

**МИНИСТЕРСТВО ПО РАЗВИТИЮ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ И КОММУНИКАЦИЙ РЕСПУБЛИКИ УЗБЕКИСТАН
ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

К защите допустить
Зав. кафедрой

« _____ » _____ 2015 г.

**Выпускная
квалификационная работа бакалавра**

на тему: **ЭФФЕКТИВНЫЕ СИСТЕМЫ СЖАТИЯ ЗВУКА
И РЕКОМЕНДАЦИИ ПО ИХ ПРИМЕНЕНИЮ**

Выпускница _____
(подпись)

Ахмедова З.А.
(Ф. И. О.)

Консультант _____
(подпись)

Носиров Х.Х.
(Ф. И. О.)

Рецензент _____
(подпись)

Юлдашев Ж.М.
(Ф. И. О.)

Консультант по БЖД и Э _____
(подпись)

Амурова Н.Ю.
(Ф. И. О.)

Ташкент – 2015

8. Консультанты по отдельным разделам выпускной работы

Раздел	Ф.И.О. руководителя	Подпись, дата	
		Задание выдал	Задание получил
1. Глава	Носиров Х.Х	12.02.15	12.02.15
2. Глава	Носиров Х.Х	12.03.15	12.03.15
3. Глава	Носиров Х.Х	30.04.15	30.03.15
ОТ и ТБ	Амурова Н.Ю	05.06.15	05.06.15

9. График выполнения работы

№	Наименование раздела работы	Срок выполнения	Отметка руководителя о выполнении
1.	Введение	24.03.15	
2.	Анализ современных и эффективных методов сжатия звука.	24.03.15	
3.	Сравнительный анализ современных форматов сжатия звука.	15.05.15	
4.	Преимущество аудиокодеков и рекомендация по их применению.	30.05.15	
5.	Безопасность жизнедеятельности и Экология	06.06.15	

Выпускник _____ « ____ » _____ 2015 г.
(подпись)

Руководитель _____ « ____ » _____ 2015 г.
(подпись)

Данная ВКР посвящается анализу эффективных систем сжатия звука и рекомендации по их применению. В работе рассмотрены методы сжатия звуковых данных с потерями и без потерь.

Приводятся результаты сравнительного анализа эффективных методов сжатия и рекомендации по их отрасли применения.

The given FQW is devoted to analysis of efficient audio compression systems and their using recommendations. The methods of compression of voice data are in-process considered with losses and without losses.

Comparative analysis of effective methods of compression and recommendation on their industry of application are given.

Ушбу БМИ товушни сиқишнинг самарали тизимларини таҳлиliga бағишланган бўлиб, уларни қўллаш бўйича тавсиялар ҳам келтирилган. Битирув ишида тавуш маълумотларини йўқотишлар билан ва йўқотишларсиз сиқиш услублари кўриб чиқилган.

Самарали сиқиш услубларининг солиштирма таҳлилининг натижалари ва уларни қўллаш соҳалари бўйича тавсиялар келтирилган.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	6
1. АНАЛИЗ СОВРЕМЕННЫХ И ЭФФЕКТИВНЫХ МЕТОДОВ СЖАТИЯ ЗВУКА.....	7
1.1. Общие положения.....	7
1.2. Структура аудио формата WAVE.....	9
1.3. Анализ методов сжатия в стандарте MP-3.....	12
1.4. Анализ эффективных форматов сжатия звука с потерями.....	15
1.5. Анализ эффективных форматов сжатия звука без потерь.....	22
Выводы.....	32
2. СРАВНИТЕЛЬНЫЙ АНАЛИЗ СОВРЕМЕННЫХ ФОРМАТОВ СЖАТИЯ ЗВУКА.....	33
2.1. Сравнение сжатие аудиоданных.....	33
2.2. Сравнение цифровых аудио форматов.....	39
2.3. Сравнение звуковых форматов со сжатием без потерь.....	41
2.4. Сравнение звуковых форматов со сжатием с потерями.....	45
Выводы.....	48
3. ПРЕИМУЩЕСТВО АУДИОКОДЕКОВ И РЕКОМЕНДАЦИЯ ПО ИХ ПРИМЕНЕНИЮ.....	50
3.1. Преимущество звуковых форматов, сжимающие с потерями и рекомендации по применению.....	50
3.2. Преимущество звуковых форматов, сжимающие без потерь и рекомендации по применению.....	60
Выводы.....	68
4. БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ И ЭКОЛОГИЯ.....	70
4.1. Влияние шума на организм человека.....	70
4.2. Взаимодействие человека и техносферы.....	76
4.3. Экологический мониторинг.....	85
Выводы.....	90
ЗАКЛЮЧЕНИЕ.....	91
СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ.....	95

ВВЕДЕНИЕ

В жизни современного общества телевидение и радиовещание стали неотъемлемой частью повседневной жизни почти каждого человека. Также активно развивается и мобильное телевидение. Однако для передачи огромного количества мультимедийной информации требуются широкополосные каналы связи, поскольку одна минута цифровой аудиозаписи с качеством лазерного диска занимает порядка 10,8 Мбайт, а для передачи его по каналу связи в реальном времени нужно иметь скорость передачи до 1,5 Мбит/с. Что при использовании обычных модемных каналов является невыполнимым. Поэтому необходимо использовать эффективные методы сжатия аудио-видео информации.

В настоящее время существует достаточно много разных стандартов: Monkey's Audio, TAK, FLAC, Apple Lossless, (ALAC), имеющих свои достоинства и недостатки. Одни обеспечивают приемлемое качество звука, но малую степень сжатия, другие, хорошую степень сжатия, но высокую сложность реализации алгоритма, сильно влияет на скорость компрессии и декомпрессии.

Все методы сжатия данных делятся на два основных класса: с потерями и без потерь. Методы сжатия с потерями намного эффективнее методов кодирования без потерь. Чем выше коэффициент сжатия, тем больше шума добавляется в данные. При использовании метода сжатия без потерь восстановленные данные идентичны исходным. Этот метод применяется для обработки многих типов данных. При этом не допускается потеря ни одного бита информации. В то же время файлы данных, представляющие изображения и другие полученные сигналы, нет необходимости хранить и передавать без потерь. Любой электрический сигнал содержит шум. Если изменения в этих сигналах схожи с небольшим количеством дополнительного шума, вреда не наносится.

1. АНАЛИЗ СОВРЕМЕННЫХ И ЭФФЕКТИВНЫХ МЕТОДОВ СЖАТИЯ ЗВУКА

1.1. Общие положения

Звуковой сигнал представляет собой сложную суперпозицию не гармонических колебаний, амплитуда которых меняется больше тысячи раз в секунду. Для упрощения обработки, звуковой (сложный) сигнал разделяют на куски гармонического сигнала.

Для передачи звуковых сигналов по каналам передачи данных, пропускная способность, которых ограничена, используют различные методы сжатия.

Существует два типа сжатия звуковых сигналов: с потерями и без потерь.

При сжатии звуковых сигналов алгоритмами сжатия без потерь сжатие получается не эффективным и не превышает 2 раз, поэтому обычно применяют методы сжатия с потерями, основанные квантовании коэффициентов спектральных преобразований и психоакустических особенностей нашего слухового восприятия. Сокращение статистической избыточности основано на учете свойств самих звуковых сигналов. Она определяется наличием корреляционной связи между соседними отсчетами цифрового звукового сигнала, устранение которой позволяет сокращать объем передаваемых данных на 15...25% по сравнению с их исходной величиной. Важными условиями применения такого метода преобразования являются:[1]

- возможность восстанавливать исходный сигнал без искажений;
- способность обеспечивать наибольшую концентрацию энергии в небольшом числе коэффициентов преобразования быстрый вычислительный алгоритм.

Существуют различные понятия звукового формата.

Формат представления звуковых данных в цифровом виде зависит от способа квантования аналогово-цифровым преобразователем (АЦП). В звукотехнике в настоящее время наиболее распространены два вида квантования:[2]

- импульсно-кодовая модуляция
- сигма-дельта-модуляция

Зачастую разрядность квантования и частоту дискретизации указывают для различных звуковых устройств записи и воспроизведения как формат представления цифрового звука (24 бита/192 кГц; 16 бит/48 кГц).

Формат файла определяет структуру и особенности представления звуковых данных при хранении на запоминающем устройстве ПК. Для устранения избыточности аудио данных используются аудиокодеки, при помощи которых производится сжатие аудиоданных. Выделяют три группы звуковых форматов файлов:

- аудиоформаты без сжатия, такие как WAV, AIFF
- аудиоформаты со сжатием без потерь (APE, FLAC)
- аудиоформаты, с применением сжатия с потерями (mp3, ogg)

Форматы носителей цифрового звука применяют как для массового распространения звуковых записей (CD, SACD), так и в профессиональной звукозаписи (DAT, минидиск).

Также форматом называют количество каналов в системах многоканального звука (5.1; 7.1). Изначально такая система была разработана для кинотеатров, но впоследствии была расширена для систем домашнего кинотеатра.

При этом часто исходные оцифрованные аудиоданные хранятся в формате WAVE.

1.2. Структура аудио формата WAVE

Формат WAVE – подмножество спецификации RIFF Майкрософта для хранения мультимедийных файлов. RIFF (ResourceInterchangeFileFormat - Формат файла обмена ресурс) файл начинается с заголовка файла, за которым последовательно следуют отсчеты данных. WAVE файл – в большинстве случаев только RIFF файл, единственный тип "WAVE" файлов, который состоит из двух подкусков – "fmt", определяющий формат данных и "data", содержащий фактически типовые данные [3]. Этот тип файла имеет структуру, представленный на рисунке 1.1. и таблице 1.1.



Рис. 1.1. Структура аудио файла WAVE типа RIFF.

Таблица 1.1

Описание составляющих заголовка WAVE файла

Смещение файла	Размер	Название	Описание
0	4	ChunkID	Содержит надпись “RIFF” в форме ASCII (0x52494646 в форме big-endian).
4	4	ChunkSize	36 + SubChunk2Size, или более точно: 4 + (8 + SubChunk1Size) + (8 + SubChunk2Size) где SubChunk1Size - размер остальной части куска после этого числа. SubChunk2Size - размер всего файла в байтах минус 8 байтов для две области, не включенные в этот счет: ChunkID и ChunkSize.
8	4	Format	Содержит надпись “WAVE” (0x57415645 в форме big-endian)
Параметры аудио файла – область “fmt”			
12	4	Subchunk1ID	Содержит буквы “fmt” (0x666d7420 в форме big-endian)
16	4	Subchunk1Size	16 для РСМ. Это размер остальной части куска после этого числа.
20	2	AudioFormat	РСМ = 1 (то есть. Линейное квантование)
22	2	NumChannels	Количество каналов Моно = 1, Стерео = 2, и т.д.
24	4	SampleRate	Частота дискретизации 8000, 44100, и т.д.
28	4	ByteRate	= SampleRate * NumChannels * BitsPerSample/8
32	2	BlockAlign	= NumChannels * BitsPerSample/8
34	2	BitsPerSample	Бит, приходящий на один отсчет. 8 битов= 8, 16 битов = 16, и т.д.
Область “data” содержит размер данных и аудио данные:			
36	4	Subchunk2ID	Содержит буквы “data” (0x64617461 в форме big-endian)
40	4	Subchunk2Size	= NumSamples * NumChannels * BitsPerSample/8 NumSamples – количество отсчетов
44	***		Данные

Представленный вид аудио файла принят для картотек данных WAVE, имеющих форму little-endian. У файлов, использующих форму big-endian, используется идентификатор RIFX вместо RIFF.

Структуры данных в файле одинаковы, независимо от того, 8 или 16 битные отсчеты содержит аудио файл. При этом 8-битовые файлы сохранены как байты без знака, в пределах от 0 до 255, а 16-битовые данные сохранены в пределах от -32768 до 32767.

В потоке данных WAVE могут быть дополнительные фрагменты (название альбома, автор, и.т.д.). Если они есть, то каждый дополнительный фрагмент будет иметь 4х байтные переменные: SubChunkID - название, SubChunkSize без знака и количество данных SubChunk2Size.

Обозначения файлов RIFF.

Мультимедийные приложения требуют хранения и управления широким разнообразием данных, включая битовые массивы, звуковые данные, видео данные, и периферийную информацию для правильной работы воспроизводящего устройства. Поэтому RIFF обеспечивает способ хранения всех этих разнотипных данных. Тип данных, которые содержат файл RIFF, обозначается следующими расширениями файла:

- Звуковые/визуальные чередованные данные (Audio/visualinterleaveddata) (.AVI)
- Данные формы WAVE (.WAV)
- Данные с побитовым отображением (.RDI)
- Информация MIDI (.RMI)
- Цветная палитра (.PAL)
- Мультимедийное кино (.RMN)
- Оживляемый курсор (Animated cursor) (.ANI)
- Связка других файлов RIFF (.BND)

1.3. Анализ методов сжатия в стандарте MP-3

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.

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.1.2.

Входные 16 разрядные отсчеты звукового сигнала (ЗС) с частотой дискретизации 48 кГц поступают на вход аудиокодека, где разделяется на 32 субполосные составляющие, причем обработка ведется аудиокадрами, содержащими 1152 отсчётов звукового сигнала. Каждый аудиокадр длится 24 мс. Все 32 субполосы имеют постоянную ширину $F = f_d / (2n)$, где f_d — частота дискретизации ЗС, n — число субполос, равную 750Гц.



Рис. 1.2. Общая структурная схема кодера MP3

После фильтрации отсчёты ЗС в каждой субполосе собираются в блоки, а затем масштабируются (нормируются), и обрабатываются модифицированным дискретно-косинусным преобразованием (МДКП), в результате которого формируются коэффициенты МДКП, характеризующие распределение энергии сигнала по его спектральным составляющим.

