

ГОСУДАРСТВЕННОЙ КОММИТЕТ СВЯЗИ ИНФОРМАТИЗАЦИИ И
ТЕЛЕКОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ РЕСПУБЛИКИ
УЗБЕКИСТАН
ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

«К защите»
Заведующий кафедрой «КГ и Д»
доц. Нуралиев Ф.М
«__» _____ 2013г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

**ТЕМА: ПРИНЦИПЫ СЖАТИЯ ЗВУКОВОЙ ИНФОРМАЦИИ НА
ОСНОВЕ АЛГОРИТМОВ MPEG**

Выпускник	_____	<u>Урунбаев Ш.Н.</u>
	(подпись)	(Ф.И.О)
Научный руководитель	_____	<u>Алимов Д.Б.</u>
	(подпись)	(Ф.И.О)
Рецензент	_____	<u>Абдурахманов Б. Э</u>
	(подпись)	(Ф.И.О)
Консультант по БЖД	_____	<u>Қодиров Ф.М.</u>
	(подпись)	(Ф.И.О)

ТАШКЕНТ – 2013 г.

**ГОСУДАРСТВЕННОЙ КОММИТЕТ СВЯЗИ, ИНФОРМАТИЗАЦИИ И
ТЕЛЕКОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ РЕСПУБЛИКИ УЗБЕКИСТАН
ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

Факультет Телевизионных технологий , кафедра «Компьютерная графика и дизайн» Направление (специальность) 5525700 – «Технология звукозаписи»

УТВЕРЖДАЮ

Зав кафедрой Нуралиев Ф.М.
« ____ » _____ 2013 г.

ЗАДАНИЕ

на выпускную квалификационную работу

Урунбаев Шарофиддин Низомиддинович

(фамилия, имя, отчество)

1. Тема работы: Принципы сжатия звуковой информации на основе алгоритмов MPEG
2. Утверждена приказом по университету от « 4 » 02.2013 г. Б № 110
3. Срок сдачи законченной работы 27 мая 2013 г.
4. Исходные данные к работе Введение. Различные методы и способы звукозаписи. Теоритические аспекты цифровой обработки звука и форматы цифровых звуковых файлов. БЖД. Заключение. Список используемой литературы.
5. Содержание расчетно-пояснительной записи (перечень подлежащих разработке вопросов
6. Перечень графического материала презентация
7. Дата выдачи задания 5 февраль 2013 г.

Руководитель _____
(подпись)

Задание принял _____
(подпись)

8. Консультанты по отдельным разделам выпускной работы

Раздел	Ф.И.О. руководителя	Подпись, дата	
		Задание выдал	Задание получил
1. Введение	Алимов Д.Б.	05.02.2013	05.02.2013
2. Различные методы и способы звукозаписи	Алимов Д.Б.	10.02.2013	10.02.2013
3. Теоретические аспекты цифровой обработки звука	Алимов Д.Б.	1.03.2013	1.03.2013
4. Форматы цифровых звуковых файлов	Алимов Д.Б.	2.04.2013	2.04.2013
5. Безопасность жизнедеятельности	Қодиров Ф.М.	24.05.2013	24.05.2013
6. Заключение.	Алимов Д.Б.	25.05.2013	25.05.2013

9. График выполнения работы

№	Наименование раздела работы	Срок выполнения	Отметка руководителя о выполнении
1	Введение	05.02.2013-9.02.2013	
2	Различные методы и способы звукозаписи	10.02.2013-28.03.2013	
3	Теоретические аспекты цифровой обработки звука	1.03.2013-1.04.2013	
4	Форматы цифровых звуковых файлов	2.04.2013-3.05.2013	
5	Безопасность жизнедеятельности	24.05.2013-30.05.2013	
6	Заключение	25.05.2013-30.05.2013	

Выпускник _____
(подпись)

« ____ » _____ 2013г.

Руководитель _____
(подпись)

« ____ » _____ 2013г.

Данная выпускная квалификационная работа посвящена принципам сжатия звуковой информации на основе алгоритмов MPEG. Также были описаны математические методы кодирования и сжатия звуковой информации. В данной работе были рассмотрены различные форматы звуковых данных. Был сделан сравнительный анализ форматов звуковой информации. Были выявлены преимущества и недостатки формата MPEG Layer III.

Ушбу битирув малакавий иши MPEG алгоритмлари асосида товушли ахборотни сиқиш принципларига бағишланган. Шунингдек товушли ахборотнинг кодлаш ва сиқишнинг математик моделлари таърифланган. Ушбу ишда турли хил товуш ахборотлари форматлари келтирилган. Товушли ахборот форматларининг таққосий таҳлили амалга оширилган. MPEG Layer III форматининг афзалликлари ва камчиликлари аниқланган.

In given exhaust qualification work is dedicated to principle of the compression to sound information on base algorithm MPEG. Also, the mathematical methods of the coding were described and compressions to sound information is given. In given functioning were considered different formats sound data. There is made benchmark analysis format sound information. The advantage and defect of the format MPEG Layer III were revealed.

Введение

Цель работы

Изучить формирования звука в цифровом виде, алгоритмы и форматы цифровой звукозаписи

Актуальность работы

На сегодняшний день проблема избрания единого цифрового формата звукозаписи так и для озвучивания компьютерных игр и приложений так и для массового прислуживания имеет свой приоритет. Исходя из этого изучения алгоритмов и формирования формата MPEG и других форматов является актуальной

Цели и задачи

1. Рассмотреть принципы оцифровки звука
2. Изучить алгоритмы формирования звука формата MPEG и Ogg Vorbis
3. Произвести сравнительный анализ разных звуковых форматов

В наше время активного развития электронных технологий и внедрения их в бытовые изделия широкого потребления, в частности, в мультимедийную технику: цифровые плееры, фотоаппараты, камеры, остро встает вопрос об удобстве хранения цифровой информации, передачи по различным интерфейсам и протоколам и, естественно, о ее сжатии. Существует достаточно много форматов сжатия аудио сигнала. Среди них наиболее известные:

- MP3 (MPEG-1 layer 3);
- Ogg Vorbis;
- WMA (Windows Media);
- RealAudio.

Безусловно, MP3 сейчас является самым распространённым форматом сжатия аудио. Но вовсе не потому, что он – лучший по качеству звучания или компрессии. Просто исторически сложилось так, что он появился на рынке намного раньше других форматов. Еще одна весомая причина – это

инертность производителей разнообразной мультимедийной техники, не желающих вводить новые аудиоформаты. Если судить по проведенным тестам независимых исследователей, то выявляется другой лидер – Ogg Vorbis [1,2].

И MP3 и Vorbis основаны на разбиении сигнала на сегменты и применении к каждому сегменту дискретного преобразования Фурье с использованием психоакустического фильтра и алгоритмов кодирования (Huffman, VQ), но отличаются реализацией. В перспективе развития разработчики кодека Ogg Vorbis планируют привлечь вейвлет-технологии сжатия цифровых аудиоданных.

Естественным образом возникла идея изучить вейвлет-анализ (анализ всплесков), активно внедряющийся в сферы обработки одномерных и двумерных сигналов, научиться использовать его на практике, а также попробовать применить оптимизационный метод в целях решения поставленной задачи.

Для сжатия звука, можно использовать следующие методы или их комбинации.

1) Трешолдинг – отбрасывание близких к нулю коэффициентов в разложении сигнала.

2) Округление коэффициентов. Естественно, за сжатие надо оплатить потерей качества, которая возникает из-за округления. Желательно, чтобы восстановленный сигнал как можно меньше отличался от исходного. Поэтому, одной из главных задач становится оценка значимости того или иного коэффициента в разложении относительно остальных и, соответственно, определение точности его хранения.

3) Подбор преобразования, обеспечивающего наименьший суммарный объем хранимого числа разрядов.

4) Подбор разбиения сигнала, обеспечивающего наименьший суммарный объем хранимого числа разрядов. Есть гипотеза, о том, что можно сэкономить суммарное количество разрядов, подбирая сегменты

разбиения так, чтобы сигнал совпадал по фазе с некоторыми базисными функциями. Следовательно, коэффициенты при других базисных функциях будут близкими к нулю. Для нахождения оптимального разбиения можно использовать подходящий оптимизационный метод, например, динамическое программирование.

ГЛАВА 1. РАЗЛИЧНЫЕ МЕТОДЫ И СПОСОБЫ ЗВУКОЗАПИСИ

1.1 Магнитная запись

Магнитофон



Рис.1.1 Студийный катушечный магнитофон

В 1927 году немецкий инженер Фриц Пфлеймер (Fritz Pfleumer), после ряда экспериментов с различными материалами, сделал напыление

порошком оксида железа на тонкую бумагу с помощью клея. В 1928 году он получил патент за применение магнитного порошка на полоске бумаги или киноплёнке. В этом же 1928-м году он демонстрирует свой прибор магнитной записи с бумажной лентой публике. Бумажная лента хорошо намагничивалась и размагничивалась, её можно было обрезать и склеивать. В 1936-м году Национальный суд Германии признал права по патенту Пфлеймера недействительными, так как покрытие бумажной ленты железным порошком было изложено ещё в патенте Паульсена от 1898-го года.

В 1932 году компания AEG, взяв на вооружение идею Пфлеймера, начала производство прибора для магнитной записи под названием «Магнетофон-K1». Носителем в «Магнетофоне-K1» была плёнка, которую изготавливал немецкий химический концерн BASF. «Магнетофон-K1» был представлен публике в 1935-м году на радиовыставке в Берлине.

В 1939 году BASF представил публике плёнку, покрытую порошком оксида железа. Это было революционно. Параллельно этому инженер Уолтер Вебер работал над улучшением качества воспроизведения Магнитофонов, производимых AEG. Он проводил эксперименты с подмагничиванием плёнки. Опытным путем было доказано, что высокочастотное подмагничивание переменным током намного улучшает качество воспроизведения. Весной 1940-го года Вебер получает патент на технологию высокочастотного подмагничивания переменным током, и уже в 1941-м году AEG выпускает новую модель Магнитофона: Magnetophon K4-HF. Технические характеристики этой модели аппарата магнитной записи превосходили все существовавшие тогда аппараты магнитной записи: благодаря открытой Вебером технологии, уровень сигнал-шум составил 60 Дб, а воспроизводить он уже мог частоты выше 10 кГц.

В 1942-м году AEG начала проводить эксперименты по записи звука в стерео формате.

В 30—40-х гг. XX века получили развитие бобинные магнитофоны, причём имелись уже и портативные модели, с конца 50-х гг. — картриджные, а с конца 60-х гг. — кассетные. Цифровые магнитофоны появились в конце 70-х гг.

Магнитная пленка открыла массу новых возможностей для манипулирования звуком музыкантам, композиторам, продюсерам и инженерам. Магнитная пленка была относительно дешевым и очень надежным звуконосителем. И звучала она лучше, чем любые другие звуконосители того времени. Но самое важное то, что магнитная пленка была так же удобна в работе как и кинопленка.

Кассетный формат

Первый кассетный формат, использовавший ленту, склеенную в бесконечную петлю, был запатентован в 1952 году.

В 1963 году компания Philips представила новый формат кассеты для звукозаписи, вышедший на рынок под названием «компакт-кассета» (англ. *Compact Cassette*). Опасаясь ответного хода конкурентов из Sony, Philips предпочёл отказаться от платы за лицензии на производство кассет, что и привело к массовому распространению нового формата.

В 1965 году фирма Grundig выпустила на рынок очень похожую кассету DC International и магнитофоны под нее. Конкуренции с компакт-кассетой этот формат не выдержал и через несколько лет сошел со сцены.

Принципиальным преимуществом компакт-кассеты перед 4- и 8-дорожечными картриджами была простота механизма магнитофона. Бытовые 8-дорожечные магнитофоны могли только воспроизводить ленты, а для записи требовались профессиональные аппараты. Поэтому, несмотря на теоретически худшее качество звучания (из-за вдвое меньшей скорости протяжки), на практике к середине 1970-х годов компакт-кассеты захватили рынок даже в США.

В 1971 году компания Advent Corporation впервые представила кассету с магнитной лентой на основе оксида хрома(IV). Появление этих кассет кардинально изменило судьбу этого вида носителей аудиоинформации. Качество звука на них было намного выше. Это привело к появлению кассет с записанной на них (в фабричных условиях) музыкой (фонограммой), кроме того, кассеты начали использоваться для самостоятельной записи музыки.

1.2 Оптическая запись звука

Фотографическая запись

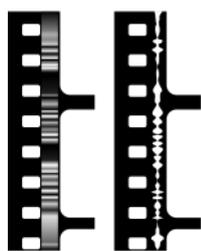


Рис.1.2 Фонограмма переменной плотности (слева) и фонограмма переменной ширины (справа)

В 1919 году американский изобретатель Ли де Фóрест подал свой первый патент на процесс озвучивания фильмов, в котором усовершенствовал разработку финского изобретателя Эрика Тигерштедта и немецкой компании Tri-Ergon, и назвал этот процесс «Фонофильм Фореста». В «Фонофильме» звук записывается непосредственно на плёнку в виде

параллельных линий различных оттенков серого цвета. Позже такой метод стал известен как метод «переменной плотности» в отличие от метода «переменной ширины» в системе «RCA Фотофон», разработанной в RCA. Эти линии кодируют электрические сигналы от микрофона и наносятся фотографическим способом на плёнку, а во время демонстрации фильма переводятся обратно в звуковые волны.

В ноябре 1922 года Форест организовал в Нью-Йорке свою компанию Фонофильм, но ни одна из голливудских студий не выразила никакого интереса к его изобретению. Тогда Форест создал 18 коротких звуковых фильмов, и 23 апреля 1923 года организовал их показ в театре Риволи в Нью-Йорке. Макс и Дэйв Флейшеры использовали процесс «Фонофильм» в своём музыкальном трюковом мультсериале «Вслед за грохочущим шаром», начиная с мая 1924 года. Форест работал вместе с Фриманом Оуэнсом и Теодором Кейсом, совершенствуя систему «Фонофильм». Однако, они потерпели неудачу. Кейс передал их патенты владельцу студии Fox Film Corporation Вильяму Фоксу, который затем усовершенствовал собственный процесс озвучивания «Мувитон». В сентябре 1926 г. компания Фонофильм подала документы на банкротство. Голливуд к тому времени внедрил новый метод озвучивания «Витафон», разработанный компанией Warner Brothers, и выпустил 6 августа 1926 г. звуковой фильм «Дон-Жуан» с Джоном Бэрримором в главной роли.

В 1927 —1928 годах, Голливуд начал использовать для озвучивания фильмов системы «Мувитон» Фокса и «Фотофон» RCA. Между тем, владелец сети кинотеатров Великобритании Шлезингер приобрёл права на «Фонофильм», и с сентября 1926 г. по май 1929 г. выпускал короткометражные музыкальные фильмы британских исполнителей.

Лазерная запись

При помощи лазерного луча на вращающийся оптический диск записываются цифровые сигналы. В результате записи на диске образуется спиральная дорожка, состоящая из впадин (питов) и гладких участков. В режиме воспроизведения лазерный луч, сфокусированный на дорожку, перемещается по поверхности вращающегося оптического диска и считывает записанную информацию. При этом впадины считываются как нули, а ровно отражающие свет участки - как единицы. Такой метод записи обеспечивает практически полное отсутствие помех и высокое качество звучания. По сравнению с механической и магнитной звукозаписью оптический диск имеет целый ряд преимуществ - очень высокую плотность записи и полное отсутствие механического контакта между носителем и считывающим устройством в процессе записи и воспроизведения.

Звуковой компакт-диск

В марте 1979 года на пресс-конференции компания Philips продемонстрировала качество звука прототипа системы компакт-диск. Через неделю в Японии заключило соглашение с фирмой Sony по созданию стандарта на аудиодиск. В 1980 году Philips и Sony после некоторых доработок выпускают их Red Book (красную книгу), которая была положена в основу стандарта на звуковой компакт-диск. В апреле 1982 года Philips представил свой первый компакт-диск проигрыватель. В этом же году на заводе в Германии началось массовое производство компакт-дисков. Совершенно новый носитель информации, заменивший грампластинки, стал основой для будущих поколений оптических дисков, применяемых не только для хранения звукозаписей, но и для видео, медиаинформации и в компьютерных системах.

1.3 Магнитооптическая запись

Запись ведется при помощи магнитной головки и лазерного луча на специальный магнитооптический слой диска. Излучение лазера разогревает участок дорожки выше температуры точки Кюри 121°C, после чего электромагнитный импульс изменяет намагниченность, создавая отпечатки, эквивалентные питам на оптических дисках. Считывание осуществляется тем же самым лазером, но на меньшей мощности, недостаточной для разогрева диска: поляризованный лазерный луч проходит сквозь материал диска, отражается от подложки, проходит сквозь оптическую систему и попадает на датчик. При этом в зависимости от намагниченности изменяется плоскость поляризации луча лазера (эффект Керра, открыт в 1875 году) что и определяется датчиком.

Минидиск

Минидиск был разработан и впервые представлен компанией Sony 12 января 1992 года. Позиционировался как замена компакт-кассетам, к тому времени уже полностью изжившим себя.

Hi-MD

В январе 2004 года, Sony анонсировала формат медианосителя Hi-MD, как дальнейшее развитие формата MiniDisc. Новый диск вмещал уже один Гигабайт данных и мог использоваться не только для звукозаписи, но и для хранения документов, видео и фотографии. Появилась возможность выбирать из трех режимов записи: высококачественного (PCM mode), позволяющего записывать 94 минуты (1 час 34 мин) звуковых данных CD-качества, 7 часов в стандартном режиме записи (Hi-SP) со сжатием ATRAC, и низкокачественного режима (Hi-LP) с записью на 34 часа, помещенной на одном диске.

1.4 Заключение по 1-й главе

История звукозаписи очень древняя и её можно изучить как отдельный предмет. Веками звук, в частности музыка была основной развлекательной культурным видом произведений. Запись и хранение музыкальных (звуковых) произведений все усовершенствовалось. Создавались все новые виды звуконосителей. Развитием науки в частности техники воспроизводились новые методы звукозаписи, совершенствовались и минимизировались микрофоны, кассеты, лазерные диски. Также не мало важную роль в этом внес создание персонального компьютера и различных звуковых карт. Не только записывался звук с концерта или оркестра, стало возможным создавать звуки музыкальных инструментов посредством компьютерных программ.

В данной главе были рассмотрена иерархия истории создания звукозаписи, в том числе такие составляющие данной иерархии как магнитная, оптическая, магнито-оптическая, лазерная запись звука. Были приведены история создания, а также особенности и техническое построение данных видов звукозаписи. С уверенностью можно сказать, что записи методов звукозаписи рассмотренные в первой главе можно оцифровать, редактировать и переформировать в различные цифровые звуковые форматы.

