об объекте, изменение его параметров (цвета, текстуры, громкости звучания или языка), исключение объекта из сцены, создание пользователем новых сцен из объектов, полученных от разных источников или хранящихся в памяти кодека.

Для характеристики MPEG-4 с технических позиций, следует отметить, что стандарт содержит весь арсенал методов кодирования и компрессии пикселей изображения, используемых в MPEG-1 и MPEG-2. Кардинальным нововведением при компрессии видео в MPEG-4 является переход от деления изображения на квадратные блоки к операциям с объектами произвольной формы. К примеру, человек,двигающийся в кадре, будет восприниматься, и обрабатываться как единый отдельный объект, перемещающийся относительно неподвижного объекта - заднего плана.

При кодировании неподвижных изображений и текстур в MPEG-4 применяется очень эффективный алгоритм на основе вейвлет-преобразований, обеспечивающий кодирование объектов произвольной формы, и плавную масштабируемость по качеству картинки.

Кроме того предусмотрена возможность создания высокоскоростных видеопотоков - до 38,4 Мбит/с, а в студийных условиях - до 1,2 Гбит/с.

Особенности кодека MPEG-4 [44]:

- более высокая степень компрессии (по сравнению с MJPEG);
- требуется меньше дискового пространства для хранения архива;
- ниже требования к пропускной способности сети;

- требуются более мощные процессоры (т.е. более дорогие кодеры и декодеры);
- более сложное восстановление отдельного кадра;
- увеличивается задержка при передаче сигнала.

Структурная схема, отображающая работу кодера стандарта H.264 представлена на рис.3.1. соответственно при этом она содержит практически те же элементы, что и кодеки MPEG-1, MPEG-2.

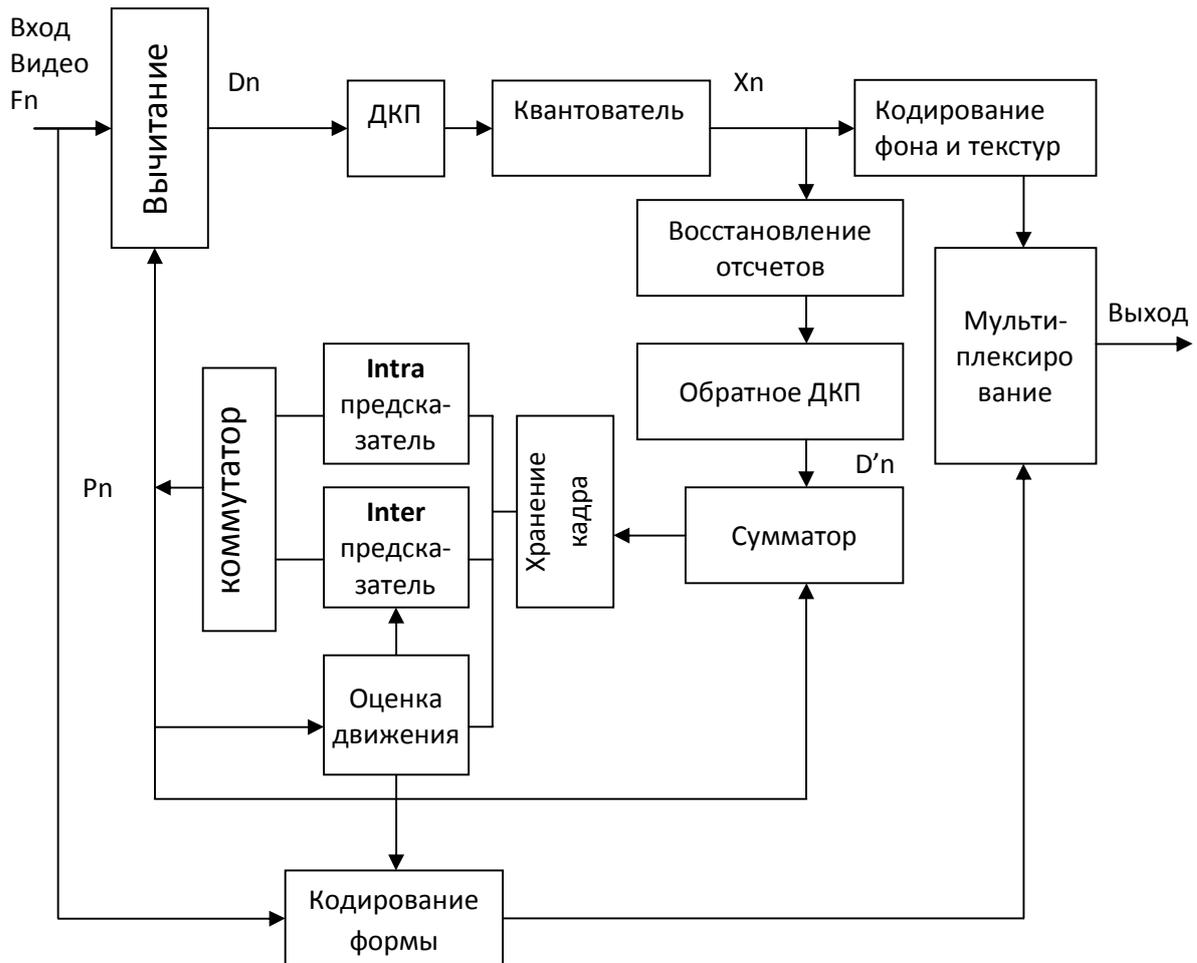


Рис.3.1. Обобщённая структурная схема видеокодера MPEG-4

Структура работы декодера приведена на рис.3.2. При декодировании сжатых изображений осуществляются обратные процедуры по отношению к кодеру. Декодер получает сжатый битовый поток и производит декодировку данных. В блоках формирования Intra- и Inter- прогнозов нет элементов, отвечающих за анализ кадра. Информация о конкретном режиме и способе его реализации передается внутри битового потока.

Основными шагами обработки аудио-видео информации в MPEG-4 являются [29]:

- разделение исходной картинки на различные элементы - "медиа объекты" (media objects);
- описание структуры и взаимосвязей этих объектов, позволяющее затем собрать их в единую видеозвуковую сцену;
- возможность для конечного пользователя внесения интерактивных изменений в сцену.
- Для достижения такого уровня гибкости представления все медиа объекты организуются в единую иерархическую структуру, включающую:
 - неподвижные изображения (например, фон);
 - натуральные видео объекты (например, человек);
 - аудио объекты (голос, связанный с этим человеком);
 - текст, связанный с данной сценой;
 - синтетические объекты, которых изначально в записываемой сцене не было, но они добавляются при демонстрации пользователю;
 - текст, связанный с синтетическим объектом, и преобразуемый в голос.

Технология сжатия видео MPEG распадается на две части: уменьшение избыточности видеoinформации во временном измерении, основанное на том, что соседние кадры, как правило, отличаются несильно, и сжатие отдельных изображений [29].

При обработке видео потока в MPEG-4 все видеоплоскости подразделяются на четыре типа (модель YUV 4:2:0):

I-кадры. Для обеспечения быстрого доступа, кодируется независимо от других кадров. Прежде всего, идёт разбиение на макроблоки 16×16 пикселей (I-макроблоки), которые разбиваются на 4 I-блока 8×8 для ДКП

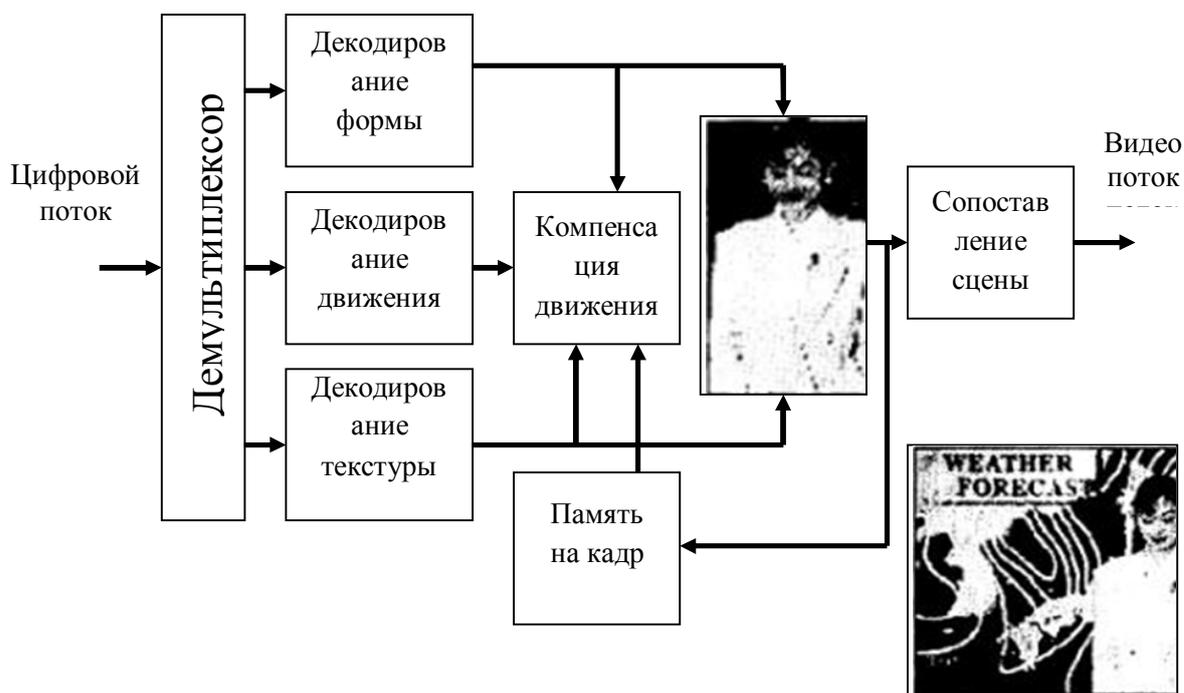


Рис.3.2. Декодирование видеокadra MPEG-4

3.2.1. Видео кодирование в стандарте MPEG-4-10 (H.264)

P-кадры кодируются с использованием информации о предшествующих I - или P - кадрах. Эти кадры аналогично дробятся на макроблоки 16×16 , однако, среди этих макроблоков присутствуют и I-макроблоки и i-макроблоки (I — intra, i — inter). *Inter-макроблоки* формируются в результате компенсации движения, где в предшествующих I- или P-кадрах ищется макроблок/блок, расположенный произвольно и максимально соответствующий текущему inter-макроблоку/блоку. При его нахождении формируется и вставляется в поток его двухкомпонентный вектор движения. Затем вычисляется попиксельная разность текущего и найденного макроблоков/блоков (ошибка предсказания) и к ней применяется ДКП. Для повышения эффективности компенсации движения вблизи границ кадра, выполняется процедура дополнения кадра с каждой стороны на величину одного макроблока (пространство заполняется цветом ближайших к границам пикселей). Таким образом, для каждого макроблока могут быть получены либо 1, либо 4 вектора движения, а выбор между вариантами

осуществляется исходя из минимума вклада текущего макроблока и его векторов движения в поток. Стандарт также предусматривает особый тип компенсации движения — компенсация с перекрытием: выполняется она только для блоков компоненты Y . Отличительная особенность этого метода состоит в составлении разностного блока по взвешенной суперпозиции трёх схожих с ним блоков в предыдущих I- или P-кадрах, а не по одному схожему блоку. Соответственно, за три вектора движения будут взяты: вектор для данного блока и два вектора для блоков, соседних к данному блоку в обрабатываемом макроблоке.

B-кадры отличаются от P-кадров тем, что при их кодировании, для компенсации движения используются не только предшествующие, но и последующие I- и P-кадры. Каждый макроблок этого кадра может быть предсказан макроблоком на предыдущий кадр, макроблоком на последующий кадр и суперпозицией этих макроблоков.

S - кадр имеет непосредственное отношение к спрайтам (в стандарте MPEG-4 спрайтом называют часть изображения, видимую на протяжении определённого интервала в видеопоследовательности). В роли спрайта чаще всего выступает фон (или задний план). При декодировании с использованием спрайтов, кадры частично восстанавливаются из отдельных областей спрайта путём отображения этих областей на ту или иную область декодируемого кадра с использованием перспективного преобразования. Спрайт хранится отдельно и кодируется подобно I-кадрам. Частным случаем использования спрайтов и являются S-кадры (GMC — global motion compensation); роль спрайта отводится в этом случае одной из предшествующих I- или P-кадров.

Как и кадры в MPEG-1, идущие подряд плоскости различных типов делятся на обособленно кодируемые группы; в начале группы должна всегда находиться I-плоскость, далее обычно B-плоскости чередуются с плоскостями P- и S-типа.

Как было указано выше, для высокоэффективного сжатия видеопотоков

MPEG-4 использует выделение объектов сцены, однако для выделения объектов на сложном неоднородном фоне довольно сложная задача, требующая больших вычислительных затрат. Поэтому в стандарте MPEG-4-10 (H264) часто применяется разбивка изображения на прямоугольные блоки переменного размера, что позволяет увеличить точность компенсации движения.

Видео кодек MPEG-4 при блочной компенсации движения использует блоки 16x16 пикселей. При кодировании, видео кодек подразделяет видео на I-слои и P-слои по системе intra. Основное их отличие заключается в том, что I-слой состоит только из макроблоков, закодированных по сэмплам этого же кадра, а P-слой может иметь макроблоки закодированные по ранее обработанным снимкам.

Когда в битовом потоке внутри P - слоя появляется сигнал пропущенного макроблока, тогда для этого макроблока дальнейшие данные в битовый поток не посылаются. Декодер реконструирует кадр из «первичных» (неизбыточных) снимков и игнорирует все лишние снимки. Однако, если первичный снимок был испорчен, декодер может попытаться заменить испорченную область декодированными данными из избыточного снимка.