Особенностью спектра МДКП является, то, что его энергия группируется в небольшой низкочастотной области, а остальные коэффициенты либо балы, либо равны 0. Затем эти коэффициенты для повышения коэффициента сжатия кодека квантуются (производится деление на определенные числа. С последующим округлением до ближайшего целого) и затем кодируются. Масштабирование в каждом субполосном канале представляет собой выбор из 18 отсчётов ЗС максимального значения, которое называется масштабным коэффициентом SCF (ScaleFactor) выборки субполосного ЗС. В кодеке имеем 32 масштабных коэффициента, причем каждый блок из 36 отсчётов ЗС делится на три подблока, называемые гранулами. В каждой грануле, включающей 18 отсчетов ЗС, также определяется максимальный отсчёт, 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 отсчёта 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 (Psychoacoustic Model). 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 (FF transform) дискретизированного входного сигнала.

Для каждой скорости передачи (или группы скоростей) существует своя таблица. Значения f_d и скорость передачи цифровых данных определяют также и максимальный номер субполосы n , отсчёты 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.

1.4. Анализ эффективных форматов сжатия звука с потерями

Формат MPEG-2 AAC. 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. 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.[4] Технология AAC позаимствовала у AC3 возможность кодирования нескольких параллельных каналов, а у MP3 - принципы психоакустической модели. 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 lowcomplexityprofile) в которой нет предсказания, а TNS ограничено двенадцатью коэффициентами, но с рабочей полосой до 18 кГц.

MPEG-2 AAC обладает высоким качеством звучания и хорошей степенью компрессии аудиокomпозиций. Так, например, аудиокomпозиция в формате AAC с bitrate 96 kbs обеспечивает качество звучания, аналогичное потоку MPEG-1 Layer III bitrate 128 kbs.

Более того, AAC (с некоторыми модификациями) является единственной высококачественной аудиотехнологией, используемой в рамках стандарта MPEG-4 - глобального мультимедийного формата

будущего.

Формат Vorbis. Vorbis — свободный формат сжатия звука с потерями. По функциональности и качеству аналогичен таким кодекам как ААС, АС3 и VQF, превосходящим MP3.Психоакустическая модель, используемая в Vorbis, по принципам действия близка к MP3 и подобным, однако математическая обработка и практическая реализация этой модели существенно отличаются, что позволило авторам объявить свой формат совершенно независимым от всех предшественников.[5]

Для хранения аудиоданных в формате Vorbis чаще всего применяется медиаконтейнер Ogg, такой файл обычно имеет расширение .ogg и называется двойным именем «Ogg/Vorbis» или «OggVorbis». Однако «OggVorbis» называют и сам кодек без контейнера, так как он является частью проекта Ogg

Формат изначально разрабатывался с возможностью потокового вещания. Это даёт формату достаточно полезный побочный эффект — в одном файле можно хранить несколько композиций с собственными тегами. При загрузке такого файла в плеер должны отобразиться все композиции, будто их загрузили из нескольких различных файлов.

Формат имеет гибкую систему тегов. Заголовок тегов легко расширяется и позволяет включать тексты любой длины и сложности (например, текст песни), перемежающиеся изображениями (например, фотографиями обложек альбомов). Текстовые теги хранятся в UTF-8, что позволяет писать на нескольких языках одновременно и исключает возможные проблемы с кодировками.

OggVorbis по умолчанию использует переменный битрейт, при этом значения последнего не ограничены какими-то жёсткими значениями, и он может варьироваться даже на 1 kbps. При этом стоит заметить, что форматом жёстко не ограничен максимальный битрейт, и при максимальных настройках кодирования он может варьировать от 500 до 1000 кбит/с. Такой же гибкостью обладает частота дискретизации — пользователям

предоставляется любой выбор в пределах от 2 до 192 кГц.

Vorbis был разработан сообществом «Xiph.Org» для того, чтобы заменить все платные запатентованные аудио форматы. Несмотря на то, что это самый молодой формат из всех конкурентов MP3, OggVorbis имеет полную поддержку на всех популярных платформах (MicrosoftWindows, Linux, AppleMacOS, Android, PocketPC, Palm, Symbian, DOS, FreeBSD, BeOS и др.), а также большое количество аппаратных реализаций.

Vorbis является частью мультимедиапроекта Ogg, в который также входят свободные кодировщики: Speex — для сжатия голоса; FLAC — для сжатия звука без потерь; Theora — для сжатия видео.

Кодек Speex. Speex — это свободный кодек для сжатия речевого сигнала, который может использоваться в приложениях «голос-через-интернет» (VoIP). С высокой вероятностью он не имеет никаких патентных ограничений и лицензирован под последней версией лицензии BSD (без третьей статьи). Сжатые кодеком Speex данные можно хранить либо в формате хранения звуковых данных Ogg, либо передавать напрямую с помощью пакетов UDP/RTP.[5]

Speex относится к классу так называемых CodeExcitedLinearPrediction (CELP)-кодеков, то есть кодеков, построенных на основе так называемого Линейного Предсказательного кодирования ЛПК. ЛПК использует для аппроксимации отрезка речевого сигнала цифровой фильтр только с обратными связями (т.н. «авторегрессионный фильтр»). Коэффициенты этого фильтра «подгоняются» под отрезок сигнала с помощью процедуры Левинсона (в западной литературе — Левинсона-Дурбина). CELP-модификация ЛПК предусматривает наличие т.н. «кодовой книги», которая содержит предопределённые наборы возбуждающих ЛПК-фильтр единичных импульсов.

По заявлению разработчиков, Speex оптимизирован для получения высококачественного речевого сигнала при низких скоростях. Кодек Speex

также позволяет использовать переменную степень сжатия сигнала и поддерживает сигналы с различной шириной полосы: сверхширокополосный (англ. ultra-wideband, частота дискретизации 32 КГц), широкополосный (англ. wideband, 16 КГц) и узкополосный (англ. narrowband, 8 КГц). Направленность на системы «голос-через-интернет» (VoIP) предопределяет, что Speex должен быть устойчив к потерям пакетов данных, но не к повреждению их, так как UDP/IP-протокол, в отличие от TCP/IP, не гарантирует доставку пакетов потребителю.

Но, при этом, вероятность повреждения данных в UDP/IP-пакете в проводных системах связи крайне мала. Следует отметить, что можно использовать для передачи данных, сжатых кодеком Speex и TCP/IP-пакеты. В последнем случае, вероятность порчи данных становится мизерной. Эта особенность определяет отсутствие в формате данных кодека Speex каких-либо средств защиты от ошибок.

Процесс кодирования речи контролируется специальным параметром, представляющим собой число в диапазоне от 0 до 10. В режиме с постоянной скоростью (англ. Constantbit-rate, CBR) параметр качества — целое число, а в режиме с переменной скоростью — число с плавающей запятой.

Кодек Opus. Opus (ранее Harmony) — аудиокодек для сжатия с потерями, разработанный сообществом InternetEngineeringTaskForce(IETF) специально для применения в приложениях реального времени в интернете. Является открытым форматом стандартизованным в RFC 6716, а эталонная реализация распространяется под 3-условной лицензией BSD. Все известные патенты под которые попадает новый кодек Opus доступны под лицензией royalty-free.

Основным преимуществом данного кодека, является низкая задержка кодирования (от 2.5 мс до 60 мс, настраиваемо), более высокая компрессия аудио данных, поддержка многоканального звука (до 255 каналов).

Формат битового потока был зафиксирован 8 января 2012 года.

Поддерживается в разработанных Mozilla программах Firefox и

Thunderbird с 15 версии. Также Skype переходит на Opus[6]

Как несложно догадаться, движок SILK используется для компрессии голоса, то есть, по сути, почти всегда в тех случаях, когда требуется эффективно использовать пропускную способность канала связи. В данном случае кодек начинает обработку аудиосигнала с того, что анализирует его на предмет наличия человеческой речи. Голосовые составляющие отделяются от прочих звуков, после чего кодек производит анализ частотной характеристики звука, понижая уровень дискретизации для данных, содержащих голосовую информацию, то есть для речи. Далее Opus исследует присутствующие шумы и оптимизирует сигнал для определенного битрейта.

На следующем этапе кодек преобразует сигнал с помощью фильтра предварительной очистки – напомним, все это происходит в режиме реального времени! Используя речевые кадры, модуль предсказания частоты аудиосигнала вносит изменения в последующие кадры, после чего частотное квантование нормирует частоты человеческой речи.

На третьем этапе обработки звука из сформированного сигнала устраняются искажения, которые неизбежно возникают при недостаточно высоком битрейте. После этого подключается модуль формирования шума квантования, который снижает шумы внутри рабочей полосы, вытесняя их за пределы рабочего диапазона.

Ну и на заключительном, четвертом этапе интервального кодирования SILK работает с дискретными величинами, которые могут принимать ограниченное число значений, – производится покадровый вывод сигнала.

Если же Opus приходится кодировать аудиоданные с высоким качеством, такие как музыка, то в дело вступает модуль CELT. Механизм его работы похож на принцип работы большинства кодеков с потерями – он завязан на дискретные косинусные преобразования, а также на «оптимизацию» звука. Ее суть состоит в том, что из сигнала удаляются составляющие, которые не несут полезной нагрузки для слуха человека, – до кодирования он их или не слышит, или слышит с большим трудом.

В программах, в которых предусмотрено управление настройками кодирования, новый кодек предлагает пользователю выбрать режим сжатия – речь или музыку. Эта настройка и определяет приоритет того или иного алгоритма кодирования Opus.

Формат WindowsMediaAudio(WMA). WindowsMediaAudio — лицензируемый формат файла, разработанный компанией Microsoft для хранения и трансляции аудиоинформации.[5]

Изначально формат WMA рекламировался как альтернатива MP3, но на сегодняшний день Apple противопоставляет ему формат AAC (используется в популярном онлайн-музыкальном магазине iTunes).

Большинство портативных аудиопроигрывателей поддерживает формат WMA наряду с MP3. Данный формат очень плохо поддерживается на альтернативных платформах (вследствие его закрытости). Популярность формата WMA была достигнута за счет его использования при воспроизведении видеофайлов на DVD-плеерах, а также на портативных устройствах и мобильных телефонах Playstation и Nokia. Обычно сжатие в формат WMA происходит с потерями, однако возможно сжатие и без потерь.

Формат WMA имеет хорошее сжатие, но на низких скоростях потока имеются проблемы, особенно в ранних версиях формата. Так же WMA формат известен своими ошибками. Если по какой-то причине некоторая часть файла .WMA повреждается, то прослушивание файла становится проблематичным, как за пару-тройку десятков секунд до места повреждения, так и после. По сравнению с mp3, где можно прослушать файл вплоть до ошибки и сразу же после.

Microsoft включила в WMA поддержку цифровой системы управления авторскими правами (DRM) (система защиты). Основным следствием её является невозможность прослушивать защищенные композиции на других компьютерах, кроме того, на котором композиция была загружена из музыкального магазина.

В последних версиях формата, начиная с WindowsMediaAudio 9.1,

предусмотрено кодирование без потери качества (англ. lossless), многоканальное кодирование объёмного звука и кодирование голоса.

Формат RealMedia. RealMedia — проприетарный стандарт на потоковое вещание и на формат медиафайлов, принадлежащий фирме «RealNetworksProductsandServices».[7]

Файлы RealMedia оптимизированы для потокового воспроизведения в сети интернет, поэтому аудио и видео информация в них кодируется с постоянным битрейтом для гарантированного соответствия пропускной способности интернет соединения. Для потокового воспроизведения файлов RealMedia используется протокол RTSP (Real Time Streaming Protocol). Однако этот стандартный протокол используется только для установления и управления соединением, аудио и видео данные передаются с использованием проприетарного протокола компании RealNetworks – RDTP (RealDataTransportProtocol).

Для поддержки аудио и видео с переменным битрейтом компания RealNetworks разработала новый формат файлов использующий расширение ".rmvb". Эти файлы в основном предназначены для хранения и воспроизведения видео на локальных компьютерах.

Формат RealMedia использует FourCC (fourcharactercodes) для идентификации элементов файла. FourCC – это 32-битный код, состоящий из четырех или менее символов в кодировке ASCII. Основным строительным блоком файлов RealMedia является 'часть' (chunk). Каждая часть состоит из следующих полей:

- 1) FourCC - 32-битный идентификатор частей
- 2) 32-битное поле определяющее размер данных хранящихся в части
- 3) данные в части

Части могут входить в состав друг друга создавая сложные структуры.(Рис 1.6.)

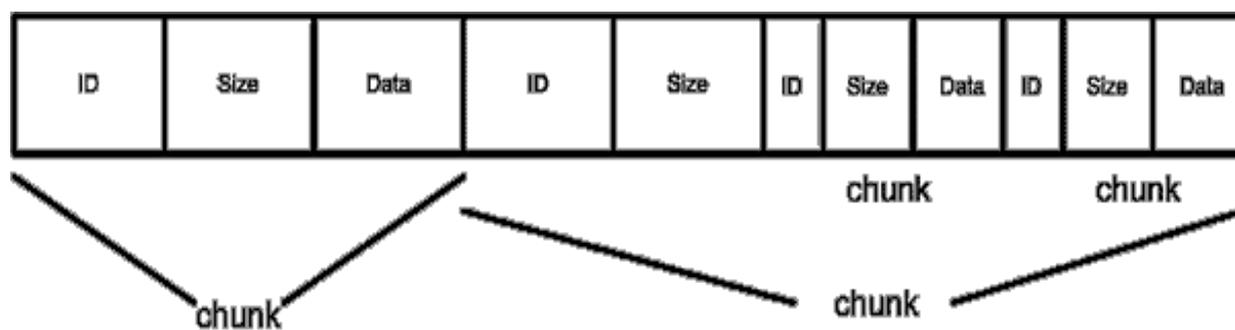


Рис.1.6. Структура файла RealMedia

1.5. Анализ эффективных форматов сжатия звука без потерь

Кодек FLAC. FLAC (англ. FreeLosslessAudioCodec) — популярный свободный кодек, предназначенный для сжатия аудиоданных без потерь.[5]

Сегодня формат FLAC поддерживается множеством аудио приложений и портативных аудио плееров, а также имеет большое количество аппаратных реализаций

Упрощенно говоря, формат FLAC похож на формат MP3, но без потерь, а это означает, что звук сжимается в FLAC без потери качества. Это подобно тому, как работает ZIP, но в случае с файлом .FLAC вы получите гораздо лучшее сжатие, поскольку она предназначена специально для аудио, и вы можете воспроизводить сжатые файлы FLAC в вашем любимом проигрывателе (или в вашем автомобиле, или в домашней стерео системе), точно так же, как вы проигрывали бы MP3-файл.