ГЛАВА 2. ТЕОРИТИЧЕСКИЕ АСПЕКТЫ ЦИФРОВОЙ ОБРАБОТКИ ЗВУКА

2.1 Цифровое преобразование звука

Цифровóй звук — кодирование аналогового звукового сигнала в виде битовой последовательности. Простейшая форма кодирования аналогового звукового сигнала состоит в представлении последовательности уровней электрических звуковых колебаний в определенные промежутки времени с применением [импульсно-кодовой модуляции](#). Также издавна известна [сигма-дельта-модуляция](#). Современные системы кодирования в цифровой звук используют более сложные подходы, некоторые из которых, но не все, основаны на изначальном незначительном искажении, обычно невоспринимаемом человеческим ухом. Кроме описания звуковых колебаний в цифровом виде, применяется также создание специальных команд для автоматического воспроизведения на различных электронных музыкальных инструментах, ярчайшим примером такой технологии является [MIDI](#).

Преимущества битового кода используются при передаче кодированного сигнала на расстояние, [криптовании](#) сигнала, цифровой подписи сигнала, восстановлении потерь, вызванной помехами при передаче, а также в прочих приложениях.

Цифровая звукозапись — технология преобразования аналогового звука в цифровой с целью сохранения его на физическом носителе для возможности последующего воспроизведения записанного сигнала.

Принцип цифровой звукозаписи методом периодической дискретизации и квантования сигнала

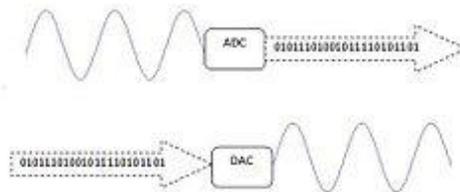


Рис.2.1 Преобразование аналогового сигнала в цифровой в АЦП и обратное восстановление его в ЦАП

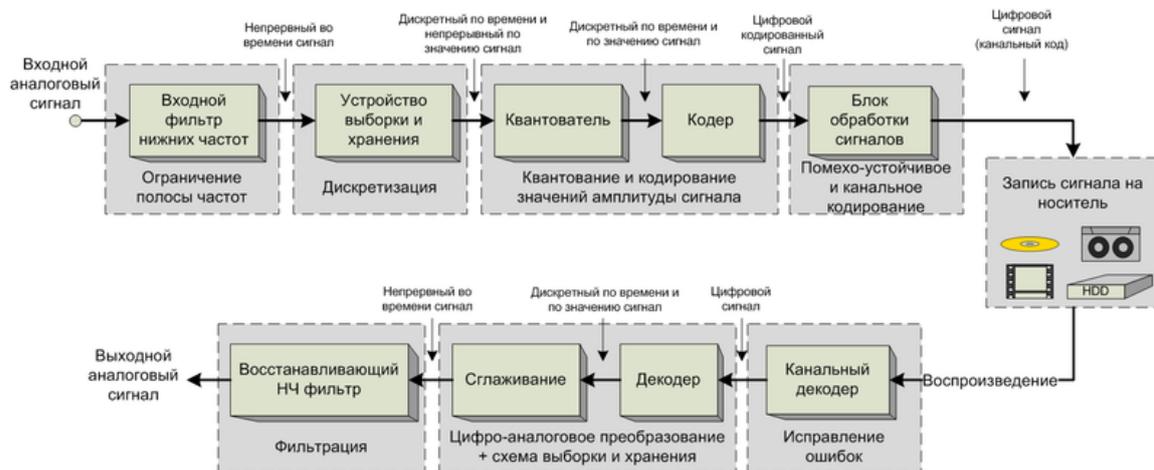


Рис.2.2 Структурная схема цифровой звукозаписи и воспроизведения

Принцип цифрового представления колебаний звукозаписи достаточно прост:

- вначале нужно преобразовать аналоговый сигнал в цифровой, это осуществляет устройство — аналого-цифровой преобразователь (АЦП)
- произвести сохранение полученных цифровых данных на носитель: магнитную ленту (DAT), жёсткий диск, оптический диск или флеш-память
- для того чтобы прослушать сделанную запись, необходимо воспроизведение сделанной записи с носителя и обратное

преобразование из цифрового сигнала в аналоговый, с помощью [цифро-аналогового преобразователя](#) (ЦАП).

Принцип действия АЦП тоже достаточно прост: аналоговый сигнал, полученный от микрофонов и электро-музыкальных инструментов, преобразовывается в цифровой. Это преобразование включает в себя следующие операции:

1. Ограничение полосы частот производится при помощи [фильтра нижних частот](#) для подавления спектральных компонент, частота которых превышает половину частоты дискретизации.
2. [Дискретизацию](#) во времени, то есть замену непрерывного аналогового сигнала последовательностью его значений в [дискретные](#) моменты времени — отсчетов. Эта задача решается путём использования специальной схемы на входе АЦП — [устройства выборки-хранения](#).
3. [Квантование](#) по уровню представляет собой замену величины отсчета сигнала ближайшим значением из набора фиксированных величин — уровней квантования.
4. [Кодирование](#) или оцифровку, в результате которого значение каждого квантованного отсчета представляется в виде числа, соответствующего порядковому номеру уровня квантования.

Делается это следующим образом: непрерывный аналоговый сигнал «режется» на участки, с частотой дискретизации, получается цифровой дискретный сигнал, который проходит процесс квантования с определенной разрядностью, а затем кодируется, то есть заменяется последовательностью кодовых символов. Для качественной записи звука в [полосе частот](#) 20-20 000 Гц применяется минимальная стандартная частота дискретизации от 44,1 кГц и выше (в настоящее время появились АЦП и ЦАП с частотой дискретизации 192,3 и даже 384,6 кГц). Для получения довольно качественной записи достаточно разрядности 16 бит, однако для расширения динамического

диапазона и повышения качества звукозаписи используется разрядность 24 (реже 32) бита.

Дискретизация сигнала во времени

В процессе формирования АИМ сигнала осуществляется дискретизация непрерывного (аналогового) сигнала во времени в соответствии с известной теоремой дискретизации (теоремой В.А.Котельникова): любой непрерывный сигнал, ограниченный по спектру верхней частотой F_B полностью определяется последовательностью своих дискретных отсчетов, взятых через промежуток времени $T_d=1/2 F_B$, называемый периодом дискретизации. В соответствии с им частота дискретизации, т.е. следования дискретных отсчетов, выбирается из условия $F_d \geq 2F_B$.

Поскольку все реально существующие непрерывные сигналы связи представляют собой случайные процессы с бесконечно широким спектром, причем основная энергия сосредоточена в относительно узкой полосе частот, перед дискретизацией необходимо с помощью фильтра нижних частот ограничить спектр сигнала некоторой частотой F_B . Для телефонных сигналов необходимо использовать ФНЧ с частотой среза $F_B=3,4$ кГц. Частота дискретизации для телефонных сигналов выбрана равной 8 кГц.

Устройство, выполняющие дискретизацию во времени, называют устройством выборки и хранения (УВХ) (Рис. 8.21). УВХ могут выпускаться в интегральном исполнении. Вид сигналов в точках 1, 2 и 3 УВХ показан, соответственно, на Рис. 8.22, Рис. 8.23 и Рис. 8.24.

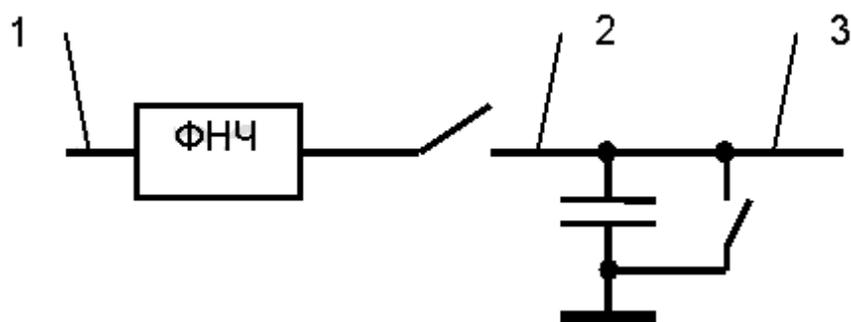


Рис. 2.3. Устройство выборки и хранения



Рис. 2.4. Аналоговый сигнал

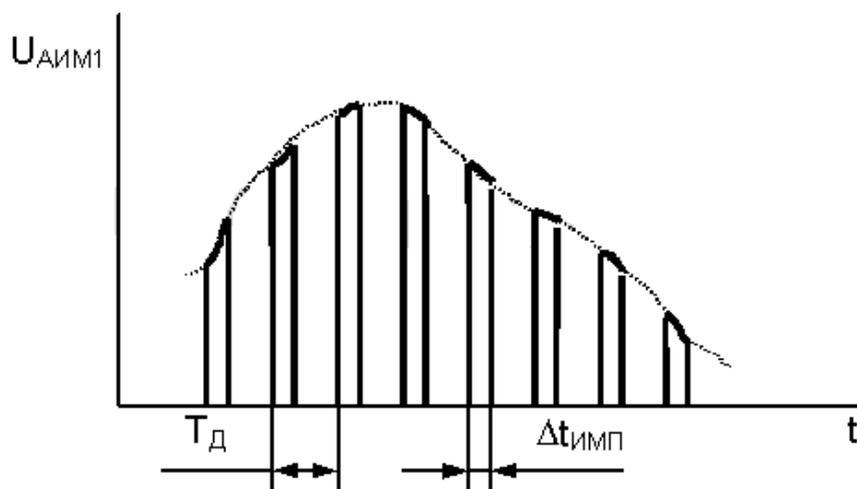


Рис. 2.4. Сигнал АИМ1

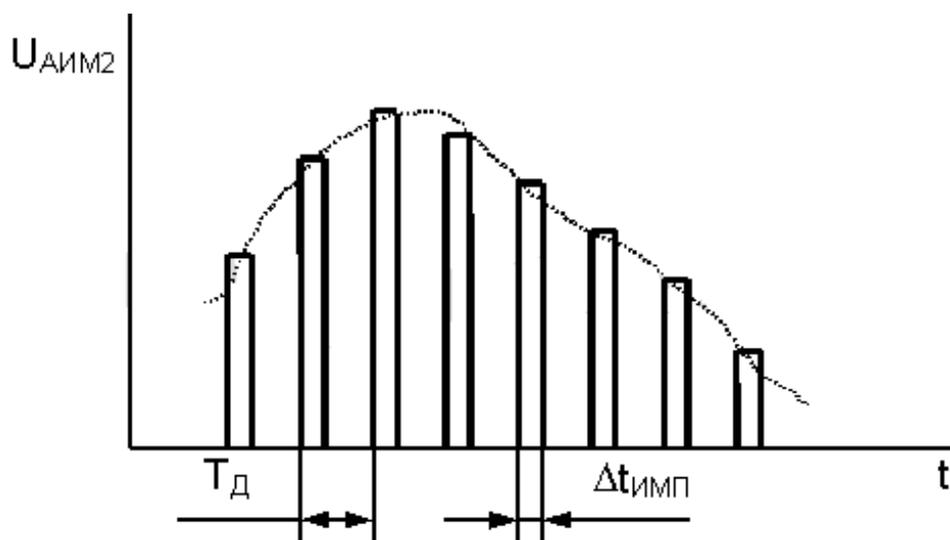


Рис. 2.5. Сигнал АИМ2

Квантование ([англ. quantization](#)) — в [информатике](#) разбиение диапазона значений непрерывной или дискретной [величины](#) на конечное число [интервалов](#). Существует также [векторное](#) квантование — разбиение [пространства](#) возможных значений векторной величины на конечное число областей. Простейшим видом квантования является деление целочисленного значения на [натуральное число](#), называемое коэффициентом квантования.

Не следует путать квантование с [дискретизацией](#) (и, соответственно, шаг квантования с [частотой дискретизации](#)). При дискретизации изменяющаяся во времени величина (сигнал) замеряется с заданной частотой (частотой дискретизации), таким образом, дискретизация разбивает сигнал по временной составляющей (на графике — по горизонтали). Квантование же приводит сигнал к заданным значениям, то есть, разбивает по уровню сигнала (на графике — по вертикали). Сигнал, к которому применены дискретизация и квантование, называется [цифровым](#).

Квантование часто используется при обработке [сигналов](#), в том числе при [сжатии звука](#) и изображений.

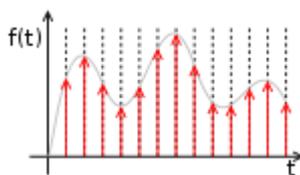


Рис.2.6. Неквантованный сигнал с дискретным временем

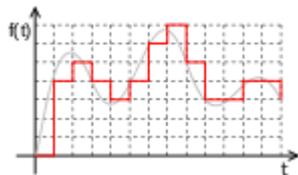


Рис.2.7. Квантованный сигнал

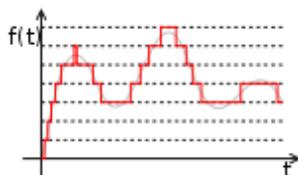


Рис.2.8. Цифровой сигнал

При [оцифровке](#) сигнала уровень квантования называют также **глубиной дискретизации** или [битностью](#). Глубина дискретизации измеряется в [битах](#) и обозначает количество бит, выражающих амплитуду сигнала. Чем больше глубина дискретизации, тем точнее цифровой сигнал соответствует аналоговому. В случае однородного квантования глубину дискретизации называют также [динамическим диапазоном](#) и измеряют в [децибелах](#) (1 бит \approx 6 дБ).

2.2 Сжатие и кодирование звука

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

Методы сокращения статистической избыточности аудиоданных также называют сжатием без потерь, а, соответственно, методы сокращения психоакустической избыточности - сжатием с потерями.

Сжатие без потерь

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. Для передачи сигнала необходимо получить более компактное его представление, что возможно осуществить с помощью [ортогонального преобразования](#). Важными условиями применения такого метода преобразования являются:

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

Этим требованиям отвечает [модифицированное дискретно-косинусное преобразование](#) (МДКП).

Уменьшить скорость цифрового потока позволяют методы кодирования, учитывающие статистику звуковых сигналов, например, вероятности появления уровней разной величины. Одним из таких методов является [код Хаффмана](#), где наиболее вероятным значениям сигнала приписываются более короткие кодовые слова, а значения отсчетов, вероятность появления которых мала, кодируются кодовыми словами большей длины. Именно в силу этих двух причин в наиболее эффективных алгоритмах компрессии [цифровых аудиоданных](#) кодированию подвергаются не сами отсчеты звукового сигнала, а коэффициенты МДКП.

Подобные методы применяются при [архивации](#) файлов.

Структура кодера сжатия аудиоданных с потерями

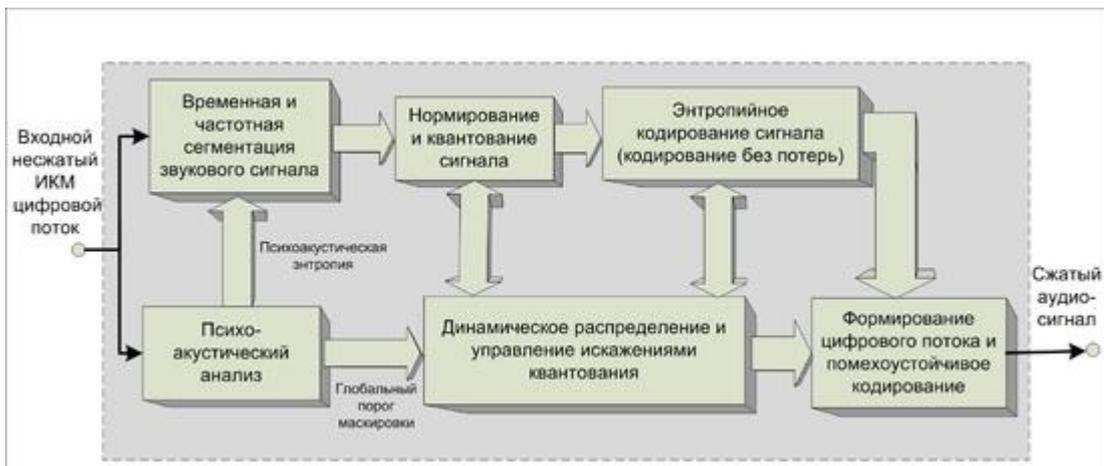


Рис.2.9.Обобщенная структура кодера звукового сигнала с компрессией цифровых аудиоданных

- Исходный цифровой звуковой сигнал разделяется на частотные поддиапазоны и сегментируется по времени в блоке временной и частотной сегментации.
- 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.
- 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 превышал порога их слышимости, рассчитанного психоакустической моделью.
- Также могут использоваться:
 - матрицирование стерео - сложение и вычитание левого и правого канала для устранения повторяющейся информации
 - специальные процедуры итерационных циклов, позволяющие управлять величиной энергии искажений квантования в поддиапазонах при недостаточном числе доступных для кодирования бит
 - процедуры [линейного](#) и обратного адаптивного предсказаний
 - техника сглаживания переходных шумов во временной области (Temporal Noise Shaping - TNS), позволяющая управлять микроструктурой искажений квантования внутри каждого поддиапазона кодирования

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.

Кодирование и обработка звуковой информации

Звуковая информация. Звук представляет собой распространяющуюся в воздухе, воде или другой среде волну с непрерывно меняющейся **интенсивностью и частотой**.

Человек воспринимает звуковые волны (колебания воздуха) с помощью слуха в форме звука различных **громкости и тона**. Чем больше интенсивность звуковой волны, тем громче звук, чем больше частота волны, тем выше тон звука (рис. 1.1).

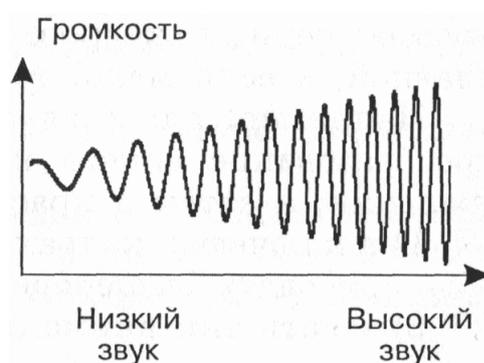


Рис. 2.10. Зависимость громкости и высоты тона звука от интенсивности и частоты звуковой волны

Человеческое ухо воспринимает звук с частотой от 20 колебаний в секунду (низкий звук) до 20 000 колебаний в секунду (высокий звук).

Человек может воспринимать звук в огромном диапазоне интенсивностей, в котором максимальная интенсивность больше минимальной в 10^{14} раз (в сто тысяч миллиардов раз). Для измерения громкости звука применяется специальная единица "децибел" (дбл) (табл. 5.1). Уменьшение или увеличение громкости звука на 10 дбл соответствует уменьшению или увеличению интенсивности звука в 10 раз.

Громкость звука

Таблица 2.1.

Звук	Громкость	В
------	-----------	---

	децибелах
Нижний предел чувствительности человеческого уха	0
Шорох листьев	10
Разговор	60
Гудок автомобиля	90
Реактивный двигатель	120
Болевой порог	140

Временная дискретизация звука. Для того чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть преобразован в цифровую дискретную форму с помощью временной дискретизации. Непрерывная звуковая волна разбивается на отдельные маленькие временные участки, для каждого такого участка устанавливается определенная величина интенсивности звука.