Яркостная составляющая каждого макроблока (размером 16x16) может быть разбита четырьмя способами (16x16, 16x8, 8x16 или 8x8) или для блоков 8x8 (8x8, 8x4, 4x8 или 4x4) как это показано на рис.3.3.

Каждая из областей меньшего размера является частью макроблока. Возможность делить макроблок на блоки, а те в свою очередь на субблоки, позволяет получить большое количество комбинаций их сочетаний в пределах каждого макроблока. Для каждой части макроблока или субблока необходимо задать отдельный вектор движения. Каждый вектор кодируется и передается по каналу связи. Кроме того, необходимо закодировать в битовом потоке выбранный метод деления. Использование большого размера частей деления (16x16, 16x8, 8x16) означает, что для передачи векторов движения

потребуется меньше бит, но это также приводит к большой ошибке компенсации движения остаточных блоков в областях изображения с мелкими деталями.

Выбор малого размера частей деления (8x4, 4x4 и т.д.) приводит к малой ошибке позиционировании остаточных блоков после компенсации движения, но это приводит к увеличению объема передаваемой информации о векторах движения и способов деления. Таким образом, выбор размера деления существенным образом влияет на степень сжатия видеоизображения.

Обычно, данные ошибки происходят при приеме информации со спутника. Когда, по причине особых метеоусловий или других факторов пропадает сигнал на определенный промежуток времени, то часть блоков от нового кадра передается, а часть отсутствует. Новые блоки размещаются согласно координатам, а вместо отсутствующих вставляются блоки предыдущего кадра.

Там, где имеется существенное расхождение между опорным и текущим кадром (например, при смене сцен), можно кодировать макроблок без компенсации движения. Таким образом, кодер выбирает между режимами intra и inter для каждого макроблока, рис.3.3.

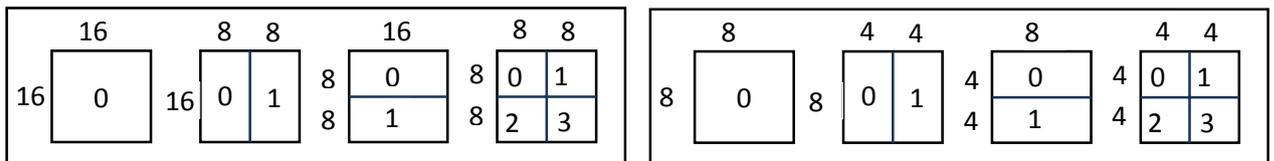


Рис.3.3. Пример разбиения макроблоков на блоки меньшего размера.

Данный кодер включает две ветви обработки видеопотока:

- канал прямого кодирования, где обработка данных происходит слева направо;
- канал реконструкции видеоизображения, где обработка данных происходит справа налево.

На вход кодера поступает кадр - F_n . Обработка кадра происходит с помощью макроблоков, размером 16x16 пикселей исходного изображения.

Каждый макроблок может быть обработан в двух режимах: *Intra* (без компенсации движения) или *Inter* (с компенсацией движения). В любом режиме прогноз макроблока - P_n формируется на основе восстановленного кадра.

В режиме *Intra* предсказание формируется из выборок текущего кадра N , предварительно закодированных и восстановленных (F''_n). В режиме *Inter* предсказание формируется с учетом изменений, произошедших в текущем кадре по сравнению с предыдущими кадрами (или последующими). Кадры, служащие для прогноза, предварительно кодируются и потом восстанавливаются. В блоках формирования *Intra*-предсказания и *Inter*-предсказания происходит выбор наиболее подходящего способа в зависимости от типа кодируемого кадра. Полученный прогноз P_n вычитается из текущего макроблока. В результате вычисляется макроблок остаточных коэффициентов D_n . Этот макроблок поступает в преобразователь, где происходит частотное преобразование остаточных коэффициентов на основе прямого ДКП. Таким образом, значительно уменьшается вычислительная сложность алгоритма, но в процессе сжатия появляются дополнительные искажения. Частотные коэффициенты квантуются (масштабируются), что позволяет регулировать сжатие видеоданных с потерями. Полученный набор преобразованных и квантованных коэффициентов X_n будет исходным для обратного канала реконструкции данных. Закодированные коэффициенты вместе с дополнительной информацией, требуемой для правильного декодирования макроблока, составляют сжатый битовый поток данных (bitstream) абстрактного сетевого уровня (NAL).

В канале реконструкции вначале происходит обратное квантование, затем обратное частотное преобразование. В итоге получают восстановленные разностные коэффициенты D'_n . Они суммируются с прогнозируемым кадром P_n , и это позволяет получить восстановленный кадр F'_n .

Для повышения эффективности сжатия видеопотка в MPEG-4 используется кодирование видеообъектов сложной формы. Например, в качестве видеообъекта может быть взята область изображения, отличающаяся от окружения яркостью или цветом. Эта область может

перемещаться и деформироваться. В этом случае при формировании предсказанного изображения с компенсацией движения смещаются не прямоугольные макроблоки, а выделенные области, которые к тому же могут изменять свою форму. При этом ошибка предсказания оказывается значительно меньше, и объем информации, содержащейся в разности предсказанного и настоящего изображений очередного кадра, существенно уменьшается. В кодере (рис.3.1), такой вариант предсказания выполняется в блоке Кодирования формы. Форма видеообъекта описывается матрицей. *При бинарном кодировании*, используемому альфа карта, элементы матрицы, соответствующие элементам видеообъект, приравниваются единице, а соответствующие элементам изображения, находящимся вне видеообъекта, - нулю. *При градационном кодировании* элементы матрицы принимают большее число значений, что позволяет описывать такие свойства объекта, как прозрачность. Элементы указанной матрицы определяются и кодируются в Кодере формы, после чего они включаются через мультиплексор в выходной поток данных.

Несмотря на то, что данные о форме видеообъекта занимают значительно больше двоичных символов, чем простой вектор движения, но общий выигрыш в уменьшении объема передаваемой информации по сравнению с MPEG-1,2 оказывается существенным.

Для сжатия изображений неподвижного фона и текстур протяженных объектов используется Кодер фона и текстур, основанный на вейвлет-преобразовании. Этот метод обеспечивает высокие степени сжатия и многоступенчатую масштабируемость по пространственному разрешению. При этом не изменяющийся или почти не изменяющийся задний план изображения может передаваться как так называемый *спрайт* (sprite). Полное изображение спрайта передается один раз. Затем передаются только 8 параметров глобального движения, описывающих панорамирование, т.е. перемещение камеры относительно заднего плана (рис.3.4).

Для того чтобы на границе вставляемого кадра не было разрывов, используется алгоритм сглаживания. Неизвестные области в панораме заполняются при помощи алгоритма пространственного заполнения. Данный метод компенсации движения является наиболее широко используемым, потому что ни одно видео не содержит сцен с абсолютно статичным фоном без движений видеокамеры. Поэтому, можно выделить данный метод среди остальных как наиболее важный и дающий существенные результаты по сжатию.

Такой метод представления данных позволяет помещать и перемещать медиа объекты в любое место сцены, трансформировать объекты и изменять их геометрические размеры, собирать из отдельных объектов составные объекты и проводить с ними различные операции, изменять вид объектов (цвет, текстуру) и манипулировать объектами (перемещать, разворачивать и т. д.), изменять композицию сцены и отображать ее под разными ракурсами.

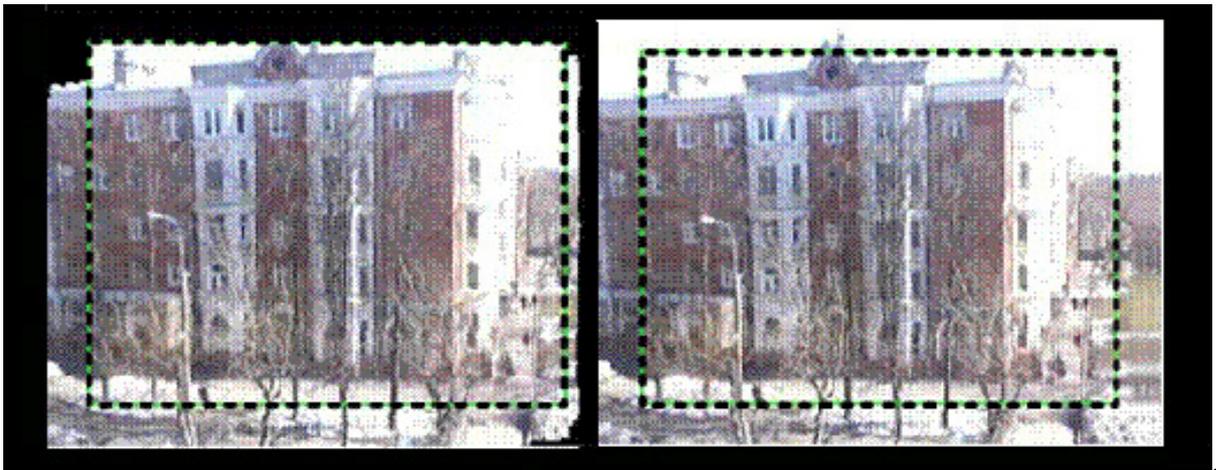


Рис.3.4.Панорамное изображение: (а) – без заполнения неизвестных областей; (б) – с заполнением неизвестных областей

3.2.3. Стандарт сжатия H.265/MPEG-HEVC

H.265 или HEVC (англ. High Efficiency Video Coding — высокоэффективное кодирование видеоизображений) — формат видеосжатия, с применением более эффективных алгоритмов по сравнению с

H.264/MPEG-4 AVC. Рекомендация стандарта разработана в связи с растущей потребностью в более высокой степени сжатия движущихся изображений для самых разных приложений, таких как потоковая передача в интернете, передача данных, видеоконференц-связь, цифровые запоминающие устройства и телевизионное вещание [45].

При кодировании видео в HEVC применяется такой же «гибридный» подход, что и во всех современных кодеках, начиная с H.261. Он заключается в применении внутри- и межкадрового (Intra-/Inter-) предсказания и двумерного кодирования с преобразованием (рис.3.5).

В кодере HEVC каждый видеокادر делится на блоки. Первый кадр видеопоследовательности кодируется с использованием только внутрикадрового предсказания, то есть применяется пространственное предсказание ожидаемого уровня отсчёта внутри кадра по соседним отсчётам, при этом отсутствует зависимость от других кадров. Для большинства блоков всех остальных кадров последовательности, как правило, используется режим межкадрового временного предсказания. В режиме межкадрового предсказания на основании данных о величине отсчётов опорного кадра и вектора движения оцениваются текущие отсчёты каждого блока. Кодер и декодер создают идентичные межкадровые предсказания путем применения алгоритма компенсации движения с помощью векторов движения и данных выбранного режима, которые передаются в качестве дополнительной информации.

Разностный сигнал предсказания, который представляет собой разницу между опорным блоком кадра и его предсказанием, подвергается линейному пространственному преобразованию. Затем коэффициенты преобразования масштабируются, квантуются, применяется энтропийное кодирование, и затем передаются вместе с информацией предсказания.

Кодер в точности повторяет цикл обработки декодером так, что в обоих случаях будут генерироваться идентичные предсказания последующих данных. Таким образом, преобразованные квантованные коэффициенты

подвергаются обратному масштабированию и затем обратному преобразованию, чтобы повторить декодированное значение разностного сигнала. Разность затем добавляется к предсказанию, и полученный результат фильтруется для сглаживания артефактов, полученных делением на блоки и при квантовании. Окончательное представление кадра (идентичное кадру на выходе декодера) хранится в буфере декодированных кадров, которое будет использоваться для прогнозирования последующих кадров. В итоге, порядок кодирования и декодирования обработки кадров часто отличается от порядка, в котором они поступают из источника.

Предполагается, что видеоматериал на входе кодера HEVC имеет прогрессивную развёртку. В HEVC не представлено явных функций кодирования чересстрочной развёртки, так как чересстрочная развёртка не используется в современных дисплеях и имеет всё меньшее распространение. Тем не менее, в HEVC были представлены метаданные, позволяющие указать кодеру, что было закодировано видео с чересстрочной развёрткой в одном из двух режимов: в виде отдельных изображений, как два поля (четные или нечетные строки кадра) или весь кадр целиком. Этот эффективный метод обеспечивает кодирование видеосигнала с чересстрочной развёрткой, минуя необходимость нагружать декодеры поддержкой специального процесса декодирования.

Кодек H.265 HEVC имеет достаточно серьезные перспективы. Существующие линии передачи со скоростью 24Mbit, которые работали с AVCHD смогут без, каких-либо изменений и доработок доставлять видео с качеством 4K и ProRes 4444 на рекордеры с обычными картами SD.