Файлы FLAC находятся в свободном доступе и поддерживаются большинством операционных систем, включая Windows, «Unix» (Linux, *BSD, Solaris, OS X, IRIX), BeOS, OS/2, и Amiga. Созданы системы для Autotools, MSVC, Watcom C и Project Builder.

FreeLosslessAudioCodec, или FLAC, хорошо известен и профессионалам в области звукозаписи, и рядовым меломанам. Этот кодек используется для сжатия аудио: благодаря компрессии музыкальные коллекции занимают меньше места, и их хранение становится более удобным. При этом FLAC формат относится к так называемым lossless-форматам, которые при декомпрессии обеспечивают звучание, абсолютно

аналогичное оригинальному. В файлах с расширением .flac музыка соответствует оригинальной записи вплоть до бита, за счет того, что при сжатии кодек не отбрасывает информацию, которую считает «лишней» (как это характерно для lossy-форматов, к которым относится и всем известный MP3). Естественно, качество FLAC ценят поклонники хорошего звука, но в полной мере насладиться им можно лишь при наличии высококлассной аудиотехники.

Разработчик FLAC, некоммерческая организация Xiph.Org Foundation, выпустила новый релиз популярного кодека в середине 2013 года (до этого пользователи могли лишь скачать FLAC образца 2006 года). FLAC 2013 продолжает радовать своей кроссплатформенностью и бесплатностью. Для рядового пользователя это означает, что бесплатная музыка FLAC будет доступна для него на самых различных устройствах, включая планшеты, смартфоны и мультимедийные проигрыватели. Новый релиз имеет такую же широкую аппаратно-программную поддержку, которая обеспечила популярность и предыдущей версии.

Аудиокодек AppleLossless. AppleLossless (также известный как AppleLosslessEncoder, ALE, или AppleLosslessAudioCodec, ALAC) — открытый аудиокодек для сжатия без потерь качества цифровой музыки, разработанный Apple.[5]

Данные, сжатые AppleLossless, хранятся в контейнере MP4 с расширением m4a. Хотя AppleLossless имеет такое же расширение файла, как AAC, это не AAC: данный кодек схож с другими lossless-кодеками, такими, как FLAC и др. Плееры iPod с док-разъемами и свежей прошивкой могут воспроизводить звук формата AppleLossless (iPodshuffle 4-го поколения с прошивкой 1.01 тоже воспроизводит AppleLossless). В кодеке не используются какие-либо специфические средства цифрового управления правами (DRM), но использование DRM возможно, поскольку предусмотрено форматом контейнера.

Тесты показали, что сжатые ALAC файлы получаются размером

примерно от 40 % до 60 % размера оригиналов в зависимости от вида музыки, подобно другим lossless-форматам. Кроме того, скорость, с которой он может быть декодирован, делает его полезным для устройств с низкой производительностью, таких как iPod и подобных.

AppleLosslessEncoder был представлен в качестве одного из компонентов QuickTime 6.5.1 28 апреля 2004 года и как функция iTunes 4.5. Кодек используется также в AirPort Express в приложении AirTunes.

Декодер для формата AppleLossless есть в свободной библиотеке libavcodec. Это означает, что любой мультимедийный проигрыватель на основе этой библиотеки, включая Медиапроигрыватель VLC и MPlayer, может проигрывать AppleLossless-файлы.

Формат Monkey'sAudio. Monkey'sAudio — популярный формат кодирования цифрового звука без потерь.[8] Кодек распространяется в виде бесплатного одноимённого программного обеспечения для MicrosoftWindows, а также плагинов к популярным медиаплеерам. Файлы Monkey'sAudio используют следующие расширения: .are для хранения аудио и .apl для хранения метаданных. Несмотря на то, что кодек распространяется бесплатно вместе с исходным кодом, Monkey'sAudio не является формально свободным, так как его лицензия накладывает некоторые ограничения на использование. Первый пункт лицензионного соглашения: исходный код и Monkey'sAudio SDK могут быть свободно использованы для добавления поддержки воспроизведения, кодирования или редактирования тегов в любой продукт, свободный или коммерческий.

Официально кодек Monkey'sAudio выпускается только для платформы MicrosoftWindows, хотя существует ряд неофициальных кодеков для GNU/Linux, BeOS и Mac OS X. FFmpeg поддерживает данный формат.

Программное обеспечение:

- MPlayer
- VLC
- Winamp

- AIMP
- JetAudio
- foobar2000

Monkey'sAudio - это быстрый и мощный звуковой компрессор, который может без потерь сжать цифровую музыку. В отличие от традиционных методов сжатия, например, таких как mp3, ogg, или lqt, при использовании которых приходится навсегда отказываться от качества звука для того, чтобы уменьшить размер аудио файла, Monkey'sAudio только улучшает копии ваших аудио записей бит за битом.

Это означает, что сжатые звуковые файлы будут казаться такими же качественными, как и оригинальные. Даже качественные аудио записи могут занимать совсем немного места. (Представьте, что эта программа является своего рода архиватором наподобие Winzip, но только для музыки).

Аудиокодек TrueAudio. TrueAudio (ТТА) — свободный аудиокодек, осуществляющий сжатие аудиофайлов без потерь, способный работать в режиме реального времени. Кодек основан на адаптивных предсказывающих фильтрах и обладает такими же или лучшими характеристиками, как и большинство современных кодеков без потерь.[5]

ТТА обеспечивает сжатие мультисканальных 8-, 16- и 24-битных WAV-файлов без потерь. Термин «беспотерьный» означает, что звуковые данные качество не будут потеряны при сжатии и распакованные данные не будут отличаться от оригинальных. Показатели сжатия ТТА-кодека зависят от содержания сжимаемого музыкального файла, но сжатый размер в основном будет в пределах от 30 до 70 % от оригинального. Формат ТТА поддерживает оба формата информационных тэгов ID3v1/ID3v2 и теги APEv2.

ТТА-кодек позволяет хранить до 20 звуковых компакт-дисков (CD-Audio) на одном DVD-R, с поддержкой тэгов ID3 и APEv2.

Исходные коды ТТА-аудиокодека без потерь и дистрибутивы проекта свободно доступны и распространяются под универсальной общедоступной лицензией LGPL.

Таблица 1.2.

ПО с поддержкой TTA

Наименование	Примечание
<u>CorePlayer</u>	мультиплатформенный мультимедиа плеер, встроенная поддержка
<u>K-Multimedia</u>	мультимедиа плеер, встроенная поддержка
VLC	медиа плеер, встроенная поддержка
Foobar2000	аудиоплеер, поддержка через <u>плагин</u>
<u>Winamp</u>	аудиоплеер для Windows, поддержка через <u>плагин</u>
<u>JetAudio</u>	многоформатный проигрыватель аудио- и видеофайлов, встроенная поддержка
AIMP	аудиоплеер, встроенная поддержка
<u>DeaDBeeF</u>	аудиоплеер, поддержка через плагин (расширение)
XRecode II	Многоформатный аудиоконвертер, встроенная поддержка
Плееры с поддержкой <u>DirectShow</u>	например, <u>WindowsMediaPlayer</u>
Плееры на движке <u>Gstreamer</u>	например, <u>amaroK</u>

Аудиокодек TAK. TAK (нем. Tom's verlustfreier Audio Kompressor) — аудиокодек и формат сжатия цифрового звука без потерь. Отличается высокой степенью сжатия и скоростью кодирования и декодирования. Использует теги APEv2.

Распространяется бесплатно вместе с набором программного обеспечения для кодирования и воспроизведения, а также плагинами к популярным плеерам: Winamp, foobar2000 и др.

Разработчик — ThomasBecker, Германия. Первая финальная версия 1.0

была опубликована 26 января 2007 года. Формат продолжает активно развиваться. Согласно проводимому опросу на форуме hydrogenaudio.org, входит в число четырёх (четвёртое место) наиболее популярных форматов аудиосжатия без потерь (после FLAC, WavPack и ALAC)

ТАК – этот активно разрабатываемый кодек не перестает радовать. Если брать во внимание все три параметра (сжатие, кодирование, декодирование), ТАК выглядит наиболее привлекательно. Высокая скорость работы объясняется активным использованием процессорных оптимизаций (в т.ч. SSSE3). А использование двух ядер дает почти двукратный прирост скорости кодирования! Таким образом, в случае с ТАК преимущество от использования современных процессоров наиболее ощутимо.

Аудиокодек WavPack. WavPack — свободный и универсальный lossless и lossy аудиокодек. Позволяет сжимать и восстанавливать 8-, 16-, 24- и 32-битные аудиофайлы. Поддерживает многоканальные потоки (до 255) и высокие частоты дискретизации (до 192 кГц). Эффективность сжатия зависит от исходных данных, но обычно оно лежит в диапазоне между 30 % и 70 % для обычной популярной музыки, выше для классической музыки и других аудиоданных с более широким динамическим диапазоном.[5]

Обзор

- современные и инновационные технологии сжатия звука;
- сверхбыстрое сжатие и декомпрессия;
- открытый исходный код, доступный по BSD-подобной лицензии;
- мультиплатформенность;
- устойчивость к ошибкам;
- поддержка потокового вещания;
- поддержка мультисканального аудио и высокого разрешения;
- гибридный режим;
- аппаратная поддержка:
- COWON (через прошивку RockBox);

- мультимедийные плееры ICONBIT HD390DVD/ HD400L/ HD400DVD;
- iRiver, Gigabeat (Toshiba) (через прошивку RockBox);
- iAudio (Cowon) (через прошивку RockBox);
- iPod (через прошивку RockBox);
- SANSА (через прошивку RockBox);
- Samsung (через прошивку RockBox);
- теги (ID3v1, APE);
- поддержка RIFF;
- совместимость с ReplayGain;
- возможность создания самораспаковываемых файлов для Win32;
- поддержка 32-битных потоков с плавающей точкой;
- поддержка встраиваемых Cuesheet;
- включает MD5-хеши для быстрой проверки целостности;
- может кодировать симметрично и асимметрично (означает более медленное сжатие для ускорения декомпрессии).

Формат ADX. ADX – основанный на АДИКМ проприетарный формат сжатия с потерями и хранения звукозаписи, разработанный CRI Middleware специально для использования в видеоиграх.[5] Наиболее характерная особенность – возможность зациклить звукозапись, что делает применение формата удобным для использования в качестве фоновой музыки в различных играх, поддерживающих этот медиаконтейнер. Его поддерживают множество игр для SEGA Dreamcast некоторые игры для PlayStation 2, Game Cube и Xbox. Одной из первых игр, использовавших ADX, была Burning Rangers на базе Sega Saturn. Наиболее заметно использовался в играх серии SonictheHedgehog, начиная с поколения Dreamcast, вплоть до наших дней (исключением стала лишь игра Sonic 2006), в которой этот формат использовался для записи музыки и голосов персонажей.

Поверх основного кодирования АДИКМ, набор инструментов ADX также включает родственный формат АНХ, который использует вариант спецификации звукозаписи MPEG-2 для записи голоса и архивирования, а также AFS для хранения сборки нескольких дорожек ADX и АНХ в одном файле мультимедиа контейнера.

На данный момент CRI Middleware более не лицензирует этот кодек. Ему на замену пришёл кодек CRI ADX2, также известный как HSA.

ADX — формат сжатой звукозаписи, однако, в отличие MP3, в нём не применяется психоакустическая модель уменьшения объёма данных о звуке (уменьшения его сложности). Вместо этого модель ADPCM использует для сохранения образцов запись данных относительной ошибки функции предсказания, что означает большую сохранность исходного сигнала после кодирования; по существу, сжатие ADPCM, вместо использования полных переразмерянных образцов звукозаписи, предоставляет образцы отклонения сигнала от предыдущего значения, которые имеют гораздо меньший размер, обычно — 4 бита. Для человеческого уха такое отклонение находится на уровне шума, что делает потерю качества едва заметной.

Как и другие форматы кодирования, ADX поддерживает несколько частот дискретизации, такие как 22050 Гц, 44100 Гц, 48000 Гц и т. д., тем не менее, глубина образца зафиксирована на значении 16 бит из-за первоначальной привязки к игровому оборудованию. Он поддерживает многоканальность, однако есть мнение, что под этим подразумевается стерео (2 канала), хотя сам файл поддерживает включение в него до 255 каналов.

Аудиокодек MPEG-4 ALS. MPEG-4 ALS является продолжением аудио семейства MPEG-4 кодирования для сжатия без потерь аудиоданных [9]. Основной кодек ALS на основе вперед адаптивного прогнозирования линейной, который предлагает хорошую производительность сжатия с низким уровнем сложности. Дополнительные функции включают в себя долгосрочный прогноз, многоканальное кодирование и сжатие с плавающей точкой аудио материала. ALS также предлагает хорошую гибкость с точки

зрения сжатия сложности компромисса, начиная с очень низкой сложности реализации в максимальных режимах сжатия, так что адаптироваться к различным требованиям. Его многие другие функции включают в себя:

- -Общая поддержка практически любого несжатого цифрового аудио формата, включая формат файла Sony Wave64 и формат Broadcast Wave (BWF).
- -Поддержка линейной ИКМ (импульсно-кодовая модуляция) с разрешением до 32 бит при произвольных частотах дискретизации.
- -Multichannel поддержки / многодорожечной до 65536 каналов, в том числе 5.1-, 7.1- и 22.2-канальный объемный звук.
- -Поддержка аудио данных в 754 32-битном плавающей точкой аудио-формате IEEE.
- -Быстрый произвольный доступ к закодированной информации

Возможные приложения

Lossless кодирования звука в целом и MPEG-4 ALS, в частности, есть много приложений, на обоих профессиональных и потребительских уровнях.