Таким образом, непрерывная зависимость громкости звука от времени $A(t)$ заменяется на дискретную последовательность уровней громкости. На графике это выглядит как замена гладкой кривой на последовательность "ступенек" (рис. 1.2).

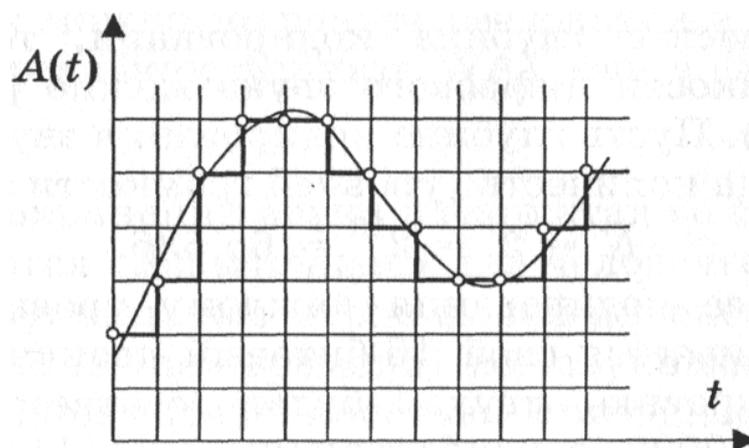


Рис. 2.11. Временная дискретизация звука

Частота дискретизации. Для записи аналогового звука и его преобразования в цифровую форму используется микрофон, подключенный к звуковой плате. Качество полученного цифрового звука зависит от количества измерений уровня громкости звука в единицу времени, т. е. **частоты дискретизации.** Чем большее количество измерений производится за 1 секунду (чем больше частота дискретизации), тем точнее "лесенка" цифрового звукового сигнала повторяет кривую аналогового сигнала.

Частота дискретизации звука - это количество измерений громкости звука за одну секунду.

Частота дискретизации звука может лежать в диапазоне от 8000 до 48 000 измерений громкости звука за одну секунду.

Глубина кодирования звука. Каждой "ступеньке" присваивается определенное значение уровня громкости звука. Уровни громкости звука можно рассматривать как набор возможных состояний N , для кодирования которых необходимо определенное количество информации I , которое называется глубиной кодирования звука.

Глубина кодирования звука - это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука.

Если известна глубина кодирования, то количество уровней громкости цифрового звука можно рассчитать по формуле $N = 2^I$. Пусть глубина кодирования звука составляет 16 битов, тогда количество уровней громкости звука равно:

$$N = 2^I = 2^{16} = 65\,536.$$

В процессе кодирования каждому уровню громкости звука присваивается свой 16-битовый двоичный код, наименьшему уровню звука будет соответствовать код 0000000000000000, а наибольшему - 1111111111111111.

2.3 Теория преобразования звука в стандартах MPEG и Orbis

Алгоритм сжатия сегмента сигнала

Рассмотрим процедуру сжатия одного сегмента сигнала. В зависимости от количества каналов в исходном звуковом файле, на вход подается один (моно) или два (стерео) массива чисел длины N . Эти числа будем называть сэмплами (от англ. sample). На выходе, после обработки, мы получаем N_b байтов закодированного сигнала. Сама же обработка массивов состоит из последовательности шагов, каждый из которых рассмотрим отдельно.

Дискретное вэйвлет-преобразование

Для понимания техники вэйвлет-разложения введем некоторые понятия. С более подробным изложением основных понятий вэйвлет-анализа можно ознакомиться в приложении или в [7]. Рассмотрим две функции ϕ и ψ из $L^2(\mathbf{R})$, удовлетворяющие некоторым свойствам, описанным в приложении, где ϕ – масштабирующая функция, а ψ – вэйвлет. Каждую функцию f из $L^2(\mathbf{R})$ можно приблизить функциями $\{\phi(2^P - k) : k \in \mathbf{Z}\}$ с любой наперед заданной точностью, выбрав достаточно большое значение

$$p: f(x) = \sum$$

$$k \in \mathbf{Z}$$

$$A^P \phi(2^P x - k).$$

Коэффициенты $\{A^P\}$ называются аппроксимирующими коэффициентами.

Будем называть функции $\phi(2^{p-1}x - k)$, соответствующие фиксированному p функциями p -го уровня.

Функции ϕ и ψ обладают таким свойством, что масштабирующие функции p -го уровня выражаются через масштабирующие функции и вэйвлет-функции $p-1$ -го уровня и наоборот:

$$\phi(2^p x - l) = \sum [a$$

$$\phi(2^{p-1} x - k) + b$$

$$\psi(2^{p-1} x - k)], l, p \in \mathbf{Z},$$

$$\phi(2^{p-1} x - l) = \sum$$

$$pk + 2l$$

$$l - 2k$$

$$\phi(2^p x - k), l, p \in \mathbf{Z},$$

$$\psi(2^{p-1} x - l) = \sum q$$

$$k + 2l \psi(2x - k), l, p \in \mathbf{Z},$$

где последовательности $\{ak\}$ и $\{bk\}$ называются последовательностями разложения, а $\{pk\}$ и $\{qk\}$ – последовательности восстановления.

Следовательно, используя (1.1.2) мы можем переписать (1.1.1) как

$$= \sum^{pp} - = \sum$$

$$p - 1$$

$$p - 1 \rightarrow \sum^{p-1}$$

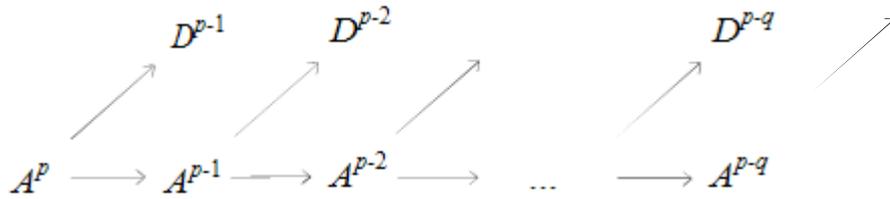
$$p - 1 _$$

p

$$\begin{aligned}
& f_p(x) \\
& k \in \mathbf{Z} \\
& A_k \phi(2^k x) \\
& A_k \\
& k \in \mathbf{Z} \\
& \phi(2^k x) \\
& D_k \\
& k \in \mathbf{Z} \\
& \psi(2^k x). \\
& p-1 \\
& k \\
& = \sum_{l=0}^{p-1} a_l - k A_l, \\
& | l \\
& \{ p-1 p \\
& | D_k \\
& = \sum_{l=0}^{p-1} b_l - k A_l, \\
& \{ A^{p-1} = A^{p-1}, \\
& | k 2 k \\
& | D^{p-1} = D^{p-1}, \\
& 2 k
\end{aligned}$$

где коэффициенты $\{D^{p-1}\}$ называются *детализирующими коэффициентами*. Отметим, что коэффициенты A^{p-1} и коэффициенты D^{p-1} можно получить из коэффициентов A^p , используя (1.1.5) и (1.1.6), причем применение (1.1.6) называется *сгущающей выборкой*, то есть, применив (1.1.5), мы оставляем лишь те коэффициенты, которые имеют четный индекс.

Продолжая этот процесс, получаем следующую схему расчетов коэффициентов:



В итоге мы получили *вэйвлет-разложение*, то есть набор последовательностей коэффициентов $D^{p-1}, D^{p-2}, \dots, D^{p-q}, A^{p-q}$.

Используя обратные соотношения между функциями ϕ и ψ (1.1.3) и (1.1.4) мы можем построить обратный процесс:

Условимся называть аппроксимирующие и детализирующие коэффициенты A^j и D^j коэффициентами j -го уровня. Следуя теории вэйвлетов, для вэйвлет-разложения, сначала необходимо аппроксимировать сигнал с требуемой точностью на верхнем уровне. Пусть этот уровень имеет индекс 0. Предположим, нам дано N сэмплов, и необходимо получить N аппроксимирующих коэффициентов для дальнейшего разложения по алгоритму, описанному выше. Аппроксимировав сигнал на нулевом уровне, мы получим аппроксимирующие сигнал коэффициенты нулевого уровня $\{A^0\}$. Предположим для простоты, что $N = 2^q$, $q \in \mathbf{N}$. В соответствии с алгоритмом разложения, описанным выше, получаем A^{-1} и D^{-1} . Заметим, что в алгоритме используется сгущающая выборка (берутся коэффициенты с четными номерами), поэтому в массивах A^{-1} и D^{-1} содержится по $N/2$ элементов. Аналогично, массивы A^{-2} и D^{-2} будут содержать по $N/4$ элементов. На последнем шаге A^{-q} и D^{-q} содержат по одному элементу. Таким образом, мы получили вэйвлет-разложение $A^{-q}, D^{-q}, D^{-q+1}, \dots, D^{-2}, D^{-1}$, в котором по-прежнему N элементов ($1+1+2+4+8+\dots+2^{q-1} = 2^q = N$).

Заметим, что последовательности разложения $\{a_k\}$ и $\{b_k\}$ могут иметь более двух элементов, и в процессе вычисления коэффициентов нижестоящего уровня соответственно (1.1.5) и (1.1.6) они могут выходить за пределы массива A^{-j} $j = 0, 1, \dots, q-1$ (рис 1.1):

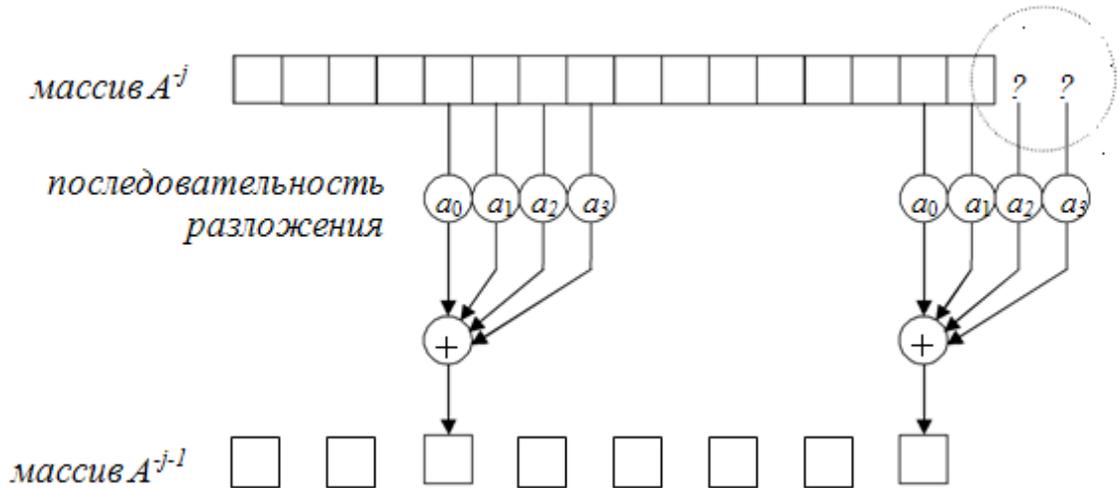


Рис. 2.12. Алгоритм разложения.

Один из способов разрешения ситуации – это замыкание массива A^{-j} в кольцо. Для этого достаточно взять остаток от деления индекса на длину массива. Так, для получения последнего элемента массива A^{-j-1} будут использованы два последних и два первых элемента массива A^{-j} . Этот метод позволяет взаимно-однозначно преобразовывать $j = 0, 1, \dots, j-1$. A^{-j} в A^{-j-1} и D^{-j-1} , и обратно, Мы рассмотрели тот случай, когда $N = 2^q$. Если же N нечетное, то на первом же шаге у нас не будет взаимно-однозначного соответствия между A^0 , где N элементов и (A^{-1}, D^{-1}) где $N-1$ элемент. Решение следующее: Если на каком-то шаге j массив A^{-j} имеет нечетное число элементов, то расширим массив путем добавления в конец одного элемента с произвольным значением, например, равным значению последнего элемента A^{-j} для обеспечения непрерывного расширения. Здесь

возникает избыточность информации, но она минимальна, так как число дополнительных коэффициентов не превышает $\log_2(N)$.

Отметим также, что трудоемкость алгоритмов вэйвлет-разложения и вэйвлет-восстановления равна $O(N)$. Даже по сравнению с быстрым преобразованием Фурье (БПФ), трудоемкость которого равна $O(N \log_2 N)$, описанный алгоритм работает значительно быстрее.

Дискретное преобразование Фурье

Определение 1.1. Дана конечная последовательность $x_0, x_1, x_2, \dots, x_{N-1}$ комплексных чисел. Дискретное преобразование Фурье (ДПФ) заключается в поиске последовательности $X_0, X_1, X_2, \dots, X_{N-1}$, элементы которой вычисляются по формуле:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i \frac{2\pi kn}{N}}$$

Определение 1.2. Дана конечная последовательность $X_0, X_1, X_2, \dots, X_{N-1}$ комплексных чисел. Обратное дискретное преобразование Фурье заключается в поиске последовательности $x_0, x_1, x_2, \dots, x_{N-1}$, элементы которой вычисляются по формуле:

$$x_k = \sum_{n=0}^{N-1} X_n e^{i \frac{2\pi kn}{N}}$$

При помощи дискретного преобразования Фурье можно получить спектр сигнала, то есть коэффициенты при синусах и косинусах в разложении Фурье. Основным свойством преобразования Фурье (см.,

например, [7]) является обратимость, то есть если из последовательности $\{x_k\}$ при прямом преобразовании получается последовательность $\{X_k\}$, то при обратном преобразовании из $\{X_k\}$ получится исходная последовательность $\{x_k\}$.

Отметим, что данное преобразование имеет трудоемкость $O(N^2)$ для набора чисел длины N , однако его можно оптимизировать так, что трудоемкость составит $M^2 2^T + NT$, где $N = M 2^T$ (см. [8]).

Преобразование и оценка коэффициентов

Пусть мы имеем массив коэффициентов $\{c_k\}$, $k = 0, 1, \dots, N-1$, и этот массив *нормализован*, в том смысле, что $|c_k| < 2$, $k = 0, 1, \dots, N-1$. Преобразуем этот массив в два новых массива $\{e_k\}$ и $\{m_k\}$ так, что $-e_k + c_k = m_k \cdot 2$, где $m_k \in [1, 2)$, $e_k \in \mathbf{Z}$, $k = 0, 1, \dots, N-1$. Заметим, что на ЭВМ стандартное представление числа с плавающей точкой в виде *мантиссы* и *экспоненты*, аналогично (1.3.1) с той лишь разницей, что в (1.3.1) экспонента берется со знаком '-'. Теперь нам нужно оценить точность хранения коэффициентов. Отметим во-первых, что коэффициенты с большим значением e_k можно отбросить, как близкие к нулю. Во-вторых, при округлении массива экспонент $\{e_k\}$ на n бит мы можем получить ошибку, равную $c_k 2^n$. Ясно, что при больших c_k , мы получаем неприемлемую ошибку, то есть массив экспонент необходимо хранить точно. А вот при округлении мантиссы m_k на n бит мы получаем ошибку

$$2^n - (e_k + l), \text{ где } l - \text{ это исходная разрядность мантиссы.}$$

Будем оценивать разрядность хранения каждого элемента m_k во-

первых, в зависимости от совокупности значений $\{e_k\}$, и во-вторых, от порядкового номера k в массиве. Поясним это на примере:

$$\text{BitCount}(k) := [K \cdot \exp\{P \cdot e_{\min} - Q \cdot e_k\} \cdot F(k)], \text{ где } e_{\min}$$

$$0 \leq k \leq N-1$$

Здесь функция $\text{BitCount}(k)$ определяет разрядность хранения мантииссы m_k . Параметр K задает верхнюю границу разрядности хранения (например $K = 16$), а функции \exp и F , могут уменьшить разрядность, т.к. $0 \leq P \leq Q$, $0 \leq F(k) \leq 1$, где P и Q – константы.

Отметим некоторые моменты. Чем больше максимальный коэффициент, тем меньше у него e_{\min} , следовательно, тем меньше бит для хранения будет выделено для остальных коэффициентов. Это отражает тот факт, что на фоне доминирующей частоты, остальные частоты менее слышны, нежели в ее отсутствие. Функция F отражает зависимость разрядности коэффициента от его положения в разложении, например, если рассматриваемые коэффициенты – это спектр разложения Фурье, то мы можем применить частотную фильтрацию. Допустим, если мы а priori знаем, что исходный сигнал содержит только низкие частоты, а остальные не важны или являются помехами, то положив мы зададим тем самым *низкочастотный фильтр (low-pass filter)*, то есть фильтр, пропускающий только низкочастотную составляющую сигнала, это позволит существенно сократить объем выходных данных. Отметим, что описанная оценка разрядности называется *психоакустическим фильтром*.

Кодирование с предсказыванием по частичному совпадению

В ходе экспериментов было выявлено, что для обоих используемых преобразований (DWT, DFT) величины $\{e_k\}$ имеют стабильное

распределение, близкое к нормальному. Это натолкнуло на мысль использовать какой-либо вероятностный метод кодирования. Мы провели анализ методов кодирования и выбрали контекстный метод, основанный на предсказывании по частичному совпадению (Prediction by Partial Matching, PPM), который является надстройкой над методом арифметического кодирования. Описание арифметического кодирования можно найти в [9]. Описание PPM можно найти в [10].

Сtereo соединение

Чаще всего, стерео сигнал содержит избыточную информацию, так как часть звуковой информации дублируется. В данной работе используется метод соединения каналов, с целью устранить избыточность информации.

Итак, пусть мы имеем два массива коэффициентов $\{c^l\}$ и $\{c^r\}$, $k = 0, 1, k, \dots, N-1$. Преобразуем каждую пару (c^l, c^r) в другую пару (c_k, α_k) при помощи полярного преобразования координат. Далее, применим рассмотренную в пункте 1.3 схему для преобразования коэффициентов c_k . А для α_k применим следующий прием: мы масштабируем интервал изменения α_k к интервалу $[1,2)$, и теперь α_k в записи (m_k, e_k) имеет значение $(\alpha_k, 0)$, где 0 нам вовсе не обязательно хранить. То есть в итоге мы получили три массива величин – $\{e_k\}$, $\{m_k\}$ и $\{\alpha_k\}$ – это массив экспонент, массив мантисс и массив фаз.

Битовые потоки

ЭВМ позволяет эффективно манипулировать данными разрядности 8, 16, 32 бита, но этот формат невыгоден для хранения коэффициентов

разложения, так как у нас появляются не используемые разряды. Например, если функция оценки разрядности коэффициента m_k выдает значение 10, то записывая этот коэффициент в 16-разрядную ячейку, мы не используем оставшиеся 6 разрядов. Проблема решается путем введения битовых потоков, в которых числа разной разрядности хранятся без зазоров. Это достигается при помощи арифметических сдвигов и логических операций.

Общая схема

Итак, рассмотрев по отдельности все шаги, поясним общую схему алгоритма сжатия сегмента звукового сигнала.