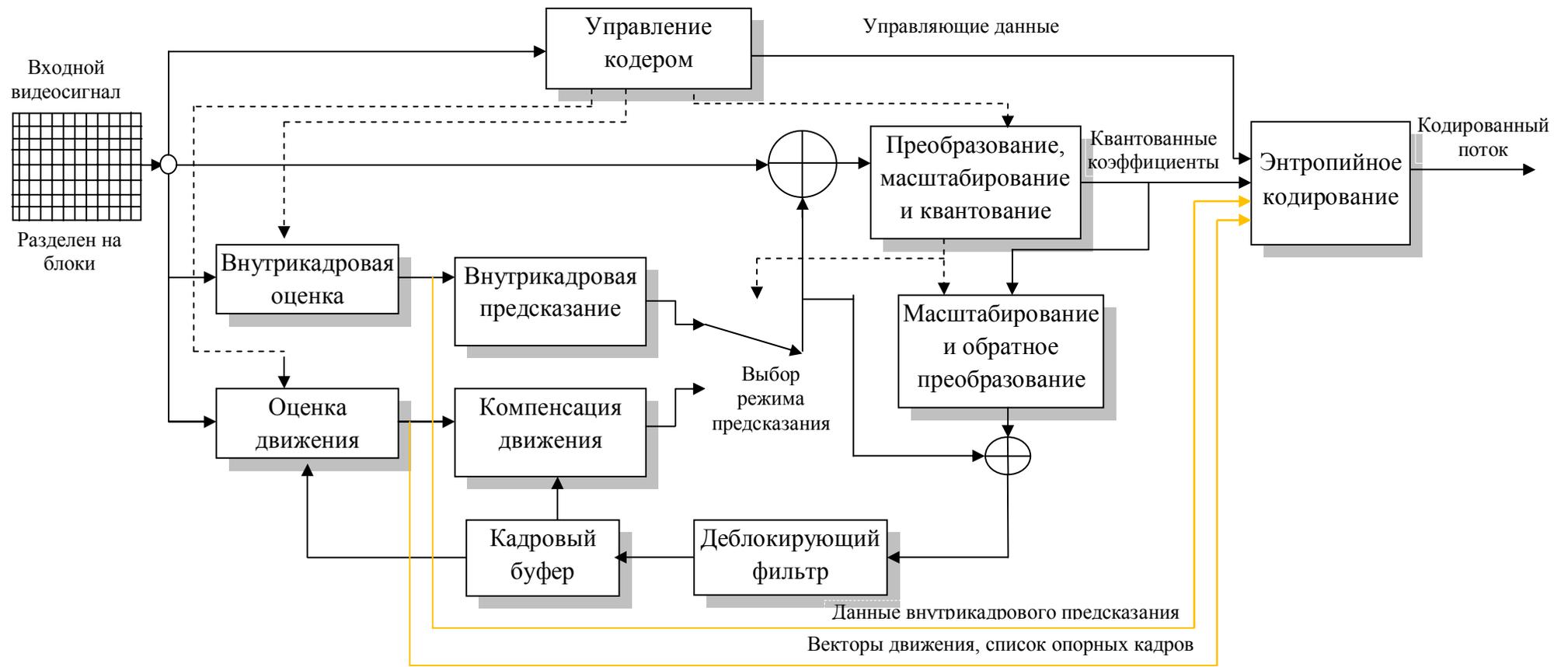


Рис.3.5. Схема типowego кодера HEVC

Новый кодек позволит также снимать видео с разрешением 4К на смартфоны без значительной потери качества при низких потоках и это мы уже предусмотрено на некоторых современных телефонах.

Маловероятно, что программное обеспечение кодека H.265 сможет работать на современных процессорах, применяемых в бытовых камерах, даже на самых мощных. Математика и алгоритмы кодека нуждаются в гораздо большем количестве энергии, чем для H.264.

Технология HEVC также может использоваться и для фотографии. Вместо массивного файла 8 бит JPEG, новый кодек может предложить сохранить снимок с гораздо меньшим размером и поднять дискретизацию с 8 бит до 10 бит, что придаст снимку плавные градации яркости и цветов.

H.265 поддерживает разрешение до 8К, которое в тестовом режиме уже используется в Японии для вещания. Новый кодек является важным ключом к переходу на более высокое качество изображения.

3. Анализ видео кодеков на основе вейвлет-преобразований

Как уже рассматривалось ранее применение вейвлет – преобразований, основанных на всплескоподобных базисных функциях, позволяет до 1,5-2 раза увеличить величину сжатия изображений без ухудшения их визуального качества [23], как показано на рис.3.6.

К настоящему времени вейвлеты применяются для сжатия опорных кадров в кодеках стандартах MPEG-4 и их модификаций, но для обработки предсказанных кадров в них используются ДКП. Также существуют кодеки работающие в стандарте MJPEG, которые используются для видеонаблюдения и других прикладных задач. Однако для вещательного телевидения вейвлет кодеки пока еще не получили широкого применения и находятся в стадии разработки. На сегодняшний день по материалам доступных источников известны 2 гибридных видеокодека с компенсацией движения: Dirac и Rududu [41],

способные конкурировать с кодеками вещательного стандарта H.264, которые рассмотрим более подробно.



а) $K_{сж} = 3,5$ (ДКП)



б) $K_{сж} = 5,2$ (ВП)

Рис.3.6. Результаты декомпрессии изображений после сжатия на основе ДКП и ВП при одинаковых параметрах кодирования

Rududu. Rududu является кодеком применяемый для современного видео-контента. Сжатия видео применяется, чтобы уменьшить размер файла (или пропускную способность канала) видео.

Цель проекта Rududu является создание быстрого и состоятельного видеокодека. Проект в настоящее время в стадии разработки, в нем можно кодировать видео, с приемлемым качеством изображения. В рамках видеокодека, была разработана утилита для частичной проверки сжатия изображения. Данная утилита осуществляет сжатие лучше, чем JPEG2000. Если данную разработку использовать для тестирования, она может быть полезна для исследования сжатия изображения.

Основное различие между Rududu и многими кодеками, используемых сегодня является то, что, к примеру, преобразования MPEG, H26x и другие для сжатия неподвижных и движущихся изображение используют ДКП (дискретно - косинусное преобразование) в то время как Rududu использует ДВП (дискретное вейвлет-преобразование).

Rududu использует ДВП, потому что оно дает лучшие результаты при сжатии естественных неподвижных изображений, по сравнению с ДКП, но для видео в них большой разницы нет.

Ограничения кодека:

- Размеры видео должно быть кратно 8 в обоих направлениях, это ограничение не будет удален;
- Кодек работает только на процессорах с MMX.

Интерфейсы кодека:

- VFW (Video For Windows) в компрессоре и режиме декомпрессии;
- DirectShow в режиме декомпрессии.

Как и кодеки MPEG, Rududu это гибридный кодек. В данном кодеке часть битов используется для отправки векторов движения для построения прогнозируемого кадра, используя последнее декодированное изображение.

Предсказанный кадр строится с помощью метода компенсации с перекрытием движения блоков: блоки имеют гладкие границы, чтобы избежать создание искусственных высоких частот на реконструированном кадре [41]

Далее вычисляется ошибка изображения, то есть разница между предсказанием и текущим кадром. Затем изображение далее преобразуется с использованием вейвлет – преобразования, а коэффициент вейвлета квантуется и кодируется.

Максимальное количество предсказываемых кадров между последовательностями для ключевых кадров должно быть установлено на 10 секунд (так 300 кадров с частотой 30 кадров/с), чтобы иметь быстрый доступ к случайной точки в файле. Меньше ключевых кадров улучшает немного коэффициент сжатия, большое количество ключевых кадров является бесполезным. Rududu вводит ключевые кадры, когда это необходимо.

Проект Dirac-Pro. Кодек *Dirac Pro* представляет собой коммерческий вариант видеокodeка Дирак, разработанного английской исследовательской группой BBC (British Broadcasting Corporation) для исследовательских целей

в области применения вейвлет преобразований [24]. Особенностью данных кодеков является применение ВП как для сжатия опорных кадров, так и для межкадрового сжатия видеоданных. Кроме того это гибридный видеокодек с компенсацией движения, соответствующий стандартам MPEG - 4. Движение изображения отслеживается в смежных кадрах, а полученная информация о движении используется для формирования предсказания последующего кадра, которое квантуется и подвергается статистическому кодированию. Термин «гибридный» отражает тот факт, что используются как вейвлет преобразование, так и компенсация движения, благодаря которой устраняется временная избыточность, что значительно увеличивает коэффициент сжатия видео потока. В кодеке Dirac применена более гибкая и эффективная разновидность статистического кодирования — так называемое арифметическое кодирование.

Профессиональный профиль кодека Dirac Pro был реализован в сентябре 2008 года [47]. Профиль рассчитан на применение в монтажных, архивных и вещательных системах, где существуют повышенные требования к качеству сжатого видео. Алгоритм сжатия имеет открытый код и распространяется по лицензии Royalty Free.

Согласно спецификации поддерживаются форматы от QSIF525 (176 × 120 пикселей) вплоть до последних UHD TV 8K-50 (7680 × 4320 пикселей, 50 кадров/с) и цифровых киноформатов 2K и 4K D-Cinema.

Кодек поддерживает значительное число форматов:

- кадровая частота — 23,97...60 Гц
- дискретизация — 4:2:0, 4:2:2, 4:4:4 и RGB
- разрядность — 8, 10, 12 (до 16)
- поддержка чересстрочного и прогрессивного форматов.

Кодек Dirac Pro был направлен в SMPTE для стандартизации как VC-2 Codec. В 2010 году SMPTE приняли в качестве стандарта видео компрессии VC-2 [24].

- SMPTE 2042-1-2009 VC-2 Video Compression

- SMPTE 2042-2-2009 VC-2 Level Definitions
- RP (Recommended Practices) 2047-1-2009 — VC-2 Mezzanine Level Compression of 1080P High Definition Video Sources
- SMPTE 2047-2-2010 Carriage of VC-2 Compressed Video over HD-SDI — Передача сжатого видео VC-2 по интерфейсу HD-SDI
- RP 2042-3-2010 — VC-2 Conformance Specification

Вторая реализация кодека Dirac, названная «Schrödinger» (Шрёдингер) была нацелена на создание переносимых библиотек, а также плагинов для использования Dirac внутри Ogg контейнеров. Проект был назван в честь австрийского физика Эрвина Шрёдингера. 22 февраля 2008 года была выпущена финальная версия Schrödinger 1.0.0. Данная версия может декодировать видео с разрешением HD720x576/25 в режиме реального времени на ноутбуке с процессором класса Core Duo [17]

Имеется версия [24], написанная с использованием технологии CUDA (Compute Unified Device Architecture), ускоряющей вычисления путём использования возможностей видеокарты, за счёт чего скорость декодирования в несколько раз больше.

CUDA— программно-аппаратная архитектура параллельных вычислений, которая позволяет существенно увеличить вычислительную производительность благодаря использованию графических процессоров фирмы Nvidia. Архитектура CUDA даёт разработчику возможность по своему усмотрению организовывать доступ к набору инструкций графического ускорителя и управлять его памятью.

Первоначально Dirac был создан как кодек, предназначенный для распределения вещательного или Интернет-контента [30]. Его эффективность не отличалась от кодека AVC/H.264 или даже превосходила его по эффективности. Однако в процессе работы возобладала идея создать очень гибкую масштабируемую систему сжатия, пригодную не только для задач распределения, но и для производства, обработки и архивного хранения контента. В такой системе должны использоваться только I-псевдокадры и

профили с длинными GOP (Group Of Pictures), обеспечивающими разрешение от 2К/ТВЧ до QSIF. Кроме того, эффективный способ формирования уровней позволяет работать с разрешением 4К посредством дополнения базового уровня 2К/ТВЧ «улучшающим» уровнем.

Также как и другие распространённые кодеки, такие как MPEG-4 Part 2 или WMV 7, Divx может сжимать любой размер картинки от низкого разрешения QCIF (176x144 пикселей) до HDTV (1920x1080) и выше [48].

4. Сравнительная оценка различных методов кодирования видео сигналов ТВ программ

Как было указано выше, существует довольно много различных видеокодеков, используемых для сжатия мультимедийного контента ТВ программ, отличающихся быстродействием, качественными показателями и техническими характеристиками. Для практической оценки эффективности работы видеокодеков проведем исследования наиболее распространенных и перспективных разработок [32].

Эксперименты проводились на ПЭВМ со следующими значимыми характеристиками:

- Процессор: AMD A8-3500 APU with Radeon (tm) HD Graphics 1.50 GHz;
- ОЗУ: 8,00 Гб, DDR3;
- ОС Windows 7 x 64.

Для оценки эффективности и качественных восстановленных изображений было использовано 4 тестовых видеосюжета формата телевидения стандартной четкости формата 720x576 (рис.3.7):

- №1 – видеосюжет «Медведь»;
- №2 – видеосюжет «Формула-1»;
- №3 – видеосюжет «Звездные войны»;
- №4 – видеосюжет «Оружие».

Данные видеопоследовательности содержат по 150 кадров и обрабатывались кодеками стандартов **H264**, **H265** и **Dirac-pro**. При этом оценка качества восстановленных изображений производилась как визуально, так и использованием объективной метрики MSE на битрейтах 2800 кб/с, 1400 кб/с, 700 кб/с.



Медведь



Формула-1



Звездные войны



Оружие

Рис.3.7. Фрагменты тестовых видеосюжетов различных жанров

В результате экспериментов оценивался коэффициент сжатия видеопотока, средне квадратичной ошибки (MSE) восстановленного изображения и временных затрат на кодирование изображений. Полученные результаты исследований сведены в таблицы 3.1, 3.2 для кодека H.264, 3.3, 3.4 для кодека H.265 и 3.5, 3.6 для кодека Dirac-Pro. При этом визуальное качество исходных и восстановленных после декомпрессии изображений

можно оценить по рис.3.8-3.11 для H.264, рис.3.12-3.15 для H.265 и рис.3.16-3.19 для кодека Dirac-pro. Кроме того, сравнительное быстродействие кодеков и их коэффициенты сжатия представлены в гистограммах на рис.3.20, 3.21.

Следует отметить, что помимо объективной оценки качества изображений использовалась субъективная оценка на основе нашего восприятия.

Международный союз электросвязи (ITU) разработал "Рекомендацию ITU-R BT.500" на формальное субъективное тестирование качества изображения - "Методика субъективной оценки качества телевизионного изображения" [49].