К ним относятся:

- -Интернет Распределение аудио файлов (потокное, онлайн музыкальных магазинах, и загрузка)
- Формат диска Высокая разрешения
- -Portable Музыкальные плееры
- -Archival Системы (для вещания, студий, звукозаписывающих и цифровой передачи)
- -Студия операции (для хранения, совместного труда, цифровой резервного копирования и передачи цифровых)

В этих приложениях, MPEG-4 ALS используется в качестве кодирования двигателя для сжатия без потерь данных аудио. Для архивных систем и операций студии, инструмента архивирования, основанных на MPEG-4 ALS, описанных в разделе 3 будет более подходящим.

Выводы

Как было установлено в результате проведенного анализа, на сегодняшний день существует довольно много методов и стандартов сжатия аудио сигналов, к которым можно отнести: MP-3 (MPEG – 1 audiolayer 3), WMA (windows media audio), OggVorbis, Musicam, ATRAC, MPEG -2, AAC, MPEG -4ALS, AMP3

Один из перечисленных кодеков Monkey's Audio — популярный, но ресурсоёмкий кодек. Дает действительно высокое сжатие, но имеет проблемы с прокруткой. А также есть еще кодек TAK — этот активно разрабатываемый кодек не перестает радовать. Если брать во внимание все три параметра (сжатие, кодирование, декодирование), TAK выглядит наиболее эффективнее. Высокая скорость работы объясняется активным использованием процессорных оптимизаций (в т.ч. SSSE3). А использование двух ядер дает почти двукратный прирост скорости кодирования! Таким образом, в случае с TAK преимущество от использования современных процессоров наиболее ощутимо.

TrueAudio (TTA) — тут надо отметить разве что очень высокую скорость кодирования и приемлемую степень сжатия. При этом скорость декодирования нельзя назвать очень высокой. Степень сжатия чуть выше чем FLAC. FLAC — степень сжатия средняя, но вот скорость декодирования порадовала. Правда, главной причиной лидерства этого кодека среди общественности является открытый исходный код, и, как следствие, широчайшая аппаратная/сотовая поддержка.

Apple Lossless (ALAC) — один из тех кодеков, которые ничем не примечательны, но продолжают активно пользоваться разработчиками (в данном случае Apple). Низкая степень сжатия и скорость декодирования. Скорость сжатия — средняя. Подойдет разве что пользователям iPod.

В следующей главе будем рассматривать различие форматов, достоинства и недостатки

2. СРАВНИТЕЛЬНЫЙ АНАЛИЗ СОВРЕМЕННЫХ ФОРМАТОВ СЖАТИЯ ЗВУКА

2.1 Сравнение сжатие аудиоданных

Сжатие (компрессия) аудиоданных представляет собой процесс уменьшения скорости цифрового потока за счет сокращения статистической и психоакустической избыточности цифрового звукового сигнала.

Структура кодера сжатия аудиоданных с потерями

Исходный цифровой звуковой сигнал разделяется на частотные поддиапазоны и сегментируется по времени в блоке временной и частотной сегментации.[10]

Длина кодируемой выборки зависит от формы временной функции звукового сигнала. При отсутствии резких выбросов по амплитуде используется так называемая длинная выборка, обеспечивающая высокое разрешение по частоте. В случае же резких изменений амплитуды сигнала длина кодируемой выборки резко уменьшается, что дает более высокое разрешение по времени. Решение об изменении длины кодируемой выборки принимает блок психоакустического анализа, вычисляя значение психоакустической энтропии сигнала (Рис.2.1).



Рис. 2.1. фывфывфыв

После сегментации сигналы частотных поддиапазонов нормируются,

квантуются и кодируются. В наиболее эффективных алгоритмах компрессии кодированию подвергаются не сами отсчеты выборки звукового сигнала, а соответствующие им коэффициенты МДКП.

Учет закономерностей слухового восприятия звукового сигнала выполняется в блоке психоакустического анализа. Здесь по специальной процедуре для каждого частотного поддиапазона рассчитывается максимально допустимый уровень искажений (шумов) квантования, при котором они еще маскируются полезным сигналом данного поддиапазона.

Блок динамического распределения бит в соответствии с требованиями психоакустической модели для каждого поддиапазона кодирования выделяет такое минимально возможное их количество при котором уровень искажений, вызванных квантованием, не превышает порога их слышимости, рассчитанного психоакустической моделью.

Многие другие приёмы могут послужить способом сократить объём данных звуковой информации. Даже простое сужение полосы частот сигнала вместе с уменьшением динамического диапазона может уже называться сжатием аудиоданных. Например, в стандарте сжатия звука в сотовой связи используется и то и другое. Стремясь удалить избыточность из звука, кодек при плохом качестве сигнала становится избирателен к определённым словам, упорно проглатывая их.

Субъективная оценка качества.

Для сжатых аудиоданных существует субъективная оценка качества, оцениваемая как процент людей, почувствовавших разницу с оригиналом.

Соответствие битрейта кодека MP3 в режиме стерео и процента людей заметивших разницу с оригиналом приведены в таблице 2.1.

Таблица 2.1.

Субъективная оценка стерео режима кодека MP3

Приблизительное количество людей, услышавших разницу между оригинальной и сжатой записями, %	Битрейт сжатой записи, кбит/сек
0...1	320
5...30	256
30...40	192
40...70	128

Следует учесть тот факт, что качество получившегося материала зависит от характера сжимаемых данных, от жанра, наличия фона, помех. После сжатия, например MP3, на средних битрейтах, слушатели отмечают оловянность перкуSSIONных. А на голосе сжатие (даже сильное) отражается мало.[5]

Таблица 2.2.

Сравнение аудиоформатов

Название формата	Квантование, бит	Частота дискретизации, кГц	Число каналов	Величина потока данных с диска, кбит/с	Степень сжатия/упаковки
CD	16	44,1	2	1411,2	1:1 без потерь
Dolby Digital (AC3)	16-24	48	6	до 640	~12:1 с потерями
DTS	20-24	48; 96	до 8	до 1536	3:1 с потерями

DVD-Audio	16; 20; 24	44,1; 48; 88,2; 96	6	6912	1:1 без потерь
DVD-Audio	16; 20; 24	176,4; 192	2	4608	1:1 без потерь
MP3	16-24	до 48	2	до 320	~11:1 с потерями
AAC	16-24	до 96	до 48	до 512	с потерями
AAC+ (SBR)	16-24	до 48	2	до 320	с потерями
OggVorbis	до 32	до 192	до 255	до 500	с потерями
WMA	до 24	до 96	до 8	до 768	2:1, есть версия без потерь

Как видно из таблицы 2.2 кодек DVD-Audio имеет большую частоту дискретизации, величину потока данных с диска и степень сжатия, а кодек OggVorbis имеет большое квантование и число каналов по сравнению с другими форматами.

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.

Таблица 2.3.

Сравнение и возможности аудиокодеков

Возможности	FLAC	WavPack	TAK	Monkey's	Optim FROG	ALAC	WMA	Shorten	LA	TTA	MPEG-4 ALS	MPEG-4 SLS	RealLossless
Кодирование	Быстро	очень быстро	очень быстро	быстро	медленно	Быстро	средне	очень быстро	очень медленно	быстро	средне	медленно	медленно
Декодирование	очень быстро	быстро	очень быстро	средне	медленно	Быстро	быстро	очень быстро	медленно	быстро	быстро	медленно	очень быстро
Сжатие*	55.7%	55.7%	53.9%	53.7%	53.2%	56.7%	57.1%	60.9%	52.1%	55.3%	55.1%	?	55.4%
Гибкость**	очень хорошо	Плохо	плохо	Плохо	средне	плохо	очень хорошо	плохо	плохо				

Error handling	Да	да	да	нет	да		да	нет	нет	да	да	да	
Seeking	Да	да	да	да	да	Да	да	да	да	да	да	да	да
Теги	Vorbis tags	ID3/APE	APEv2 (exp.)	ID3/APE	ID3/APE	iTunes	ASF	нет	ID3 v1	ID3	да	да	proprietary
Аппаратная поддержка	очень хорошо	ограничено	нет	ограничено	нет	Хорошо	ограничено	ограничено	нет	ограничено	нет	нет	нет
Программная поддержка	очень хорошо	хорошо	средне	хорошо	средне	Средне	хорошо	очень хорошо	плохо	средне	плохо	плохо	плохо
Гибридный (без потерь)	Нет	да	нет	нет	да	Нет	нет	нет	нет	нет	нет	да	нет

Replay Gain	Да	да	да	нет	да	Sortof	нет	нет	нет	да	да	да	нет
RIFF chunks	Да	да		да	да		нет	да	да	нет			
Поддержка потока	Да	да	да	нет	да	Да	да	нет		нет	да	да	да
Pipesupport	Да	да	да	да	да	Да	да	да	да	нет			нет
Открытый исходный код	Да	да	нет	да	нет	Да	нет	да	нет	да	да	да	нет
Многоязычный	Да	да	да	нет	нет	Да	да	нет	нет	да	да	да	нет
Highres	Да	Да	да	да	да	Да	да	нет	нет	да	да	да	нет

olution													
Поддер жка ОС	Кросспла тформенн ый	Кросспла тформенн ый	Win/Li nuxWi ne	Кросспла тформенн ый	Win/M ac/Lin ux	Кросспла тформенн ый	Win/ Mac	Кросспла тформенн ый	Win/ Linu x	Кросспла тформенн ый	Кросспла тформенн ый	Кросспла тформенн ый	Win/M ac/Lin ux

Из таблицы 2.3 видно какие возможности имеют форматы сжатия. Перечислим форматы имеющие большую возможность:

- WavPack, TAK - кодирование очень быстрое.
- FIAC, TAK, Real Lossless – декодирование очень быстрое.
- TAK, Monkey`s, Optim FROG – сжатие меньше чем у других.
- WavPack, TAK, FIAC, Monkey`s, Optim FROG, MPEG-4ALS - очень хорошая гибкость.
- FIAC - Vorbis tags, WavPack - ID3/APE, Monkey`s –ID3/APE, Optim FROG, MPEG-4ALS – тоже есть теги.
- FIAC – аппаратная поддержка очень хорошая.
- FIAC – программная поддержка очень хорошая, но еще можно WavPack, Monkey`s, WMA.
- WavPack, TAK, FIAC, Optim FROG, ALAC, WMA, MPEG-4ALS – есть поддержка потока, открытый исходный код, многоканальный, Highresolution, Поддержка ОС.

Коэффициент сжатия вычисляется как отношение размера после сжатия к размеру до сжатия, умноженное на 100%. Таким образом, чем меньше значение, тем лучше.

Скорость кодирования и декодирования и коэффициент сжатия, приведённые здесь, получаются при использовании стандартных настроек кодировщика. Под гибкостью понимается число вариантов кодирования, доступных пользователям (быстро/низкое сжатие, медленно/высокое сжатие и все промежуточные варианты)

2.2 Сравнение цифровых аудио форматов

В этой таблице сравниваются основные технические данные различных цифровых аудиоформатов. [11] Сравниваются как форматы файла без сжатия, так и с применением сжатия. Сравнение звуковых форматов без потерь. (2.4)

Таблица 2.4.

Сравнение звуковых форматов без сжатия

Название формата	Расширение файла	Квантование, бит	Частота дискретизации, кГц	Число каналов	Битрейт (на канал), Мбит/с	Степень сжатия/упаковки	Назначение	Выпуск
WAVE (WAV)	.wav	8; 16; 24; 32	любая	1; 2; 3; 4; 6	до 6,144	1:1; без сжатия	хранение звуковых данных на ПК	1991, Microsoft и IBM
DSD	.dff; .dsf	1	2822,4 (64f _s); 5644,8 (128f _s)	2, 5.1	2,8224; 5,6448	1:1; без сжатия, возможно сжатие DST	SACD	1998, Sony и Philips
DigitaleXtremeDefinition (DXD)	DXD	24; 32	352,8	2, 5.1	8,4672; 11,2896	1:1; без сжатия	профессиональное производство SACD	2004, Sony
AIFF	.aiff; .aif	8; 16; 24; 32	11,025; 22,05; 24; 32; 44,1; 48; 96; 192	1; 2; 3; 4; 6	до 6,144	1:1; без сжатия	хранение звуковых данных на ПК	1988, Apple

Из таблицы 2.4 следует ,что звуковой формат WAVE (WAV) более эффективнее.

WAVE – подмножество спецификации RIFF Майкрософта для хранения мультимедийных файлов. RIFF файл начинается с заголовка файла, за которым последовательно следуют отсчеты данных.

В потоке данных WAVE могут быть дополнительные фрагменты (название альбома, автор, и.т.д.). Если они есть, то каждый дополнительный фрагмент будет иметь 4х байтные переменные: SubChunkID - название, SubChunkSize без знака и количество данных SubChunk2Size.

По сравнению DSD чем WAVE частота дискретизации и число каналов больше

2.3 Сравнение звуковых форматов со сжатием без потерь

Сжатие без потерь применяется в тех случаях, когда информацию нужно восстановить с точностью до бита. Сжатие данных без потерь используется когда важна идентичность сжатых данных оригиналу. Обычный пример — исполняемые файлы, документы и исходный код. Программы использующие форматы сжатия без потерь называются архиваторы, все знают популярные файловые форматы ZIP, или RAR, Unix-утилиту Gzip и т.д.