- На вход процедуры сжатия подается два массива $\{x^l\}$ и $\{x^r\}$, $k = 0, k_1, \dots, N-1$ (рассмотрим случай со стерео сигналом).
- Применяем к каждому из массивов выбранное нами преобразование (будь то DWT или DFT) и получаем массивы коэффициентов $\{c^l\}$ и $\{c^r\}$.
- Применяем стерео соединение коэффициентов получаем три массива $\{e_k\}$, $\{m_k\}$ и $\{\alpha_k\}$.
- Оцениваем разрядность хранения величин $\{m_k\}$ и $\{\alpha_k\}$, используя массив экспонент $\{e_k\}$.
- Выполняем RPM-кодирование массива $\{e_k\}$.
- Выполняем битовую упаковку массивов $\{m_k\}$ и $\{\alpha_k\}$ с нужным числом разрядов.

В результате получается блок, состоящий из 3-х подблоков.

PPM-код массива экспонент	битовый массив мантисс	битовый массив фаз
---------------------------------	------------------------------	--------------------------

2.13. Алгоритм восстановления сигнала симметричен алгоритму сжатия, но с использованием обратных преобразований.

Устранение помех на границах сегментов

Так как в качестве базиса, как правило, берутся непрерывные функции, то, несмотря на округление коэффициентов, при восстановлении сегмента мы получаем непрерывную функцию. Но между сегментами может возникнуть скачек, т.к. в конце текущего сегмента восстановленный сигнал является суперпозицией одних компонент, а в начале следующего сегмента – суперпозицией других компонент. Под компонентами здесь понимаются базисные функции, умноженные на соответствующие коэффициенты разложения. И если значения сигнала на стыке двух сегментов совпадали до округления, то после округления они могут не совпадать. Таким образом, непрерывность исходного сигнала может быть утеряна.

Если рассматривать спектр восстановленного сигнала, то разрыв интерпретируется как высокочастотная осцилляция, и на слух он воспринимается, как щелчок. Такие щелчки сильно выделяются на фоне остальных погрешностей и выливаются в шум на протяжении всего восстановленного сигнала.

Эта проблема решается перекрыванием сегментов. При сжатии сегменты берутся с некоторым заступом на следующие, а при

восстановлении применяется следующий прием: на протяжении участка перекрытия, амплитуда текущего восстановленного сегмента непрерывно уменьшается до нулевой, а амплитуда следующего восстановленного сегмента непрерывно увеличивается от нулевой амплитуды до исходного значения. Чтобы описать этот процесс, возьмем неубывающую функцию

$$w(x) \in C[0,1], \text{ такую, что } w(0) = 0, w(1) = 1.$$

Пусть область перекрытия на временной оси – есть интервал $[a, b]$. Будем умножать амплитуду сигнала текущего сегмента в области перекрытия на чего, эти амплитуды складываются. Этот процесс называется *кроссфэйдингом* (crossfading).

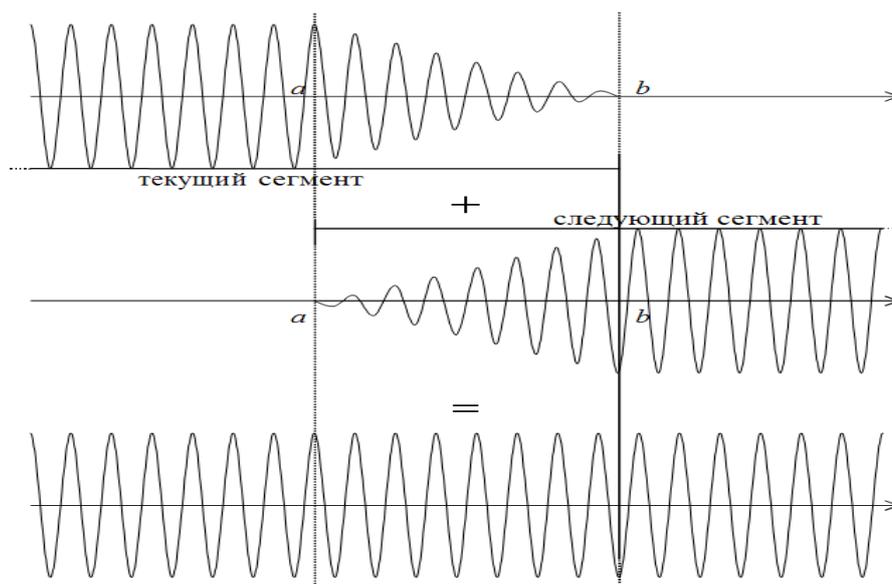


Рис. 2.14. Кроссфэйдинг в области перекрытия сегментов.

На рисунке 2.14 изображен восстановленный сигнал, который был сжат с использованием метода перекрытия сегментов; сигнал представляет собой синусоидальную волну, а в качестве $w(x)$ взята линейная функция. Таким образом, возможный разрыв устраняется, так как сигнал сегмента теперь представляет собой непрерывную функцию на временной оси, а

восстановленный сигнал является композицией сигналов сегментов.

Деление сигнала на сегменты: Статическая реализация

Первый и самый простой способ заключается в выборе фиксированной длины сегмента, то есть на вход процедуру сжатия последовательно подаются сегменты одинаковой длины, а затем сегмент-остаток.

Деление сигнала на сегменты: Адаптивная реализация

Второй способ заключается в выборе длины текущего сегмента на основе характеристик предыдущего. В качестве таких характеристик можно взять коэффициенты разложения предыдущего сегмента. В случае преобладания высоких частот можно уменьшать длину текущего сегмента для лучшей локализации высоких частот. В случае преобладания низких частот мы, наоборот, увеличиваем длину текущего сегмента для лучшей локализации низких частот. Этот метод используется почти во всех доминирующих форматах сжатия звука на сегодняшний день, таких как MP3 и Ogg Vorbis.

Динамическая реализация

Третий способ заключается в подборе такого разбиения, которое минимизирует суммарное количество байтов закодированного сигнала, то есть предоставление программе возможности адаптироваться к структуре сигнала. Это возможно в случае, если преобразование чувствительной к разбиению сигнала на сегменты. Поясним это свойство на примере.

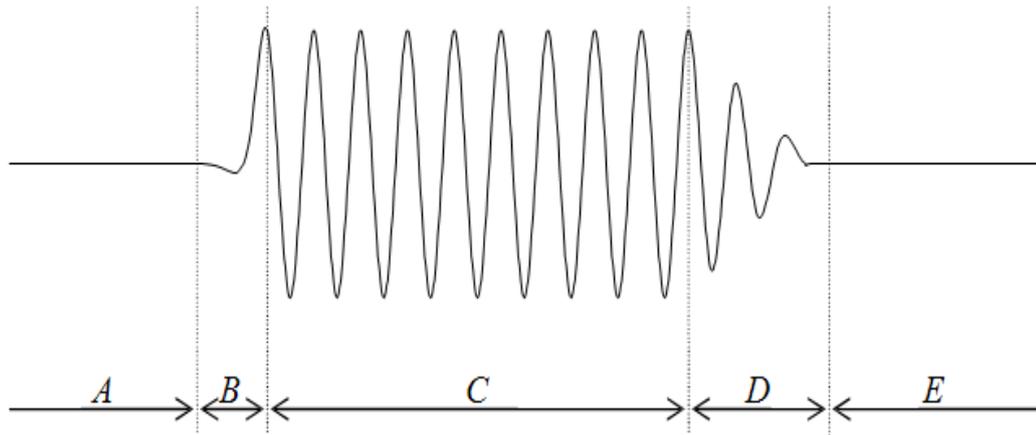


Рис. 2.15. Разбиение на сегменты.

На рисунке 2.15 изображена часть сигнала, разбитая на сегменты. Отметим, что для преобразования Фурье этот вариант разбиения будет самым эффективным среди всех возможных с точки зрения сжатия, так как сегмент *C* будет иметь всего один ненулевой коэффициент разложения, в *A* и *E* будут нули, и лишь сегментах *B* и *D* будут присутствовать разные составляющие спектра. Чтобы определить наиболее эффективное с точки зрения сжатия разбиение, можно использовать динамическое программирование.

Итак, нам необходимо определить разбиение, минимизирующее объем выходных данных – такую последовательность точек, что интервалы между соседними точками принимаются за сегменты звукового сигнала, подлежащие сжатию. Во-первых, отметим, что слишком большие длины сегментов брать бессмысленно, так как в случае преобразования Фурье, с ростом длины сегмента растут трудоемкость и погрешность. Во-вторых, мы должны зафиксировать шаг для динамического программирования, который определяет длину минимального возможного сегмента в разбиении так, что длина каждого сегмента кратна этому шагу. Если фиксировать шаг равный 1, то при сжатии больших объемов аудио информации алгоритм становится очень трудоемким.

Введем некоторые обозначения:

h – длина минимально возможного сегмента.

n – количество шагов в сегменте максимальной возможной длины.

Таким образом, мы получили равномерную сетку на сигнале с шагом h :

$$w_h = \{x_k := kh \mid k = 0, 1, \dots, m\}, \quad \text{где } m = \frac{\lceil N \rceil}{h}$$

Введем понятие пути P^k из x_0 в x_k :

$$P^k = \{p_i \in w_h \mid i = 0, 1, \dots, M\}, \quad \text{причем } \begin{cases} p_0 = x_0, \\ h \leq p_j - p_{j-1} \leq nh, j = 1, \dots, M \\ p_M = x_k, \end{cases}$$

где $M = |P^k| - 1$ – число сегментов в пути P^k . Обратим внимание на то, что в определении пути присутствуют три свойства: длина каждого сегмента (p_{j-1}, p_j) пути P^k кратна шагу h , не равна нулю и не превышает длины максимально возможного сегмента nh .

Пусть вес сегмента (x_i, x_j) – это количество байт после сжатия сегмента сигнала, ограниченного этими точками, обозначим вес сегмента за $d(x_i, x_j)$.

Тогда вес пути P^k равен сумме весов всех его сегментов:

$$d(P^k) = \sum_{i=1}^M d(p_{i-1}, p_i), \quad \text{где } M = |P^k| - 1.$$

Путь P^m из x_0 в конечную точку x_m можно интерпретировать как разбиение исходного сигнала. Отметим, что путь из x_0 в x_m минимального веса и будет искомым разбиением. Итак, можно ввести эквивалентную

задачу на следующем графе:

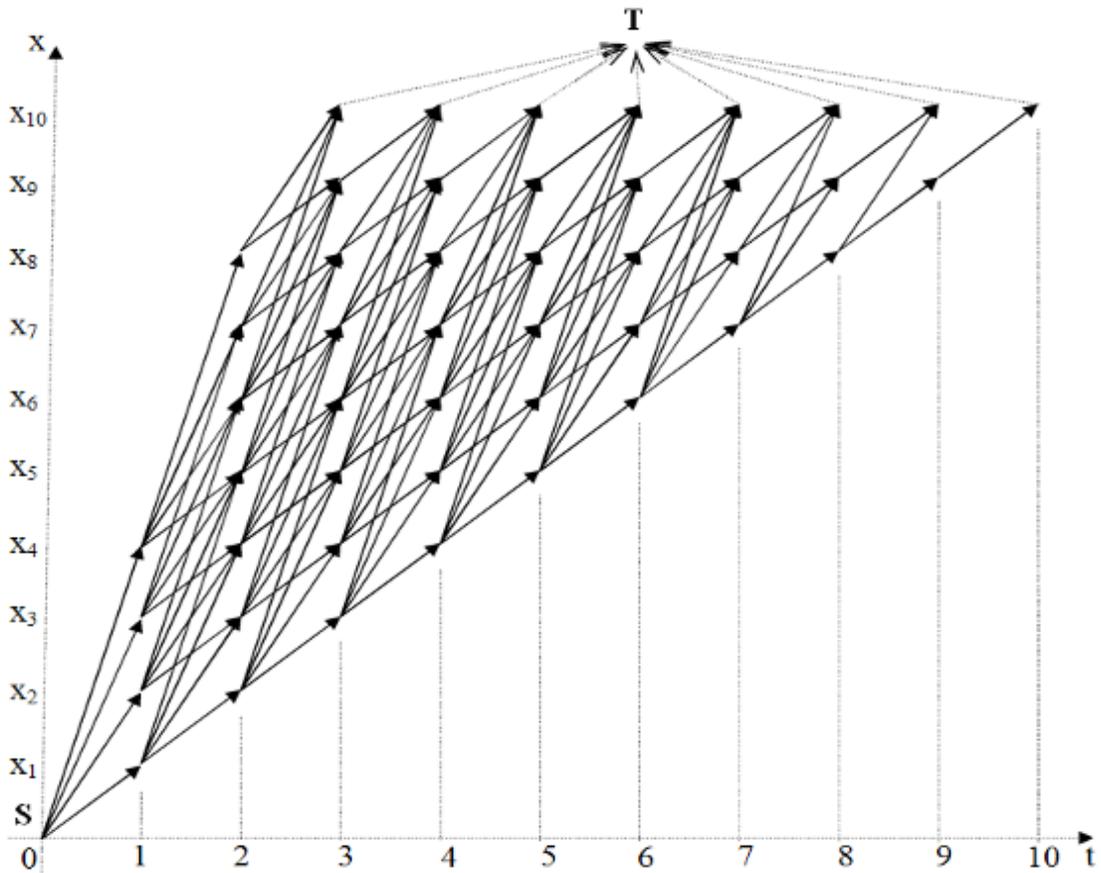


Рис. 2.16. Граф состояний.

На рисунке 2.16 изображен *граф состояний*, соответствующий конкретной задаче сжатия, в которой число шагов $m = 10$, количество шагов в максимально возможном сегменте $n = 4$.

Каждая вершина графа (x_k, t) характеризует множество путей из x_0 в x_k , которые состоят из t сегментов. Вес любого ребра

$$c[(x_{k_1}, t_1), (x_{k_2}, t_2)] = d(x_{k_1}, x_{k_2})$$

в графе определим как вес сегмента

Нам необходимо найти путь минимального веса до любой из вершин (x_{10}, t) . Можно ввести фиктивные ребра, вес которых равен нулю (изображены пунктиром) до фиктивной вершины T , тогда задача сводится к нахождению пути минимального веса на графе (рис. 2.16) из S в T .

Вес ребер не зависит от числа сегментов t , и, по сути, нам не важно, сколько сегментов в пути, а важен его вес, поэтому все состояния, соответствующие фиксированному x_k и разным t можно считать тождественными, таким образом, можно рассматривать проекцию графа (рис. 2.16) на ось x :

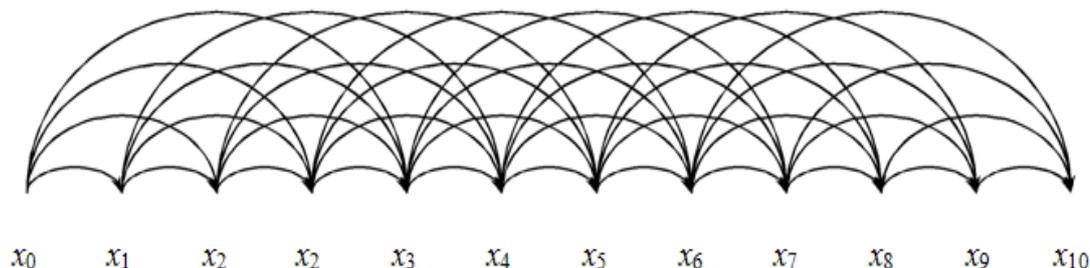


Рис. 2.17. Проекция графа состояний на ось x .

Задача сводится к нахождению пути минимального веса на графе (рис. 2.3) из x_0 в x_m . Будем решать задачу методом динамического программирования. Разобьем задачу на этапы. На k -м этапе необходимо найти путь минимального веса из x_0 в x_k , обозначим его через P^{k*} , при условии, что P^{k-1*} уже найдены. Запишем уравнение Беллмана:

$$d(P^{k*}) = \min_{i=1, \dots, n} \{d(P^{k-i*}) + d(x_{k-i}, x_k)\}.$$

На k -м этапе у нас имеется n управлений, где i -е управление отвечает за выбор пути

$$i = \arg \min_{j=1, \dots, n} \{d(P^{k-j*}) + d(x_{k-j}, x_k)\}$$

Выбор оптимального управления осуществляется очевидным образом,

в соответствии с уравнением Бэллмана, то есть $i = \arg \min \{d(P_{j=1, \dots, nk-j^*}) d(x_{k-j}, x_k)\}$.

Для определения величин $d(x_{k-j}, x_k)$, где $j = 1, 2, \dots, n$, нам необходимо провести сжатие n сегментов $(x_{k-1}, x_k), (x_{k-2}, x_k), \dots, (x_{k-n}, x_k)$.

Заметим, что на каждом шаге алгоритма Дейкстры при добавлении вершины x_k мы также должны сжать n сегментов $(x_k, x_{k+1}), (x_k, x_{k+2}), \dots, (x_k, x_{k+n})$.

Следовательно, так как количество шагов алгоритма Дейкстры и описанного алгоритма равно m , то трудоемкости этих алгоритмов совпадают. Но описанный алгоритм более удобен в реализации, так как в алгоритме Дейкстры вершины могут добавляться непоследовательно.

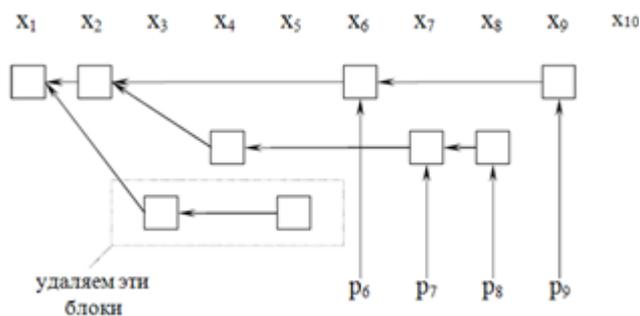
Сделаем несколько замечаний относительно реализации описанного алгоритма динамического программирования:

1. При сжатии сегментов, на каждом шаге формируются n блоков, содержащих сжатые сегменты сигнала, из которых выбирается один, а остальные удаляются.

2. В пути, фактически, не содержится точек разбиения, а содержится несколько таких блоков, связанных указателями, и мы храним только указатели на последние блоки путей.

3. При формировании нового пути, нам не нужно копировать все блоки из $k-j$ -го пути, нам достаточно установить указатель i -го блока, на последний блок $k-j$ -го пути, а указатель k -го пути на j -й блок.

4. На k -м шаге нам уже не нужны пути $P^{0^*}, P^{1^*}, \dots, P^{k-n-1^*}$, и после k -го шага, путь P^{k-n^*} уже не понадобится. Удалим те его блоки, на которые нет ссылок из других путей. Следовательно, на k -м шаге можно хранить только указатели на последние блоки последних n путей.



5. На каждом шаге блоки со ссылками образуют дерево и являются его узлами.

2.4 Заключение по 2-й главе

Оцифрование звука и любой другой аналоговой информации стало доступно после изобретения теоремы Котельникова и применением преобразования Фурье, а также развитием полупроводниковый приборов в частности микросхем для осуществления данной цели.

Во второй главе данной работы были рассмотрены и глубоко изучены основные виде дискретизации, кодирования, сжатия. А также были изучены все теоретические аспекты функционирования форматов MPEG и Ogg Vorbis кодирования и сжатия звука.