В Рекомендации описано несколько методов субъективного измерения качества видеоизображения. В данном исследовании использовалась один из методов субъективного измерения под названием Double Stimulus Impairment Scale ("Шкала деградации с двумя стимулами") заключается в следующем. Участники просматривают эталонную сцену, а затем - ту же сцену после ее обработки в тестируемой системе. Каждый участник оценивает деградацию изображения сцены после обработки одним числом по пятибалльной шкале:

- 1 - очень раздражает;
- 2 - раздражает;
- 3 - слегка раздражает;
- 4 - заметно, но не раздражает;
- 5 - незаметно.

Баллы выставляются для некоторого количества различных сцен.

Таблица 3.1

Сравнительные затраты времени на сжатие тестовых сюжетов Кодеком H.264

	Сюжет	Исходный размер файла (Кб)	Число кадров	Битрейт кадра, Кбит/с								
				700			1400			2800		
				Ткод (с)	MSE (дБ)	DSIS	Ткод (с)	MSE (дБ)	DSIS	Ткод (с)	MSE (дБ)	DSIS
1.	Белый медведь	186538	150	7	18,6	4	8	18,2	3	9	18,5	3
2.	Звездные войны	197898	150	10	19,3	4	10,5	16,7	4	11	16,6	4
3.	Оружие	197898	150	10	24,4	3	10,5	32,9	3	9	34,4	3
4.	Формула - 1	197898	150	8,5	24,9	4	9	23,5	4	9,5	25	4

Таблица 3.2

Сравнительные коэффициенты сжатия тестовых сюжетов Кодеком H.264

	Сюжет	Исходный размер файла (Кб)	Битрейт кадра, Кбит/с					
			700		1400		2800	
			Коэф. сжат. (раз)	Получ. Объем (Кб)	Коэф. сжат. (раз)	Получ. Объем (Кб)	Коэф. сжат. (раз)	Получ. Объем (Кб)
1.	Белый медведь	186538	266	700 кб	150	1 240	79,3	2 350
2.	Звездные войны	197898	291,4	679 кб	163	1 210	85	2 330
3.	Оружие	197898	296	668 кб	170	1 160	92	2 160
4.	Формула - 1	197898	298	663 кб	170	1 160	90	2 190



а)



б)



в)



г)

Рис.3.8. а) Исходное изображение сюжета «Белый медведь» и результаты сжатия кодеком H.264 при битрейтах, Кбит/с:
б) 700; в) 1400; г) 2800



а)



б)

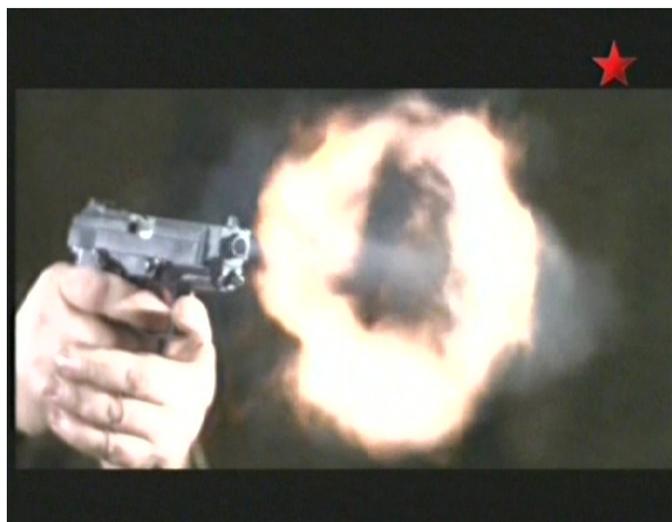


в)



г)

Рис.3.9. а) Исходное изображение сюжета «Звездные войны» и результаты сжатия кодеком H.264 при битрейтах, Кбит/с:
б) 700; в) 1400; г) 2800



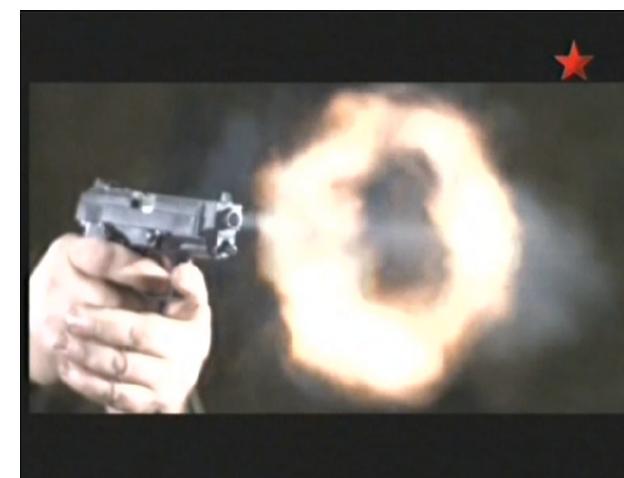
а)



б)



в)



г)

Рис.3.10. а) Исходное изображение сюжета «Оружие» и результаты сжатия кодеком H.264 при битрейтах, Кбит/с: б) 700; в) 1400; г) 2800



а)



б)



в)



г)

Рис.3.11. а) Исходное изображение сюжета «Формула - 1» и результаты сжатия кодеком H.264 при битрейтах, Кбит/с:
б) 700; в) 1400; г) 2800

Таблица 3.3

Сравнительные затраты времени на сжатие тестовых сюжетов кодеком H.265

	Сюжет	Исходный размер файла (Кб)	Число кадров	Битрейт кадра, Кбит/с								
				700			1400			2800		
				Ткод (мин)	MSE (дБ)	DSIS	Ткод (мин)	MSE (дБ)	DSIS	Ткод (мин)	MSE (дБ)	DSIS
1.	Белый медведь	186538	150	1,42	16,7	4	2,04	17,8	3	2,34	18,4	4
2.	Звездные войны	197898	150	1,31	16,5	3	1,50	16,6	4	2,14	16,6	4
3.	Оружие	197898	150	1,40	27,9	4	1,59	34,3	4	2,20	22,7	4
4.	Гонщик F - 1	197898	150	1,40	23,9	4	2,04	24,6	4	2,34	24,9	5

Таблица 3.4

Сравнительные коэффициенты сжатия тестовых сюжетов кодеком H.265

	Сюжет	Исходный размер файла (Кб)	Битрейт кадра, Кбит/с					
			700		1400		2800	
			Коэф. сжат. (раз)	Получ. объем (Кб)	Коэф. сжат. (раз)	Получ. объем (Кб)	Коэф. сжат. (раз)	Получ. объем (Кб)
1.	Белый медведь	186538	280	664	162	1 150	86,3	2 160
2.	Звездные войны	197898	300	661	175,1	1 130	94 раз	2 100
3.	Оружие	197898	295	670	172	1 150	92	2 150
4.	Гонщики F-1	197898	305	648	176,7	1 120	94,2	2 100



а)



б)



в)



г)

Рис.3.12. а) Исходное изображение сюжета «Белый медведь» и результаты сжатия кодеком H.265 при битрейтах,

Кбит/с: б) 700; в) 1400; г) 2800



а)



б)

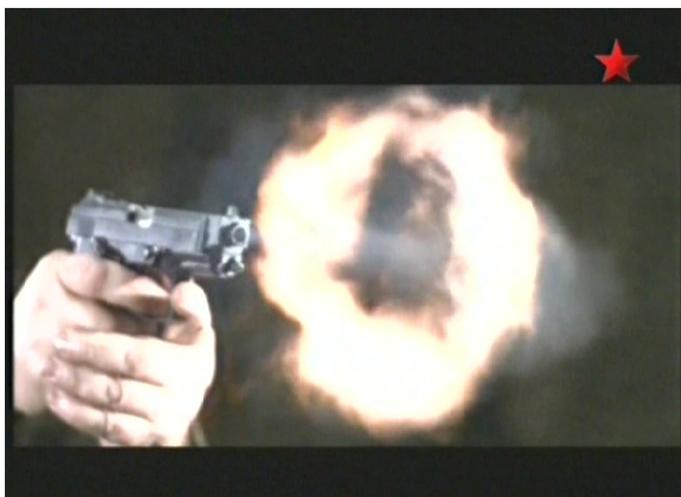


в)

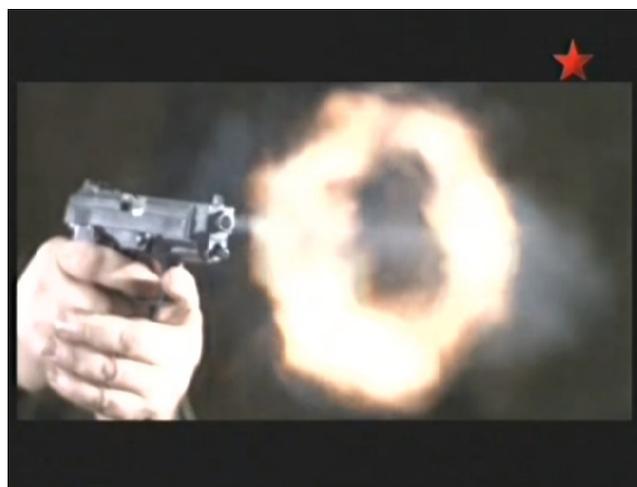


г)

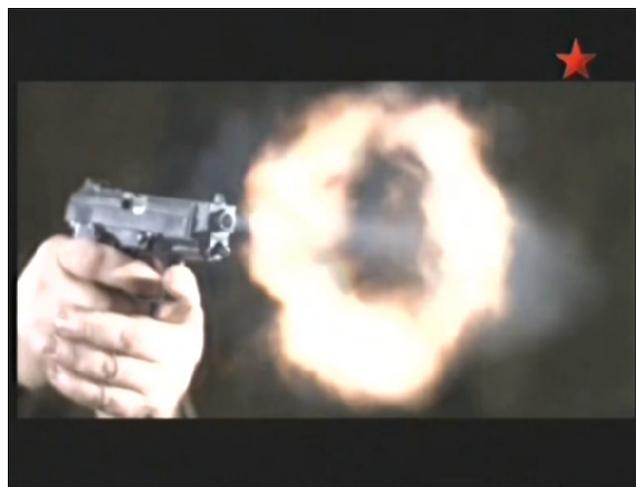
Рис.3.13. а) Исходное изображение сюжета «Звездные войны» и результаты сжатия кодеком H.265 при битрейтах, Кбит/с: б) 700; в) 1400; г) 2800



а)



б)



в)



г)

Рис.3.14. а) Исходное изображение сюжета «Оружие» и результаты сжатия кодеком H.265 при битрейтах, Кбит/с: б) 700; в) 1400; г) 2800



а)



б)



в)



г)

Рис.3.15. а) Исходное изображение сюжета «Формула - 1» и результаты сжатия кодеком H.265 при битрейтах, Кбит/с:

б) 700; в) 1400; г) 2800

Таблица 3.5

Сравнительные затраты времени на сжатие тестовых сюжетов Кодеком Dirac

	Сюжет	Исходный размер файла (Кб)	Число кадров	Битрейт кадра, Кбит/с								
				700			1400			2800		
				Ткод (мин)	MSE (дБ)	DSIS	Ткод (мин)	MSE (дБ)	DSIS	Ткод (мин)	MSE (дБ)	DSIS
1.	Белый медведь	186538	150	0,7	27.33	3	0,9	26.3	3	1,2	24.5	4
2.	Звездные войны	197898	150	0,6	16.9	3	0,7	28,6	4	1,05	26,5	4
3.	Оружие	197898	150	0,7	25,7	3	0,9	26,2	3	1,05	26.2	3
4.	Гонщик F - 1	197898	150	0,6	25.0	4	0,8	26.9	4	0,9	23,2	4

Таблица 3.6.

Сравнительные коэффициенты сжатия тестовых сюжетов Кодеком Dirac

	Сюжет	Исходный размер файла (Кб)	Битрейт кадра, Кбит/с					
			700		1400		2800	
			Коэф. сжатия (раз)	Получ. объем (Кб)	Коэф. сжатия (раз)	Получ. объем (Кб)	Коэф. сжатия (раз)	Получ. объем (Кб)
1.	Белый медведь	186538	86,1	2 167	86,6	2 153	85,1	2 194
2.	Звездные войны	197898	92,7	2 147	92,7	2 ,148	90,7	2 195
3.	Оружие	197898	119	1 667	119	1 669	92,6	2 149
4.	Формула - 1	197898	88,2	2 250	88,2	2 200	88	2 200



а)



б)



в)



г)

Рис.3.16. а) Исходное изображение сюжета «Белый медведь» и результаты сжатия кодеком Dіgас при битрейтах, Кбит/с:
б) 700; в) 1400; г) 2800



А)



Б)

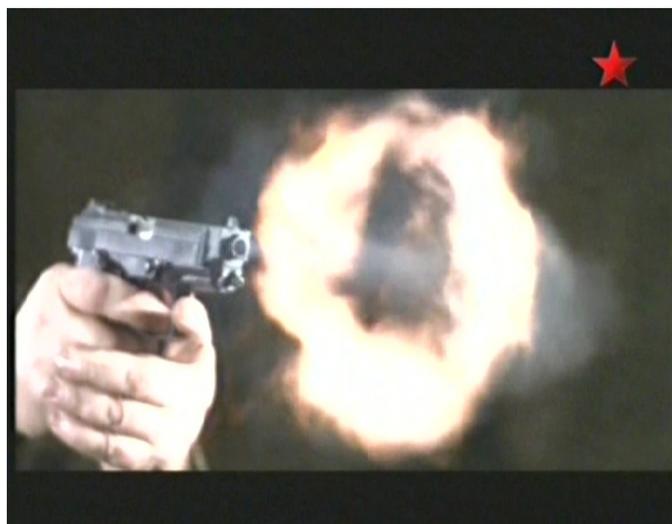


В)