Все эти программы отличаются примененными алгоритмами (одним или несколькими) и потому разными свойствами сжатия разных файлов. Есть и другие варианты — обычное архивирование (ZIP, RAR) или специальные программы типа WavPack, Monkey's Audio, RK Audio, FLAC, LPAC Archiver, Shorten.

Рассмотрим сравнение звуковых форматов со сжатием без потерь по таблице 2.5 и выделим более эффективные форматы.

Таблица 2.5

Сравнение звуковых форматов со сжатием без потерь

Название формата	Лицензия	Расширение файла	Квантование, бит	Частота дискретизации, кГц	Число каналов	Степень сжатия/упаковки	Назначение	Выпуск
Shorten		.shn	16	44.1	2	3:1 — 5:1, сжатие без потерь	хранение звуковых данных на ПК	1994, TonyRobinson
WavPack	свободный, реализация под BSD	.wv	8; 16; 24; 32	6 — 192	1 — 256	1.4:1 — 3.3:1, сжатие без потерь	хранение звуковых данных на ПК	1998, ConiferSoftware
MeridianLosslessPacking (MLP)		.mlp	до 24	до 192	1; 2; 5.1; 6.0; 8.0	~2:1, сжатие без потерь	DVD-Audio	1998, MERIDIAN
RK Audio (RKAU)^[2]		.rka	16	44.1	2	2:1, сжатие без потерь, сжатие с незначительными потерями	хранение звуковых данных на ПК	2000, MalcolmTaylor, RK Software
FLAC	свободный; реализация:	.flac	4 — 32	1Гц — 655.350	1 — 8	1.4:1 — 4:1, сжатие без	хранение звуковых данных	2000, JoshCoalson

	утилиты подGPL, библиотеки подBSD			кГц с шагом 1 Гц		потерь	на ПК, звуковое сопровождение к HD-видео, медиаплееры	
Monkey's Audio		.ape	16; 24	8; 11.025; 12; 16; 22.05; 24; 32; 44.1; 48; 96	2	1.4:1 — 4:1, сжатие без потерь	хранение звуковых данных на ПК	2000+, Matthew T. Ashland
OptimFROG ^[2]		.ofs	до 32	до 192	2	1.4:1 — 4:1, сжатие без потерь	хранение звуковых данных на ПК	2001, Florin Ghido
Lossless Predictive Audio Coder (LPAC) ^[4]		.pac	8; 16; 20; 24	до 192	2	1.5:1 — 4:1, сжатие без потерь	хранение звуковых данных на ПК	2002, Tilman Liebchen, Marcus Purat, Peter Noll
Lossless Audio (LA) ^[2]		.la	16	48	2	1.4:1 — 3.3:1, сжатие без потерь	хранение звуковых данных на ПК	2002, Michael Bevin
Windows Media Audio 9 Lossless		.wma	16; 24	8; 11.025; 16; 22.05; 32; 44.1; 48; 88.2; 96	до 6	1.7:1 — 3:1, сжатие без потерь	хранение звуковых данных на ПК	2003, Microsoft
Apple Lossless	Apache 2.0	.m4a	16; 24	44.1; 48;	до 6	1.7:1 —	хранение	2004, Apple Inc.

(ALAC, ALE)				88.2; 96; 192		2.5:1, сжатие без потерь	звуковых данных на ПК, плееры iPod	
RealAudioLoss less (RAL, ralf)		.rmvb	16 и др.	44.1 и др.	2	н/д, сжатие без потерь	Потоковое мультимедиа	2004, RealNetworks
TrueAudio (TTA)		.tta	8; 16; 24	0—4 ГГц	65535	1.4:1 — 3.3:1, сжатие без потерь	хранение звуковых данных на ПК	2004, Александр Джурик
DTS-HD Master Audio (DTS++, DTS HD)		-	до 24	до 192	до 8	2:1 — 4:1 , сжатие без потерь	Blu-ray Disc, HD DVD ,PlayStation 3	2004, DigitalTheaterSy stem.
DirectStreamT ransfer (DST)		DSD	1	64*44.1; 128*44.1; 256*44.1	2; 5.1	сжатие без потерь	DSD-поток при производствеSAC D	2005, MPEG-4 ISO/IEC 14496- 3:2001/Amd 6:2005
DolbyTrueHD		.mlp	до 24	до 192	до 14	2:1 — 4:1 , сжатие без потерь	Blu-ray Disc, HD DVD	2005, DolbyLaboratori es
ATRAC AdvancedLossl ess (AAL)		.aa3; .oma; .at3	16	44.1	2	1.25:1 — 3:1, сжатие без потерь	Минидискплееры, PlayStation Portable , Playstation 3	2006, SonyCorporation
MPEG-4		.mp4	8; 16; 20; 24;	44.1; 48;	до 65536	1.5:1 — 4:1,	музыкальные	2006, MPEG

Audio Lossless Coding (ALS)		.als	32	88.2; 96; 192; (384)		сжатие без потерь	интернет магазины, Потокое мультимедиа, дисковые форматы высокого разрешения, плееры, архивные системы, профессиональное студийное производство, звуковое сопровождение к MP4-видео	ISO/IEC 14496-5:2001/Amd 10:2007/Cor 3:2009
Tom's lossless Audio Kompressor (TAK)		.tak	16; 24	до 192	2	1.4:1 — 3.3:1, сжатие без потерь	хранение звуковых данных на ПК	2007, ThomasBecker

Из этой таблицы можно выделить форматы WavPack, MPEG-4 Audio Lossless Coding (ALS), TrueAudio (TTA), WindowsMediaAudio, AppleLossless, FLAC, Monkey'sAudio, Monkey'sAudio, т.к. эти форматы более эффективны, далее в следующей главе рассмотрим их преимущества.

Формат Monkey's Audio не поддерживает настроек кодирования с потерями. Не рекомендую использование с уровнем кодирования "Insane", а так же кодирование версией 4.0.6 для различных девайсов отличных от компьютера. Используемые расширения файлов — .ape, .apl. У FLACa нет настроек для работы с потерями, это гарантия того, что материал не будет испорчен на стадии сжатия. Кодирование лучше делать с настройками максимального сжатия: -8. Используемые расширения файлов — .flac, .fla

Формат WAV 8-, 16-, 24- и 32-битные аудиофайлы в .WAV формате. Он также поддерживает многоканальные потоки и высокие частоты дискретизации (sampling rate).

2.4 Сравнение звуковых форматов со сжатием с потерями

При редактировании, во избежание постепенного ухудшения качества при каждом сохранении, следует использовать форматы сжатия с потерями. Это относится и к записям, которые могут быть объединены с видеозаписями, так как в одном файле не всегда можно объединить видео и звук в произвольных форматах.

Рассмотрим по таблице 6 сравнение звуковых форматов со сжатием с потерями и выделим более эффективные форматы.

Таблица 2.6.

Сравнение звуковых форматов со сжатием с потерями

Название формата	Расширение файла	Алгоритм	Частота дискр., кГц	Число каналов	Битрейт, кбит/с	Задержка-ка, мс	Пост. битрейт	Перем. битрейт	Обл. прим.	Выпуск
АС-3	.ac3, .vob и др.	МДКП	32; 44.1; 48	до 6 (АС-3), до 13.1 Е-АС-3	32 — 640 (АС-3), до 6144 Е-АС-3	40.6	Да	Нет	Звуковое сопровождение к фильму, DVD, Blu-rayDisc, HD DVD	1992, Dolby Laboratories
MP3 (MPEG-1, 2, 2.5 Audio Layer III)	.mp3	МДКП, Субполосное кодирование	8; 11.025; 12; 16; 22.05; 24; 32; 44.1; 48	до 2	8 — 320	>100	Да	Да	Хранение звуковых данных на ПК, CD-, DVD-носителях, передача через Интернет, цифровые проигрыватели	1993, MPEG ISO/IEC 11172-3
ААС	.aac, .m4a, и др.	МДКП, Субполосное кодирование	8 — 192	до 48 каналов	8 — 529 (стерео)	20–405	Да	Да	iTunes, DVB-H, iPhone, iPod, iPad, Nintendo, PlayStation 3, PlayStation Portable, Sony Walkman, мобильные телефоны SonyEricsson, Nokia и с ОС Android, аудиокниги, подкасты,	1997, ИСО/МЭК Комитет по звуку MPEG
Windows	.wma	МДКП	8; 11.025;	до 8 и	4 - 768	>100	Да	Да	Хранение звуковых данных	1999,

Media Audio			16; 22.05; 32; 44.1; 48; 88,2; 96	выше					на ПК, CD-, DVD-носителях, передача через Интернет, Интернет-радио, цифровые проигрыватели	Microsoft
Vorbis (Ogg)	.ogg	МДКП	8 — 192	до 255	до 1000	>100	Да/ABR	Да	Хранение звуковых данных на ПК, CD-, DVD-носителях, передача через Интернет, цифровые проигрыватели	2000, Xiph.Org Foundation

Из таблицы 6 следует, что форматы Vorbis(Ogg), AAC, MP3, более эффективнее и используемые в наше время, в следующей главе рассмотрим их преимущества.

OggVorbis — это полностью открытый и свободный от патентов звуковой формат, позволяющий хранить и передавать звуковую информацию с высоким качеством звука (частота дискретизации 44.1-48.0 КГц, 16+ бит, полифония (многоканальный звук)) и битовой частотой, варьирующейся от 16 до 512 Кбит/с на канал. При этом количество обрабатываемых каналов может достигать 255. Это позволяет поставить Vorbis в один ряд с MPEG-4 audio (AAC и TwinVQ), и заметно выше таких кодеков сжатия, как WMA, AAC и MPEG-1 audiolayer 3 (MP3).

Из-за популярности iPod (портативный проигрыватель музыкальных файлов) и iTunes (программный проигрыватель, также работающий с iPod) сравнительно большую популярность получил формат AAC, продвигаемый автором обеих разработок, компанией Apple, как стандарт распространения музыки. Как и MP3, этот формат несвободен.

Вывод

При проведении сравнительного анализа было выявлено, что можно выделить аудиокодеки которые имеют больше возможности чем другие форматы и более эффективнее. Это такие аудиокодеки, как Vorbis (Ogg), AAC, MP3 сжатие с потерями и WavPack, MPEG-4 Audio Lossless Coding (ALS), TrueAudio (TTA), WindowsMediaAudio, AppleLossless, FLAC, Monkey'sAudio, Monkey'sAudio сжатие без потерь. Например MP3 по сравнению AAC имеют критическое значение битрейта, выше которого качество звука улучшается очень медленно (по мере дальнейшего увеличения потока), зато ниже - ухудшается очень быстро.

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.

По сравнению с другими форматами ,форматы WavPack, TAK, FLAC, Optim FROG, ALAC, WMA, MPEG-4ALS – есть поддержка потока, открытый исходный код, многоканальный, Highresolution, Поддержка ОС. Следует что у формата FLAC нет настроек для работы с потерями, это гарантия того, что материал не будет испорчен на стадии сжатия. Кодирование лучше делать с настройками максимального сжатия: -8. Используемые расширения файлов — .flac, .fla. А вот WavPack позволяет сжимать и восстанавливать 8-, 16-, 24- и 32-битные аудиофайлы. Поддерживает многоканальные потоки (до 255) и высокие частоты дискретизации (до 192 кГц). Эффективность сжатия зависит от исходных данных, но обычно оно лежит в диапазоне между 30 % и 70 % для обычной популярной музыки, выше для классической музыки и других аудиоданных с более широким динамическим диапазоном.

В современных форматах, таких как: MP3, Vorbis, AAC, AC3, WMA

производится второе dct-преобразование, что позволяет им добиться лучшего качества на средних и низких битрейтах, но не позволяет добиться высоких результатов на более высоких. MusePack не производит второго dct-преобразования, что позволяет достичь непревзойденного качества на битрейтах выше 180.

В следующей главе рассмотрим преимущества и рекомендацию по их применению этих аудиокодеков.

3. ПРЕИМУЩЕСТВО АУДИОКОДЕКОВ И РЕКОМЕНДАЦИЯ ПО ИХ ПРИМЕНЕНИЮ

3.1. Преимущество сжатие звуковых форматов с потерями

MPEG-2 AAC обладает высоким качеством звучания и хорошей степенью компрессии аудиокomпозиций. Так, например, аудиокomпозиция в формате AAC с bitrate 96 kbs обеспечивает качество звучания, аналогичное потоку MPEG-1 Layer III bitrate 128 kbs. 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.

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 AM-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

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. АЦП производит заменой или иной характеристики аналогового сигнала. В плане критического битрейта многоканальное кодирование еще более эффективно.

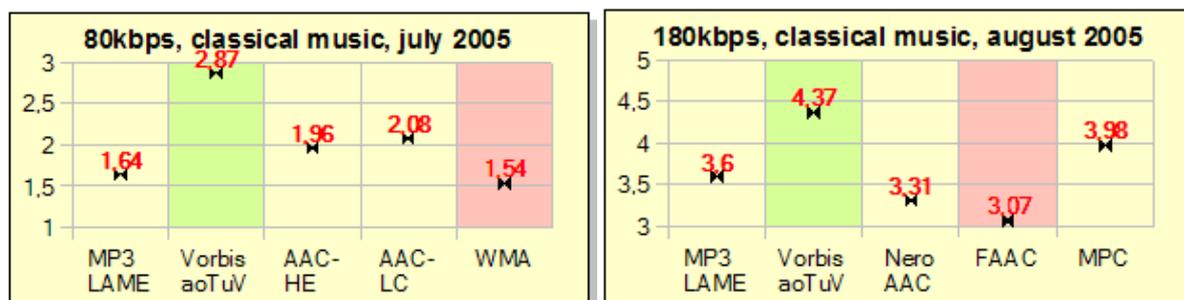
Так же как в MP3, сжатие в AAC основано на психоакустических особенностях восприятия звука. Вместе с тем между ними имеются существенные различия: Вместо гибридного (каскадного) банка фильтров AAC использует модифицированное дискретное косинусное преобразование в частотную область (MDCT) с импульсным 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. [12]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 (TemporalNoiseShaping - 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.