Преобразование звука в форматах MPEG и Ogg Vorbis довольно схожи, но имеют различия в самих кодах используемых для кодирования звуковой информации. Процесс создания MPEG или Ogg Vorbis файла, или преобразование звука в данные форматы, требует прохождения различных процессов описанных во второй главе. На сегодняшний день качество и «вес» каждого звукового файла зависит от теоретических основ, математических методов используемых в различных форматах. Но кодирование и сжатие не происходит без потерь. Для обеспечения высокого качество нужен больший размер. При минимизации размера теряется качество. Самые оптимальные критерии размера и качества приведены в формате MPEG Layer III. Теоритическое различие создания разных методов, схем и кодов для кодирования и сжатия обуславливается применением

данного звукового формата в разных областях применения оцифрованных звуковых файлов. Это может быть как озвучивание компьютерных игр и приложений, так и озвучивание интернет сайтов. Также отдельно подразделяют файлы для прослушивания музыки. В настоящее время формат MPEG Layer III популярен не только для прослушивания музыки и различных аудио данных, но и для озвучивания интернет сайтов, веб-приложений и в некоторых случаях для озвучивания компьютерных приложение.

ГЛАВА 3. ФОРМАТЫ ЦИФРОВЫХ ЗВУКОВЫХ ФАЙЛОВ

3.1 Звуковые стандарты MPEG

MPEG-1 Audio Layer I (сокр. MP1) — один из трёх форматов (уровень 1) сжатия звука с потерями, определённых в стандарте MPEG-1[5]. Хотя MPEG-1 Audio Layer I поддерживается большинством современных медиаплееров, этот формат считается устаревшим. Вместо него в основном используются форматы MP2 и MP3.

Для обозначения файлов, в которых содержится только аудиоинформация формата MP1, используется расширение .mp1 или, иногда, .m1a.

Технология сжатия MP1 использует сравнительно простую схему полосного кодирования с 32 поддиапазонами.

Layer I (слой 1) рекомендуется для применения в профессиональной области, в системах записи-перезаписи с высоким студийным качеством с достаточной ёмкостью памяти. Он характеризуется небольшой сложностью и невысокой степенью редукции аудиоданных.

MPEG-1 layer I также использовался в Digital Compact Cassette, как часть аудиокодека PASC (англ. Precision Adaptive Sub-band Coding — прецизионное адаптивное полосное кодирование). В связи с необходимостью передачи постоянного потока отдельных блоков данных (кадров) на ленточный носитель, в PASC использовался дополнительный бит в MPEG заголовке для определения начала кадра. Битрейт всегда 384 кбит/с.[6]

Расширение формата было представлено в MPEG-2 Audio Layer I и определено в стандарте ISO/IEC 13818-3 (MPEG-2 Part 3)[8], первая редакция была опубликована в 1995[4]

дополнительные частоты дискретизации: 16, 22.05 и 24 кГц

дополнительные битрейты: 48, 56, 80, 112, 144 и 176 кбит/с.

Кодирование и декодирование звуковых сигналов MPEG-1 Layer I

Метод кодирования

Входной цифровой звуковой сигнал разделяется на кадры (фреймы), каждый из которых кодируется и декодируется независимо от других кадров. Размер кадра для уровня Layer I составляет 384 отсчёта.

Полоса аудиосигнала с помощью цифровых полосовых фильтров разбивается на 32 поддиапазона. Все поддиапазоны имеют одинаковую ширину, которая зависит от частоты дискретизации входного сигнала. После деления частота дискретизации уменьшается в 32 раза, так что число отсчётов в кадре в каждом поддиапазоне равно 12.

Затем выполняется квантование данных. Предварительно определяются масштабные множители, которые зависят от максимального значения сигнала. При этом масштабный множитель определяется для каждого поддиапазона в кадре, то есть для 12 отсчётов сигнала поддиапазона. Перед квантованием значения сигнала делятся на соответствующие масштабные множители. [7]

Затем в блоке квантования и кодирования выполняется квантование данных. В основе сжатия звуковой информации на уровне Layer I лежит метод, называемый адаптивным распределением битов. Этот метод заключается в выполнении квантования с различным числом двоичных разрядов квантования для разных частотных поддиапазонов. При этом используется равномерное квантование. Полное число битов, выделяемых на все поддиапазоны в данном кадре, зависит от частоты дискретизации входного сигнала и от заданной выходной скорости передачи двоичных символов, то есть от требуемой степени сжатия звуковой информации. Распределение битов по поддиапазонам осуществляется блоком психоакустической модели.

Чтобы выполнить распределение битов в блоке психоакустической модели, анализируется спектр исходного звукового сигнала (не разложенного на поддиапазоны). Для этого производится быстрое преобразование Фурье участков этого сигнала по 512 отсчётов, после чего вычисляется спектр

мощности звукового сигнала и величины звукового давления в каждом частотном поддиапазоне.

Затем анализируются тональные (синусоидальные) и нетональные составляющие звукового сигнала, определяются локальные и глобальные пороги маскировки и вычисляются отношения сигнал/маскирующий сигнал для всех поддиапазонов, на основании которых производится распределение битов по поддиапазнам.

В тех поддиапазнах, в которых искажения звука, вызываемые квантованием, менее заметны для слушателя или маскируются большим уровнем сигнала в других поддиапазнах, квантование делается более грубым, то есть для этих поддиапазонов выделяется меньше битов. Для полностью маскируемых поддиапазонов битов совсем не выделяется. Благодаря этому, удаётся существенно уменьшить количество передаваемой информации при сохранении достаточно высокого качества звука.

Декодирование

Данные, содержащиеся в кадре, декодируются в соответствии с порядком их следования и таблицами кодов, которые содержатся в программе работы декодера. Декодированные данные о распределении битов и о масштабных множителях используются для декодирования и деквантования звуковых данных. После деквантования отсчёты сигналов поддиапазонов умножаются на соответствующие масштабные множители.

После декодирования и деквантования отсчёты сигналов всех поддиапазонов объединяются в выходной цифровой звуковой сигнал.

MPEG-1 Audio Layer II (сокр. MP2, иногда называется Musicam) — один из трёх форматов (уровень 2) сжатия звука с потерями, определённых в стандарте MPEG-1[3]. Применяется в цифровом радиовещании DAB и устаревшем стандарте Video CD, который в 90-е годы использовался для распространения фильмов на оптических компакт-дисках и существовал до широкого распространения DVD [15].

Кодер MPEG-1 Audio Layer 2 развился из аудиокодека MUSICAM (англ. Masking pattern adapted Universal Subband Integrated Coding And Multiplexing — универсальное полосное кодирование и мультиплексирование с адаптацией к шаблону маскировки), разработанного CCETT, Philips и IRT в 1989 как часть исследований EUREKA 147 европейских межправительственных разработок для систем цифрового радиовещания для стационарных, портативных и мобильных приёмных устройств (основан в 1987). Основные параметры MPEG-1 Audio были унаследованы из MUSICAM, включая банк фильтров, обработку во временной области, размер аудиокадра и т. д. Однако, после дополнительного усовершенствования, алгоритм MUSICAM не был использован в финальной версии стандарта MPEG-1 Layer II.

MPEG-1 Audio Layer II определён в стандарте ISO/IEC 11172-3 (MPEG-1 Часть 3)

Частота дискретизации: 32, 44.1 и 48 кГц

Битрейты: 32, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256, 320 и 384 кбит/с.

Расширение формата было представлено в MPEG-2 Audio Layer II и определено в стандарте ISO/IEC 13818-3 (MPEG-2 Part 3)[5][6]

дополнительные частоты дискретизации: 16, 22.05 и 24 кГц

дополнительные битрейты: 8, 16, 24, 40 и 144 кбит/с, для формата 5.1 — около 1 Мбита/с.

поддержка мультисканальности — до 5 полных каналов и канала низкочастотных эффектов.

Поддерживается переменный битрейт (VBR)

Кодирование и декодирование звуковых сигналов MPEG-1 Layer II

Метод кодирования

Входной цифровой звуковой сигнал разделяется на кадры (фреймы), каждый из которых кодируется и декодируется независимо от других кадров. Размер кадра для уровня Layer II составляет 1152 отсчёта.

Полоса аудиосигнала с помощью цифровых полосовых фильтров разбивается на 32 поддиапазона. Все поддиапазоны имеют одинаковую ширину, которая зависит от частоты дискретизации входного сигнала. После деления частота дискретизации уменьшается в 32 раза, так что число отсчётов в кадре в каждом поддиапазоне равно 36.

Затем выполняется квантование данных. Предварительно определяются масштабные множители, которые зависят от максимального значения сигнала. При этом масштабный множитель определяется для групп по 12 отсчётов в каждом поддиапазоне, причём множитель может быть общим для двух или трёх групп. Таким образом, для каждого поддиапазона в кадре определяется до трёх масштабных множителей. Перед квантованием значения сигнала делятся на соответствующие масштабные множители.

Затем в блоке квантования и кодирования выполняется квантование данных. В основе сжатия звуковой информации на уровне Layer II лежит метод, называемый адаптивным распределением битов. Этот метод заключается в выполнении квантования с различным числом двоичных разрядов квантования для разных частотных поддиапазонов. При этом используется равномерное квантование. Полное число битов, выделяемых на все поддиапазоны в данном кадре, зависит от частоты дискретизации входного сигнала и от заданной выходной скорости передачи двоичных символов, то есть от требуемой степени сжатия звуковой информации. Распределение битов по поддиапазонам осуществляется блоком психоакустической модели.

После квантования выполняется кодирование полученных данных. Квантованные отсчёты сигнала в каждом поддиапазоне объединяются по три, и полученные последовательности битов кодируются с использованием таблиц кодов с переменной длиной. Кроме того с помощью соответствующих таблиц кодируются данные о распределении битов по поддиапазонам и данные о масштабных множителях [17].

Чтобы выполнить распределение битов в блоке психоакустической модели анализируется спектр исходного звукового сигнала (не разложенного на поддиапазоны). Для этого производится быстрое преобразование Фурье участков этого сигнала по 1024 отсчётов, после чего вычисляется спектр мощности звукового сигнала и величины звукового давления в каждом частотном поддиапазоне.

Затем анализируются тональные (синусоидальные) и нетональные составляющие звукового сигнала, определяются локальные и глобальные пороги маскировки и вычисляются отношения сигнал/маскирующий сигнал для всех поддиапазонов, на основании которых производится распределение битов по поддиапазнам.

В тех поддиапазонах, в которых искажения звука, вызываемые квантованием, менее заметны для слушателя или маскируются большим уровнем сигнала в других поддиапазонах, квантование делается более грубым, то есть для этих поддиапазонов выделяется меньше битов. Для полностью маскируемых поддиапазонов битов совсем не выделяется. Благодаря этому удаётся существенно уменьшить количество передаваемой информации при сохранении достаточно высокого качества звука.

Декодирование

Данные, содержащиеся в кадре, декодируются в соответствии с порядком их следования и таблицами кодов, которые содержатся в программе работы декодера. Декодированные данные о распределении битов и о масштабных множителях используются для декодирования и деквантования звуковых данных. После деквантования отсчёты сигналов поддиапазонов умножаются на соответствующие масштабные множители.

После декодирования и деквантования отсчёты сигналов всех поддиапазонов объединяются в выходной цифровой звуковой сигнал.

MP3 (более точно, англ. MPEG-1/2/2.5 Layer 3; но не MPEG-3) — это кодек третьего уровня, разработанный командой MPEG, лицензируемый формат файла для хранения аудиоинформации.

MP3 является одним из самых распространённых и популярных форматов цифрового кодирования звуковой информации с потерями. Он широко используется в файлообменных сетях для оценочной передачи музыкальных произведений. Формат может проигрываться практически во всех популярных операционных системах, на большинстве портативных аудиоплееров, 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 звука очень близкого к оригинальному (по мнению большинства слушателей), хотя аудиофилы говорят об осязаемом различии. При создании MP3 со средним битрейтом 128 кбит/с в результате получается файл, размер которого примерно равен 1/11 от оригинального файла с CD-Audio. Само по себе несжатое аудио формата CD-Audio имеет битрейт 1411,2 кбит/с. MP3-файлы могут создаваться с высоким или низким битрейтом, который влияет на качество файла-результата. Принцип сжатия заключается в снижении точности некоторых частей звукового потока, что практически неразличимо для слуха большинства людей. Данный метод называют кодированием восприятия.[1] 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.

Существует три версии MP3 формата для различных нужд: MPEG-1, MPEG-2 и MPEG-2.5. Отличаются они возможными диапазонами битрейта и частоты дискретизации:

32—320 кбит/с при частотах дискретизации 32000 Гц, 44100 Гц и 48000 Гц для MPEG-1 Layer 3;

16—160 кбит/с при частотах дискретизации 16000 Гц, 22050 Гц и 24000 Гц для MPEG-2 Layer 3;

8—160 кбит/с при частотах дискретизации 8000 Гц и 11025 Гц для MPEG-2.5 Layer 3.

Режимы управления кодированием звуковых каналов

Так как формат MP3 поддерживает двухканальное кодирование (стерео), существует 4 режима:

Стерео — двухканальное кодирование, при котором каналы исходного стереосигнала кодируются независимо друг от друга, но распределение бит между каналами в общем битрейте может варьироваться в зависимости от сложности сигнала в каждом канале.

Моно — одноканальное кодирование. Если закодировать двухканальный материал этим способом, различия между каналами будут полностью стёрты, так как два канала смешиваются в один, он кодируется и он же воспроизводится в обоих каналах стереосистемы. Единственным плюсом данного режима может являться только выходное качество по сравнению с режимом Стерео при одинаковом битрейте, так как на один канал приходится вдвое большее количество бит, чем в режиме Стерео.

Двухканальное стерео (англ. Dual Channel) — два независимых канала, например звуковое сопровождение на разных языках. Битрейт делится на два канала. Например, если заданный битрейт 192 кбит/с, то для каждого канала он будет равен только 96 кбит/с.

Объединённое стерео (англ. Joint Stereo, M/S Stereo) — по мнению некоторых, самый оптимальный способ двухканального кодирования. Например, в одном из режимов Объединённое стерео левый и правый каналы преобразуются в их сумму (L+R) и разность (L-R). Для большинства звуковых файлов насыщенность канала с разностью (L-R) получается намного меньше канала с суммой (L+R). Также тут свою роль играет восприятие звука человеком, для которого различия в направлении звука

намного менее примечательны. Поэтому объединённое стерео позволяет либо сэкономить на битрейте канала разности ($L-R$), либо улучшить качество на том же битрейте, поскольку на канал суммы ($L+R$) отводится большая часть битрейта. Бытует мнение, что данный режим не подходит для звукового стереоматериала, в котором в двух каналах воспроизводится субъективно абсолютно различный материал, так как он стирает различия между каналами. Но современные кодеки используют различные схемы в разных фреймах (включая чистое стерео) в зависимости от исходного сигнала.

CBR

CBR расшифровывается как Constant Bit Rate, то есть постоянный битрейт, который задаётся пользователем и не изменяется при кодировании произведения. Таким образом, каждой секунде произведения соответствует одинаковое количество закодированных бит данных (даже при кодировании тишины). CBR может быть полезен для потоков мультимедиа данных по ограниченному каналу; в таком случае кодирование использует все возможности канала данных. Для хранения данный режим кодирования не является оптимальным, так как он не может выделить достаточно места для сложных отрезков исходного произведения, при этом бесполезно тратя место на простых отрезках. Повышенные битрейты (выше 256 кбит/с) могут решить данную проблему, выделив больше места для данных, но зато и пропорционально увеличивая размер файла.

VBR

VBR расшифровывается как Variable Bit Rate, то есть изменяющийся битрейт или переменный битрейт, который динамически изменяется программой-кодером при кодировании в зависимости от насыщенности кодируемого аудиоматериала и установленного пользователем качества кодирования (например, тишина закодируется с минимальным битрейтом). Этот метод MP3-кодирования является самым прогрессивным и до сих пор развивается и улучшается, так как аудиоматериал разной насыщенности

может быть закодирован с определённым качеством, которое обычно выше, чем при установке среднего значения в методе CBR. Плюс к тому, размер файла уменьшается за счёт фрагментов, не требующих высокого битрейта. Минусом данного метода кодирования является сложность предсказания размера выходного файла. Но этот недостаток VBR-кодирования незначителен в сравнении с его достоинствами. Также минусом является то, что VBR считает «незначительной» звуковой информацией более тихие фрагменты, таким образом получается, что если слушать очень громко, то эти фрагменты будут некачественными, в то время как CBR делает с одинаковым битрейтом и тихие, и громкие фрагменты.

Формат VBR постоянно улучшается, благодаря постоянному совершенствованию математической модели кодеков, в частности, после выхода обновлённой версии свободного MP3-кодека LAME (версия 3.99.3), кодирование с переменным битрейтом, по заявлению самих разработчиков, качественно лучше CBR и тем более ABR.

ABR

ABR расшифровывается как Average Bit Rate, то есть усредненный битрейт, который является гибридом VBR и CBR: битрейт в кбит/с задаётся пользователем, а программа варьирует его, постоянно подгоняя под заданный битрейт. Таким образом, кодек будет с осторожностью использовать максимально и минимально возможные значения битрейта, так как рискует не вписаться в заданный пользователем битрейт. Это является явным минусом данного метода, так как сказывается на качестве выходного файла, которое будет немного лучше, чем при использовании CBR, но намного хуже, чем при использовании VBR. С другой стороны, этот метод позволяет наиболее гибко задавать битрейт (может быть любым числом между 8 и 320, против исключительно кратных 16 чисел метода CBR) и вычислять размер выходного файла.

Технические недостатки. Количество каналов звука ограничено двумя, в отличие от AAC и Vorbis.

Юридические ограничения. Патентом на MP3 владеет компания Alcatel-Lucent, которая требует лицензирования некоторых способов использования формата. Срок действия связанных с MP3 патентов истекает в 2007—2017 годах. Почти полный стандарт появился в открытом доступе 6 декабря 1991 года. В США изобретения публично раскрытые более года не могут быть запатентованы. Однако для патентов, оформленных до 8 июня 1995 года, существовала возможность увеличить их сроки действия. Известные патенты, касающиеся расшифровки MP3, прекратили действие в США к декабрю 2012; по другим данным, если учитывать только патенты, заявка на которые была подана до декабря 1992 года, это может произойти в сентябре 2015 года.

MPEG-3 - предназначался для использования в системах телевидения высокой чёткости (high-definition television, HDTV) со скоростью потока данных 20-40 Мбит/с, но позже стал частью стандарта MPEG-2 и отдельно теперь не упоминается. Кстати, формат MP3, который иногда путают с MPEG-3, предназначен только для сжатия аудиоинформации и полное название MP3 звучит как **MPEG-Audio Layer-3**.

MP3.