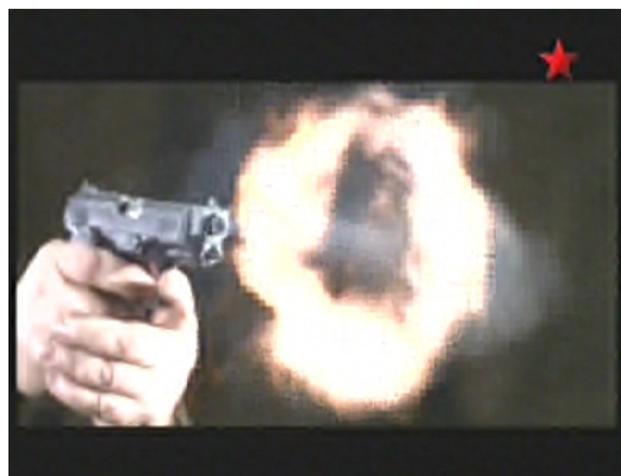


Г)

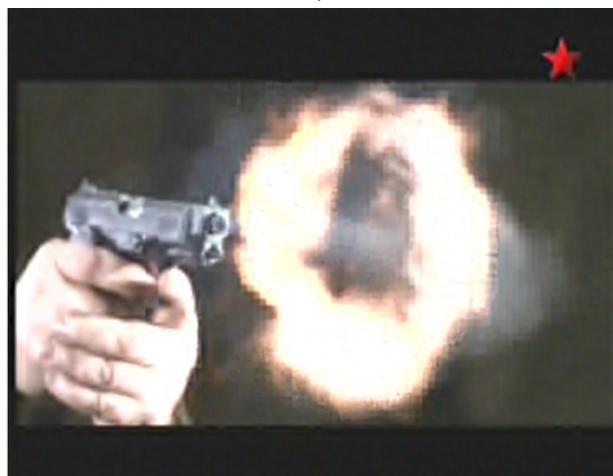
Рис.3.17. а) Исходное изображение сюжета «Звездные войны» и результаты сжатия кодеком Divac при битрейтах, Кбит/с: б) 700; в) 1400; г) 2800



а)



б)



в)



г)

Рис.3.18. а) Исходное изображение сюжета «Оружие» и результаты сжатия кодеком Divac при битрейтах, Кбит/с: б) 700; в) 1400; г) 2800



а)



б)



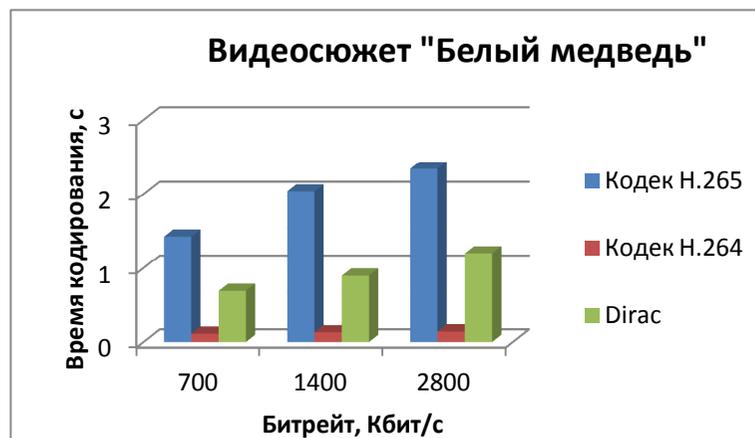
в)



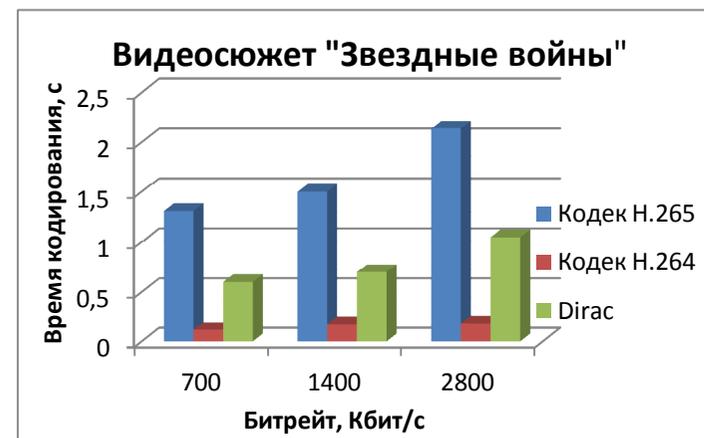
г)

Рис.3.19. а) Исходное изображение сюжета «Формула - 1» и результаты сжатия кодеком H.264 при битрейтах, Кбит/с:
б) 700; в) 1400; г) 2800

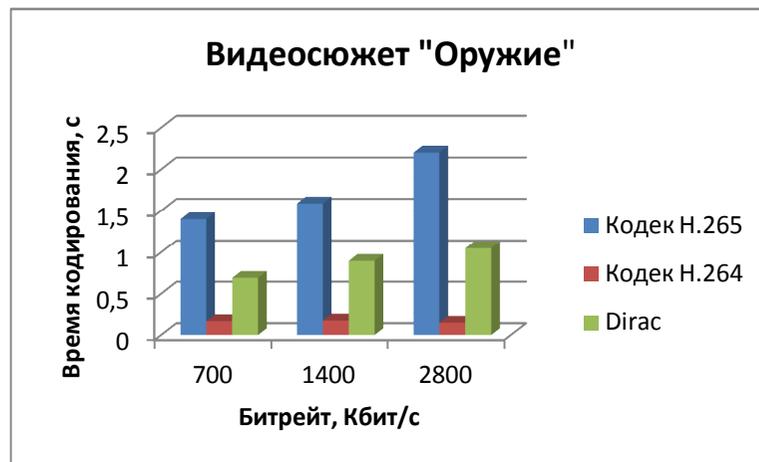
Гистограммы сравнение временных затрат и коэффициентов сжатия для четырех видеосюжетов



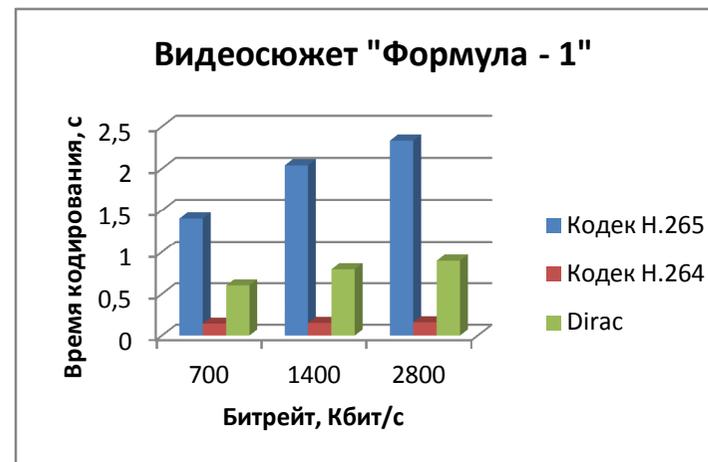
а)



б)

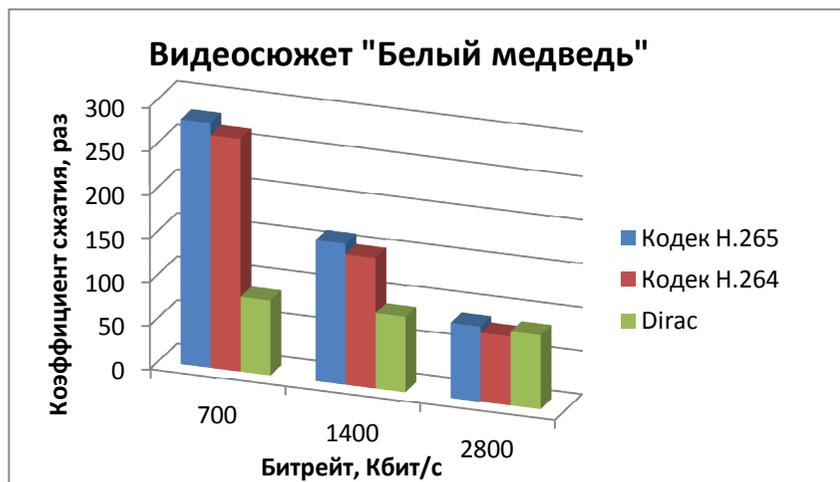


в)

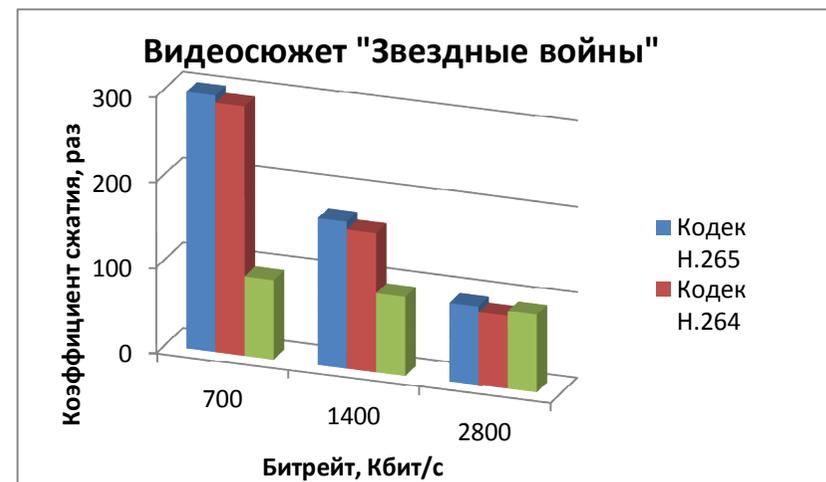


г)

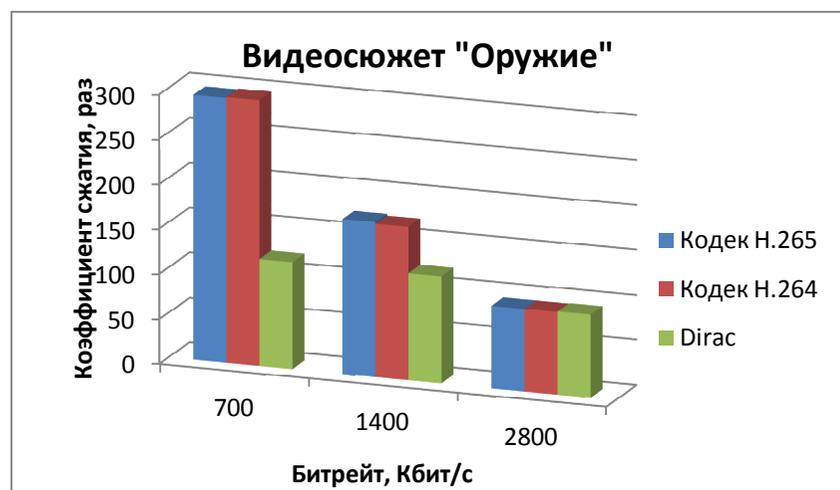
Рис.3.20. Сравнение временных затрат сжатия исследованных кодеков для соответствующих сюжетов: а) Белый медведь, б) Звездные войны, в) Оружие, г) Формула - 1



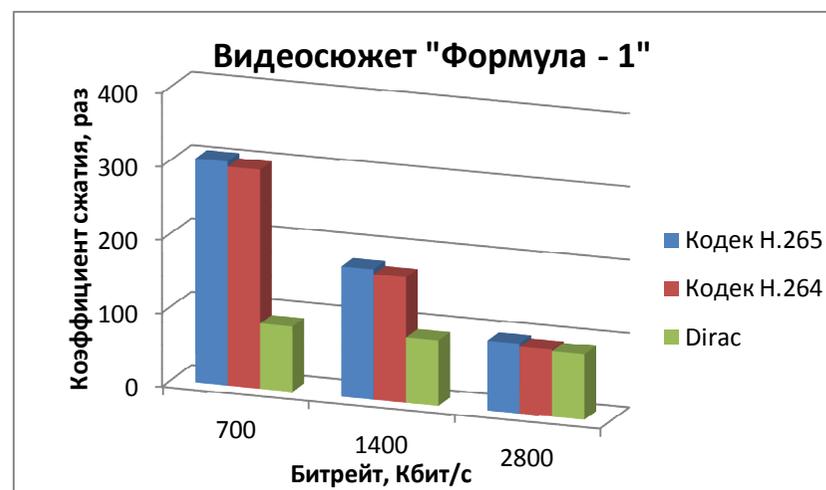
а)



б)



в)



г)

Рис.3.21. Сравнение коэффициентов сжатия исследованных кодеков для для соответствующих сюжетов: а) Белый медведь, б) Звездные войны, в) Оружие, г) Формула - 1

Таким образом, в результате эксперимента сжатия 4-х тестовых видеосюжетов на исследуемых битрейтах 2800 кб/с, 1400 кб/с, 700 кб/с можно сделать вывод, что:

- 1) При сжатии кодером H.265/HEVC происходит большая нагрузка на процессор по сравнению с кодером H.264/ MPEG-AVC и Dirac;
- 2) Наибольшим быстродействием обладают кодек стандарта H.264/MPEG-AVC и Дирак, а H.265 имеет самое низкое быстродействие по сравнению с остальными;
- 3) Визуальное качество восстановленных изображений на всех режимах кодирования приблизительно одинаковое.

5. Исследование эффективности сжатия аудио данных различными кодеками MP3, AAC, Ogg Vorbis

Для практической оценки эффективности работы кодеков было проведено исследование коэффициентов сжатия 4 фонограмм различных жанров, представленных в таблице 3.7, при различных значениях параметров качества кодирования. А также производилась оценка качества восстановленных сигналов при различных значениях скорости выходного потока. При этом результаты исследований были сведены в таблицы 3.8, 3.9. соответственно.