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 96 kHz. 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.

Формата Vorbis. На сегодняшний день основными игроками на арене аудио форматов помимо MP3, есть OggVorbis, WMA и AAC. Почему же именно OggVorbis мы считаем наиболее оптимальным выбором? Для начала предлагаем взглянуть (таблица 3.1) на результаты наиболее свежих тестов на слух, проведенных участниками наиболее авторитетного в этой области ресурса [HydrogenAudio](#). [5]

Таблица 3.1.



Результаты тестов на битрейтах 80 и 180 kbps

Безоговорочная победа на на наиболее востребованных битрейтах 80 и 180 kbps - весьма весомый показатель. Если желаете, можете взглянуть на подробности тестирования: 80kbps и 180kbps (на английском языке). Высокое качество - не единственное преимущество формата. OggVorbis значительно более продуман со стороны технической реализации. Более того, среди всех рассматриваемых - это единственная бесплатная и свободная альтернатива.

Основные возможности Vorbis являются:

- Отсутствие патентных ограничений.
- До 255 каналов.
- «Sampleaccurate» — звуковые данные не будут иметь смещений, дополнительных или потерянных семплов относительно друг друга.
- «Streamable» — поддержка поточного воспроизведения.
- Эффективные алгоритмы переменногобитрейта.
- Частота дискретизации до 192 кГц.
- Разрядность до 32 бит.
- Гибкий Jointstereo.
- Гибкая психоакустическая модель.
- Теги хранятся в Юникоде, а не в национальной кодировке.

- Большая степень компрессии по сравнению с MP3 при равном качестве (размер файла меньше)

Используется в компьютерных играх, для подкастов.

Vorbis идеален для применения в качестве звуковых дорожек фильмов, так как не изменяет их длину при переменном битрейте, что позволяет сохранять синхронность с видеодорожкой и применим для многоканального звука (например 6-канальный звук).

Используется для звуковой дорожки файлов в формате WebM вместе с видеокодеком VP8 (с VP9 в формат добавили поддержку нового свободного кодека Opus).

Реализации. Благодаря открытой модели разработки и распространения существуют несколько вариантов программных кодеров и декодеров Vorbis.

Официальные программы и библиотеки. Официальная реализация от разработчиков стандартов Ogg и Vorbis, фонда Xiph. Org, распространяется под лицензиями типа BSD и GNU LGPL.

Официальной программой-кодировщиком является свободная oggdropXPd, существующая в разных вариантах исполнения: используются библиотеки libVorbis или aoTuVb, и оптимизация под разные процессоры.

AoTuV. Версия от японского разработчика Aoiumi. Благодаря улучшенной психоакустической модели достигается значительно более высокое качество звука, при этом aoTuV не нарушает совместимость с официальным стандартом. Хотя Aoiumi не является официальным разработчиком кодека Vorbis, его заслуги оказались настолько высоки, что в официальные библиотеки версии 1.1.0 от 22 сентября 2004 года был добавлен разработанный им код (beta 2). Последняя версия — Beta 6.03 от 25 апреля 2011 года.Общепринятой практикой является использование последних версий утилит и библиотек от Aoiumi (особенно при работе с битрейтом ниже 64 кбит/с), а не Xiph.Org.

Lancer. Японский разработчик Nyaochi работает над увеличением

скорости кодирования в формат Vorbis. Его версия, называемая Lancer, основывается на актуальных версиях aTuV. Благодаря оптимизациям кода под современные процессоры достигается существенное увеличение скорости кодирования и декодирования (на десятки процентов, в некоторых условиях в несколько раз). Качество при этом страдает весьма незначительно. Его последняя разработка датирована 10 ноября 2006 года.

Формата Speex. Речевой сигнал в кодеке Speex разбивается на неперекрывающиеся отрезки длительностью 20 мс (160 отсчётов при 8 КГц). В отличие от других кодеков, с целью избежать патентных ограничений, Speex не использует алгебраическое кодирование, а только векторное. Возбуждения двух предыдущих подотрезков складываются с переменными весами, в отличие от ряда других кодеков, где используются переменные положения по времени.[13]

Основные характеристики кодека:

- Свободное и открытое программное обеспечение, не имеет патентных ограничений;
- Интеграция широко- и узкополосного канала в одном потоке данных
- Динамическое переключение скорости и переменная скорость (англ. Variable bit-rate, VBR);
- Обнаружитель речи (англ. Voice Activity Detection, VAD, объединено VBR);
- Настройка степени сжатия.

Разработчики противопоставляют свою разработку другим открытым кодекам, например, кодеку Vorbis, утверждая, что именно кодек Speex лучше всего подходит для передачи голоса по сети с ненадёжной доставкой пакетов данных. При этом авторы разработки специально подчёркивают, что кодек подходит для использования в сетях с ненадёжной передачей пакетов, то есть либо пакет пришёл, либо нет. При этом возможность искажения

содержимого пакета должна быть исключена, поэтому Speex не подходит для передачи голоса, например, в радио- и сотовой связи.

Speex поддерживает большое количество приложений, от потоковых приложений (телеконференции) до видеоигр и программ обработки звука. Большинство из них используют фильтр DirectShow. Также для проигрывателей Winamp и XMMS имеются соответствующие дополнения. Кроме того, KSP SoundPlayer начиная от версии 2006.0.0.2 и foobar2000 поддерживают Speex.

MIME тип для Speex — audio/x-speex. В ближайшем будущем он будет заменён на audio/speex.

Последние версии движка Half-Life 1 и основанные на нём моды используют кодек voice_speex.dll для реализации внутри игровых VoIP функций. Speex обеспечивает значительно лучшее качество, чем используемый по умолчанию кодек Miles.

Система The United States Army's Land Warrior, разработанная General Dynamics, также использует Speex.

В Sid Meier's Civilization 4 описания технологий, озвученные Леонардом Нимом, закодированы в формат Speex. VoIP-приложение Teamspeak предлагает Speex в числе 3-х доступных кодеков. Открытая VoIP-программа Mumble использует исключительно Speex. VoIP-сервис flaphone использует Speex кодек. Система управления предприятием TeamWox для внутренних голосовых коммуникаций использует кодек Speex.

Поддержка кодека Speex есть в программном коммутаторе РТУ для сетей NGN. Speex позволяет изменять степень сжатия сигнала, которая в описании кодека именуется «сложностью». Процесс управления поиском осуществляется с помощью целого числа от 1 до 10. В обычных условиях, уровень шума при сложности 1 на 1-2 дБ выше чем при сложности 10, но загрузка процессора при сложности 10 выше приблизительно в 5 раз.

На практике лучшим выбором вероятно будет диапазон чисел от 2 до 4, но для кодирования неречевых сигналов (например DTMF —

двухтональный многочастотный набор телефонного номера), могут оказаться полезными и более высокие значения.

Формата Opus. Основные возможности Opus:

- дискретизация от 8 до 48 кГц; битрейт от 6 до 510 кбит/с;
- поддержка режимов постоянного и переменного битрейта;
- задержка кодирования от 2,5 до 60 мс, настраиваемо.

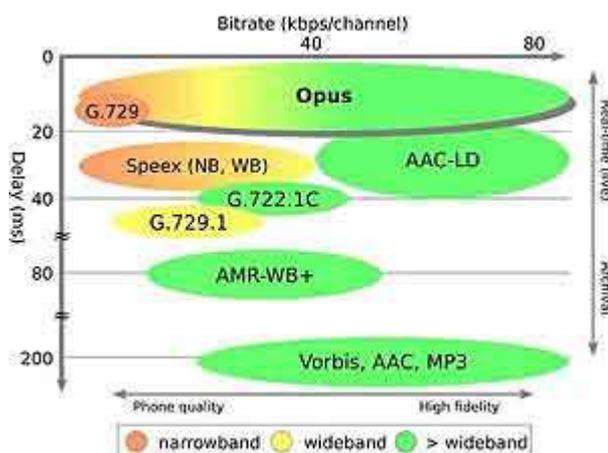


Рисунок 3.1.

По мнению большинства специалистов в области цифрового звука (рис. 3.1.), существующий сейчас в альфа-версии 1.1 открытый кодек Opus избавлен он наиболее серьезных недостатков MP3, но при этом его разработчикам удалось сохранить и приумножить все достоинства кодека-«ветерана».[6]Opus продемонстрировал высокое качество на битрейте 64 кбит/с по сравнению с Apple HE-AAC, Nero HE-AAC, Vorbis и AAC LC.

Значительную поддержку при разработке и тестировании кодека оказала корпорация Google,, так как развивает ряд собственных проектов в области IP-телефонии – от мессенджера GoogleTalk и сервиса GoogleVoice до VoIP-функционала в устройствах линейки GoogleNexus и мобильной ОС Android.

Такой серьезный подход и состав разработчиков позволил создать кодек, структура которого позволяет эффективно справляться со звуковыми артефактами. В частности, для этого используется многоступенчатая архитектура обработки аудиосигнала. Но главное – в движке кодека Opus в

равной мере использованы два независимых стандарта – они предложены, соответственно, Xiph.Org Foundation и SkypeTechnologies. В результате новый кодек оказался гибридным решением, которое сочетает в себе технологии кодеков CELT (ConstrainedEnergyLappedTransform) и SILK, уже знакомый нам по реализации связи в Skype. Поступивший сигнал кодируется SILK или CELT избирательно.

Благодаря своей гибридной структуре Opus получил универсальность, которая позволила этому кодеку на невысоком битрейте обойти по качеству и параметрам задержки своих главных конкурентов – Apple HE-AAC, Nero HE-AAC, Vorbis и AAC LC. Результаты разнообразных тестов показывают, что аудиосигнал, декодированный при помощи Opus, в большинстве случаев наиболее полно восстанавливает исходную картину звука – на разных битрейтах и на разной частоте.

Специалисты также отмечают стабильность работы нового кодека в разных условиях, что оказывается особенно важным при передаче данных в беспроводных сетях. Opus обладает гибким алгоритмом адаптации к изменению пропускной способности канала связи, поэтому качество звука остается неизменным, а сам кодек частично компенсирует потери, обеспечивая трансляцию без сбоев.

Формат MP3 все же слишком популярен, чтобы пользователи легко от него отказались. Вряд ли кто-то возьмется переводить свою домашнюю коллекцию музыки из MP3 в Opus только из-за мало ощутимого улучшения качества звука. Другое дело – разработчики решений для VoIP и производители разнообразных гаджетов. Последние обеспечат аппаратную поддержку Opus, ведь основные показатели у нового кодека столь хорошие, что в скором времени наверняка можно ожидать появления новой технологии в беспроводных наушниках и портативных плеерах.

На официальном сайте Opus сказано буквально следующее: «Кодек может использоваться для любых целей, за исключением Lossless-сохранения (для этого используйте FLAC) и за исключением кодирования с

ультранизким битрейтом (для этого используйте codec2)»).

Формата WindowsMediaAudio. Номинально формат WMA характеризуется хорошей способностью сжатия, что позволяет ему «обходить» формат MP3 и конкурировать по параметрам с форматами OggVorbis и AAC.[14]

Но как было показано независимыми тестами, а также при субъективной оценке, качество форматов всё-таки не является однозначно эквивалентным, а преимущество даже перед MP3 однозначным, как это утверждается компанией Microsoft. Особенно стоит отметить, что ранние версии формата (или его реализации) имели проблемы на низких скоростях потока. Также многие меломаны и владельцы цифровых плееров недолюбливают формат WMA за низкую стойкость к ошибкам. Если при кодировании/передаче файла WMA некоторая часть его повреждается, то воспроизведение файла становится невозможным, как после места повреждения, так и за несколько десятков секунд до него. (Для сравнения: при повреждении файла формата MP3, его всё ещё можно воспроизвести от начала до самого места повреждения, затем пропустить несколько секунд и воспроизвести дальше до конца; иногда же ошибки в несколько байт в файле MP3 бывают на слух малозаметны или не заметны вообще.) Однако данный формат постоянно развивается, так что можно предполагать, что качество будет оптимизироваться.

Достоинства:

- По утверждению разработчика, обеспечивает большую эффективность сжатия аудиоинформации, нежели MP3. Так, 96 kbps WMA по качеству звучания примерно соответствует 128 kbpsMP3. Далеко не все согласны с такой оценкой.
- Наряду с MP3, поддерживается большинством аппаратных аудиопроигрывателей.

Главным преимуществом плеера считаю то, что он предустановлен в системе. При этом вам не придется скачивать и устанавливать сторонние плееры для разового прослушивания композиции либо просмотра фильма.

Формат RealMedia. Основные преимущества:

- С версии 6 начиная, из всех существующих стандартов вещания долго оставался единственным, позволявшим осуществлять произвольную «перемотку» по оси времени лежащих на http сервере файлов, в том числе и при работе через прокси.[15] Это обусловило активное выкладывание пользователями RealMedia файлов на дешёвых и бесплатных хостингах. По мере развития других форматов вещания и, в особенности, FlashPlayer'a, а также по мере развития предоставляемых услуг хостинга, данное преимущество полностью сошло на нет;

- Приемлемые качество изображения и разборчивость речи при сверхнизких битрейтах видеопотока. Маленький размер выходного файла — это прежде всего низкий трафик, требуемый для его трансляции по каналам связи, поэтому в качестве стандарта де-факто RealMedia держался весьма долго. Появление кодеков h.264 и AAC, обеспечивающих при том же битрейте существенно лучшее качество изображения и звука, сделало RealMedia неактуальным;

- Низкие процессорные мощности, необходимые для воспроизведения потока с низким битрейтом;

- 100 % совместимость старых файлов и потоков с новыми версиями плеера;

- Возможность включения в поток слайд-шоу, текстовой информации, интерактивных элементов. Функциональность RealMedia потока ранних версий значительно превышала таковую у конкурентов и, теоретически, позволяла создавать не менее развитые интерфейсно проекты, нежели нынешний (версии 8+) Flash.