MP3 - это звуковые файлы с компрессией по технологии MPEG (MPEG 1.0, 2.0 и 2.5, MP3, MPEG 1 Layer 3, Audio MPEG), позволяющей обеспечить наилучшее качество звука при минимальном объеме файла. Это достигается учетом особенностей человеческого слуха, в том числе эффекта маскирования слабого сигнала одного диапазона частот более мощным сигналом соседнего диапазона, когда он имеет место, или мощным сигналом предыдущего фрейма, вызывающего временное понижение чувствительности уха к сигналу текущего фрейма (попросту, удаляются второстепенные звуки, которые не слышатся человеческим ухом из-за наличия в данный/предыдущий момент другого - более громкого). Также учитывается

неспособность большинства людей различать сигналы, по мощности лежащие ниже определенного уровня, разного для разных частотных диапазонов. Данный процесс называется адаптивным кодированием и позволяет экономить на наименее значимых с точки зрения восприятия человеком деталях звучания. Степень сжатия (следовательно и качество), определяются не форматом, а шириной потока данных при кодировании в MP3.

Аудиоинформация, сжатая по данной схеме, может передаваться потоком (streaming), а может храниться в файлах формата MP3 или WAV-MP3. Отличие второго от первого состоит в наличии дополнительного заголовка WAV-файла, что позволяет при наличии MP3 кодека в системе использовать для работы с таким файлом стандартные средства Windows. Параметры компрессии при кодировании файла можно варьировать в широких пределах. Так, например, наивысшее качество, не отличимое от качества CD, достигается при скорости передачи (bitrate) 112...128 kb/s (обеспечиваемой ISDN-модемами), при этом сжатие составляет примерно 14:1 относительно исходного объема (вспомним: на 650 Мб компакт-диска помещается всего 74 минуты звука, т.е. 1 минута "стоит" почти 9 Мб!). Для Интернет, впрочем, и такие объемы великоваты, поэтому чаще используется кодирование с качеством "Подобно УКВ-вещанию" (MPEG 1.0, bitrate 56 kb/s, 44.1 sampling frequency, bandwidth 11 kHz, stereo, 1 min=415 kb), при котором воспроизведение в реальном масштабе времени может быть обеспечено 56К-модемами. Что же касается классического рок-н-ролла, то для него, поскольку он изначально записывался в моно режиме и с неширокой полосой частот, оптимальным можно считать кодирование с качеством "Лучше ДВ/СВ радио" (MPEG 2.0, bitrate 32 kb/s, 22.05 samp. freq., bandwidth 7.5 kHz, mono, 1 min=237 kb).

Принципиальной особенностью MPEG-кодирования является компрессия с потерями. После упаковки и распаковки звукового файла

методом MP3 результат не идентичен оригиналу "бит в бит". Напротив, упаковка целенаправленно исключает из упаковываемого сигнала несущественные компоненты, что приводит к чрезвычайному возрастанию коэффициента сжатия. В зависимости от необходимого качества звука метод MP3 способен сжать звук в десять и более раз (См. Степень сжатия и качество).

Преимущества MP3.

MP3 на сегодня имеет два огромных преимущества перед другими доступными форматами. Правда, Microsoft пытается потеснить MP3 со своим новым форматом WMA, а также есть альтернативные форматы VQF и AAC, но они еще не получили должного распространения, хотя качество WMA, если судить по отзывам, несколько лучше. Однако WMA пока, фактически, закрыт для свободного использования, поэтому у него есть проблемы с различным софтом для кодирования / прослушивания / обслуживания (хотя, кто же сомневается в мобилизационных возможностях Microsoft :-). Первое преимущество состоит в том, что ни про один из существующих подобных форматов нельзя пока сказать, что он полностью гарантирует устойчивое сохранение качества звучания на достаточно высоких битрейтах, кроме MP3, который достойно выдержал проверку временем. Второе, не менее важное преимущество - на ближайшие годы, а возможно, и на все десятилетие, MP3 стал стандартом де факто, поскольку много сделано в него вложений пользующимися им сторонами, в том числе и цифровыми радиостанциями. Для MP3 также написано множество удобного программного обеспечения. Сейчас уже налажено производство аппаратных MP3 плееров, и карманных, и для автомобилей. Таким образом, MP3 стал первым массово признанным форматом хранения аудио после CD-Audio (пусть, часто и нелегальным).

Степень сжатия и качество.

128 kbps, в принципе, нормальный формат, если Вы имеете обычную аудиотехнику и рядовой слух. Но, чтобы быть уверенным, лучше кодировать на 160 kbps, т.к. это даст запас (весьма существенный) по сравнению с 128 kbps на тот случай, если улучшится качество аппаратуры. А если хотите быть совсем уверены в том, что Ваша музыка ничего не потеряет от кодирования, то берите 192 kbps.

128 kbps (11:1) Самый популярный на сегодня битрейт. Степень сжатия 11:1 - это, конечно, аргумент, особенно для Интернета, где каждый килобайт на счету. Однако при этом не очень хорошо сохраняются высокие частоты и имеют место некоторые искажения звука. При этом можно точно сказать, что на обычной аппаратуре, например используя обычную звуковую карту, компьютерные колонки, пусть и неплохого качества, разницы не будет заметно, если только Вы не эксперт по звуку. Однако на качественных колонках отсутствие высоких частот проявляется довольно заметным образом - музыка звучит приглушенно. На 128 кб/с лучше себя ведут кодеры от Fraunhofer-IIS - они оставляют больше высоких частот, чем кодеры ISO и, тем более, Xing. Все ISO кодеры как один добавляют характерный звон, наличие которого связано с особенностью психоакустических фильтров ISO. На этом битрейте отлично себя показал MP3 Producer, специально оптимизированный под низкие битрейты.

160 kbps (8:1) Лучше, чем 128 кб/с. На неплохой (хорошей непрофессиональной) аппаратуре разницы уже не заметно. Но все равно присутствует недостаток высоких частот. На этом битрейте себя лучше проявили кодеры от Fraunhofer, т.к. ISO кодеры продолжают "звенеть", хотя и меньше. Выбирать приходится между кодерами от Fraunhofer, поэтому неплохой выбор - новый MP3 Producer.

192 kbps (7:1) Последнее время, с ростом пропускной способности каналов Интернета, этот битрейт становится более популярным на его просторах. Заметная часть новых записей оцифровывается именно в 192 kbps, во всяком случае теми, кто занимается MP3 серьезно. Кодеры от Fraunhofer убирают высокие больше, чем кодеры ISO, которые уже не "звонят". На этом битрейте кодеры от ISO чуть-чуть опережают Fraunhofer.

Но ISO-оптимизированные - гораздо быстрее. 256 kbps (5:1) Вот тут практически все совсем перестают ощущать разницу между записью с CD и MP3-файлами, даже на очень качественной аппаратуре, при кодировании ISO-кодерами. Все ISO-кодеры обеспечивают на слух одинаковое качество, которое выше, чем у других кодеков [13].

Используя стерео эффекты и ограничивая ширину полосы звуковых частот, кодирование схем может достигнуть приемлемого надежного качества в более низких частотах. Некоторые типичные данные для **MPEG Layer-3** приведены в таблице 3.1

Типичные данные для **MPEG Layer-3**

Таблица 3.1.

sound quality	bandwidth	mode	bitrate	reduction ratio
telephone sound	2.5 kHz	mono	8 kbps *	96:1
better than shortwave	4.5 kHz	mono	16 kbps	48:1
better than AM radio	7.5 kHz	mono	32 kbps	24:1
similar to FM radio	11 kHz	stereo	56...64 kbps	26...24:1
near-CD	15 kHz	stereo	96 kbps	16:1
CD	>15 kHz	stereo	112..128kbps	14..12:1

Во всех международных тестах слушания, MPEG Layer-3 впечатляюще доказывал высокое исполнение, поддерживая исходное качество при сжатии данных 1:12 (около 64 kbit/s). Если материал допускает ограниченную ширину полосы частот около 10 kHz, разумное надежное качество для стерео сигналов может быть достигнуто при сжатии 1:24. Степени сжатия приведены в таблице 3.2.

Степени сжатия

Таблица 3.2.

1:4	by Layer 1 (corresponds with 384 kbps for a stereo signal),
------------	--

1:6...1:8	by Layer 2 (corresponds with 256..192 kbps for a stereo signal),
1:10...1:12	by Layer 3 (corresponds with 128..112 kbps for a stereo signal),

3.2. Сравнительный анализ цифровых звуковых стандартов

MP3

Формат сжатия MPEG Layer III был изобретен и запатентован институтом Фраунхоффера в начале 90-х годов XX века. Принцип кодирования основан на том, что человеческое ухо воспринимает не все звуки, идущие из динамика, а значит, большое количество информации, записанной на обычных аудио-CD, является избыточным. Например, если возникает сильный звук на определенной частоте, а на частоте чуть выше – слабый, то человек слышит только сильный звук, а слабый при этом «выпадает». Также человек не воспринимает другие звуки за 5 мс до и в течение 100 мс после сильного звука (инертность слуха). Эти и другие особенности позволяют уменьшить звуковые файлы в десятки раз. Правда, чем качественнее звук на выходе надо получить, тем более «объемистым» будет MP3-файл. Одной из важных особенностей MP3-файлов является наличие так называемого ID3-Tag'a. Он включает в себе информацию о той музыке, которая записана в файл. Чаще всего ее вносят перед кодировкой файла. На данный момент существует две версии тага: ID3v1 Tag и ID3v2 Tag. Оба они представляют собой несколько полей, в которые вносятся номер трека, название композиции, альбома, композитор (группа), жанр, год, комментарии. Во второй версии тага добавлены еще несколько дополнительных полей.

MPEG (от Motion Picture Expert Group – группа экспертов по движущимся изображениям) 1 Layer III (реже MPEG 2 Layer III) уже долгие годы является для многих пользователей единственной ассоциацией со словосочетанием "компьютерная музыка". Разработанный в конце 80х годов, нетребовательный к ресурсам (воспроизведение MP3 файлов возможно даже

на компьютерах с процессорами 486) формат, позволявший сжимать музыку до 10 раз без катастрофических потерь качества быстро прижился на домашних компьютерах. Хотя еще недавно большинство кодировщиков были платными, сейчас несложно найти проигрыватели и кодировщики, распространяемые по лицензии freeware. Через некоторое время стало ясно, что "CD качество" при битрейте в 128 Кб/с невозможно, по крайней мере с данным стандартом, так как с оснащением компьютеров более совершенными звуковыми картами и акустическими системами позволяло выявить недостатки подобного кодирования. Вполне закономерным стало повышение битрейта и совершенствование кодеков: технологии VBR и Joint Stereo (комбинированное стерео) позволяли значительно сократить размер файла при повышении качества. Современные кодировщики позволяют достичь качества звучания, на слух неотличимого от компакт диска на битрейтах в диапазоне 192-256 Кб/с даже на высококачественной аппаратуре. Тем не менее, в некоторых редких случаях (при наличии хорошего слуха и аппаратуры) даже битрейта 320Кб/с бывает недостаточно. Трудность заключается в том, что сам по себе формат MP3 имеет недостатки, от которых практически невозможно избавиться. Одним из них является так называемый эффект преэхо, из-за которого кодирование определенных сигналов сопряжено со значительными трудностями. На практике же использование постоянного битрейта 320Кб/с зачастую оказывается избыточным и чаще всего приводит к бессмысленной трате места. Качество звучания MP3 файла может сильно зависеть от выбранных кодировщика и проигрывателя. Для создания MP3 файлов идеально подходит бесплатный LAME (кодировщики Fraunhofer являются платными и позволяют достичь сравнимого с LAME качества, а кодеры Xing, Blade и большинство других не заслуживают внимания), а для воспроизведения – одна из последних версий Winamp 2.

MP3pro

Кодек MP3Pro был анонсирован в июле 2001. Появления кодека ожидали с интересом и некоторым нетерпением по двум причинам. Во-первых, этот кодек по своим возможностям обещал быть если не революционным, то очень развитым. Во-вторых, MP3Pro анонсировался и задумывался как продолжение знаменитого «MP3» (MPEG-1 Layer III), а то, что разработкой MP3Pro занимался «отец» MP3 (институт Fraunhofer), внушало доверие и уважение. Несмотря на солидность разработчиков и проекта, MP3Pro довольно долго оставался в тени. Сначала не было ни одной программы не только для кодирования в MP3Pro, но даже для воспроизведения файлов в этом формате. Потом, хотя программы, использующие MP3Pro, и стали появляться, процесс их интеграции проходил вяло и малозаметно для пользователя. И только сегодня MP3Pro стал развиваться и приобретать аппаратную основу, что незамедлительно сказалось на потребительском рынке. В частности, на рынке начали появляться аппаратные плееры, способные воспроизводить аудио в формате MP3Pro. Этот факт является хорошим сигналом для пользователя, говорящим о том, что MP3Pro после всех неурядиц все-таки принят производителями. В данной статье мы кратко рассмотрим историю появления MP3Pro, популярно и доступной форме разберемся в его возможностях, области применения, а также в проблемах, возникающих при использовании кодека. Механизм «упрощения» аудио можно пояснить следующим образом. Исходный аудио поток раскладывается на частотные составляющие спектра. Ясно, что чем меньше «нюансов звучания» имеются в сигнале (чем проще частотный спектр), тем лучше он поддается компрессии. Упростить сигнал можно по-разному. Например, существует ряд способов упрощения информации о стерео панораме сигнала; можно также отфильтровать все частоты выше определенной границы, что автоматически упростит сигнал в высоких частотных областях (но при этом, естественно, заметно испортит звучание). Но основным способом является психоакустический анализ (с последующей соответствующей обработкой сигнала), когда кодер анализирует аудио

информацию и, опираясь на указанный пользователем битрейт, «решает» какие тонкости звучания можно выбросить. В качестве справки: при сжатии в MP3 пользователь указывает желаемый битрейт (или границы изменения битрейта) для сжатого выходного потока (битрейт – количество бит, используемых для хранения одной секунды аудио). Чем ниже битрейт, тем меньше бит позволяет кодеру отводить для хранения информации об одной секунде аудио и, таким образом, тем «глубже» кодер упрощает сигнал, что соответственно влияет на качество звучания получаемого сжатого потока аудио. Наиболее распространенное среднее значение битрейта для MP3 колеблется в пределах от 128 до 192 Kbps («килобит в секунду»). Здесь следует заметить, что применение психоакустики приводит к тому, что процесс декодирования уже не способен восстановить утраченные во время компрессии данные (нюансы звучания, отфильтрованные частоты и проч.).

Кодек MP3Pro анонсирован в июле 2001 года компанией Coding Tech. вместе с Tomson Multimedia и институтом Fraunhofer. Формат MP3Pro является прямым продолжением, или точнее, развитием старого MP3. MP3Pro совместим с MP3 «вперед» полностью и «назад» частично. То есть файлы, закодированные с помощью MP3Pro, можно воспроизводить в обычных проигрывателях, однако качество звучания при этом заметно хуже, чем при воспроизведении в специальном MP3Pro-проигрывателе. Это связано с тем, что файл (контейнер) в формате MP3Pro несет в себе два потока аудио: один – стандартный в формате MPEG-1 Layer III, а другой – специальный, дополнительный поток. Обычные проигрыватели MP3 распознают в новом формате только один поток - обычный MPEG-1 Layer 3, это и обуславливает лишь частичную совместимость «назад». В MP3Pro использована новая технология SBR (Spectral Band Replication). Эта технология предназначена для передачи верхнего частотного диапазона. Идея технологии и ее предпосылки следующие. Дело в том, что как и алгоритмы компрессии данных без потерь ограничены в своей эффективности, так и технологии, основанные на использовании психоакустической модели, имеют один

общий недостаток: все они работают качественно до битрейта 128 Kbps. На более низких битрейтах начинают проявляться сильные искажения звучания, так называемые «артефакты кодирования»: бульканье, позвякивания, заметные скачки уровня сигнала на различных частотах и проч. Единственный найденный способ борьбы с этим явлением заключается в урезании частотного диапазона сигнала перед компрессией. Однако эта процедура также отрицательно сказывается на восприятии звучания. Все это показывает, что для получения более-менее качественного звучания на низких битрейтах использования психоакустической модели недостаточно. Новая технология SBR от Coding Tech. дополняет использование психоакустической модели. Идея технологии состоит в следующем: в файле передается (кодируется) чуть более узкий диапазон частот, чем обычно (с «урезанными верхами»), а верхние частоты воссоздаются уже самим декодером на основе имеющейся небольшой дополнительной информации о более высоких частотных составляющих. Таким образом, технология SBR применяется фактически не столько на стадии сжатия, сколько на стадии декодирования. Технология SBR в MP3Pro задействована следующим образом. Основные данные кодируются в несколько более зауженном, чем обычно, диапазоне частот в MP3 и записываются в первый основной поток контейнера MP3Pro. Второй, параллельный поток данных, невидимый для обычного MP3-проигрывателя, несет ту минимальную необходимую информацию, которая используется при воспроизведении для восстановления верхних частот. Исследования показывают, что эта информация есть усредненная мощность сигнала в верхнем (урезанном) диапазоне частот. Точнее, не одна усредненная мощность для всего диапазона вырезанных частот, а информация о средней мощности в нескольких полосах частот верхнего диапазона. Такой вывод был сделан на основе следующего несложного, но очень показательного теста. Возьмем тестовый сигнал протяженностью 13 секунд, состоящий из белого шума от 0 до 10 КГц, плюс тон, плавно переходящий от 10 до 19 КГц. Закодируем этот

сигнал с помощью MP3Pro (в режиме VBR со средним битрейтом 142 Kbps) и, для сравнения, в обычный MP3 (в режиме CBR на битрейте 140 Kbps). Во время сжатия кодер закодировал в MP3 весь частотный диапазон сигнала ниже 10 КГц, сигнал выше этого рубежа был нарезан на полосы, в каждой такой полосе кодер просто следил за средним уровнем сигнала и сохранял результат в выходном потоке. Пока плавно поднимающийся тон находился в пределах одной подполосы, для кодера средний уровень сигнала в этой подполосе оставался неизменным. Как только тон «уходил» из подполосы, измеряемый кодером уровень сигнала падал до нуля. Во время декодирования декодер, не имея данных о том, что находилось в каждой подполосе, просто брал известный сигнал в нижнем диапазоне (в какой-то определенной области), умножал его на известный средний уровень сигнала в подполосе и размещал в ней полученный после умножения сигнал. Конечно, это лишь искусственный тест, однако он как никакой другой наиболее ярко выявляет механизм работы MP3Pro. Стоит обратить внимание на то, что «старичок» MP3 на битрейте даже чуть ниже MP3Pro достаточно точно и четко «отработал» и сохранил информацию о всем частотном диапазоне.