Для экспериментальных исследований использовались следующие кодеки:

1. Lame Encoder v3.92 (MP3 кодер) и Nullsoft Audio Decoder 2.8a (MP3 декодер)
2. AAC (MPG-4 AAC)
3. OggEnc v0.9 RC3 (OGG кодер) и Nullsoft Vorbis Decoder v0.2 beta7 (OGG декодер)

Для сравнения эффективности работы кодеков проводилось сжатие исходного несжатого файла WAV-фрагмента каждым из кодеков, а после обратного преобразования сравнение результирующих АЧХ и частотно-временных характеристик оригинального и полученного .WAV-файлов. Так как Ogg Vorbis кодирует с переменным битрейтом и не позволяет явным образом указать битрейт, подбор параметров кодирования для всех кодеков осуществлялся таким образом, чтобы сжатые файлы были приблизительно одного размера.

Таблица 3.7.

Типы и параметры тестовых музыкальных фонограмм

Жанр	Композиция	Длительность (мин)	Исходный размер wav (байт)	Формат
Classic	Бетховен- Лунная соната	4:41	60 549 644	PCM WAV 44.1 KHz, 16 bit, stereo
Pop	Валерия - Расставание	3:34	45 545 932	
Rock	Europe - The Final Countdown	3:12	54 683 228	
Rap	Eminem- 8 mile	3:45	39 858 220	

Кроме того для оценки качества восстановленных фонограмм использовалась как субъективная оценка на слух, так и объективная по методу расчета среднеквадратического отклонения декомпрессированного файла относительно исходного по формуле (3.1) (среднеквадратичная мера):

$$difference = 100 * \sqrt{\frac{\sum(s-s')^2}{n}} \quad (3.1),$$

где

s, s' – исходный и восстановленный сигналы;

n – число отсчётов в сигнале.

Степень сжатия оценивается соотношением:

$$compression = \frac{size'}{size} \quad (3.2),$$

где

size – размер сжатого исходного файла;

size' - размер сжатого файла коэффициентов.

Все результаты расчетов занесены в таблицы погрешностей преобразования 3.8, 3.9, и по ним проводился анализ результатов работы алгоритмов сжатия

Кодирование исходных музыкальных фрагментов проводилось в 5 различных режимах, битрейтами 320, 256, 128, 64 кбит/с. То есть, как на высоких, так и на низких битрейтах..

Таблица 3.8

Результаты сжатия фонограмм при разных битрейтах кодеком MP-3

Жанр	Исходный размер (wav), байт	Битрейт Кбит/с	Коэффициент сжатия (mp3)	Сред.квдр. откл. (%)
Classic	60 549 644	320	4,41	0,01
Pop	45 545 932		4,41	0,02
Rock	54 683 228		4,41	0,03
Rap	39 858 220		4,41	0,01
Classic	60 549 644	256	5,51	0,21
Pop	45 545 932		5,51	0,21
Rock	54 683 228		5,51	0,25
Rap	39 858 220		5,51	0,39
Classic	60 549 644	128	11,02	6,04
Pop	45 545 932		11,02	6,55
Rock	54 683 228		11,02	6,34
Rap	39 858 220		11,02	8,76
Classic	60 549 644	64	22,04	11,31
Pop	45 545 932		22,04	9,24
Rock	54 683 228		22,04	10,79
Rap	39 858 220		22,04	15,1

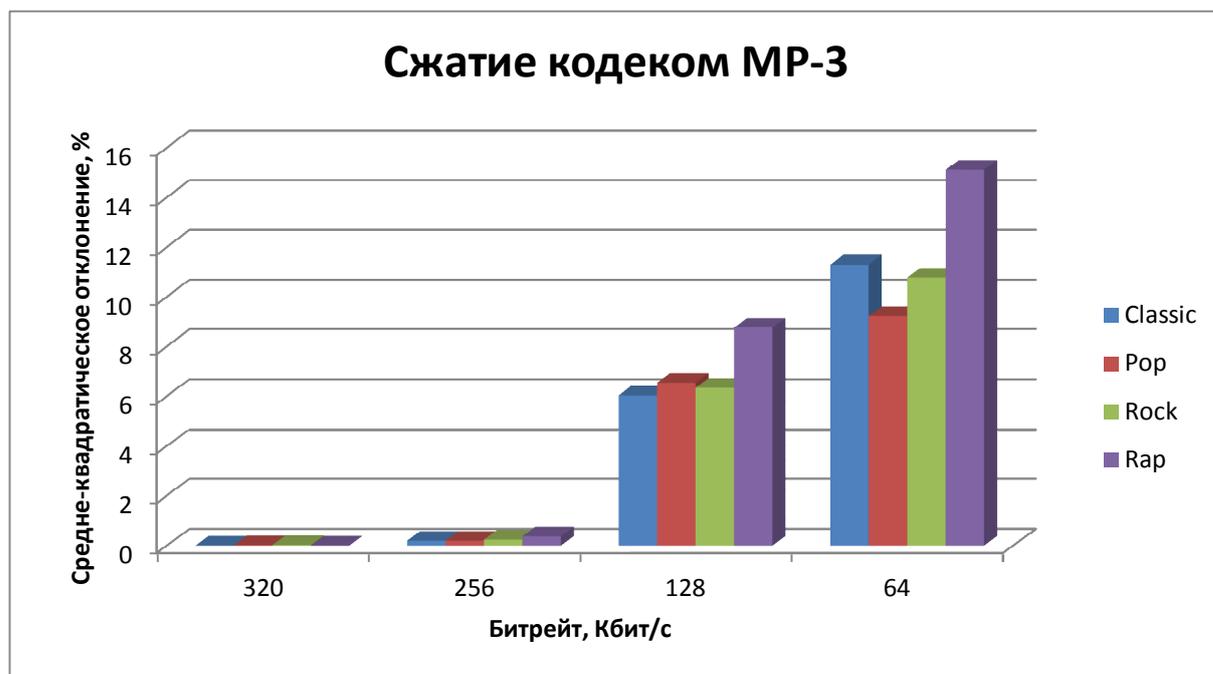


Рис.3.22. Среднеквадратичное отклонение при сжатии кодеком MP3

Таблица 3.9

Результаты сжатия фонограмм при разных битрейтах кодеком OGG

Жанр	Исходный размер (wav), байт	Битрейт Кбит/с	Коэффициент сжатия (wma)	Сред.квдр. откл.(%)
Classic	37 785 644	320	4,39	0,05
Pop	45 545 932		4,39	0,06
Rock	54 683 228		4,39	0,09
Rap	39 858 220		4,38	0,1
Classic	37 785 644	256	7,32	1,24
Pop	45 545 932		7,32	3,32
Rock	54 683 228		7,31	2,92
Rap	39 858 220		7,29	4,92
Classic	37 785 644	128	10,95	4,58
Pop	45 545 932		10,95	5,91
Rock	54 683 228		10,94	6,49
Rap	39 858 220		10,94	9,27
Classic	37 785 644	64	21,77	10,67
Pop	45 545 932		21,78	10,66
Rock	54 683 228		21,70	12,17
Rap	39 858 220		21,69	15,34

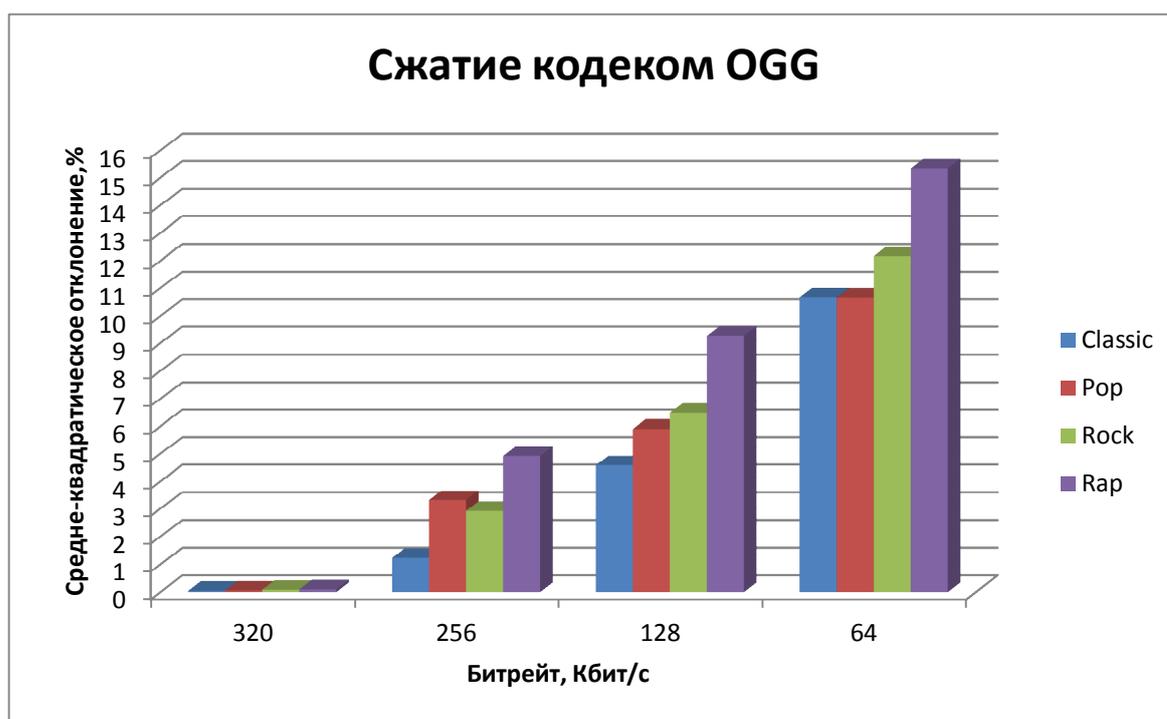


Рис.3.23. Среднеквадратичное отклонение при сжатии кодеком OGG

Таблица 3.10

Результаты сжатия фонограмм при разных битрейтах кодеком MP3-4 AAC

Жанр	Исходный размер (wav), байт	Битрейт Кбит/с	Коэффициент сжатия (AAC)	Средне-квадр. отклонение(%)
Classic	37 785 644	320	4,09	0,01
Pop	45 545 932		4,04	0,02
Rock	54 683 228		4,00	0,03
Rap	39 858 220		4,00	0,01
Classic	37 785 644	256	5,60	0,09
Pop	45 545 932		5,48	1,5
Rock	54 683 228		5,49	0,28
Rap	39 858 220		5,51	1,26
Classic	37 785 644	128	11,02	6,13
Pop	45 545 932		10,95	7,42
Rock	54 683 228		10,99	8,08
Rap	39 858 220		10,91	9,07
Classic	37 785 644	64	14,82	9,7
Pop	45 545 932		14,68	10,91
Rock	54 683 228		14,56	12,63
Rap	39 858 220		14,64	14,11