Применения

Файлы RealMedia обычно имеют расширение *.RM, *.RAM или *.RMVB. В этом формате можно встретить музыку и видео в сети Интернет или в западных кабельных каналах. В среде смартфонов и сотовых телефонов данный стандарт поддерживают аппараты Nokia и SonyEricsson.

Программ для перекодировки из realmedia в другие форматы существует много, но наиболее известные — это AdobePremiere, Quicktime PRO и Real HELIX Producer,RealProducerPlus.

3.2. Преимущество сжатие звуковых форматов без потерь

Формата Monkey's Audio.Monke's Audio является быстрый и простой способ для сжатия цифровой музыки. В отличие от традиционных методов, таких как MP3, OGG, WMA или постоянно отбрасывать качества для экономии места, Monke's Audio только делает прекрасные , бит-в-бит копии музыки. Это означает, что всегда звучит идеально - точно так же, как в оригинале. [16]

Установив Monkey's Audio на компьютер получаем программу компрессор-декомпрессор, с помощью которой можно сжимать WAV файлы в формат APE, а также, в случае необходимости, разжимать их обратно в WAV. Дополнительно с программой устанавлируются модули, позволяющие воспроизводить файлы формата APE практически любыми проигрывателями.

Особенности

- Эффективное (быстрое и большое сжатие) - Monke's Audio высоко оптимизирован и высокоэффективные
- Идеальный звук - абсолютно без потери качества, означает, что звучит отлично и восстанавливается идеально.

- Легкий - интерфейс Windows, одновременно мощный и простой в использовании

- Бесплатно - Monke's Audio совершенно бесплатно!

- Обнаружение ошибок - Monke's Audio включает избыточным КПП для обеспечения надлежащего декомпрессии данных (ошибки никогда не остаются незамеченными)

- Пометка поддержка - Monke's Audio использует собственные чрезвычайно гибкие APE теги, так что можно легко управлять и каталогизировать коллекцию Monke's Audio

- Внешняя поддержка кодер - можно использовать Monke's Audio, как фронт-енд для всех ваших потребностей кодирования

- Свободно доступны исходный код, просто SDK- другие разработчики могут легко использовать Monke's Audio в своих собственных программах, и нет ограничительные лицензионные соглашения

Особенности TTA аудиокодека

- Компрессия аудио до 30 % без потерь

- Алгоритм кодирования/декодирования в режиме реального времени

- Высокая скорость работы и минимальные системные требования

- Может быть скомпилирован на большом количестве различных платформ

- Бесплатный и открытый исходный код и документация

- Аппаратная поддержка

- Простой и открытый формат данных

- Подключаемые модули для большинства популярных программ-проигрывателей

- Tau Producer — TTA-компрессор для Windows

- DirectShow-фильтры

Используя **TrueAudio (TTA)** кодек, вы сможете хранить до 20-ти звуковых компакт-дисков (CD) на одном единственном DVD-R для воспроизведения, сохраняя всю информацию компакт-диска в популярных тегах ID3 и APE.[17]

Все исходные коды TTA аудиокодека без потерь и дистрибутивы проекта свободно доступны и распространяются под универсальной общедоступной лицензией GPL. Вы можете распространять и использовать программное обеспечение и документацию TTA аудиокодека без потерь свободно на некоммерческой основе. Пожалуйста свяжитесь с нами если вы хотите использовать или распространять программное обеспечение или документацию проекта на других условиях.

Основной задачей проекта является оптимизация алгоритмов компрессии TTA аудиокодека без потерь, в целях более легкой аппаратной поддержки, возможно даже в ущерб максимально достижимым уровням компрессии.

Сравнение TTA аудиокодека с тремя популярными звуковыми компрессорами без потерь (рис 3.2.):

- Monkeys Audio v3.97
- Wavpack v3.97
- Flac v1.1.0

Для тестов был использован персональный компьютер P4 Celeron, 2.6ГГц, SATA, с 256Мб оперативной памяти, под управлением Windows XP Homeedition.

В качестве тестовых выбраны 6 альбомов в различных музыкальных стилях:

1. Enigma - Voyageur

4. Ott - Blumenkraft

2. FranzLiszt - PianoConcertos

5. Schumann - PianoConcerto

3. Dagda - CelticTrance

6. Yello - TheEye

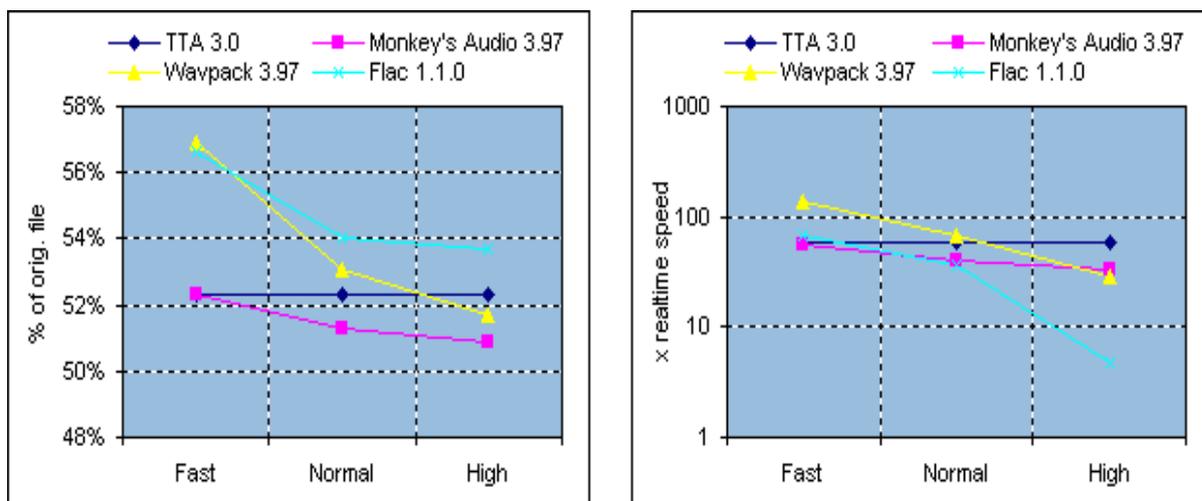


Рис. 3.2. Усредненные результаты сравнения для всех альбомов

TheTrueAudio (TTA) - в настоящее время единственный аудиокодек без потерь способный работать в режиме кодирования в реальном масштабе времени. TTA обладает лучшими характеристиками среди всех беспотерных аудиокодеков, предоставляющих аппаратную поддержку

Формат FLAC. В отличие от аудио-кодеков, обеспечивающих сжатие с потерями (MP3, AAC, WMA, OggVorbis) FLAC, как и любой другой lossless-кодек, не удаляет никакой информации из аудиопотока и подходит как для прослушивания музыки на высококачественной звуковоспроизводящей аппаратуре, так и для архивирования аудиоколлекции.[5]

Сравнение скорости

Машина используется для запуска теста имеет T9600 Intel Core2Duo с 4 Гб оперативной памяти и использует Kubuntu 12.10. Тесты выполняются с электронного диска, а жесткий диск является узким местом в некоторых случаях. Для кодеков, которые имеют только Windows-поддержку, двоичные файлы используются через WINE. Измеряется это так называемый процессорное время используется, не в режиме реального времени используется для WMA Lossless и RealAudioLossless, которые были приуроченная вручную, за исключением.

Аудио корпус в настоящее время состоит исключительно из CD альбомов. Эти 29 компакт-дисков, чьи жанры варьируются от рока до поп-музыки до металла до классики в мир. Следующие графики отображают результаты этих тестов: коэффициент сжатия в зависимости от кодирования и декодирования скорости для наиболее часто используемых в настоящее время кодеков без потерь для всех рекомендуемых режимов сжатия.

На графиках ниже можно увидеть как сжатие и производительность декомпрессии всех протестированных кодеков по их скорости по отношению к «играет скорость» на оси x и их сжатия по сравнению с волны на оси y (рис 3.3.). Подключенные точки для каждого кодека являются его пресеты. Для получения более подробной информации о тестовой среде, кодеков и настроек.

Пожалуйста, имейте в виду, что различия учитывая архивированное сжатие преувеличены: в самом деле, разница между лучшей и худшей компрессии фактором 1000 в то время, но только около 7%, используемого пространства (Рис 3.4.).

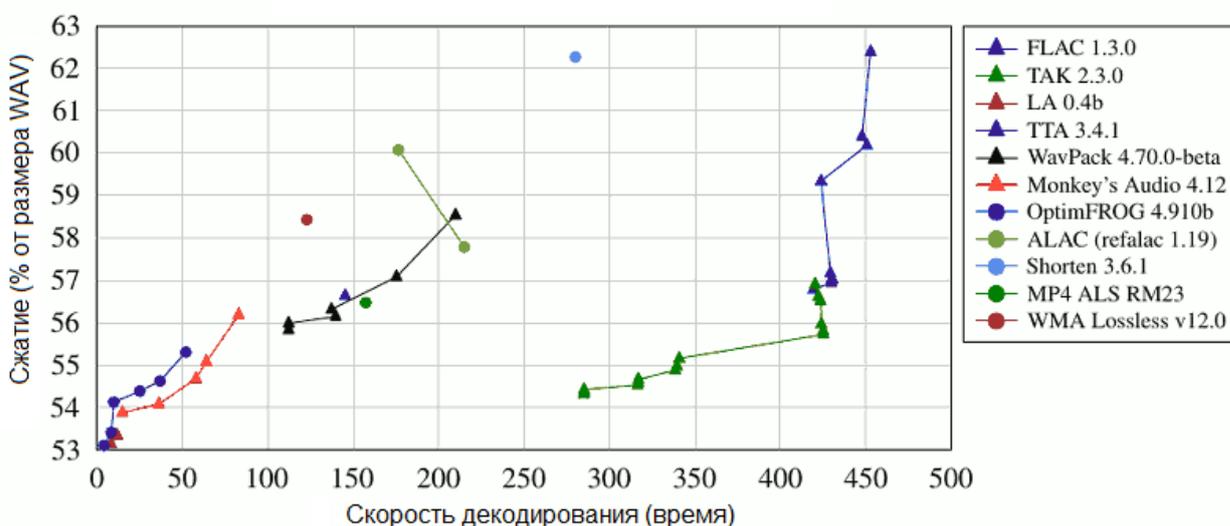


Рис. 3.3. Отношение скорости процессора при декодировании к сжатию для всех композиций.

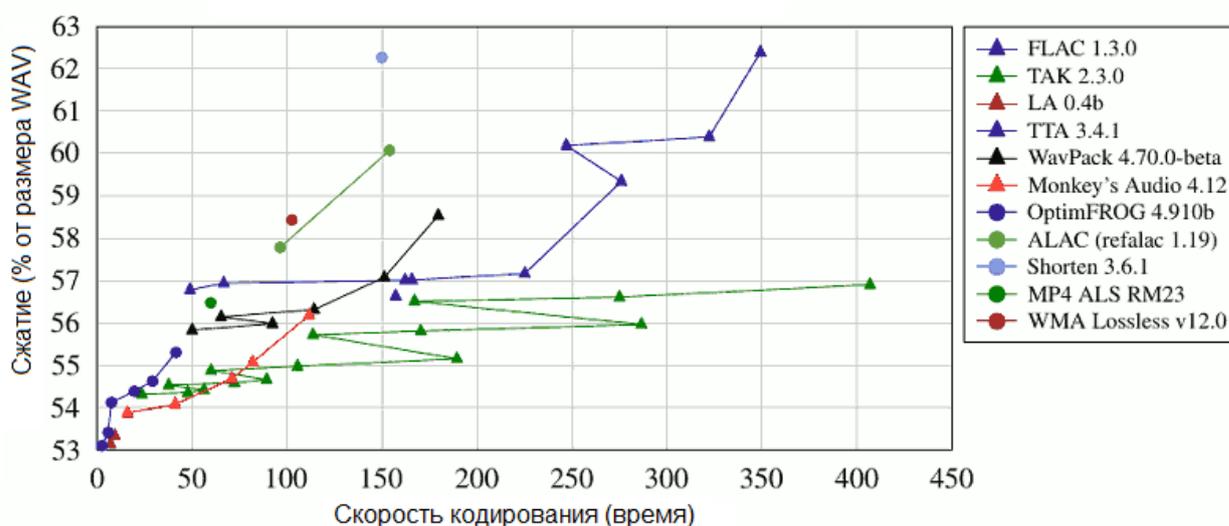


Рис. 3.4. Отношение скорости процессора при кодировании к сжатию для всех композиций.

Формат MPEG-4 ALS. Спецификации для кодированного битового потока MPEG-4 ALS и его схема декодирования были созданы, но есть еще некоторые возможности для улучшения эффективности кодирования по отношению к производительности сжатия и кодирования и декодирования раз, что должно быть сделано без потери совместимости с декодирования стандарт.

В тестах, двух реализаций в ALS кодека MPEG-4, MPEG-4 ALS RM18 и MPEG-4ALS быстро, были по сравнению с тремя из самых популярных программ для сжатия аудио без потерь FLAC: (версия 1.1.2), Monke's Audio (MAC 4.01) , и OptimFROG (OptimFROG v4.520b1) . Испытания проводились на 2.39-ГГц процессором AMD Opteron 250 с 2 Гб оперативной памяти (память произвольного доступа). Данные для испытаний была приобретена из стандартного набора звуковых последовательностей для MPEG-4 кодирования без потерь, первоначально пожертвованные Panasonic Corporation. Мы использовали в общей сложности 51 файлов стерео формы сигнала с частотой дискретизации 48 кГц и разрешением 16 бит и частотой дискретизации 48, 96 и 192 кГц и разрешением 24 бит. Длительность каждого сигнала составляла 30 секунд. Общий размер файлов был 51 682562430 байт.