MPEGplus/Musepack (MP+/MPC/MPP)

Данный кодировщик похож по принципу действия на MPEG Layer II (MP2), но использует более совершенный алгоритм. В отличие от большинства других современных кодеков целью создателей Musepack было вовсе не стремление получить максимально возможное качество на низких битрейтах: лучше всего формат показывает себя на средних и высоких битрейтах (типичный битрейт файлов обычно находится в диапазоне 160-180Кб/с). Великолепная психоакустическая модель, использующая VBR кодирование позволяет добиться прекрасного качества звучания. В результате кодек показывает результаты более высокие, чем большинство его соперников на аналогичных битрейтах. Скорость работы кодировщика достаточно высока: на создание MPC файла тратится примерно в два раза

меньше времени, чем на создание MP3 файла при помощи lame с аналогичными настройками. Качество файлов, получаемых при сжатии в MPC значительно превышает качество аналогичных файлов MP3. При использовании настройки --normal кодировщика я ни разу не смог отличить кодированный файл от оригинала в ABX тесте. Более того, MPC файлы, созданные таким образом звучат лучше, чем высококачественные MP3 файлы с битрейтом 320Кб/с (разумеется, в тех редких случаях, когда данное сравнение можно проводить). Следует сказать, что именно пресет normal является "изюминкой" формата. Данный режим, вопреки названию, дает результат, абсолютно неотличимый от оригинала на слух, причем разницу не ощущают даже люди, обладающие прекрасным слухом и качественным оборудованием.

Одним из серьезных недостатков нынешней версии Musepack является ограничение на формат файла: 44КГц, 16 бит, стерео, что делает его неприменимым для, например, сжатия звуковых дорожек к фильмам на DVD. Отсутствие аппаратной поддержки также может стать препятствием при выборе формата.

WMA 8\9

Патент на формат Windows Media Audio, принадлежит Microsoft. Прообразом WMA был небезызвестный формат VQF, разработанный компанией Voxware в составе проекта TwinVQ в 1998 году. После этого компания отделилась от проекта и уже самостоятельно разработала формат Voxware Audio Codec. После его удачного шествия по всему миру, Microsoft заинтересовалась кампанией Voxware, включив поддержку Voxware Audio Codec в Windows Media Player, после чего купила ее. Вопрос качественного отличия MP3 и WMA до сих пор остается открытым. Результаты работы кодека WMA оказались немногим лучше, чем в случае использования MP3 с битрейтом 64Кб/с.

WMA8 является наиболее распространенной на данный момент версией кодека WMA, нередко используемой для кодирования музыки дома.

Значительным преимуществом по сравнению с остальными форматами является тот факт, что зачастую кодеки WMA уже установлены в системе. WMA9 – последняя имеющаяся на данный момент версия кодека, имеющая ряд улучшений, например, режим кодирования с переменным битрейтом. MP3 на данный момент является наиболее распространенным форматом аудиосжатия. Формат Windows Media Audio был разработан всенародно любимой Microsoft в качестве очередного "заместителя" MP3. Помимо обычных для разработчиков обещаний "качество звука, как у MP3, но при вдвое меньшем размере файла" создатели позаботились о защите данных внутри музыкального файла. На практике для рядового пользователя это оборачивалось невозможностью редактировать теги уже готового WMA файла, а также множеством проблем, связанных с так называемыми защищенными WMA файлами (например, записанный дома файл нельзя было воспроизвести на рабочем компьютере). Для того чтобы проиграть защищенный файл часто требовалось загрузить из интернет специальный сертификат, обычно позволявший прослушивать композицию в течение ограниченного времени, по окончании которого вам нужно было либо заплатить за использование файла и получить возможность слушать его в дальнейшем, либо купить компакт диск с понравившейся композицией. Формат постоянно совершенствуется, постепенно появляется поддержка более высоких битрейтов (до 160 Кб/с в новых версиях кодека).

Качество звучания WMA файла вполне сравнимо с качеством MP3 файла с тем же битрейтом, иногда превышая MP3 на низких битрейтах. Немалым плюсом можно считать то, что кодировщики WMA уже встроены в Windows, а последние версии Windows Media Player позволяют кодировать компакт диски сразу же в новый формат. На некоторых интернет сайтах, посвященных сжатию музыки можно встретить призывы не использовать WMA из-за его относительно низкого качества звучания. Лично я через некоторое время отказался от использования данного формата, т.к. на битрейтах 64 и 96Кб/с искажения и артефакты отчетливо слышны даже на не

слишком качественной аппаратуре, а повышения качества по сравнению с аналогичными файлами MP3 на более высоких битрейтах я не заметил. Так что заявления Microsoft о том, что WMA 64Кб/с это «CD качество» советую рассматривать не более, чем глупую шутку. На сайте Microsoft регулярно проводит тестирование новых версий кодека (преимущественно на битрейтах до 128 Кб/с), противопоставляя им устаревшие или низкокачественные кодировщики MP3. Неудивительно, что WMA выходит в подобных тестах победителем. В данное время формат WMA наряду с MP3 поддерживается большим числом производителей портативных аппаратных проигрывателей, что является несомненным плюсом. Вряд ли файлы, сжатые по технологии WMA удовлетворят качеством звука любителей высококачественной музыки, но менее привередливых пользователей данный кодек вполне может устроить. Недавно появившийся формат WMA9, поддерживает кодирование с переменным битрейтом, а также сжатие без потерь (loseless) . Несмотря на это никаких принципиальных улучшений качества новая версия кодировщика не принесла: сжатие с потерями все еще недостаточно качественна, а результаты loseless компрессии хуже, чем у современных свободно распространяемых кодировщиков.

OGG Vorbis

Один из наиболее перспективных форматов сжатия. Он был опубликован в 2000 году. А главное - OGG Vorbis совершенно бесплатный. Принцип кодирования схож с MP3, но качество звука - значительно выше, и, в отличие от MP3, в файле, записанном в формате OGG, может содержаться до 250 каналов, а значит, можно кодировать звук Dolby Surround. Также OGG-файлы могут содержать изображение и тексты.

В качестве кодировщика и декодировщика использовались консольные программы Oggenc и Oggdec с сайта разработчика, основанные на финальной версии Xiph.Org libVorbis I.

Еще до выхода финальной версии кодировщика не утихал шум вокруг этого нового формата сжатия музыки. Заявления разработчиков выглядят

заманчивыми: полная открытость формата и его свобода от различных патентов (в отличие от MP3 и многих других форматов), поддержка широкого диапазона частот дискретизации (8-48 КГц) и битрейтов (от 16 до 256 Кб/с на канал), возможность кодирования не только стерео сигнала, но и нескольких каналов аудио данных, высокое качество звучания и многое другое. Считается, что битрейты 160-190 Кб/с достаточно для обеспечения звучания, на слух неотличимого от компакт диска.

Информация об исполнителе, названии композиции и т.д. хранится в Unicode, что позволяет избежать в тэгах проблем с символами, отличными от латиницы, да и сама по себе организация тэгов реализована более грамотно, чем в ID3V2 (достаточно новый формат тэгов MP3/AAC файлов). У разработчиков поистине наполеоновские планы: сделать OGG единственным форматом для хранения музыки и передачи ее через Интернет, вытеснив все остальные "коммерческие" форматы

AAC

Формат Advanced Audio Coding (расширенное аудио кодирование), также известный как MPEG2 nbc (not backwards compatible, не обратно совместимый) является преемником формата MP3. Сочетая в себе алгоритмы кодирования MPEG2/MPEG4 AAC имеет более широкие, чем MP3 возможности: возможность кодирования нескольких звуковых каналов с частотой дискретизации до 96КГц и более высокое, чем у MP3 соотношение качество/размер делают его весьма привлекательным как для создания музыкальной коллекции, так и для кодирования многоканальных звуковых дорожек. Качество звучания файлов, сжатых при помощи AAC оценивается как отличное. Этот формат прекрасно подойдет как для архивного хранения музыки (без заметной потери качества), так и для создания небольшой высококачественной музыкальной коллекции. На данный момент уже есть несколько моделей аппаратных проигрывателей, имеющих поддержку формата AAC.

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. Следует отметить, что существует несколько кодеков AAC, различающихся по качеству/скорости и не всегда совместимых друг с другом. По завершении работы над MP3 была начата разработка кодировщика, который обеспечивал бы лучшее качество звука при сохранении совместимости с MPEG-1. Результатом этих усилий стала технология MPEG-2 Advanced Audio Coding (AAC). AAC (Advanced Audio Coding) - это технология высококачественного кодирования аудиоинформации, используемая множеством приложений в области интернет-вещания и распространения музыки по электронным каналам. По результатам независимых тестов эффективность кодирования в формате AAC оказалась выше, чем в формате MP3, обеспечивая лучшее качество звука при меньшей скорости потока данных. Формат AAC разработан и стандартизирован как часть стандарта ISO/IEC MPEG-2 четырьмя компаниями - лидерами в области аудиокодирования (AT&T, Dolby Laboratories, Fraunhofer IIS и Sony Corporation) и поддерживается все растущим числом производителей программного обеспечения и аппаратных средств. Несмотря на то, что оба формата используют технологию трансформирования информации, они заметно различаются способом трансформации. AAC использует более современные способы кодирования. Например, на карте памяти объемом 32 МБ, используемой музыкальным проигрывателем Nokia, умещается вдове больше музыки в формате AAC, чем в формате MP3, при одинаковом качестве звука.

ADPCM

Кодек Microsoft ADPCM (Adaptive Differential Pulse Code Modulation, адаптивная дифференциальная импульсно-кодовая модуляция), некогда популярный у пиратов, не желавших тратить время на кодирование музыки в MP3 - далеко не образец качества. Для уменьшения объема файла

приходилось создавать восьми или даже четырехбитные wave файлы с частотой дискретизации 22КГц, что давало весьма слабый результат. И если MP3 файл с фиксированным битрейтом 128Кб/с звучит более-менее сносно (а для некоторых слушателей - идеально), то аналогичного размера wave файл, сжатый кодеком ADPCM звучит просто ужасно. ADPCM использует очень простой алгоритм сжатия, который обеспечивает высокую скорость кодирования на слабых компьютерах, но абсолютно неприменим для хранения музыки. Крайне низкое качество звучания при приемлемой степени сжатия привели к тому, что в данный момент для сжатия музыки данный кодек почти не используется. Исключением являются компьютерные игры, создатели которых нередко используют ADPCM WAV файлы для хранения саундтреков.

Liquid Audio

Еще один закрытый коммерческий формат. Используя современные высококачественные алгоритмы AAC и ATRAC он позволяет получить результат по качеству звучания превосходящий аналогичный MP3 файл. Кроме самого аудиотрека внутри LQT файла может храниться дополнительная информация: сведения об исполнителе и альбоме, тексты песен, графика и т.д. Содержимое файла шифруется во избежание нелегального копирования. Недостатком для домашнего использования может являться сложность декодирования файлов LQT в wave (для последующей записи на CD-R/RW) и отсутствие бесплатных кодировщиков. Для кодирования музыки дома лучше подойдет "обычный" AAC.

3.3 Заключение по 3-й главе

В третьей главе были рассмотрены и изучены различные форматы звуковых файлов, их особенность, преимущества по сравнению их аналогу, а также их недостатки. Такое разнообразие цифровых звуковых форматов

обусловлена тем, что каждый производитель звуковых форматов желает чтобы его продукт стал популярным и общепринятым. Также имеет фактор в данном вопросе первенство в индустрии аудио бизнеса. То есть, те производители которые впервые предложили рынку свои форматы и дальше усовершенствовали свои продукты. Но наряду с этими явлениями выпускались новые виды форматов других производителей, в которых учитывались все нюансы звукового воспроизведения и редактирования. Но к сожалению данные форматы не стали глобально масштабными по сравнению звуковых форматов MPEG. Форматы MPEG предназначались прежде всего для качественного оцифрования видео с звуковым сопровождением. Различные форматы MPEG для видео масштабно используется в наземном цифровом телевидении, а также в спутниковом телевидении. А формат MPEG layer III сокращенно MP3 стал общеизвестным и очень популярным. На сегодняшний день все выпускаемые сотовые телефоны, флагманские продукты и планшеты поддерживают формат MP3.

4. БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ

4.1. Требования к организации рабочего места

Привлекательность рабочего места способствует росту продуктивности и улучшению самочувствия персонала. Рабочие места менеджеров,

работников аппарата управления и всех сотрудников должны быть организованы и оборудованы с учетом целого ряда требований, основными из которых являются следующие:

- информационные,
- экономические,
- эргономические,
- гигиенические,
- эстетические,
- технические и организационные требования.

Информационные требования охватывают комплекс мероприятий по информационному обеспечению работы менеджера: определение объемов и структуры информации, которая поступает на рабочее место, обрабатывается на нем, создается и передается на другие рабочие места. Информация, которая концентрируется на рабочем месте, должна отвечать общим принципам научной организации труда, требованиям к управленческой информации и быть достаточной для выполнения служебных обязанностей.

Экономические требования предусматривают такую организацию рабочего места, при которой затраты на его содержание минимальные, однако достаточные для его нормального функционирования.

Эргономические требования изучаются и формулируются отраслью науки, называемой эргономикой. *Эргономика* изучает функциональные возможности человека в трудовых процессах с целью создания для него оптимальных условий труда, которые делают его высокопродуктивным и надежным, одновременно обеспечивают человеку необходимые удобства и сохраняют его силу, здоровье и трудоспособность. Таким образом, все, что окружает работающего человека, – помещение, мебель, оборудование, машины, механизмы, должно отвечать требованиям эргономики и быть максимально приспособленным к человеку, к его физической, физиологической и эстетической природе.

Гигиенические требования – это требования к освещенности, воздухообмену, температурному режиму, влажности, шуму и другим факторам среды, которые влияют на здоровье и трудоспособность человека.

Эстетические требования. Продуктивность - труда человека в значительной степени зависит от внешнего оформления среды, в которой он трудится. Это внешний вид помещения и орудий труда, их цветовая гамма, наличие живых цветов.

Технические требования. Для любой работы необходим определенный простор для размещения мебели, оборудования, проходов и самого работника. Должна учитываться особенность работы управленческого персонала: необходимость общения в процессе работы с другими людьми, для которых также необходима площадь. Санитарными нормами предусмотрены размеры рабочих площадей и мебели для различных категорий служащих.

Организационные требования. Должна быть определена сфера компетенции работника на определенном рабочем месте, его права, обязанности, подчиненность, вертикальные и горизонтальные связи с другими рабочими местами, формы и методы стимулирования эффективной работы. Эти вопросы решаются путем разработки положений о структурных подразделениях и должностных инструкций.

В последнее время на многих предприятиях интенсивно внедряются *японские методы управления производительностью труда*, включающие, в том числе, и повышение культуры производства. Применительно к организации рабочего места эти методы предусматривают следующие требования:

- убрать ненужные предметы с рабочего места;
- правильно располагать и хранить необходимые предметы и инструменты;
- постоянно поддерживать чистоту и порядок на рабочем месте;

– постоянно поддерживать рабочее место в готовности к проведению работ;

– каждому работнику усвоить и неукоснительно соблюдать перечисленные требования.

4.2. Защита от электромагнитных полей и лазерного излучения

Электромагнитные волны возникают при ускоренном движении электрических зарядов. Электромагнитные волны – это взаимосвязанное распространение в пространстве изменяющихся электрического и магнитного полей. Совокупность этих полей, неразрывно связанных друг с другом, называется *электромагнитным полем*. Несмотря на то, что длина электромагнитных волн и их свойства различны, все они, начиная от радиоволн и заканчивая гамма-излучением, – одной физической природы. Исследованный в настоящее время диапазон электромагнитных волн состоит из волн с длинами, соответствующими частотам от 10^3 до 10^{24} Гц. По мере убывания длины волны в диапазон включаются радиоволны, инфракрасное излучение, видимый свет (световые лучи), ультрафиолетовое излучение, рентгеновское излучение и гамма-излучение.

Источниками электромагнитных полей являются атмосферное электричество, космические лучи, излучение солнца, а также искусственные источники: различные генераторы, трансформаторы, антенны, лазерные установки, микроволновые печи, мониторы компьютеров и др. На предприятиях источниками электромагнитных полей промышленной частоты являются высоковольтные линии электропередач (ЛЭП), измерительные приборы, устройства защиты и автоматики, соединительные шины и др. В зависимости от длины волны электромагнитное излучение делят на ряд диапазонов

Скорость распространения электромагнитных волн в вакууме не зависит от длины волны и равна: $C = 2,997925 \cdot 10^8$ м/с.

Электромагнитная волна, распространяясь в неограниченном пространстве со скоростью света, создает переменное электромагнитное поле, которое способно воздействовать на заряженные частицы и токи, в результате чего происходит превращение энергии поля в другие виды энергии. Как уже сказано выше, переменное электромагнитное поле представляет собой совокупность магнитного и электрического полей, количественной характеристикой которых являются напряженность электрического поля E (размерность – вольт на метр, или, сокращенно, В/м) и напряженность магнитного поля H (размерность – ампер на метр, или, сокращенно, А/м). Величины E и H – векторные, их колебания происходят во взаимно перпендикулярных плоскостях.

При распространении в воздухе или в вакууме $E = 377 H$.

Плотность потока энергии (I) может быть записана (в векторной форме) как $\vec{I} = \vec{E}\vec{H}$. Эти величины показывают, какое количество энергии протекает за 1 с через площадку, расположенную перпендикулярно движению волны.

Если сформировавшаяся электромагнитная волна имеет сферическую форму, то справедливо следующее равенство:

$$I = \frac{P_{ист}}{4\pi r^2} = \vec{E}\vec{H} = \frac{E^2}{377}, \quad (4.1.)$$

где $P_{ист}$ – мощность источника излучения, Вт;

r – расстояние от источника излучения, м.

Отсюда можно определить напряженность электрического поля по формуле:

$$E = \frac{\sqrt{30P_{ист}}}{r}. \quad (4.2.)$$

Начиная от источника излучения всю область распространения электромагнитных волн принято условно разделять на три зоны: ближнюю, промежуточную и дальнюю. Радиус ближней зоны приблизительно составляет 1/6 волны от источника излучения, а дальняя зона начинается на

расстоянии, равном примерно 6 длинам волн; промежуточная зона находится между ними.

Переменные электромагнитные поля способны оказывать негативное воздействие на организм человека, последствия которого зависят от напряженности электрического и магнитного полей, частоты излучения, плотности потока энергии, размера облучаемой поверхности тела человека и индивидуальных способностей его организма. Ткани человеческого организма поглощают энергию электромагнитного поля¹, в результате этого происходит нагрев тела человека. Интенсивнее всего электромагнитные поля воздействуют на органы и ткани с большим содержанием воды: мозг, желудок, желчный и мочевой пузырь, почки. При воздействии электромагнитного излучения на глаза человека возможно помутнение хрусталика (катаракта).

Как известно, человеческий организм обладает свойством терморегуляции, т. е. поддержания постоянной температуры тела. При нагреве человеческого организма в электромагнитном поле происходит отвод избыточной теплоты до плотности потока энергии $I = 10 \text{ мВт/см}^2$. Эта величина называется *тепловым порогом*, начиная с которого система терморегуляции не справляется с отводом генерируемого тепла, происходит перегрев организма человека, что негативно сказывается на его здоровье.

Воздействие электромагнитных полей с интенсивностью, меньшей теплового порога, также небезопасно для здоровья человека. Оно нарушает функции сердечно-сосудистой системы, ухудшает обмен веществ, приводит к изменению состава крови, снижает биохимическую активность белковых молекул. При длительном воздействии на работающих электромагнитного излучения различной частоты возникают повышенная утомляемость, сонливость или нарушение сна, боли в области сердца, торможение рефлексов и т.д.