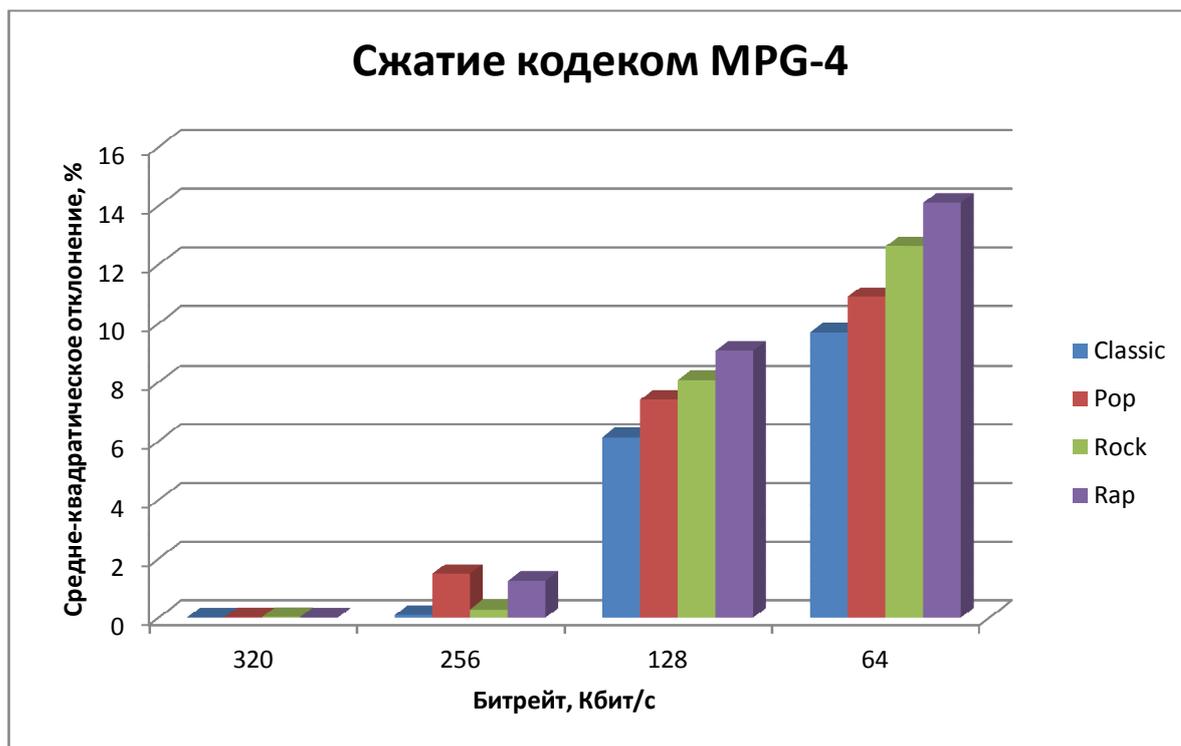


Рис.3.24. Среднеквадратичное отклонение при сжатии кодеком MPG-4

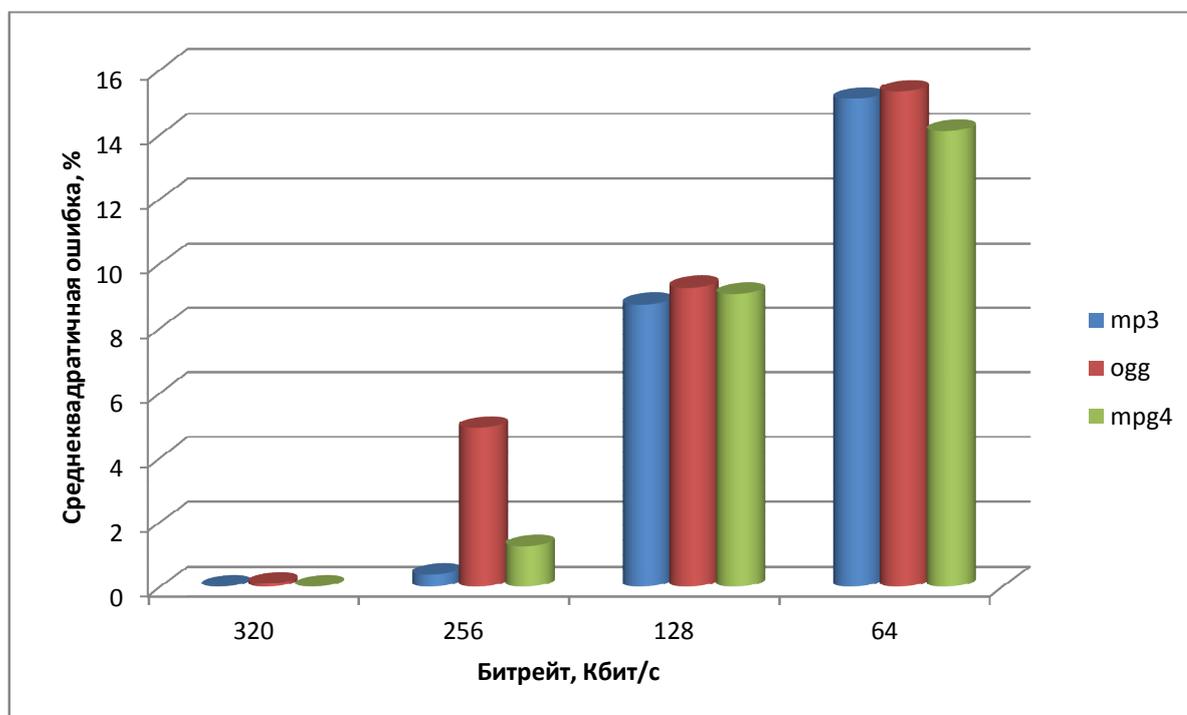


Рис.3.24. Сравнение среднеквадратичного отклонения при сжатии разными кодеками

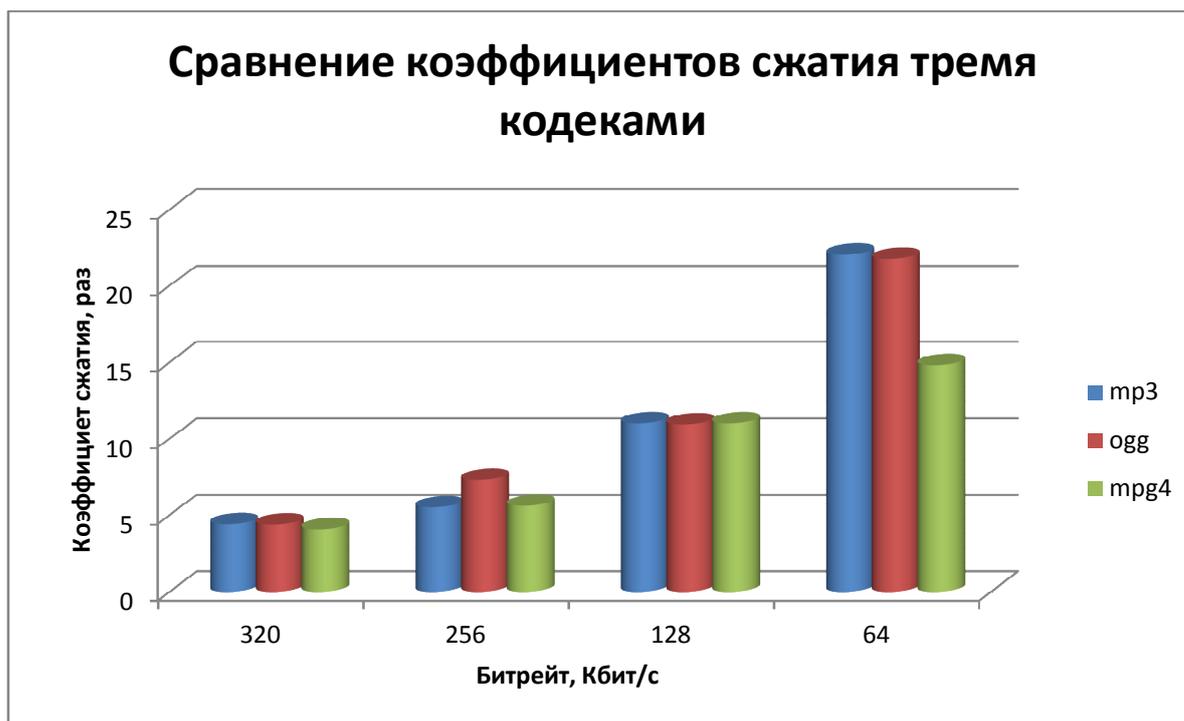


Рис.3.24. Сравнение коэффициентов сжатия разными кодеками

В результате проведенного эксперимента были получены данные тестирования сжатия 4-х музыкальных фрагментов тремя кодеками различных стандартов. На основе кодеков MP3, Ogg Vorbis, WMA было проведено исследование существующих стандартов и алгоритмов для следующих битрейтов: 320, 256, 128 и 64 Кбит/с.

Выводы по III главе

В результате проделанной работы был осуществлен анализ принципов построения и работы кодеков используемых в мультимедийном контенте ТВ программ. В ходе работы было проведено исследование эффективности методов и алгоритмов обработки ТВ изображений. Существующее разнообразие методов и стандартов сжатия ТВ изображений обладают различными возможностями. Были рассмотрены и проанализированы различные методы компрессии видео и аудио сигналов.

При анализе построения кодеков стандарта MPEG-4 были рассмотрены его модификации. Были представлены основные особенности кодека MPEG-4 и представлены структурные схемы работы кодера и декодера. Именно данный стандарт был разработан для передачи мультимедийной информации: аудио и видео. Было выявлено, что технология MPEG сжатия позволяет уменьшить избыточность видеoinформации во временном измерении и сжатие отдельных изображений. Вся обработка видео потока осуществляется при помощи четырех типов кадров: I-, P-, B- и S – кадров.

Также была рассмотрена и представлена структура аудио кодирования используемая в стандарте MPEG-4 и используемые алгоритмы компрессии, позволяющие обрабатывать звуковой сигнал в стандарте.

Анализируя построение кодеков, также было выявлено бурное развитие кодеков на основе вейвлет преобразований. Следовательно, было дано определение понятия «вейвлет». В качестве примера были приведены кодеки: «Rududu», Dirac codec и его модификация Dirac Pro.

Как было указано выше, существует довольно много различных видеокодеков, используемых для сжатия мультимедийного контента ТВ программ, отличающихся быстродействием, качественными показателями и техническими характеристиками. Для практической оценки эффективности работы видеокодеков было проведено исследование наиболее распространенных и перспективных разработок.

В ходе исследования было выявлено, что стандарт H.264/MPEG-AVC предназначен для обработки и кодирования видео, как на низких, так и на высоких скоростях передачи данных. Перспективный же стандарт H.265/MPEG-HEVC, является усовершенствованной версией стандарта H.264/MPEG-AVC, вобравший в себя улучшенные алгоритмы кодирования видео высокого разрешения для более эффективного сжатия видеопотока.

Для оценки эффективности исследуемых кодеков использовались 4 тестовых видеосюжета различных жанров с форматом изображения 720x576 пикселей, содержащих 150 кадров. Качество восстановленных изображений

и временные затраты на кодирование оценивалось на битрейтах 2800, 1400 и 700 кбит/с.

В результате проведенных экспериментов были получены временные затраты для различных кодеков, оценен коэффициент сжатия и выявлены качественные показатели с использованием объективной метрики MSE и субъективной метрики DSIS для восстановленного изображения.

По полученным экспериментальным данным видеосжатия можно сделать вывод, что:

1. При сжатии кодером H.265/HEVC происходит большая нагрузка на процессор по сравнению с кодером H.264/ MPEG-AVC и Dirac;
2. Наибольшим быстродействием обладают кодек стандарта H.264/MPEG-AVC и Дирак, а H.265 имеет самое низкое быстродействие по сравнению с остальными;
3. При обработке видеосюжетов Формула 1 и Звездные войны все тестируемые кодеки дают примерно одинаково хорошее визуальное качество восстановленных изображений на всех битрейтах. Но, на мелкоструктурных изображениях кодек Дирак уступает по качеству другим кодерам, хотя и они дают изображения со смазанной детальностью.

Однако, учитывая меньшую сложность алгоритма обработки кодека Дирак и его большее быстродействие можно рекомендовать алгоритм Дирака для практической реализации видео кодека, поскольку его можно реализовать на относительно не дорогой элементной базе. А использование перспективного стандарта H.265/HEVC в настоящее время не целесообразно.

По результатам исследования установлено, что кодек MP-3 имеет лучшие качественные показатели при битрейтах более 64 кбит/с, но среднеквадратическое отклонение от оригинала на битрейтах ниже 64 кбит/с составляет от 20-50%, что является не приемлемым. При этом кодеки *WMA* и *OGG* показывают величину ошибки не более 10%, однако проигрывают на средних битрейтах кодеку MP-3.

Заключение

В современном мире телевидение, как средство массовой информации играет очень важную роль в жизни каждого человека, не меньшую роль играет внедрение в эту отрасль новейших компьютерных технологий.

Именно поэтому, Президент Республики Узбекистан Ислам Каримов, на заседании Кабинета Министров, посвященном итогам социально-экономического развития страны в 2009 году отметил: «Говоря о развитии систем коммуникаций, особо следует отметить то стратегически важное значение, какое имеет для нас развитие высокотехнологичной телекоммуникационной отрасли. Сегодня уже невозможно представить жизнь без компьютерной техники, информационных технологий, Интернета, сотовой телефонной связи»[4].

Целью диссертационной работы являлось исследование существующих и перспективных методов сжатия мультимедийного потока ТВ программ кодеков применяемых в телевидении, обеспечивающих лучшие качественные показатели по сравнению с традиционными кодеками стандарта MPEG-4.

В настоящее время мультимедийные технологии являются бурно развивающейся сферой информационных технологий.

Во время работы был проведен анализ существующих и перспективных методов и алгоритмов сжатия ТВ изображений. Были сделаны выводы, что для сжатия мультимедийной информации существует несколько различных стандартов, методов кодирования.

Анализируя все виды информации было установлено, что любая информация обладает большим объемом избыточности, за счет которой осуществляется компрессия телевизионного сигнала для уменьшения скорости цифрового потока для более эффективного использования каналов передачи. Были установлены следующие виды избыточности: кодовая,

статистическая или межэлементная, структурная, временная или межкадровая, психовизуальная и психоакустическая.

При использовании ВП изображения обычно не разбиваются на блоки, а обрабатывается целиком, что позволяет в 1,5-2 раза увеличить сжатие статического изображения без заметного ухудшения его качества, и устраняет возникновение блочных искажений, однако без блочная обработка изображений не позволяет использовать методы блочной компенсации движение. Поэтому вейвлет – кодеки уступают по эффективности сжатия видеопотока кодекам стандартов MPEG.

При использовании фрактальных методов сжатия изображений, основанных на поиске самоподобных видеообъектов, возможно получение высоких коэффициентов сжатия, порядка 200-2000 раз. Данный метод сжатия хорошо работает на изображениях технических чертежей, текста, карт местности и т.д. На реальных ТВ изображения фрактальное сжатие дает хорошие результаты, но на поиск фракталов с учетом масштабирования и взаимной ориентации требуются большие временные затраты, что не позволяет их использовать в вещательном телевидении, где сжатие видеопотока должно происходить в режиме реального времени.

Для звуковых сигналов важной проблемой при цифровом представлении является сокращение имеющейся в них статической и психоакустической избыточности. Устранение статической избыточности не приводит к значительному уменьшению скорости цифрового потока, а благодаря психофизической избыточности можно обеспечить сжатие аудиоданных в 10-12 раз без существенных потерь в качестве. Основное устранение избыточности в аудиоданных достигается с помощью различных видов маскировок, что не приводит к потере в качестве.

Для сжатия текстовой и служебной информации необходимо использовать методы сжатия без потерь данных, поскольку нарушение структуры информации не позволит правильно восстановить исходные изображения, звуковое сопровождение и текстовую информацию. В

настоящее время для этой цели активно используются алгоритмы RLE, Хаффмана, LZW, BWT, являющиеся основой архиваторов ARJ, ZIP, RAR.

Как было установлено в ходе работы, на сегодняшний день существует довольно большое количество разнообразных кодеков используемых при сжатии мультимедийного потока. Было выявлено, что существующие кодеки, применяемые для звуковых и визуальных данных, используют сжатие с потерями, для получения приемлемого размера файла и сжатие без потерь. Но для большинства применений выгоднее кодеки с потерями информации, так как малозаметное ухудшение качества оправдывается значительным уменьшением объема данных.

В результате была представлена классификация видеокодеков. По типу исполнения кодеры делятся:

Программный видеокодек – это специализированная компьютерная программа, кодек, которая производит компрессию или декомпрессию цифровых видео данных в соответствии с файловым видео форматом или потоковым видео форматом.