Результаты для исследованных кодеров и декодеров приведены в рис. 3.5 и 3.6. Вертикальная ось показывает коэффициент сжатия, определяется как, где меньшие значения означают лучшее сжатие. Горизонтальная ось показывает среднее время обработки для кодирования или декодирования 30-х в файл. Меньшие значения означают более быструю обработку.

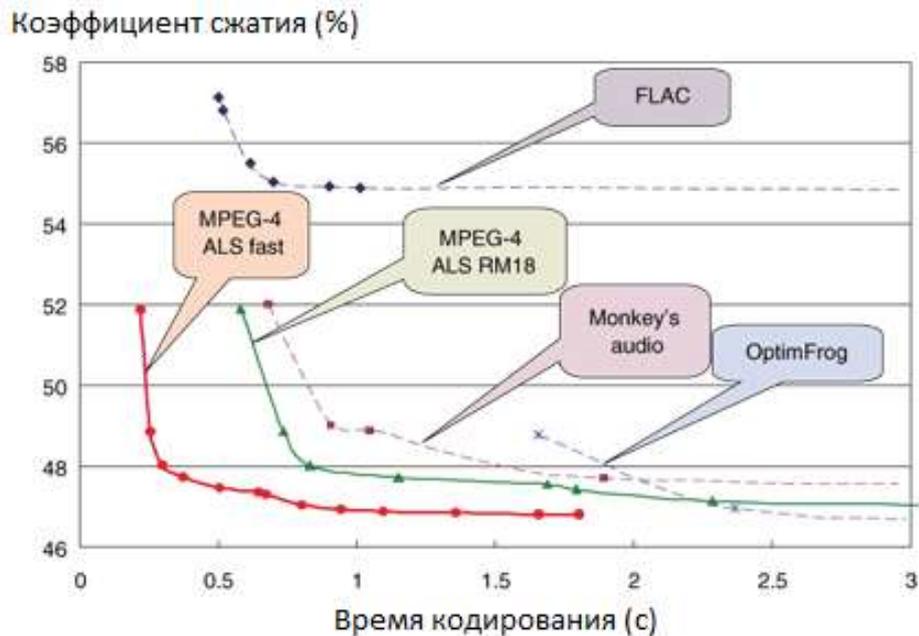


Рис. 3.5. Производительность кодера.

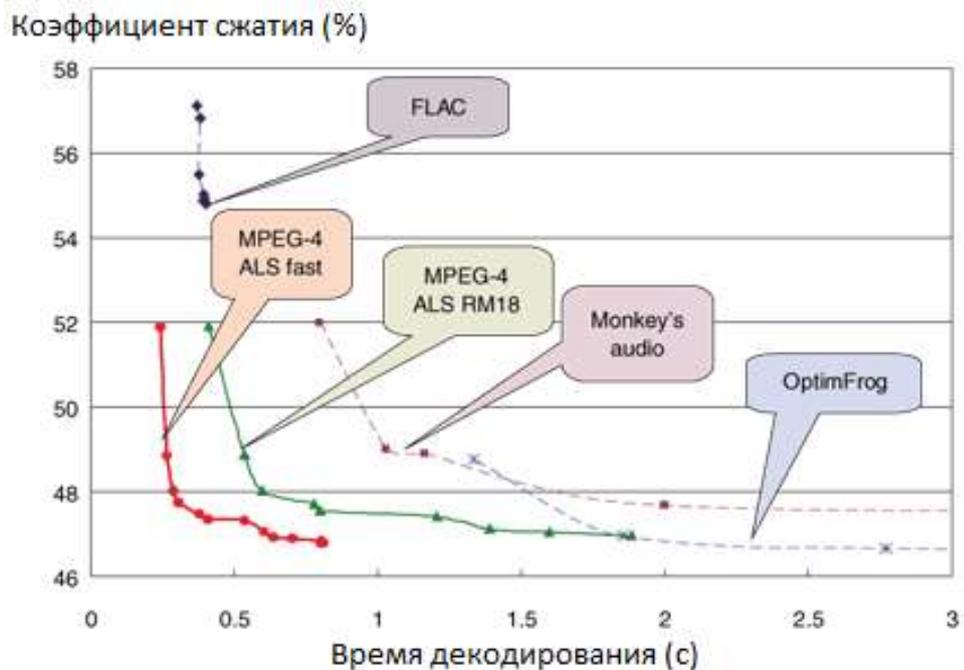


Рис. 3.6. Производительность декодера.

Результаты показывают, что MPEG-4 ALS RM18 и MPEG-4 ALS быстро может достичь лучшую производительность, чем другие без потерь

кодеков с точки зрения наилучшего баланса производительности сжатия и скорости кодирования / декодирования, с MPEG-4 ALS быстро предлагать лучшую производительность. Степень сжатия была от 7 до 8% лучше, чем у FLAC, даже когда скорость кодирования / декодирования был в два раза быстрее. Результаты также показывают, что процессор (центральный процессор) нагрузки для кодирования и декодирования в реальном времени с помощью MPEG-4 ALS быстро была приблизительно от 1 до 2%, что на Opteron процессор 250, когда был использован самый быстрый режим работы.

Выводы

В этой главе было рассмотрено эффективные форматы сжатия их преимущества и применения. Один из форматов это MPEG-2 AAC обладает высоким качеством звучания и хорошей степенью компрессии аудиоконпозиций. 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). . 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, DolbyLaboratories, FraunhoferIIS, LucentTechnologies, SonyCorporation и другие

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 96 kHz. 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.

.Со стороны технической реализации OggVorbis значительно более продуман. Безоговорочная победа на на наиболее востребованных битрейтах 80 и 180 kbps - весьма весомый показатель. Более того, среди всех рассматриваемых - это единственная бесплатная и свободная альтернатива. Еще Vorbis идеален для применения в качестве звуковых дорожек фильмов, так как не изменяет их длину при переменном битрейте, что позволяет сохранять синхронность с видеодорожкой и применим для многоканального звука (например 6-канальный звук).

Благодаря своей гибридной структуре Opus получил универсальность, которая позволила этому кодеку на невысоком битрейте обойти по качеству и параметрам задержки своих главных конкурентов – Apple HE-AAC, Nero

HE-AAC, Vorbis и AAC LC. Результаты разнообразных тестов показывают, что аудиосигнал, декодированный при помощи Orus, в большинстве случаев наиболее полно восстанавливает исходную картину звука – на разных битрейтах и на разной частоте.

Специалисты также отмечают стабильность работы нового кодека в разных условиях, что оказывается особенно важным при передаче данных в беспроводных сетях. Orus обладает гибким алгоритмом адаптации к изменению пропускной способности канала связи, поэтому качество звука остается неизменным, а сам кодек частично компенсирует потери, обеспечивая трансляцию без сбоев.

Файлы RealMedia обычно имеют расширение *.RM, *.RAM или *.RMVB. В этом формате можно встретить музыку и видео в сети Интернет или в западных кабельных каналах. В среде смартфонов и сотовых телефонов данный стандарт поддерживают аппараты Nokia и SonyEricsson.

FLAC, как и любой другой lossless-кодек, не удаляет никакой информации из аудиопотока и подходит как для прослушивания музыки на высококачественной звуковоспроизводящей аппаратуре, так и для архивирования аудиокolleкций

Результаты показывают, что MPEG-4 ALS RM18 и MPEG-4 ALS быстро может достичь лучшую производительность, чем другие без потерь кодеков с точки зрения наилучшего баланса производительности сжатия и скорости кодирования / декодирования, с MPEG-4 ALS быстро предлагать лучшую производительность. Степень сжатия была от 7 до 8% лучше, чем у FLAC, даже когда скорость кодирования / декодирования был в два раза быстрее.

ЗАКЛЮЧЕНИЕ

Как было установлено в результате проведенного анализа, на сегодняшний день существует довольно много методов и стандартов сжатия аудио сигналов, к которым можно отнести: MP-3 (MPEG – 1 audiolayer 3), WMA (windows media audio), OggVorbis, Musicam, ATRAC, MPEG -2, AAC, MPEG -4 ISO/IEC 14496-3, AMP3.

Один из перечисленных кодеков Monkey's Audio — популярный, но ресурсоёмкий кодек. Дает действительно высокое сжатие, но имеет проблемы с прокруткой. А также есть еще кодек TAK — этот активно разрабатываемый кодек не перестает радовать. Если брать во внимание все три параметра (сжатие, кодирование, декодирование), TAK выглядит наиболее эффективнее. Высокая скорость работы объясняется активным использованием процессорных оптимизаций (в т.ч. SSSE3). А использование двух ядер дает почти двукратный прирост скорости кодирования! Таким образом, в случае с TAK преимущество от использования современных процессоров наиболее ощутимо. очень высокую скорость кодирования и приемлемую степень сжатия имеет TrueAudio (TTA) При этом скорость декодирования нельзя назвать очень высокой. Степень сжатия чуть выше чем FLAC. FLAC — степень сжатия средняя, но вот скорость декодирования порадовала. Правда, главной причиной лидерства этого кодека среди общественности является открытый исходный код, и, как следствие, широчайшая аппаратная/сотовая поддержка. FLAC, как и любой другой lossless-кодек, не удаляет никакой информации из аудиопотока и подходит как для прослушивания музыки на высококачественной звуковоспроизводящей аппаратуре, так и для архивирования аудиокolleкций

Apple Lossless (ALAC) — один из тех кодеков, которые ничем не примечательны, но продолжают активно пользоваться разработчиками (в данном случае Apple). Низкая степень сжатия и скорость декодирования. Скорость сжатия — средняя. Подойдет разве что пользователям iPod.

При проведении сравнительного анализа было выявлено, что можно выделить аудиокодеки которые имеют больше возможности чем другие форматы и более эффективнее. Это такие аудиокодеки, как Vorbis (Ogg), AAC, MP3 сжатие с потерями и WavPack, MPEG-4 Audio Lossless Coding (ALS), TrueAudio (TTA), WindowsMediaAudio, AppleLossless, FLAC, Monkey'sAudio, сжатие без потерь. Например MP3 по сравнению AAC имеют критическое значение битрейта, выше которого качество звука улучшается очень медленно (по мере дальнейшего увеличения потока), зато ниже - ухудшается очень быстро.

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.

По сравнению с другими форматами, форматы WavPack, TAK, FLAC, Optim FROG, ALAC, WMA, MPEG-4ALS – есть поддержка потока, открытый исходный код, многоканальный, Highresolution, Поддержка ОС. Следует что у формата FLAC нет настроек для работы с потерями, это гарантия того, что материал не будет испорчен на стадии сжатия. Кодирование лучше делать с настройками максимального сжатия: -8. Используемые расширения файлов — .flac, .fla. А вот WavPack позволяет сжимать и восстанавливать 8-, 16-, 24- и 32-битные аудиофайлы. Поддерживает многоканальные потоки (до 255) и высокие частоты дискретизации (до 192 кГц). Эффективность сжатия зависит от исходных данных, но обычно оно лежит в диапазоне между 30 % и 70 % для обычной популярной музыки, выше для классической музыки и других аудиоданных с более широким динамическим диапазоном.

В современных форматах, таких как: MP3, Vorbis, AAC, производится второе dct-преобразование, что позволяет им добиться лучшего качества на

средних и низких битрейтах, но не позволяет добиться высоких результатов на более высоких.

Один из форматов это MPEG-2 AAC обладает высоким качеством звучания и хорошей степенью компрессии аудиоконпозиций. 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 AM-диапазоне (<30 МГц). . 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, DolbyLaboratories, FraunhoferIIS, LucentTechnologies, SonyCorporation и другие

Номинально формат WMA характеризуется хорошей способностью сжатия, что позволяет ему «обходить» формат MP3 и конкурировать по параметрам с форматами OggVorbis и AAC. Главным преимуществом плеера считаю то, что он предустановлен в системе. При этом вам не придется скачивать и устанавливать сторонние плееры для разового прослушивания композиции либо просмотра фильма.

.Со стороны технической реализации OggVorbis значительно более продуман. Безоговорочная победа на на наиболее востребованных битрейтах 80 и 180 kbps - весьма весомый показатель. Более того, среди всех рассматриваемых - это единственная бесплатная и свободная альтернатива. Еще Vorbis идеален для применения в качестве звуковых дорожек фильмов, так как не изменяет их длину при переменном битрейте, что позволяет сохранять синхронность с видеодорожкой и применим для многоканального звука (например 6-канальный звук).

Благодаря своей гибридной структуре Opus получил универсальность, которая позволила этому кодеку на невысоком битрейте обойти по качеству и параметрам задержки своих главных конкурентов – Apple HE-AAC, Nero HE-AAC, Vorbis и AAC LC. Результаты разнообразных тестов показывают,

что аудиосигнал, декодированный при помощи Orus, в большинстве случаев наиболее полно восстанавливает исходную картину звука – на разных битрейтах и на разной частоте.

Специалисты также отмечают стабильность работы нового кодека в разных условиях, что оказывается особенно важным при передаче данных в беспроводных сетях. Orus обладает гибким алгоритмом адаптации к изменению пропускной способности канала связи, поэтому качество звука остается неизменным, а сам кодек частично компенсирует потери, обеспечивая трансляцию без сбоев.

Файлы RealMedia обычно имеют расширение *.RM, *.RAM или *.RMVB. В этом формате можно встретить музыку и видео в сети Интернет или в западных кабельных каналах. В среде смартфонов и сотовых телефонов данный стандарт поддерживают аппараты Nokia и SonyEricsson.

Результаты показывают, что MPEG-4 ALS быстро может достичь лучшую производительность, чем другие без потерь кодеков с точки зрения наилучшего баланса производительности сжатия и скорости кодирования / декодирования, MPEG-4 ALS может быстро предлагать лучшую производительность. Степень сжатия была от 7 до 8% лучше, чем у FLAC, даже когда скорость кодирования / декодирования был в два раза быстрее.