Произошедшие под действием электромагнитных полей нарушения в организме обратимы, если в нем не произошло патологических изменений.

Для этого необходимо либо прекратить контакт с излучением, либо разработать мероприятия по защите от него.

При воздействии на организм человека постоянных магнитных и электростатических полей с интенсивностью, превышающей безопасный уровень, могут развиваться нарушения в деятельности сердечно-сосудистой системы, органов дыхания и пищеварения, возможно изменение состава крови и др. Электрические поля промышленной частоты ($f = 50$ Гц) воздействуют на мозг и центральную нервную систему.

Между человеком, находящимся в таком поле и обладающим определенным потенциалом, и металлическим проводником с меньшим потенциалом может возникнуть электрический заряд, приводящий к судорожным сокращениям мышц или иным, более тяжелым последствиям (см. гл. 20).

Предельно допустимые уровни облучения в диапазоне радиочастот определяются ГОСТом 12.1.006-84 «Электромагнитные поля радиочастот. Допустимые уровни на рабочих местах и требования к проведению контроля». В соответствии с этим нормативным документом установлена предельно допустимая напряженность электрического поля ($E_{пд}$, В/м) в диапазоне 0,06 – 300 МГц и предельно допустимая энергетическая нагрузка за рабочий день [$ЭН_{E_{пд}}$, (В/м) 2 ·ч]. Между этими величинами

существует следующая связь:

$$E_{пд} = \sqrt{\frac{ЭН_{E_{пд}}}{T}}, \quad (4.3.)$$

где T – время воздействия в течение рабочего дня, ч.

Для частот 0,06-3,0 МГц: $E_{пд} = 500$ В/м, $ЭН_{E_{пд}} = 20\,000$ (В/м) 2 ·ч

Для частот 3,0–30 МГц: $E_{пд} = 300$ В/м, $ЭН_{E_{пд}} = 7000$ (В/м) 2 ·ч

Для частот 30-300 МГц: $E_{пд} = 80$ В/м, $ЭН_{E_{пд}} = 800$ (В/м) 2 ·ч

Предельно допустимая напряженность магнитного поля в диапазоне частот 0,06 – 3 МГц в соответствии с названным выше ГОСТом должна

составлять $H_{ПД} = 50$ А/м. Между этой характеристикой и предельно допустимой энергетической нагрузкой за рабочий день [$E H_{ПД}$, (А/м) $^2 \cdot$ ч] существует следующая зависимость:

$$H_{ПД} = \sqrt{\frac{E H_{ПД}}{T}}, \quad (4.4.)$$

где T – время воздействия, ч (величина $E H_{ПД}$ не должна превышать 200 А/м 2).

Рассмотрим *основные методы защиты от электромагнитных излучений*. К ним следует отнести рациональное размещение излучающих и облучаемых объектов, исключая или ослабляющее воздействие излучения на персонал; ограничение места и времени нахождения работающих в электромагнитном поле; защита расстоянием, т. е. удаление рабочего места от источника электромагнитных излучений; уменьшение мощности источника излучений; использование поглощающих или отражающих экранов; применение средств индивидуальной защиты и некоторые др.

Из перечисленных выше методов защиты чаще всего применяют экранирование или рабочих мест, или непосредственно источника излучения. Различают отражающие и поглощающие экраны. Первые изготавливают из материалов с низким электросопротивлением, чаще всего из металлов или их сплавов (меди, латуни, алюминия и его сплавов, стали). Весьма эффективно и экономично использовать не сплошные экраны, а изготовленные из проволочной сетки или из тонкой (толщиной 0,01–0,05 мм) алюминиевой, латунной или цинковой фольги. Хорошей экранирующей способностью обладают токопроводящие краски (в качестве токопроводящих элементов используют коллоидное серебро, порошок графит, сажу и др.), а также металлические покрытия, нанесенные на поверхность защитного материала. Экраны должны заземляться.

Защитные действия таких экранов заключаются в следующем. Под действием электромагнитного поля в материале экрана возникают вихревые

токи (токи Фуко), которые наводят в нем вторичное поле. Амплитуда наведенного поля приблизительно равна амплитуде экранируемого поля, а фазы этих полей противоположны. Поэтому результирующее поле, возникающее в результате суперпозиции (сложения) двух рассмотренных полей, быстро затухает в материале экрана, проникая в него на малую глубину.

Эффективность действия экрана, или эффективность экранирования (\mathcal{E}), может быть рассчитана по формуле:

$$\mathcal{E} = \frac{I_0}{I}, \quad (4.5.)$$

Где I_0 – плотность потока энергии в данной точке при отсутствии экрана Вт/м²; I – плотность потока энергии в той же точке при наличии экрана, Вт/м²; или выражена в децибелах:

$$\mathcal{E} = 10 \lg \frac{I_0}{I}, \text{ дБ}. \quad (4.6.)$$

Например, замкнутый экран, сваренный из листовой стали непрерывным швом, имеет эффективность экранирования в диапазоне частот 0,15–10 000 МГц примерно 100 дБ.

Другой вид экранов – поглощающие. Их действие сводится к поглощению электромагнитных волн. Эти экраны изготавливаются в виде эластичных и жестких пенопластов, резиновых коврик, листов поролон или волокнистой древесины, обработанной специальным составом, а также из ферромагнитных пластин. Отраженная мощность излучения от этих экранов не превышает 4%. Например, радиопоглощающий материал «Луч», изготовленный из древесных волокон, в диапазоне длин волн излучения 0,15–1,5 м имеет отраженную мощность 1–3%.

Существуют и другие типы экранов, например, многослойные.

Экранами могут защищаться оконные проемы и стены зданий и сооружений, находящихся под воздействием электромагнитного излучения (ЭМИ). Строительные конструкции (стены, перекрытия зданий), а также

отделочные материалы (краски и т.д.) могут либо поглощать, либо отражать электромагнитные волны.

Для защиты от электрических полей промышленной частоты, возникающих вдоль линий высоковольтных электропередач (ЛЭП), необходимо увеличивать высоту подвеса проводов линий, уменьшать расстояние между ними, создавать санитарно-защитные зоны вдоль трассы ЛЭП на населенной территории. В этих зонах ограничивается длительность работ, а также заземляются машины и оборудование.

Особым видом электромагнитного излучения является лазерное излучение, которое генерируется в специальных устройствах, называемых оптическими квантовыми генераторами или лазерами. Эти устройства широко применяются в различных областях науки и техники, в том числе для обработки различных материалов (получение отверстий, резка и т.д.), в медицине (проведение различных операций), в системах связи для передачи сигналов по лазерному лучу, для измерения расстояний, для получения объемных изображений предметов – голограмм и в ряде других областей.

Рубиновые лазеры излучают в оптической части спектра. Длительность импульсов составляет от нескольких миллисекунд (мс) до сотен наносекунд (нс). Энергия одного импульса может достигать сотен джоулей при мощности в сотни мегаватт ($1\text{МВт} = 10^6\text{Вт}$). В настоящее время разработан ряд оптических квантовых генераторов, использующих различные оптические среды (фтористый кальций, вольфрамат кальция, различные газы и др.). Эти лазеры могут работать как в импульсном, так и в непрерывном режимах.

Лазерное излучение – электромагнитное излучение, генерируемое в диапазоне волн 0,2–1000 мкм. Этот диапазон делится на следующие области спектра в соответствии с биологическим действием лазерного луча: 0,2–0,4 мкм – ультрафиолетовая область, 0,4–0,75 – видимая, 0,75–1,4 мкм – ближняя инфракрасная, свыше 1,4 мкм – дальняя инфракрасная область. Наиболее

часто используют в технике лазеры с длинами волн, мкм: 0,34, 0,49-0,51, 0,53, 0,694, 1,06 и 10,6.

Воздействие излучения лазера на организм человека до конца не изучено. При работе лазерных установок на организм человека могут воздействовать следующие опасные и вредные производственные факторы: мощное световое излучение от ламп накачки, ионизирующее излучение, высокочастотные и сверхвысокочастотные электромагнитные поля, инфракрасное излучение, шум, вибрация, возникающие при работе лазерных установок, и др.

При воздействии лазерного излучения на организм человека возникают различные биологические эффекты, которые зависят от энергетических и временных параметров излучения и в первую очередь от энергетической экспозиции в импульсе, длины волны и времени воздействия лазерного излучения, вида облучаемой ткани человеческого организма и ряда других факторов. Энергетическая экспозиция может быть рассчитана по формуле:

$$H = E_e t, \quad (4.7.)$$

где H - энергетическая экспозиция;

E_e - энергетическая освещенность (отношение энергии излучения, падающей на рассматриваемый участок поверхности, к его площади); t - время воздействия лазерного излучения.

Таким образом, с физической точки зрения энергетическая экспозиция – это отношение энергии излучения, падающей на рассматриваемый участок поверхности, к площади этого участка, умноженное на длительность облучения.

Различают первичные и вторичные биологические эффекты, возникающие под действием лазерного излучения. Первичные изменения происходят в тканях человека непосредственно под действием излучения (ожоги, кровоизлияния и т.д.), а вторичные (побочные явления) вызываются различными нарушениями в человеческом организме, развившимися вследствие облучения.

Наиболее чувствителен к воздействию лазерного излучения глаз человека. Воздействие на него лазерного излучения может привести к ожогам сетчатки и даже к потере зрения. Опасно попадание лазерного луча и на кожу человека, в результате чего могут возникнуть ожоги различной степени тяжести и даже обугливание кожи. Лазерные лучи высокой интенсивности могут вызвать не только повреждения кожи, но и поражение различных внутренних тканей и органов человека, что выражается в виде кровоизлияний, отеков, а также свертывания или распада крови.

Основным нормируемым параметром является *энергетическая экспозиция* (H , Дж/см²) облучаемых тканей за определенное время воздействия лазерного излучения. Если нормируемая величина H (предельно допустимый уровень) не превышена, то у работающих под воздействием лазерного излучения не будут вызываться первичные и вторичные биологические эффекты. Величина предельной энергетической экспозиции зависит от длины волны лазерного излучения и длительности его воздействия на работающего. Пример такого нормирования для лазерного излучения с длиной волны от 0,2 до 0,4 мкм представлен в табл. 18.3. Общее время облучения в этом случае составляет рабочий день. Энергетическая экспозиция нормируется на роговице глаза и коже.

Предельно допустимые уровни лазерного излучения (энергетической экспозиции) относятся к длинам волн от 0,2 до 20 мкм. Кроме того, в Санитарных нормах для длин волн от 0,4 до 1,4 мкм установлены предельно допустимые уровни энергетической экспозиции сетчатки глаза. Для видимой части спектра (0,4–0,75 мкм), кроме рассмотренных характеристик, дополнительно нормируется энергия излучения (Q , Дж) на сетчатке глаза.

К основным *коллективным средствам защиты от лазерного излучения* относятся применение защитных экранов и кожухов; использование телевизионных систем наблюдения за ходом технологического процесса с использованием лазера, а также систем блокировки и сигнализации; ограждение лазерно-опасной зоны, размеры которой определяют или

расчетным, или экспериментальным путем. Следует защищаться не только от прямого излучения лазера, но и от рассеянного и отраженного излучений.

Напряженность постоянного магнитного поля может быть измерена отечественными приборами Ш1-8 или Ф-4355. Магнитное поле промышленной частоты при напряженности до 15 кА/м измеряют отечественным прибором Г-79, а в диапазоне частот 0,01–30 МГц – приборами ПЗ-15, ПЗ-16и ПЗ-17. Три последних прибора могут быть рекомендованы и для измерения напряженности электрического поля в диапазоне частот 0,01–300 МГц. Для измерения плотности потока энергии электромагнитного поля применяют отечественные приборы ПЗ-9, ПЗ-18, ПЗ-19 и ПЗ-20, которые перекрывают частотный диапазон 0,3–400 ГГц.

Для измерения характеристик лазерного излучения применяются дозиметры типа ИЛД-2М и ЛДМ-2. Первый обеспечивает измерение параметров лазерного излучения в спектральных диапазонах 0,49–1,15 и 2–11 мкм, он дает прямые показания измеряемых параметров при работе на длинах волн 0,53; 0,63; 0,69; 1,06 и 10,6 мкм. На остальных длинах волн (0,49– 1,15 мкм) дозиметр обеспечивает косвенные измерения. Прибор ЛДМ-2 предназначен для определения параметров лазерного излучения в спектральных диапазонах 0,49–1,15 и 2–11 мкм. Прямые измерения этот дозиметр осуществляет на длинах волн 0,53; 0,63; 0,69; 0,91; 1,06 и 10,6 мкм.

Для *индивидуальной защиты от электромагнитного излучения* применяют специальные комбинезоны и халаты, изготовленные из металлизированной ткани (экранируют электромагнитные поля), а для защиты от действия лазера обслуживающий персонал должен работать в технологических халатах, изготовленных из хлопчатобумажной или бязевой ткани светло-зеленого или голубого цвета.

Заключение

Во время написания дипломной квалификационной работы был рассмотрен большой объём теоретической и документальной информации, суть которой изложена в её начальной или обзорной части.

В данной работе проведен обзор основных концепций вэйвлет-анализа, запрограммированы алгоритмы разложения и восстановления, исследованы особенности вэйвлет-разложений.

Реализован способ разбиения сигнала на сегменты основанный на методе динамического программирования. Предложена схема сжатия сегмента сигнала, особенность которой состоит в разложении массива коэффициентов на массив мантисс и массив экспонент.

Также были рассмотрены методы звукозаписи, теория цифрового преобразования звука, математические алгоритмы стандартов Vorbis и MPEG и различные стандарты сжатия и кодирования звука. Среди которых явное лидирующее место имеют стандарты Vorbis и MPEG. Был сделан сравнительный анализ стандартов цифрового сжатия звука. Были изучены особенности построения оцифрованного звука, особенности стандарта MPEG и его иерархия.

При изучении различных стандартов звуковой информации были сделаны выводы что формат MPEG Layer III имеет высокую оценку по качеству сжатия и воспроизведения и по размещению его на носители. Также стоит отметить другие форматы цифрового «звука» такие как Ogg и Wav имеют специфическое предназначение и используются для озвучивания компьютерных приложений и игр. Формат MPEG Layer III популярный как

mp3 имеет характер для воспроизведения на mp3 плеерах, сотовых телефонах, различных планшетах. Формат mp3 признан как наиболее подходящим не только для пользовательского слушания, но и для воспроизведения его на радиостанциях и в телевидении. Данный формат легко импортируется в различные форматы видео, а также удобен для озвучивания при монтаже видео файлов. Практически все операционные системы персональных компьютеров имеют кодеки для воспроизведения и редактирования MP3 файлов.

По словам экспертов, формат MPEG Layer III разработан так совершенно, что ближайшие десять лет не выйдут форматы звуковых файлов которые могли бы превзойти его. Поэтому изучение MPEG кодирования аудио и видео считается актуальным, и его дальнейший рост в аудио и видео индустрии обусловлен его преимуществами. Но немаловажен и недостатки стандарта MP3, на низких битрейтах звуковые файлы данного формата теряют качество. Но учесть «объем» данных файлов можно утверждать что формат MPEG Layer III имеет лидирующее место среди аналогичных форматов.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ:

1. Нагорный А. Vorbis против всех? Или какой кодек выбрать для сжатия аудио // портал Hardvision, 2010
(http://www.hardvision.ru/?dir=soft&doc=ogg_vorbis).
2. Alexander C., Strauss N., The Ogg Vorbis CODEC project, Xiph.Org., 2003 (<http://www.xiph.org/ogg/vorbis/>).
3. Hardle W., Kerkyacharian G., Picard D., Tsybakov. A. Wavelets, Approximation, and Statistical Applications (Lecture Notes in Statistics, Vol 129). New York: Springer-Verlag, 1997.
4. Алексеев К.А. Теория и практика шумоподавления в задаче обработки сейсмоакустических сигналов // Обработка сигналов и изображений. Wavelet Toolbox, Консультационный центр Matlab, 2010. (<http://matlab.exponenta.ru/wavelet/book5/index.php>).
5. Алексеев К.А. Вейвлеты, аппроксимация и статистические приложения // Обработка сигналов и изображений. Wavelet Toolbox, Консультационный центр Matlab, 2010.
(<http://matlab.exponenta.ru/wavelet/book6/index.php>).
6. Таха Х. Введение в исследование операций. М.: изд. дом “Вильямс”, 2001.
7. Чуи К. Введение в вейвлеты М. Мир, 2001.
8. Войнаровский М. Психологика // Быстрое преобразование Фурье, 2002-2003 (<http://psi-logic.narod.ru/fft/fft.htm>).
9. Кантор И. Алгоритмы и методы
(<http://algotlist.manual.ru/compress/standard>).
10. Смирнов М. Введение в PPM
(http://www.compression.ru/download/articles/ppm/smirnov_2000_ppm_faq.htm
1).
11. Александр Колганов, Системы мультимедиа сегодня // HARD&SOFT №4 апрель 1995г.

12. Антон Веснушкин, «Живое» видео на PC // HARD&SOFT №6 декабрь 1994г.
13. Андрей Борзенко, Программное обеспечение для мультимедиа // HARD&SOFT №2 февраль 1995г.
14. Роман Косячков, Властелины Пеллинора // Компьютерра №38 (316) сентябрь 1999г.
15. Сергей Бобровский, Стратегии // PC WEEK №21 июнь 2001г.
www.ixbt.ru, раздел SOFT.
16. Мультимедиа — синтез трех стихий. С. Новосельцев // Компьютер-Пресс №7 1991г.
17. Мультимедиа-ПК. В. Дьяконов // Домашний Компьютер №1 1996г.
18. Рош У.Л. Библия мультимедиа. Киев: ДИАСофт, 1998, 800 с.
19. Информационная культура: Кодирование информации. Информационные модели. 9–10-е классы. М.: Дрофа, 2000, 208 с.
20. Угринович Н.Д. Информатика и информационные технологии. Учебное пособие для 10–11-х классов. Углубленный курс. М.: Лаборатория Базовых Знаний, 2000, 440 с.
21. Кенцл Т. Форматы файлов Internet. СПб.: Питер, 1997, 320 с.
22. Симаненков Д. Тема номера. Компьютерра № 30–31, с. 20–33; № 32/1998, с. 20–28 (<http://www.computerra.ru/offline/1998/258> и <http://www.computerra.ru/offline/1998/260>).
23. Еремин Е.А. Представление звуковой информации в ЭВМ. Информатика 2004, № 45, с. 16–17.
24. Экология и безопасность жизнедеятельности: Учебное пособие для студентов ВУЗов / ред. Л. А. Муравий, 2002.
25. Белов С.В. Безопасность жизнедеятельности М.: Высшая школа. 2003.
26. Ёрматов Ғ.Ё., Исамухамедов Ё.У. Меҳнатни муҳофаза қилиш. Дарслик. Ўзбекистан нашриёти. Тошкент 2002.