Кодек на аппаратном уровне обозначает отдельную микросхему, которая кодирует и декодирует аналоговый аудио - и/или видео сигнал в цифровой сигнал и, наоборот, при помощи аналогово-цифрового и цифро-аналогового преобразователей.

По типу обрабатываемой информации существуют:

Аудиокодек — компьютерная программа или аппаратное средство, предназначенное для кодирования или декодирования аудиоданных.

Видеокодек — программа/алгоритм сжатия видеоданных (видеофайла, видеопотока) и восстановления сжатых данных. Также возможно кодирование кроме видео и аудиоинформации, добавления субтитров, векторных эффектов и т. п.

Проанализировав построение кодеков выявлено, что широкое распространение получили кодеки на основе вейвлет преобразований и дискретно-косинусном преобразовании.

При анализе построения кодеков стандарта MPEG-4 были рассмотрены его модификации. Были представлены основные особенности кодека MPEG-4 и представлены структурные схемы работы кодера и декодера. Именно данный стандарт был разработан для передачи мультимедийной информации: аудио и видео. Было выявлено, что технология MPEG сжатия позволяет уменьшить избыточность видеoinформации во временном измерении и сжатие отдельных изображений. Вся обработка видео потока осуществляется при помощи четырех типов кадров: I-, P-, B- и S – кадров.

Также была рассмотрена и представлена структура аудио кодирования используемая в стандарте MPEG-4 и используемые алгоритмы компрессии, позволяющие обрабатывать звуковой сигнал в стандарте.

Анализируя построение кодеков, также было выявлено бурное развитие кодеков на основе вейвлет преобразований. Следовательно, было дано определение понятия «вейвлет» и перечислены главные признаки вейвлета. В качестве примера были приведены кодеки: «Rududu», Dirac codec и его модификация Dirac Pro.

Как было указано выше, существует довольно много различных видеокодеков, используемых для сжатия мультимедийного контента ТВ программ, отличающихся быстродействием, качественными показателями и техническими характеристиками. Для практической оценки эффективности работы видеокодеков было проведено исследование наиболее распространенных и перспективных разработок.

В ходе исследования было выявлено, что стандарт H.264/MPEG-AVC предназначен для обработки и кодирования видео, как на низких, так и на высоких скоростях передачи данных. Перспективный же стандарт H.265/MPEG-HEVC, является усовершенствованной версией стандарта H.264/MPEG-AVC, вобравший в себя улучшенные алгоритмы кодирования видео высокого разрешения для более эффективного сжатия видеопотока.

Выявлено, так как существует большое количество стандартов, также существует большое количество кодеков, предназначенных для сжатия ТВ

изображений, отличающиеся быстродействием, качественными показателями и техническими характеристиками. Для практической эффективности работы видеокодеков было проведено исследование наиболее распространенных и перспективных разработок.

В результате проведенных экспериментов были получены временные затраты для различных кодеков, оценен коэффициент сжатия и выявлены качественные показатели с использованием объективной метрики MSE и субъективной метрики DSIS для восстановленного изображения.

Полученные результаты исследований приведены в таблице 3.1, 3.2 для кодека H.264; 3.3, 3.4, для кодека H.265 и 3.5, 3.6 для кодека Dirac-pro. При этом визуальное качество исходных и восстановленных после декомпрессии изображений можно оценить по рис.3.8-3.11 для H.264, рис.3.12-3.15 для H.265 и рис.3.16-3.19 для кодека Dirac-pro. Кроме того, сравнительное быстродействие кодеков и их коэффициенты сжатия представлены в гистограммах на рис.3.20, 3.21.

В результате эксперимента сжатия 4-х тестовых видеосюжетов на исследуемых битрейтах 2800 кб/с, 1400 кб/с, 700 кб/с можно сделать вывод, что:

- 1) При сжатии кодером H.265/HEVC происходит большая нагрузка на процессор по сравнению с кодером H.264/ MPEG-AVC и Dirac;
- 2) Также сильное повышение температуры процессора при использовании кодера H.265/HEVC по сравнению с кодером H.264/MPEG-AVC и Dirac;
- 3) Большая разница во времени сжатия видеосюжетов данными кодерами, наиболее эффективным по времени сжатия является кодер стандарта H.264/MPEG-AVC по сравнению с остальными;
- 4) Визуальное качество восстановленных изображений на всех режимах кодирования приблизительно одинаковое.

Таким образом, по результатам проведенных исследований можно рекомендовать алгоритм Dirac для практической реализации видео кодека,

поскольку его реализация возможна на относительно недорогой элементной базе. А использование перспективного стандарта H.265/HEVC в настоящее время не целесообразно.

Для практической оценки эффективности работы кодеков было проведено исследование коэффициентов сжатия 4 фонограмм различных жанров, представленных в таблице 3.7, при различных значения параметров качества кодирования. А также была проведена оценка качества восстановленных сигналов при различных значениях скорости выходного потока. При этом результаты исследований были сведены в таблицы 3.8, 3.9, 3.10. соответственно.

В результате проведенных экспериментов аудиосигналов было выявлено, что:

- 1) Кодек MP-3 имеет лучшие качественные показатели при битрейтах более 64 кбит/с по сравнению с OGG и MPG-4
- 2) Среднеквадратическое отклонение у MP-3 составляет 20-50 % от оригинала на битрейтах ниже 64 кбит/с, что является не приемлемым.
- 3) Кодеки WMA и OGG показывают величину ошибки не более 10%, однако проигрывают на средних битрейтах кодеку MP-3.

Список литературы

I. Закон Республики Узбекистан

1. Закон Республики Узбекистан «Об информатизации» от 11.12.2003 г. №560-II, ст. 4 (Ведомости Олий Мажлиса Республики Узбекистан, 2004 г., №1-2, ст. 10)

II. Постановления и указы Президента Республики Узбекистан

2. Постановление Президента Республики Узбекистан «О мерах по дальнейшему внедрению и развитию современных информационно-коммуникационных технологий». 21 марта 2012 г., №ПП-1730. Собрание законодательства Республики Узбекистан, 2012 г., №13, ст. 139.
3. Программа дальнейшего внедрения и развития информационно-коммуникационных технологий в Республике Узбекистан на 2012–2014 годы. Постановление Президента от 21.03.2012 г., №ПП-1730.
4. Каримов И.А. Наша главная задача – дальнейшее развитие страны и повышение благосостояния народа./Доклад Президента РУз на заседании Кабинета Министров, посвященном итогам социально-экономического развития страны в 2009 г. и важнейшим приоритетам экономической программы на 2010 г. 30 января 2010 г.

III. Основная литература

5. Пилко И.С. Информационные и библиотечные технологии: учебное пособие СПб., 2008. – С. 342.
6. Рудометов Е. Аппаратные средства и мультимедиа. Справочник. — СПб.: Питер, 1999. – 352 с.
7. Огнев И.В., Огнев А.И., Горьков А.Г. Алгоритм формирования базы данных для фрагментарного метода сжатия видеопотока без потерь, ВС/NW 2012; №2 (21):10.2.

8. В.Е.Джакония. Телевидение: Учебник для вузов/В.Е.Джакония, А.А.Гоголь, Я.В.Друзин и др.; Под ред. В.Е.Джаконии. 4-е изд., стереотип. – М.: Горячая линия – Телеком. 2007. – 616 с:
9. Шелухин О.И. Сжатие аудио и видео информации, МТУСИ, 2012. – с. 88
10. Смирнов, А.Е.Пескин. Цифровое телевидение: от теории к практике. Горячая Линия – Телеком. 2005. 340 с.
11. Б.А. Локшин Цифровое вещание: - от студии к телезрителю. от студии к телезрителю - М.: Компания САЙРУС СИСТЕМС, 2001. 446 с
12. Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет-преобразования. СПб: Изд-во ВУС, 1999. 204 с.
13. Ричардсон Я. Видеокодирование H.264 и MPEG-4 – стандарты нового поколения.// - М. – Техносфера. – 2005. - 368 с.
14. Солодовщиков А.Ю., Платонов А.К. Исследование метода Карунена-Лоэва., Москва, 2006
15. А. И. Бенилов, С. Д. Погорелый . Киевский национальный университет им. Т.Г. Шевченко/ Вейвлет-анализ и его применение для сжатия мультимедийной информации. 2002 г.
16. Мандельброт Б. Фрактальная геометрия природы. – М: Институт компьютерных исследований, 2002, 656 стр.
17. Савицкая Д.А, Пузий А.Н, “Сжатие ТВ изображений на основе сегментации объектов сцены”. Тезисы доклада в сборнике трудов Республика илмий-техник анжуман «Ахборот технологиялари ва телекоммуникация муаммолари», проходившей 21-22 апреля 2011 в Ташкенте. С.96-99.
18. Стрельников К.Н., Куликов Д.Л., “Panorama usage for multiframe motion compensation based video codecs - Использование панорамы для оптимизации кодеков на основе многокадровой компенсации движения”, Международный журнал, Программные продукты и системы. № 4 2009 г.

19. Мобильное телевидение — модель будущего: DVB
www.telemultimedia.ru
20. Майстренко, Н.В. Мультимедийные технологии в САПР : учебное пособие / Н.В. Майстренко, А.В. Майстренко, И.Л. Коробова. – Тамбов: Изд-во Тамб. гос. техн. ун-та, 2009. – Ч. II. – 80 с. – 100 экз.
21. Д. Ватолин, А. Ратушняи, М. Смирное, В. Юкин. Методы сжатия данных/ Устройство архиваторов, сжатие изображений и видео. - М.: ДИАЛОГ-МИФИ, 2003. - 384 с.

IV. Дополнительная литература

22. Сергей Швец. Видеокодеки и их роль в воспроизведении медиа-файлов на компьютере. 21.03.2014
23. Отто С. Э. Моделирование системы сжатия цифрового видеосигнала в реальном времени на основе вейвлет преобразований. Статья в сборнике докладов Республиканской научно-технической конференции молодых ученых, исследователей, магистрантов и студентов “Проблемы информационных технологий и телекоммуникаций”, Ташкент 15-16 марта 2012г. I том. стр.82-83
24. Питер Уилсон, Тим Борер, Томас Дэвид. Подготовил Валерий Хлебородов Семейство систем цифрового сжатия DIRAC расширяется // «625» : журнал. — 2007. — № 03.
25. Александр Железняков. Grass Valley ProCoder. Январь 2012.
26. А. Железняков. Файловый медиатранскодер ProMedia Carbon. 2012
27. Семен Макаров/ Кодеры StreamZ компании Digital Rapids/ Февраль 2012
28. Николай Азин. Кодеры Teleview. Февраль 2012
29. Семенов Ю.А. (ИТЭФ-МФТИ). Стандарт MPEG-4
30. Dirac Specification Version 2.2.3 Issued: September 23, 2008
31. Кирилл Кочетков, MP3-кодек, 28 ноября 2003 г.
32. Э.Г.Кнеллер. Анализ параметров речевого сигнала создающих восприятие элементарных звуков речи. <http://www.istrasoft.ru/Analysis.htm>

33. Денис Кубасов, Дмитрий Ватолин, “Обзор методов компенсации движения”, Компьютерная графика и мультимедиа, сетевой журнал. <http://www.cgm.computergraphics.ru/content/view/76>

V. Публикации в научно-технических журналах

34. Ким Е.А. Обзор методов сжатия мультимедийной информации. Статья в сборнике докладов Республиканской научно-технической конференции молодых ученых, исследователей, магистрантов и студентов “Информационные технологии и проблемы телекоммуникаций”, Ташкент 14-15 марта 2013г. 4 том, стр.192-193.
35. Ким Е.А. Сравнение эффективности сжатия тв изображений кодеками стандартов H.264/MPEG-AVC И H.265/MPEG-HEVC. Статья в сборнике докладов Республиканской научно-технической конференции молодых ученых, исследователей, магистрантов и студентов “Перспективы эффективного развития информационных технологий и телекоммуникационных систем”, Ташкент 15-16 марта 2014г. 3 том, стр.162-163.

VI. Интернет ресурсы

36. Игорь Сколотнев. Технологии MPEG для цифровой обработки аудио и видео информации. Ист: <http://www.statya.ru/index.php?op=view&id=1571>
37. Oleg Grebenyow. Алгоритмы архивации данных. [structur@freemail.ru/](mailto:structur@freemail.ru)
<http://www.getinfo.ru/article473.html>
38. Обзор методов сжатия данных.
<http://mindspring.narod.ru/alg/rcompress.html>
39. МЕТОД ХАФФМАНА И РОДСТВЕННЫЕ МЕТОДЫ.
<http://mindspring.narod.ru/alg/huffman.html>
40. Климов А. С. Форматы графических файлов. - С.-Пб.: ДиаСофт. 1995.
41. Методы сжатия информации/ <http://www.compression.ru>

42. New in DivX 10.2: HEVC Meets DivX Live.
<http://www.divx.com/en/software/divx>
43. По материалам Ericsson. Кодеры Ericsson. <http://www.ericsson.com/ru>.
Январь 2012
44. Источник: <http://daily.sec.ru/2013/10/29/CHto-takoe-horoshiy-kodek.html>
45. Родион Жабрев. HEVC (H.265) - новый кодек. 25.12.12.
<http://robureau.ru/blog/hevc-novyy-kodek-vdvoe-luchshe-vdvoe-menshe/>
46. Rududu video codec/ <http://rududu.berlios.de/>
47. Dirac 1.0.0 released. [LWN.net].
48. Dirac/ Video Compression. <http://diracvideo.org/>
49. Джим Фили. Новые методы оценки цифрового видео.
<http://www.digitalvideo.ru/archiv/993/99304.htm>